Genetic and Genomic Characterization of Composite Beef Breeds with an Insight into the Canadian 'Hays Converter'

by

Razie Khorshidi

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Animal Science

Department of Agricultural, Food and Nutritional Science

University of Alberta

Abstract

Crossbreeding is a widely used strategy in animal breeding to benefit from complementarity and heterosis. Composite breeds are one of the products of crossbreeding in which animals maintain a stabilized combination of genetic characteristics of two or more pure breeds. Hays Converter (HC), the first registered Canadian beef breed, developed by Harry Hays in the late 1950's is defined as a composite of beef and dairy breeds. The breed is well adapted to the Western Canadian climate, calves reach market weight early, convert feed to gain efficiently, and possess qualified carcass.

Before dealing with HC, in order to understand the impact of crossbreeding on livestock production, the first study investigated key concepts of developing composite cattle under (sub) tropical environments to evaluate their efficiency in productivity. A simulation study was designed with an interest in the indigenous Afrikaner cattle, a specialized dam line known for limited calving difficulties and improved performance of progeny when crossing with exotic terminal sires like Charolais. The results demonstrated the simulated composite dams were more fit producing 7.8% more calves and their progeny performance was improved by reducing feed intake (−24.4%) and increasing meat production (+11.7%).

Following that, the work was focused on the HC to evaluate their current genetic characterization. Since the breeding objective for HC was to create a beef breed that excelled in growth, selection mainly emphasized weaning and yearling weights. However, there was no selection index by genetic values from its inception. Therefore, the second study was to evaluate the genetic parameters and trends in birth, weaning and yearling weights of HC through comparison of different multiple trait models (MTM) with a random regression model (RRM). Also, in MTM scenarios, both adjusted and unadjusted data were examined besides considering

contemporary groups (GC) as fixed or random. The results indicated similar changes along the growth trajectory for estimates of variance components, heritability and genetic correlations from the two approaches and fixed CG were preferred. Although there was a considerable reduction in genetic trends from 2004 to 2008 due to weak sire selection, trends generally increased through 2016.

HC is a mixture of primarily Hereford (HER), Holstein (HOL) and Brown Swiss (BSW) with a later introgression of Angus (AN). Therefore, the third study was conducted to estimate the HC genomic breed composition based on the entirety of its genome and each chromosomal segments of equal intervals. Admixture and regression methods were used with both 6K and 50K SNP panels. The results indicated that the regression method generated similar estimates with both SNP datasets. However, the admixture method analyzed HC composition differently, depending on SNP panel size and mainly due to the ratio of number of admixed to unmixed individuals. Avoiding these constraints resulted in uniformity between admixture and regression estimates. Use of the regression-6K was recommended as it is more cost effective (for genotyping) and avoids issues that arise with the admixture method. Overall, HC genomic composition was predicted in percentage as 8±0.2 AN, 51±0.2 HER, 15±0.1 BSW and 26±0.1 HOL. Diversity of breed proportions in chromosomal segments relative to whole genome was used to imply signatures of selection from HC founders.

The founder breeds were chosen for inclusion in HC to capture benefits from fertility and carcass traits (HER), milk production and growth potential (HOL) and strong feet and udders (BSW). The presence of AN was due to occasional usage to control calving difficulty in first calf heifers and it was not considered to be a HC founder breed. Except for body weight, phenotypic data characterizing the HC breed is scarce. Therefore, the aim of the fourth study was to explore

indicators of selection across the genome using Fst and runs of homozygosity (ROH). Twenty eight chromosomal segments showing over-representation of ancestral breeds relative to the entire genome were identified using Grubbs' test. Only three were detected to be under positive trend for ROH length from 1973 to 2015. Several numbers of HOL, BSW or AN origin fragments were found in these chromosomal segments through an Fst ranking approach. They overlapped to QTLs associated with traits of body weight and milk production. However, the effects of such trends were not meaningful because, similar to the remaining 25 chromosomal segments, they still showed a lower Fst with HER. Moreover, they were mainly as a result of reduction in herd size after the year 2000 and not selection. As HER comprised the highest percentage across the composite genome, too, these findings might imply its sustainable role for weight traits, body features, milk production, fertility and carcass standards. The use of Fst, ROH, the analysis of breed proportions and the AnimalQTL database helped to interpret signatures of selection of breeds contributing to composite animals.

Preface

This thesis is the original PhD work by Ms Razie Khorshidi. The whole thesis has been drafted by her and finalized with inputs from Dr M.D. MacNeil of Delta G, Miles City, Montana, United States, Dr G. Plastow and Dr E.C. Akanno, both of Livestock Gentec, Department of Agricultural, Food and Nutritional Science, Faculty of Agricultural, Life and Environmental Sciences, University of Alberta, Edmonton, Canada. The animals used in this study were part of a commercial enterprise (Red Bow Ranching Ltd, Calgary, AB) using industry standard production practices which were generally consistent with the guidelines of the Canadian Council on Animal Care (CCAC).

Chapter two of this thesis has been published as Khorshidi Razie, Michael D. MacNeil, John J. Crowley, Michiel M. Scholtz, Annette Theunissen, and Graham S. Plastow. "Evaluating breed complementarity and sexed semen with maternal use of Afrikaner germplasm." Agricultural Sciences 8, no. 07 (2017): 507. Also, chapter three of this thesis has been published as Khorshidi Razie, Michael D. MacNeil, Daniel P. Hays, Mohammed K. Abo-Ismail, John J. Crowley, Everestus C. Akanno, Zhiquan Wang, and Graham Plastow. "Estimation of genetic parameters and trends for growth traits in Hays Converter cattle using multiple-trait and random regression models." Livestock Science 241 (2020): 104245. In both studies, she performed all the data analyses, interpreted the results, and drafted the manuscript. Michael D. MacNeil had the key role in designing and supervising the two studies. All authors read, commented and approved the final manuscript.

Dedication

این پایان نامه با سرافرازی پیشکش می گردد:

به پدر، مادر و خواهر گرامی تر از جانم،

به همگی بزرگان گرامی دانش، ادب و هنر پارسی از گذشته تاکنون که دانش خویش را وام دار آنانم به ویژه

زکریای رازی، ابوالقاسم فردوسی، محمود حسابی، احسان یارشاطر، محمد علی ندوشن و ایران درودی

و به همگی هم میهنانی که همواره برای سربلندی ایران کوشیدند، می کوشند و خواهند کوشید.

راضیه خورشیدی

شهریورماه ۱۴۰۱ هجری شمسی

Acknowledgements

I would like to express my sincere thanks to Dr. Michael MacNeil and Dr. Graham Plastow for providing their expertise and guidance and helping me gain my knowledge and skills necessary to be successful in my projects throughout my PhD in Animal Science. I am also thankful of Dr. Everestus Akanno for being a part of my committee and providing valuable feedbacks on all aspects of my thesis.

I would like to thank Dan Hays and his family for their support in provision of HC data required to do this work. I would also like to thank the members of Livestock Gentec for being always available to share their knowledge, advice, kindness and friendship. This work was supported in part by funding from Red Bow Ranching Ltd.

Finally, and most importantly, I am immensely thankful of my dear parents and my lovely identical twin sister for their endless love, support and encouragement.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AIC	Akaike information criterion
AN	Angus
BIC	Bayesian information criterion
BLUP	Best Linear Unbiased Prediction
BSW	Brown Swiss
BW	Birth weight
CCAC	Canadian Council on Animal Care
CG	Contemporary groups
CV	Cross validation
EBV	Estimated breeding values
HC	Hays Converter
HER	Hereford
HOL	Holstein
IBD	Identical by descent
LD	Linkage Disequilibrium
MP	Market progeny

MTM	Multiple trait models
QTL	Quantitative Trait Loci
REML	Restricted maximum likelihood method
RF	Replacement females
RH	Retained heterozygosity
ROH	Runs of homozygosity
RRM	Random regression model
SD	Selection differential
TFC	Total feed consumption
TP	Total productivity
WW	Weaning weight
YW	Yearling weight

Chapter 1. General Introduction

The content presented in this chapter is a summary of literature reviews regarding the importance of crossbreeding and development of composite breeds, the effect of genetic improvement on production efficiency and the role of genomic tools in estimating breed composition, inbreeding and evaluation of signatures of selection. It is then followed by a brief introduction of Hays Converter, its current situation and the objectives of this thesis.

1.1 The importance of crossbreeding in beef production

Crossbreeding is one of the most powerful tools to improve productivity and efficiency in a herd. It allows producers to take advantage of heterosis (hybrid vigor), breed complementarity and biological breed type differences so that they will be able to match their cattle to specific production resources (Cartwright, 1970; Dickerson, 1973; Gregory *et al*. 1982).

Heterosis is defined as the superior performance of a crossbred offspring over the average of its parents and shows the greatest amount when two animal parents with completely different breed backgrounds are mated to each other (Gregory *et al*. 1965). Appearance of hybrid vigor can be found through a variety of traits such as increased survivability or growth in crossbred calves or higher reproduction rate in crossbred cows. Implementation of the type of crossbreeding system along with the number of breeds incorporated into that system affect the amount of heterosis that is maintained in a herd (Cundiff and Gregory 1999).

Although all breeds may demonstrate some superiority in several economically important traits, none of them are excellent in all traits (Gregory and Cundiff 1980; Weaber 2008). In fact, breed complementarity helps breeders match the genetic potential of a hybrid offspring for economic traits such as growth rate and carcass composition with climate, feed resources,

fertility, disease resistance and market preferences of a specific production environment. This means that the strengths of one breed can complement or mask the weaknesses of the other breed when mating together (Cartwright, 1970; Cartwright, 1974; Gregory and Cundiff 1980). Therefore, to generate hybrid progeny with superior performance relative to their parents in a specific production environment, breeders should design a crossbreeding system that capitalizes on those traits that are appropriate for that environment and could be brought to the mix by parents (Gosey 1991). It is possible to get negative results from breed complementarity if crossbreeding programs are poorly conceived and designed. For example, if a large paternal sire breed with high milk potential is mated to small framed heifers on a limited forage system, this may lead to dystocia and also replacement animals will not be compatible with the breeder's resources (Garcia *et al*. 2019).

Breeds of different biological types exhibit different levels of various production characteristics (Koger *et al*. 1975). For example, crossing a Hereford or Angus bull with a Brahman cow will generate medium framed, moderate milking F1 females that are more resistant to heat and parasites than their *Bos taurus* sire and will produce calves with higher carcass quality than their *Bos indicus* dam (Garcia *et al*. 2019).

In any crossbreeding system, retaining high levels of hybrid vigor for multiple generations will be as important as how to maximize it (Cundiff and Gregory 1999). In other words, in planning a crossbreeding program, it is important to resolve problematic issues about required number of breeding pastures, how to generate replacement heifers, optimum size of the herd, source of breeds, feed resources requirements, labor availability and potential use of artificial insemination (Gosey 1991). Accordingly, compared to rotational crossbreeding systems, management of composite populations is simple, especially when producers have

limitations on herd size and number of breeding pastures. Composite breeds are developed based on mating among crossbred animals with a defined proportion of two or more breeds in their backgrounds. Although their development is complex, after that the herd can be managed as a purebred population. Thus, breeders can avoid management problems associated with small herd size and fluctuations in additive genetic composition between generations in rotational crossbreeding systems (Gregory and Cundiff 1980). Moreover, they will require only one breeding pasture, replacement females will be generated within the system and there is no need for identification of females by sire (Gosey 1991). Composite populations can maintain a high level of heterosis without further crossbreeding. This depends on maintaining an adequate size of the composite population to select for replacements and new sires. Providing this condition, inbreeding will be avoided and the initial advantage of increased heterozygosity will not be dissipated through early re-inbreeding of composite animals (Gregory and Cundiff 1980). Although composite populations do not sustain as high level of heterosis as rotational systems, they allow for effective use of additive breed effects and complementarity between breeds in addition to heterosis in order to achieve increased productivity (Cartwright 1970; Gregory and cundiff 1980; Garcia et al. 2019).

The retention of heterozygosity favors the inclusion of an optimum number of breeds in composite populations (Gregory and Cundiff 1980; Kinghorn 1982; MacNeil 1987). However, the increased retained heterozygosity resulting from additional contributing breeds must be balanced against possible loss of average additive genetic merit from the inclusion of additional breeds (Gregory *et al*. 1982; Kinghorn 1982; MacNeil 1987). Thus, it is required to determine linearity of association of loss of heterosis with loss of heterozygosity for successful composite breeding programs (Gregory *et al*. 1982). Also, the additive genetic variation needs to be

determined in composite populations relative to the contributed parental breeds, especially for fitness-related traits (Gregory and Cundiff 1980). To find practical solutions, there have been methods proposed by Kinghorn and MacNeil for obtaining the optimal proportions of breeds in the formation of composite populations (Kinghorn 1982; MacNeil 1987; Kinghorn *et al*. 1989).

Besides, systematic crossbreeding systems are not always possible or acceptable in the (sub) tropical areas. Therefore, formation of composites through crossing unrelated breeds (i.e. *Bos taurus* × *Bos indicus*) with desired production features may offer a quick and simple way of utilizing between-breed genetic variation to increase productivity. However, there should be some consideration borne in mind when breeding cattle in such environments. Namely, for optimal beef production, a percentage of 25% to 75% of the adapted breed genes is required. This will depend on the environment severity and also the level of challenge from stressors such as parasites, poor nutrition, high heat and humidity and endemic disease (Burrow 2006). Adapted genes can be derived from *Bos indicus* and/or tropically adapted taurine breeds which the latter has been best used in breeding programs to improve fertility, temperament, and carcass and beef quality without reducing resistance to stressors and adaptation to (sub) tropical environments (Frisch and O'Neill 1998a,b; Prayaga 2003a,b).

For (sub) tropical areas, real limitations exist due to financial and/or environmental constraints that limit experimental evaluation of breeds and crossbreeding systems. Therefore, systems analysis could provide a means for simultaneous consideration of many more factors that could feasibly be included in one experiment and may be a realistic method of developing or testing recommendations to improve production efficiency (Long *et al*. 1975; Cardoso and Templeman 2004; Theunissen *et al*. 2013).

Finally, the success of any crossbreeding system depends on obtaining replacement females efficiently. Use of sexed semen may contribute to increased productivity through allocating fewer females to generation of specialized maternal daughters and devoting more of the remainder to production of highly efficient crossbred sons (Hohenboken 1999; Hall and Glaze 2014).

1.2 From genetic improvement to implementation of genomic selection in beef cattle

Genetic improvement has been applied as a tool to improve production efficiency in beef cattle. Breeders attempt to identify genetically superior animals for use as parents of the next generation. Systems of genetic evaluation help them make effective decisions regarding the choice of replacement animals. The intuitive desire to increase the accuracy of prediction for traits of interest has resulted in milestone improvements through intersections among several disciplines such as animal breeding and genetics, numerical methods, computer science, economics and statistics (Golden *et al*. 2009).

Since 1973, mixed model procedures have become widely adopted as the standard statistical approach to predict additive genetic merit in livestock. Best Linear Unbiased Prediction (BLUP) of breeding values represents a substantial improvement in accuracy of estimated breeding values especially by considering records made in different contemporary groups and facilitating use of all available data (Henderson 1973; Henderson 1975). These procedures have mostly been applied for traits with large scale recording that are supported with genetic evaluation programs implementing the animal model. They have substantially resulted in successful documentation of genetic progress from past selection through estimating genetic trends. In fact, these traits are primarily the ones that are beneficial for cattle management and obtained relatively easily at low costs such as birth weight, weaning weight, yearling weight and

mature weight (Miller 2010). Two-trait and multi-trait models are widely used to predict breeding values for growth traits in beef cattle (Caetano *et al*. 2013). In such models, animals are typically evaluated at certain standard ages for which the weight records have to be adjusted and/or discarded if they are out of the accepted range of age. However, as both these editing procedures may lower the accuracy of genetic evaluation, random regression models have recently been under consideration for analyzing growth traits as repeated measures along a continuum. Moreover, they are able to estimate covariance between any two ages across the growth trajectory and calculate breeding values for the complete growth curve with higher accuracies (Tier and Meyer 2004; Baldi *et al*. 2010).

Beyond the traits which are convenient to measure in either progeny tests or routine herd recording, there are some other important traits with considerable genetic variation. These traits may be difficult or expensive to measure on an industry-wide basis. Two well-known examples are feed efficiency and beef tenderness (Schemkel *et al*. 2004; Zwambag 2007).

From the last two decades, there has been increasing growth in the development of DNA-based technologies. These genomic tools facilitate improved rates of genetic progress for several traits through increased accuracy of evaluations and in some cases reduced generation intervals (Wiggans *et al*. 2017). These technologies started with a few markers for beef quality traits and were advanced to SNP panels that contain orders of magnitude more markers. The SNP panel technology and sequencing of the bovine genome have resulted in dramatic reductions in costs of genotyping and evaluating many more traits (Elsik *et al*. 2009; Matukumalli *et al*. 2009).

Implementation of these technologies such as Illumina Bovine SNP50 Beadchip has enabled selection based on genomic regions influencing the traits, termed as Quantitative Trait Loci (QTL). In some instances the causative mutation in a gene (Gene Assisted Selection, GAS)

is used or the prediction of breeding values may use genetic markers that are in linked with the QTL (Marker Assisted Selection, MAS) (Miller, 2010). This linkage, known as Linkage Disequilibrium (LD), occurs when two loci on the genome are physically close on a chromosome that they segregate together during meiosis and the segment is passed from parent to offspring. Hence, selecting the linked marker will result in a correlated response in the trait of interest (Goddard and Meuwissen 2005). However, the MAS approach was effective for traits with simple genetic determinism while provided unsatisfactory results for more complex conditions. There were two main reasons for this low efficiency; as a small number of significant markers were used in MAS, they could only explain limited and always overestimated part of the genetic variance and secondly, at the level of population, it was possible to get low association or LD between markers and QTL (Biochard *et al*. 2016).

Contrary to MAS, the availability of large and cheap SNP panels has allowed the establishment of "genomic selection" strategies which are based on the relationship between all SNPs of the genome and important traits. In fact, genomic selection enables to estimate breeding value of an animal through summation of estimated genetic effects for all the markers spanning entire the genome (Meuwissen *et al*. 2001). In this approach, marker effects are estimated within a reference population which contains a large group of animals with both phenotypes and marker genotypes. Then, these estimates are applied to candidates for selection with only marker genotype information. Both the number of genotyped individuals and markers available can considerably determine the effectiveness of this approach (Meuwissen *et al*. 2016).

The overall purpose of genomic selection is to enhance genetic improvement through 1. increased accuracy of breeding value estimates, 2. decreased generation interval and 3. increased selection intensity; the first depends on the size of the reference population, heritability of the

trait and the structure of the genome based on the LD between SNP and causal variants and the latter is possible when a large number of candidates are genotyped than can be phenotyped (Miller 2010; Biochard *et al*. 2016).

In dairy cattle, genomic selection has produced promising results (Hayes *et al*. 2009). Accuracy of prediction has exceeded 0.8 for production traits and 0.7 for fertility, longevity somatic cell count and other traits (Lund *et al*. 2011; Wiggans *et al*. 2011). This success results from there being large reference populations for each breed that contain progeny tested bulls with highly accurate estimated breeding values based on average performances of many daughters. With such reference populations, genomic breeding values are felt to be sufficiently accurate to replace progeny testing and allow dissemination of semen of young bulls with only a genomic evaluation. This has substantially simplified the process of selection and decreased its cost due to a large reduction in generation interval (Colleau *et al*. 2015; Wiggans *et al*. 2017). Moreover, genotyping is now sufficiently inexpensive to be used to rank female candidates for selection and control inbreeding through the use of planned matings (Pryce and Hayes 2012; Kemper *et al*. 2015; Upperman *et al*. 2019).

Although genomic selection is now employed in some beef breeds like Angus on a large scale, in general, the accuracies of genomic predictions have been lower than dairy breeds and in a range from 0.3 to 0.7 (Van Eenennaam *et al*. 2014). This is because large reference populations, containing animals with very high-accuracy evaluations, are lacking for most breeds. In addition, individuals in the target population may be less closely related to the reference animals than dairy breeds (Meuwissen *et al*. 2016).

1.3 Using genomics to estimate breed proportions in cross-bred animals

Many beef and some dairy production systems use crossbreeding to capture benefit from breed complementarity and heterosis (Gregory *et al*. 1982; Rincon *et al*. 1982). Features like coat color and body conformation phenotypes are often employed to identify cross-bred animals. However, it is not possible to distinguish the breed proportions without complete pedigree information (Crum *et al*. 2019). Inaccurate or incomplete pedigrees compromise their reliability in estimating breed composition. Actual breed composition may also be different from the pedigree-based expectation due to Mendelian sampling during gametogenesis (Kuehn *et al*. 2011; VanRaden and Cooper 2015). Using genomic data, as the roles of SNP markers become more pervasive in animal breeding, can lead to more accurate determination of breed proportions. They are capable of measuring realized parental contributions at the genomic level and therefore, correcting pedigree errors (VanRaden and Cooper 2015; He et al. 2018). Precise knowledge of breed composition in crossbred animals can facilitate evaluating their adaptability to a given production environment (Kuehn et al. 2011). Genomic breed composition can also be utilized for a variety of downstream applications, such as independent authentication of breed in breed-labeled beef products, estimating genomic breeding values for crossbred animals, designing crossbreeding programs to exploit heterosis and breed complementarity and estimation of retained heterozygosity in advanced generation composite animals (Akanno *et al*. 2017; Gobena *et al*. 2018; Crum *et al*. 2019).

1.4 Utilizing genomics to estimate inbreeding for populations with incomplete pedigree

Inbreeding, as a consequence of mating among closely related individuals, is usually unavoidable in small size populations. It can lead to a reduction in genetic variation and an increase in the frequency of homozygous individuals. Moreover, inbreeding can result in less

desirable performance in fertility and production traits which will ultimately affect the profitability and sustainability of farms (Allaire and Henderson 1965; Smith et al. 1998; González-Recio *et al*. 2007).

Inbreeding coefficients have usually been estimated from pedigree records. Therefore, in order to restrict inbreeding in mating strategies, breeders need to have access to an accurate, deep and complete pedigree file. However, this requirement may not always be satisfied, especially for local breeds. The recent availability of high density SNP panels has opened new opportunities to represent an improved estimation of inbreeding for livestock populations even when there is no pedigree provided (Marras *et al*. 2015). Typically, there are two approaches to calculate genomic inbreeding coefficients using high density SNP genotypes. The first is using a genomic relationship matrix (GRM) to examine the identical by state (IBS) information marker by marker (VanRaden *et al*. 2011). The second relies on runs of homozygosity (ROH), DNA segments that harbour contiguous stretches of homozygous loci. These segments are expected within an animal when both identical haplotypes transmitted from parents to offspring and hence, share a common ancestor. This means they should be correlated to the inbreeding coefficient definition, the proportion of genome that is identical by descent (IBD) (Keller et al. 2011; Kim et al. 2013).

Measuring inbreeding through ROH seems to be particularly appealing in cattle production. This is because the extent and frequency of these segments may provide useful information on the history of practices heavily featured in some breeds. Intense selection of sires, artificial insemination and embryo transfer have actually led to a reduction in effective population size and genetic diversity. As a result, they can affect the levels of homozygosity in cattle population (Purfield *et al*. 2012). To take some examples, ROH have been used to analyze

the history of population following recent selection (Purfield, et al. 2012), study the deleterious effects of inbreeding on traits important in farm profitability (Bjelland, et al. 2013) and control the increase of inbreeding in genome-assisted breeding schemes (Pryce, et al. 2012).

1.5 Utilizing genomics to evaluate signatures of selection in cattle

It is possible to test for the effects of DNA polymorphisms on phenotypes through ongoing developments of high throughput genotyping and bioinformatic tools. They provide a genomic profile of variation for all individuals of a population and help decipher the genetic basis of phenotypic diversity as a fundamental aim in genetics. Selection signature analysis, a new paradigm offered by population genomics, is a genome to phenotype approach in which population genomic data is statistically evaluated regardless of phenotype to identify possible targets of past selection (Qanbari and Simiaber 2014). Selection signatures are defined as genomic regions harbouring functionally important variants that have been under selection and resulted in leaving specific patterns of genomic sequence (Qanbari and Simiaber 2014).

Basically selection can act in three ways including; positive selection, which occurs when a newly arisen mutation increases in frequency due to selective advantage, or negative selection which tends to remove disadvantageous variants from the population and balancing selection in which polymorphism is favoured to maintain genetic variability (Gouveia *et al*. 2014). Various statistical approaches have been used to detect selection signatures such as Tajima's D-statistic (Tajima 1989), extended haplotype homozygosity (EHH) (Sabeti *et al*. 2002), integrated haplotype score (iHS) (Voight *et al*. 2006) and Fst (Wright, 1949). The latter statistic measures the degree of genetic differentiation among populations due to differences in allele frequencies. In fact, contrary to the preceding approaches, Fst is able to quantify genomic variation at each locus between populations rather than just within populations. Therefore, it can be used as the

evidence of local positive selection when showing high values while negative or balancing selection with low values (Zhao *et al*. 2015).

In cattle, employing artificial selection strategies to develop divergent dairy and beef breeds has imposed selection pressure on particular regions of genome that control milk or meat production and also other important traits such as reproduction, body formation or disease resistance (Zhao *et al*. 2015). For example, signatures of selection in beef cattle have been found in the centromeric region of BTA14 which control marbling and fatness traits (Bovine HapMap Consortium 2009; Veneroni *et al*. 2010). There are also at least three QTLs in BTA6 that have been under selection for milk traits in dairy breeds (Bovine HapMap Consortium 2009; Weikard et al. 2012). Selection signatures across and within beef and dairy breeds indicate that the genomic region around *DGAT1* (Diacylglycerol O-acyltransferase 1) on BTA 14, which has a major effect on milk fat percentage, has been under selection (Grisart et al. 2002; Hosokawa et al. 2012). Furthermore, the regions surrounding *SIGLEC-5* (Sialic Acid Binding Ig-Like Lectin 5) and *ZNF577* (Zinc Finger Protein 577) on BTA 18 have been found to be under selection and associated with Net Merit, conformation, longevity and calving ease in dairy cattle (Cole et al. 2009). As well, on BTA 14, *PLAG1* (pleiomorphic adenoma gene 1) has been implicated in the regulation of stature and weight in dairy and beef cattle (Utsunomiya *et al*. 2017).

1.6 Hays Converter origin, characteristics and current situation

The Hays Converter (HC) is a beef breed that was developed by the late Canadian Senator Harry Hays, a livestock producer, former Mayor of Calgary, Alberta and a Canadian Minister of Agriculture in the 1950s. The purpose of creation was to develop a breed that converted feed to lean meat as efficiently as possible in order to address a divide he believed

existed between breeders and the market needs and to focus mainly on performance rather than appearance or fads (Fleming *et al*. 2016).

After meeting the requirements for a new breed to be eligible for registration, HC was the first Canadian beef breed officially recognized under the provisions of Livestock Pedigree Act in 1975. This meant HC demonstrated their ability to meet the criteria defined by the Agriculture Canada committee. They produced progeny with good performance from identified parents in a population that descended from recognizable founders. In addition, the population had been closed to outside breeding for more than three generations and providing a meaningful advance over and above breeds already existed (Fleming *et al*. 2016). The breed was mainly a synthetic (composite) descending from Holstein, Hereford and Brown Swiss germplasm. They are generally black in color with white markings on the face, legs and underbelly, however, red and white combinations also occur.

According to its history, the breed is claimed to benefit from having carcass with excellent quality and yield grades, efficient conversion of feed to gain to reach the preferred market weight early and is well adapted to Western Canada climate conditions. It is also known for high fertility and early maturity with potential pregnancy at around one year of age (Fleming 2013). Red Bow Ranching Ltd in Alberta is the main owner of HC cattle. They have recently donated animals to the Kinsella Research Centre at University of Alberta. At Kinsella, the HC cattle will be used for research in support of maintaining and expanding the herd size. There are also a few other producers in Alberta, British Columbia, Manitoba and Quebec that use HC commercially. Additionally, Australian breeders have expressed interest in using HC germplasm in a new composite with Brahman cattle to improve their crossbred production.

In recent years, the size of the nucleus herd has been substantially reduced. This reduction in population size may compromise its future improvement and expose HC to greater risk of extinction due to an increase in inbreeding. Historically, the nucleus herd is also not very complete in pedigree and except for growth traits, there is no considerable data available for research from other economically important traits. Except for some initial work on evaluating the genetic improvement of weight traits (Fleming 2013), there is need to examine the current situation of breed for specially weaning and yearling weights, the two essential traits that HC was selected on. Moreover, there are no investigations about the changes in contributions from its founder breeds over time and identification of genomic regions that may elucidate their signature of selection in creation of a composite.

1.7 Research objectives

Overall, five objectives were proposed for research to be reported in this thesis. They were:

(i) Investigate potential opportunities for improvement of the production efficiency of tropically adapted taurine breeds (straightbred Afrikaner) through incorporation of *Bos taurus* genetics (Hereford, Simmental and Charolais) by application of crossbreeding and sexed semen technologies.

(ii) Estimate genetic parameters and genetic trends of birth, weaning and yearling weights in the HC composite population using random regression and comparing the results with those obtained from more traditional analyses of weights at prescribed ages.

(iii) Predict the genomic breed composition of HC with 6K and 50K SNP panels using both ADMIXTURE (Alexander *et al*. 2009) and regression methods (Kuehn *et al*.

2011); and to dissect each autosome to explore the potential for founder breed proportions to vary across the genome.

(iv) Estimate inbreeding in HC using runs of homozygosity to provide an evaluation of genetic diversity within this small population with an incomplete pedigree.

(v) Identify signatures of selection between HC and each of its founders in order to localize the effects of selection for desired traits in developing a composite like HC.

1.8 References

Alexander, D.H., Novembre, J. and Lange, K., 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome research, 19(9), pp.1655-1664.

Allaire, F.R. and Henderson, C.R., 1965. Inbreeding within an artificially bred dairy cattle population. Journal of dairy science, 48(10), pp.1366-1371.

Akanno, E.C., Chen, L., Abo-Ismail, M.K., Crowley, J.J., Wang, Z., Li, C., Basarab, J.A., MacNeil, M.D. and Plastow, G., 2017. Genomic prediction of breed composition and heterosis effects in Angus, Charolais, and Hereford crosses using 50K genotypes. Canadian Journal of Animal Science, 97(3), pp.431-438.

Baldi, F., Albuquerque, L.G. and Alencar, M.M., 2010. Random regression models on Legendre polynomials to estimate genetic parameters for weights from birth to adult age in Canchim cattle. Journal of Animal Breeding and Genetics, 127(4), pp.289-299.

Bjelland, D.W., Weigel, K.A., Vukasinovic, N. and Nkrumah, J.D., 2013. Evaluation of inbreeding depression in Holstein cattle using whole-genome SNP markers and alternative measures of genomic inbreeding. Journal of dairy science, 96(7), pp.4697-4706.

Boichard, D., Ducrocq, V., Croiseau, P. and Fritz, S., 2016. Genomic selection in domestic animals: principles, applications and perspectives. Comptes rendus biologies, 339(7-8), pp.274-277.

Bovine HapMap Consortium, 2009. Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. Science, 324(5926), pp.528-532.

Burrow, H.W., 2006. Utilization of diverse breed resources for tropical beef production. Proceedings of the 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte, Brasil, August 13–18

Caetano, S.L., Savegnago, R.P., Boligon, A.A., Ramos, S.B., Chud, T.C.S., Lôbo, R.B. and Munari, D.P., 2013. Estimates of genetic parameters for carcass, growth and reproductive traits in Nellore cattle. Livestock Science, 155(1), pp.1-7.

Cardoso, F.F. and Templeman, R.J., 2004. The value of hierarchical Bayes models on genetic evaluation of multiple-breed beef cattle populations. J. Anim. Sci, 83, pp.62-72.

Cartwright, T.C., 1970. Selection criteria for beef cattle for the future. Journal of Animal Science, 30(5), pp.706-711.

Cartwright, T.C., 1974. Net effects of genetic variability on beef production systems. Genetics, 78(1), p.541.

Cole, J.B., VanRaden, P.M., O'Connell, J.R., Van Tassell, C.P., Sonstegard, T.S., Schnabel, R.D., Taylor, J.F. and Wiggans, G.R., 2009. Distribution and location of genetic effects for dairy traits. Journal of dairy science, 92(6), pp.2931-2946.

Colleau, JJ, Fritz, S., Guillaume, F., Baur, A., Dupassieux, D., Boscher, MY, Journaux, L., Eggen, A. and Boichard, D., 2015. Simulation of the potential of genomic selection in dairy cattle. Animal productions , 28 (3), pp. 251-258.

Crum, T.E., Schnabel, R.D., Decker, J.E., Regitano, L.C. and Taylor, J.F., 2019. CRUMBLER: a tool for the prediction of ancestry in cattle. PloS one, 14(8).

Cundiff, L.V. and Gregory, K.E., 1999. What Is Systematic Crossbreeding? BEEF Cow/Calf Management 1999. Intertec Publ. Corp. Overland Park, KS, pp.8-16.

Dickerson, G.E., 1973. Inbreeding and heterosis in animals. Journal of Animal Science, 1973(Symposium), pp.54-77.

Elsik, C.G., Tellam, R.L. and Worley, K.C., 2009. The genome sequence of taurine cattle: a window to ruminant biology and evolution. Science, 324(5926), pp.522-528.

Fleming, A. 2013. A genetic characterization of the Hays Converter. M.Sc. thesis. University of Guelph, ON, Canada. 129 pp.

Fleming, A.L., Hays, D.P., Moore, S., Plastow, G., and Miller, S.P. 2016. History of the Hays Converter. Can. J. Anim. Sci. 96: 471-477.

Frisch, J.E. and O'neill, C.J., 1998a. Comparative evaluation of beef cattle breeds of African, European and Indian origins. 1. Live weights and heterosis at birth, weaning and 18 months. Animal Science, 67(1), pp.27-38.

Frisch, J.E. and O'Neill, C.J., 1998b. Comparative evaluation of beef cattle breeds of African, European and Indian origins. 2. Resistance to cattle ticks and gastrointestinal nematodes. Animal Science, 67(1), pp.39-48.

Garcia, M.D., Chapman, C.K. and Despain, D., 2019. Applying Principles of Crossbreeding to Maximize Hybrid Vigor. https://digitalcommons.usu.edu/cgi/viewcontent.cgi?article=2994&context=extension_curall (accessed, October, 4, 2019).

Gobena, M., Elzo, M.A. and Mateescu, R.G., 2018. Population structure and genomic breed composition in an angus–brahman crossbred cattle population. Frontiers in Genetics, 9, p.90.

Goddard M. E., Meuwissen T. H. E. 2005. The use of linkage disequilibrium to map quantitative trait loci. Australian Journal of Experimental Agriculture 45, 837-845.

Golden, B.L., Garrick, D.J. and Benyshek, L.L., 2009. Milestones in beef cattle genetic evaluation. Journal of animal science, 87(suppl_14), pp.E3-E10.

González-Recio, O., De Maturana, E.L. and Gutiérrez, J.P., 2007. Inbreeding depression on female fertility and calving ease in Spanish dairy cattle. Journal of dairy science, 90(12), pp.5744-5752.

Gosey, J. A. 1991. Crossbreeding systems and the theory behind composite breeds. In: Proceedings of the Range Beef Cow Symposium XII, Ft. Collins, CO. p 33-35.

Gouveia, J.J.D.S., Silva, M.V.G.B.D., Paiva, S.R. and Oliveira, S.M.P.D., 2014. Identification of selection signatures in livestock species. Genetics and molecular biology, 37(2), pp.330-342.

Gregory, K.E. and Cundiff, L., 1980. Crossbreeding in beef cattle: evaluation of systems. Journal of animal science, 51(5), pp.1224-1242.

Gregory, K.E., Swiger, L.A., Koch, R.M., Sumption, L.J., Rowden, W.W. and Ingalls, J.E., 1965. Heterosis in preweaning traits of beef cattle. Journal of Animal Science, 24(1), pp.21-28.

Gregory, K.E., Trail, J.C.M., Koch, R.M. and Cundiff, L.V., 1982. Heterosis, crossbreeding and composite breed utilisation in the tropics. In Proceedings of the 2nd World Congress on Genetics Applied to Livestock Production (p. 279).

Grisart, B., Coppieters, W., Farnir, F., Karim, L., Ford, C., Berzi, P., Cambisano, N., Mni, M., Reid, S., Simon, P. and Spelman, R., 2002. Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. Genome research, 12(2), pp.222-231.

Hall, J.B. and Glaze Jr, J.B., 2014. System application of sexed semen in beef cattle. The Professional Animal Scientist, 30(3), pp.279-284.

Hayes, B.J., Bowman, P.J., Chamberlain, A.J. and Goddard, M.E., 2009. Invited review: Genomic selection in dairy cattle: Progress and challenges. Journal of dairy science, 92(2), pp.433-443.

He, J., Guo, Y., Xu, J., Li, H., Fuller, A., Tait, R.G., Wu, X.L. and Bauck, S., 2018. Comparing SNP panels and statistical methods for estimating genomic breed composition of individual animals in ten cattle breeds. BMC genetics, 19(1), p.56.

Henderson, C.R., 1973. Sire evaluation and genetic trends. Journal of Animal Science, 1973(Symposium), pp.10-41.

Henderson, C.R., 1975. Best linear unbiased estimation and prediction under a selection model. Biometrics, pp.423-447.

Hohenboken, W.D., 1999. Applications of sexed semen in cattle production. Theriogenology, 52(8), pp.1421-1433.

Hosokawa, D., Ishii, A., Yamaji, K., Sasazaki, S., Oyama, K. and Mannen, H., 2012. Identification of divergently selected regions between Japanese Black and Holstein cattle using bovine 50k SNP array. Animal science journal, 83(1), pp.7-13.

Keller, M.C., Visscher, P.M. and Goddard, M.E., 2011. Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. Genetics, 189(1), pp.237-249.

Kemper, K.E., Reich, C.M., Bowman, P.J., Vander Jagt, C.J., Chamberlain, A.J., Mason, B.A., Hayes, B.J. and Goddard, M.E., 2015. Improved precision of QTL mapping using a

nonlinear Bayesian method in a multi-breed population leads to greater accuracy of across-breed genomic predictions. Genetics Selection Evolution, 47(1), p.29.

Kim, E.S., Cole, J.B., Huson, H., Wiggans, G.R., Van Tassell, C.P., Crooker, B.A., Liu, G., Da, Y. and Sonstegard, T.S., 2013. Effect of artificial selection on runs of homozygosity in US Holstein cattle. PloS one, 8(11).

Kinghorn, B., 1982. Genetic effects in crossbreeding: II. Multibreed selection indices. Zeitschrift für Tierzüchtung und Züchtungsbiologie, 99(1-4), pp.315-320.

Kinghorn, B.P., Shepherd, R.K. and Banks, R.G., 1989. A note on the formation of optimal composite populations. Theoretical and applied genetics, 78(3), pp.318-320.

Koger, M., Peacock, F.M., Kirk, W.G. and Crockett, J.R., 1975. Heterosis effects on weaning performance of Brahman-Shorthorn calves. Journal of Animal Science, 40(5), pp.826-833.

Kuehn, L.A., Keele, J.W., Bennett, G.L., McDaneld, T.G., Smith, T.P.L., Snelling, W.M., Sonstegard, T.S. and Thallman, R.M., 2011. Predicting breed composition using breed frequencies of 50,000 markers from the US Meat Animal Research Center 2,000 Bull Project. Journal of animal science, 89(6), pp.1742-1750.

Lewontin, R.C. and Krakauer, J., 1973. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. Genetics, 74(1), pp.175-195.

Long, C.R., Cartwright, T.C. and Fitzhugh Jr, H.A., 1975. Systems analysis of sources of genetic and environmental variation in efficiency of beef production: cow size and herd management. Journal of Animal Science, 40(3), pp.409-420.

Lund, M.S., de Roos, A.P., de Vries, A.G., Druet, T., Ducrocq, V., Fritz, S., Guillaume, F., Guldbrandtsen, B., Liu, Z., Reents, R. and Schrooten, C., 2011. A common reference

population from four European Holstein populations increases reliability of genomic predictions. Genetics Selection Evolution, 43(1), p.43.

MacNeil, M.D., 1987. Formation of optimal composite populations. Theoretical and applied genetics, 74(6), pp.837-840.

Marras, G., Gaspa, G., Sorbolini, S., Dimauro, C., Ajmone-Marsan, P., Valentini, A., Williams, J.L. and Macciotta, N.P., 2015. Analysis of runs of homozygosity and their relationship with inbreeding in five cattle breeds farmed in Italy. Animal genetics, 46(2), pp.110-121.

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

Meuwissen, T.H., Hayes, B.J. and Goddard, M.E., 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics, 157(4), pp.1819-1829.

Meuwissen, T., Hayes, B. and Goddard, M., 2016. Genomic selection: A paradigm shift in animal breeding. Animal frontiers, 6(1), pp.6-14.

Miller, S., 2010. Genetic improvement of beef cattle through opportunities in genomics. Revista Brasileira de Zootecnia, 39, pp.247-255.

Prayaga, K.C., 2003a. Evaluation of beef cattle genotypes and estimation of direct and maternal genetic effects in a tropical environment. 1. Growth traits. Australian Journal of Agricultural Research, 54(10), pp.1013-1025.

Prayaga, K.C., 2003b. Evaluation of beef cattle genotypes and estimation of direct and maternal genetic effects in a tropical environment. 2. Adaptive and temperament traits. Australian Journal of Agricultural Research, 54(10), pp.1027-1038.

Pryce, J. and Hayes, B., 2012. A review of how dairy farmers can use and profit from genomic technologies. Animal Production Science, 52(3), pp.180-184.

Pryce, J.E., Hayes, B.J. and Goddard, M.E., 2012. Novel strategies to minimize progeny inbreeding while maximizing genetic gain using genomic information. Journal of dairy science, 95(1), pp.377-388.

Purfield, D.C., Berry, D.P., McParland, S. and Bradley, D.G., 2012. Runs of homozygosity and population history in cattle. BMC genetics, 13(1), p.70.

Qanbari, S. and Simianer, H., 2014. Mapping signatures of positive selection in the genome of livestock. Livestock Science, 166, pp.133-143.

Rincon, E.J., Schermerhorn, E.C., Mc Dowell, R.E. and Mc Daniel, B.T., 1982. Estimation of genetic effects on milk yield and constituent traits in crossbred dairy cattle. Journal of Dairy Science, 65(5), pp.848-856.

Sabeti, P.C., Reich, D.E., Higgins, J.M., Levine, H.Z., Richter, D.J., Schaffner, S.F., Gabriel, S.B., Platko, J.V., Patterson, N.J., McDonald, G.J. and Ackerman, H.C., 2002. Detecting recent positive selection in the human genome from haplotype structure. Nature, 419(6909), pp.832-837.

Schenkel, F.S., Miller, S.P. and Wilton, J.W., 2004. Genetic parameters and breed differences for feed efficiency, growth, and body composition traits of young beef bulls. Canadian Journal of Animal Science, 84(2), pp.177-185.

Smith, L.A., Cassell, B.G. and Pearson, R.E., 1998. The effects of inbreeding on the lifetime performance of dairy cattle. Journal of dairy science, 81(10), pp.2729-2737.

Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics, 123(3), pp.585-595.

Theunissen, A., Scholtz, M.M., Neser, F.W.C. and MacNeil, M.D., 2013. Crossbreeding to increase beef production: additive and non-additive effects on weight traits. South African Journal of Animal Science, 43(2), pp.143-152.

Tier, B. and Meyer, K., 2004. Approximating prediction error covariances among additive genetic effects within animals in multiple-trait and random regression models. Journal of Animal Breeding and Genetics, 121(2), pp.77-89.

Upperman, L.R., Kinghorn, B.P., MacNeil, M.D. and Van Eenennaam, A.L., 2019. Management of lethal recessive alleles in beef cattle through the use of mate selection software. Genetics Selection Evolution, 51(1), p.36.

Utsunomiya, Y.T., Milanesi, M., Utsunomiya, A.T.H., Torrecilha, R.B.P., Kim, E.S., Costa, M.S., Aguiar, T.S., Schroeder, S., Do Carmo, A.S., Carvalheiro, R. and Neves, H.H.R., 2017. A PLAG1 mutation contributed to stature recovery in modern cattle. Scientific reports, 7(1), pp.1-15.

Van Eenennaam, A.L., Weigel, K.A., Young, A.E., Cleveland, M.A. and Dekkers, J.C., 2014. Applied animal genomics: results from the field. Annu. Rev. Anim. Biosci., 2(1), pp.105-139.

VanRaden, P.M. and Cooper, T.A., 2015. Genomic evaluations and breed composition for crossbred US dairy cattle. Interbull bulletin, (49).

VanRaden, P.M., Olson, K.M., Wiggans, G.R., Cole, J.B. and Tooker, M.E., 2011. Genomic inbreeding and relationships among Holsteins, Jerseys, and Brown Swiss. Journal of Dairy Science, 94(11), pp.5673-5682.

Veneroni, G.B., Meirelles, S.L., Grossi, D.A., Gasparin, G., Ibelli, A.M.G., Tizioto, P.C., Oliveira, H.N.D., Alencar, M.M. and Regitano, L.C.A., 2010. Prospecting candidate

SNPs for backfat in Canchim beef cattle. Genetics and Molecular Research, pp.1997-2003.

Voight, B.F., Kudaravalli, S., Wen, X. and Pritchard, J.K., 2006. A map of recent positive selection in the human genome. PLoS biology, 4(3).

Weaber, Bob, 2008. Crossbreeding for Commercial Beef Production. Beef Sire Selection Manual. National Beef Cattle Evaluation Consortium. http://www.nbcec.org/producers/sire_selection/chapter5.pdf. pp. 50-57. Accessed October 10, 2019.

Weikard, R., Widmann, P., Buitkamp, J., Emmerling, R. and Kuehn, C., 2012. Revisiting the quantitative trait loci for milk production traits on BTA6. Animal genetics, 43(3), pp.318-323.

Wiggans, G.R., Cole, J.B., Hubbard, S.M. and Sonstegard, T.S., 2017. Genomic selection in dairy cattle: the USDA experience. Annual review of animal biosciences, 5, pp.309-327.

Wiggans, G.R., VanRaden, P.M. and Cooper, T.A., 2011. The genomic evaluation system in the United States: Past, present, future. Journal of dairy science, 94(6), pp.3202-3211.

Wright, S., 1949. The genetical structure of populations. Annals of eugenics, 15(1), pp.323-354.

Zhao, F., McParland, S., Kearney, F., Du, L. and Berry, D.P., 2015. Detection of selection signatures in dairy and beef cattle using high-density genomic information. Genetics Selection Evolution, 47(1), p.49.

Zwambag, A.W., 2007. Genetic improvement of beef tenderness in a multi-breed population utilizing genetic markers. M.Sc. Thesis – University of Guelph, Guelph, Ontario.

Chapter 2. Evaluating breed complementarity and sexed semen with maternal use of Afrikaner germplasm

2.1 Abstract

Livestock production impacts food security of developing countries, especially where efficiency of production is compromised by environmental stressors. In South Africa, breeding with indigenous Afrikaner cattle that are genetically well adapted to subtropical environments is considered an essential strategy for sustainable beef production. Today, there is a potential for farmers to participate in commercial systems that join adapted Afrikaner germplasm, used in a specialized maternal role, with exotic terminal sires to optimize production. The objective of this study was to assess productivity of five simulated production systems; 1) straightbred Afrikaner mated naturally, 2) a straightbred Afrikaner cow herd with two sections; one section to produce replacement females and the other to cross with Charolais terminal sires, both using natural mating, 3) similar to 2, but applying sexed semen to produce replacement females, 4) similar to 2, but using a multi-breed composite dam line with a breed combination of 50% Afrikaner, 25% Hereford and 25% Simmental and 5) similar to 4, but again applying sexed semen to produce replacement females. Parameter estimates needed to compare these systems were extracted from the scientific literature. Relative to straightbred Afrikaner dams, the simulated composite dams were more fit producing 7.8% more calves and their progeny performance was improved by reducing feed intake (−24.4%) and increasing meat production (+11.7%). The potential benefit of allocating more cows to the terminal sire was insufficient to offset the reduction in pregnancy rate that results with the use of sexed semen. Thus, system 4 had the greatest productivity (+23.1%) while requiring 22.8% less feed for finishing the progeny to be harvested relative to the purebred Afrikaner system. The combination of increased productivity and reduced feed

requirement made use of a Charolais terminal sire in conjunction with multi-breed composite females bred by natural service the most efficient system among those studied.

2.2 Introduction

Livestock production impacts food security of millions of people all over the world, and especially in developing countries (Delgado *et al*. 1999; McLeod 2011). Moreover, animal products are the most important income generator for smallholders in such countries. There is potential for emerging farmers to enhance productivity and alleviate poverty by entering a viable market for their animals (Ehui *et al*. 1998; Hazell *et al*. 2007; Scholtz and Theunissen 2010).

In subtropical environments stressors including heat, humidity, disease, parasites and poor nutrition seriously compromise livestock production efficiency and under these conditions use of adapted straightbred germplasm is potentially beneficial (Burrow 2012). In South Africa, the indigenous Afrikaner, Nguni, and Drakensberger beef breeds exhibit specific adaptive features that allow them to thrive in such environments where exotic breeds are less fit (Pienaar 2014).

Achieving optimal levels of reproduction, growth, milk production, food consumption, and carcass merit challenges beef producers as they may be environmentally constrained (Burrow 2012) and antagonistically correlated (MacNeil *et al*. 1984; MacNeil *et al*. 1991; MacNeil *et al*. 2011) while also influencing both productivity and profitability. Thus, implementing technologies that simultaneously reduce costs and carbon footprint while improving beef production efficiency is considered increasingly important (Greiner 2002; Scholtz *et al*. 2012). Crossbreeding can accomplish these goals by providing benefits from heterosis and breed complementarity (Cartwright 1970; Gregory and Cundiff 1980).

Indigenous Afrikaner cattle may be well suited for use as a specialized dam line because they benefit from limited calving difficulties and are well adapted to subtropical environments, while joining them with sires from an exotic terminal sire can improve feedlot performance and meat yield of F1 progenies (Scholtz and Theunissen 2010; MacNeil and Matjuda 2007). Use of Charolais as a terminal sire simplifies management of the crossbreeding system due to its mutation in the SILV gene (Gutierrez-Gil *et al*. 2007) that facilitates sorting maternal progenies from their Charolais sired counterparts using only coat color. Creation of a multi-breed composite dam line allows exploitation of maternal heterosis with potentially little effect on environmental adaptation (Dadi *et al*. 2002; Theunissen 2012). Use of a multi-breed composite dam line not only takes advantage of heterosis, but also by fixing different breed proportions further simplifies management compared to rotational crossbreeding (MacNeil 2005).

Successful implementation of crossbreeding requires that at least part of the herd be allocated to production of replacement females. Therefore, use of sexed semen could contribute to increased productivity by allocating fewer cows to generation of replacements while utilizing the remainder for production of terminal cross progeny (Hall and Glaze 2013). This might mitigate the inefficiency that originates from less beneficial male offspring produced by the dam line (Hohenboken 1999). However, use of sexed semen depends on the ability to successfully implement artificial insemination and opportunities for emerging farmers to use this technology are only now being explored in South Africa (Maqhashu *et al*. 2016).

Systems analysis provides a means for simultaneous consideration of many more factors than could feasibly be included in one experiment and may be a realistic method of developing or testing recommendations to improve production efficiency (Long *et al*. 1975). The aim of this paper is to investigate opportunities for use of crossbreeding and sexed semen technologies to

improve the production efficiency relative to a system that is based on straightbred Afrikaner cattle. Herein experimental results taken from the scientific literature are summarized using simulation.

2.3 Materials and methods

The data used for the simulations described herein came from previously published literature. No animals were used directly in the performance of this research. Thus, approval from the Institutional Animal Care and Use Committees of the authors was not sought. The original research upon which these simulations were based came from experimentation carried out at the Vaalharts Research Station, near Jan Kempdorp, South Africa. The Vaalharts Research Station is located at 27˚51' south and 24˚50' east at an altitude of 1175 meters. Carcass data were collected at the Animal and Dairy Research Institute, Irene, South Africa. Thus, parameters for the simulation experiments described hereafter were drawn from previously published literature (Theunissen 2012; Theunissen *et al*. 2013; Theunissen *et al*. 2014a&b).

Alternative self-sustaining systems were simulated with a deterministic model to evaluate their potential for improving productivity relative to a system based on straightbred Afrikaner cattle. Two of the alternative systems employed Afrikaner females as a specialized dam line and two considered a multi-breed composite female comprised of 50% Afrikaner, 25% Hereford and 25% Simmental as the specialized dam line. In two of the systems sexed semen was used to increase the proportion of females that could be allocated to the terminal sire. The systems were as follows:

1) A straightbred production system using Afrikaner cattle and natural service;

2) A specific-cross production system using Afrikaner females, Charolais terminal sires and natural service;

3) Similar to system 2, but with sexed semen used to produce the straightbred Afrikaner replacement females;

4) Similar to system 2, but with the Afrikaner females being replaced with the multi-breed composite females; and

5) Similar to system 4, but with sexed semen used in generation of the replacement females.

Structural equations used in the simulation were as follows:

$$\mathrm{RR} = 1 - \mathrm{PR}$$

where RR = replacement rate and PR = pregnancy rate, assuming that non-pregnant females were culled at the time of pregnancy testing (approximately coincident with weaning).

$$\mathrm{WR} = \mathrm{PR} \times \mathrm{SR}$$

where WR = weaning rate and SR = calf survival rate.

$$\mathrm{P} = \mathrm{WR} \times 440 \times \mathrm{DP} \times \mathrm{MY}$$

P = Production per cow, where a weight-constant endpoint of 440 kilograms was assumed, DP = dressing percentage and MY = meat yield from the carcass.

$$\mathrm{DOF} = (440 - \mathrm{WW})/\mathrm{ADG}$$

where DOF = Days from weaning to harvest, WW = weaning weight of the calf and ADG = its postweaning average daily gain.

$$\mathrm{TFI} = \mathrm{DOF} \times \mathrm{DFI}$$

TFI = Feed consumed by the calf from weaning to harvest and DFI = average daily feed intake.

Throughout a 1:1 sex ratio was assumed for calves resulting from natural mating. Thus, with natural mating the proportion of the cows that was allocated for production of replacement

females (RF) = RR/0.5. When use of sexed semen was simulated the expected 1:1 sex ratio was shifted to 90% heifers (Hohenboken 1999; Seidel 2003; Hall and Glaze 2013) and PR was reduced by 10% (Hall and Glaze 2013; Lardner and Damiran 2015). Thus, when use of sexed semen was simulated RF = RR/0.9. There was no effect of having used sexed semen on SR (Seidel 2003). For female calves, P was reduced by 20% relative to their male counterparts (Anderson 2012). Total productivity (TP) and total feed consumption (TFC) for the ith system were calculated as sum of the meat produced from male calves in the replacement section ($P_{RF_{i_m}}$) or the amount of feed they consumed ($TFI_{MP_{i_m}}$) as well as male and female calves in the market progeny section (MP = 1 − RF) with equal proportions, respectively:

$$\mathrm{TP_i} = \mathrm{HP_{RF_{i_m}}} \times \mathrm{RF_i} \times \mathrm{P_{RF_{i_m}}} + 1/2 \times \mathrm{MP_i} \times \mathrm{P_{MP_{i_m}}} + 1/2 \times \mathrm{MP_i} \times \mathrm{P_{MP_{i_f}}}$$

$$\mathrm{TFC_i} = \mathrm{HP_{RF_{i_m}}} \times \mathrm{RF_i} \times \mathrm{TFI_{RF_{i_m}}} + 1/2 \times \mathrm{MP_i} \times \mathrm{TFI_{MP_{i_m}}} + 1/2 \times \mathrm{MP_i} \times \mathrm{TFI_{MP_{i_f}}}$$

where HP represents the harvest proportion for male calves in the replacement section and corresponds to 1 2 or 0.1 for natural service or sexed semen, respectively.

Breed-specific genetic effects, following Dickerson (1973), on the phenotypes used in the simulations for this study were taken from the previously published work of (Theunissen *et al*. 2013; Theunissen *et al*. 2014a&b). Specifically, phenotypes were simulated for four breed groups as follows:

$$\text{Straightbred Afrikaner} = \mathrm{G_0} + \mathrm{G_A^I} + \mathrm{G_A^M}$$

$$\text{Charolais sired progeny from Afrikaner dams} = \mathrm{G_0} + 1/2\,\mathrm{G_A^I} + 1/2\,\mathrm{G_C^I} + \mathrm{H_{CA}^I} + \mathrm{G_A^M}$$

$$\text{Straightbred composite} = \mathrm{G_0} + 1/2\,\mathrm{G_A^I} + 1/4\,\mathrm{G_H^I} + 1/4\,\mathrm{G_S^I} + 1/4\,\mathrm{H_{HA}^I} + 1/4\,\mathrm{H_{SA}^I} + 1/8\,\mathrm{H_{HS}^I} + 1/2\,\mathrm{G_A^M} + 1/4\,\mathrm{G_H^M} + 1/4\,\mathrm{G_S^M} + 1/4\,\mathrm{H_{HA}^M} + 1/4\,\mathrm{H_{SA}^M} + 1/8\,\mathrm{H_{HS}^M}$$

$$\text{Charolais sired progeny from composite dams} = G_0 + 1/2\,G_C^I + 1/4\,G_A^I + 1/8\,G_H^I +$$

$$+ 1/8\,G_S^I + 1/2\,H_{CA}^I + 1/4\,H_{CH}^I + 1/4\,H_{CS}^I + 1/2\,G_A^M + 1/4\,G_H^M + 1/4\,G_S^M + 1/4\,H_{HA}^M +$$

$$1/4\,H_{SA}^M + 1/8\,H_{HS}^M$$

where A, C, H and S designate the Afrikaner, Charolais, Hereford and Simmental breeds, G^I and G^M represent individual and maternal breed additive effects, and H^I and H^M represent individual and maternal heterosis effects. Following Theunissen *et al*. (2013; 2014a&b), G_0 = the Afrikaner breed mean for the trait of interest and G^I_A and G^M_A were assumed to be zero. Other breed-specific effects were expressed as deviations from G_0. In calculating PR, SR and WR for AHS dams, only the breed specific maternal genetic effects were available and the direct effects due to the genotype of the calf were assumed to be nil.

Finally, selection index weights for improvement of a ratio (Lin 1980) were used to predict improvement in efficiency of the alternative systems (TFC_i/TP_i) relative to the straightbred Afrikaner system. Specifically, the ratio of TFC/TP in the straightbred Afrikaner system was the economic weight for TP_i and the economic weight for TFC_i was −1.0. Thus,

$$\text{Merit}_i = (TFC_1/TP_1) \times TP_i - TFC_i$$

2.4 Results

Simulation results for the fitness traits are presented in Table 2.1. Greater PR was achieved by composite dams due to favorable breed additive effects and heterosis. Thus, a lower proportion of females were required to produce replacements with systems 4 and 5. Differences in calf SR between two dam breeds were small and thus most of the advantage in WR of composite dams was attributable to their greater PR. The decrease in PR resulting from use of sexed semen resulted in similarly reduced WR and greater RR for systems 3 and 5.

Simulated results for productivity and the component traits for the breed groups simulated in this study are presented in Table 2.2. The greatest WW resulted from mating the Charolais terminal sire to composite dams with the resulting progeny being 25% heavier than straightbred Afrikaner calves. Straightbred composite calves also approached this level of performance. Although Charolais sired calves from Afrikaner dams also had increased WW (18.4%) relative to straightbred Afrikaner, the increase was not as great as for progeny of composite dams (25.2%). Similarly, Charolais sired calves from Composite dams had the greatest ADG and DFI. Further, due to the decreased number of days on feed required to attain 440 kilograms, Charolais × Composite calves consumed the least amount of feed during the time they were in the feedlot; approximately 29% less than straightbred Afrikaner. Both Charolais sired calves from Afrikaner dams and straightbred composite calves were intermediate between these breed groups for ADG, DFI, and TFI. There were no differences in DPs of the Afrikaner and composite straightbreds. However, use of the Charolais terminal sire resulted in a greater ratio of hot carcass weight to live weight. All crosses, which were approximately similar, improved meat yield from the carcass relative to straightbred Afrikaner. Thus, simulated P was greatest when composite females were joined with Charolais sires.

Dividing the cow herd into two sections, producing replacement females (RF) and the second producing market progeny (MP) illustrates commitments of resources to components of the production system (Table 2.3). In each replacement section, females are allocated for production of replacements and contemporary males that are produced as a by-product are harvested and thus contribute to P and require feed. The number of by-product male calves is decreased when using sexed semen. Use of a Charolais terminal sire on a portion of the Afrikaner cows (System 2) increased P by 7% over System 1 due to effects of breed

complementarity and individual heterosis. Using sexed semen further improved P in production of market progeny in System 3, relative to System 2. However, this advantage was not sufficient to offset the loss in output that resulted from the reduced PR. Thus, total P of System 3 was less than that either System 1 or 2. With Systems 4 and 5, P (TFI) of the RF section was reduced (increased) relative to Systems 2 and 3, respectively, due to the lesser number of females that were required in this section. Corresponding changes in P and TFI of the MP sections also result.

Using merit as a unified assessment of the systems, all systems that employed crossbreeding had greater merit than the straightbred Afrikaner production system (it was selected as the base to which the other systems were relative). Further, use of the composite specialized dam line increased merit relative to the corresponding system with a straightbred Afrikaner cow herd. Although using sexed semen increased the merit of System 3 slightly greater (3%) than System 2, it did not result in any additional advantage in System 5 relative to System 4 that employed natural service.

2.5 Discussion

This study contemplated alternative production systems that could potentially improve beef production when making use of Afrikaner germplasm. It is based, in large part, on data collected in central South Africa near Jan Kempdorp at 27˚51’ south and 24˚50’ east, and an altitude of 1175 m. Climatic conditions are classified as semi-arid with hot summers (average maximum temperature in December of 32˚C) and cold winters (average minimum temperature in July of −0.5˚C). Average precipitation is approximately 450 mm per annum, of which 88% is experienced between October and April. One concern about the data used herein regards its potentially being dated. However, similar crossbreeding studies that evaluate breeds in challenging environments such as this one have not been conducted more recently.

Here, the principal finding establishes the substantial benefit of a composite specialized dam line relative to either a straightbred production system or a production system that employs a terminal sire on straightbred dams. This is consistent with the substantial volume of literature on the benefits of crossbreeding (MacNeil *et al*. 1988; Skrypzeck *et al*. 2000; Dadi *et al*. 2002). However, this finding must be tempered by the recognition that use of non-adapted germplasm in stressful environments may not always be warranted (Barwick *et al*. 2009; Burrow 2012). The superior adaptation of Afrikaner cattle relative to British breeds to hot, semi-arid, subtropical bushveld regions has long been recognized (Bonsma 1949). Further, the composite envisioned herein has similar breed composition to the Belmont Red that was developed by large pastoral companies in northern Australia to meet market specifications and maintain tropical adaptation (Rudder *et al*. 1976; Seifert and Rudder 1984). Growth performance, carcass merit and adaptive potential of the Belmont Red have been shown to be similar to that of Bonsmara, but with Belmont Red having a shorter calving interval (Corbet *et al*. 2006). As pointed out by (Barwick *et al*. 2009), the environment for market progeny may be less severe than for females destined to become replacements and this may indicate a role for crossbreeding even in production environments where adaptation is of utmost importance for the cow herd. It has been recommended that adapted maternal lines should be crossed with large exotic terminal sire breeds to exploit breed complementarity for efficient gain, carcass quality and meat yield (Scholtz *et al*. 1990; Scholtz and Theunissen 2010). The work of Moyo (1990) supports this recommended paradigm.

At the outset, it seemed logical that if sexed semen would allow more females to be joined with a terminal sire that technology would support greater production efficiency. However, sex sorting sperm by flowcytometry is slow and along with post-processing damage

limits the number of live sperm per dose produced. This results in fewer sperm then being used per insemination and a consequent reduction of fertility (Garner and Seidel 2008). Recent literature supports the implicit assumption of similar pregnancy rates being achieved by natural service and artificial insemination (Lardner and Damiran 2015) and the explicit assumption of a 10% decrease in PR when using sexed semen (Hall and Glaze 2013). For System 3, which employs straightbred Afrikaner females, overall merit is slightly greater (+3%) than for System 2 with a 10% reduction in pregnancy rate and 90% efficiency in sorting. If the sorting were perfect and sex-sorted semen was equivalent to unsorted semen in pregnancy rate, then the advantage of System 3 over System 2 would increase to 20%. However, diverting additional females to the terminal sire by increasing the pregnancy rate and efficacy of sorting in use of sex-sorted semen only allows System 5 to marginally approach the merit of System 4. Thus, the results suggest that when using highly fertile and productive composite females, there would be little if any additional advantage from breed complementarity.

2.6 Conclusion

This simulation study demonstrates the utility of systems analysis techniques for summarizing a diverse body of data than it would be feasible to address experimentally. The results suggest that crossbreeding could be applied as a commercially appropriate technology to potentially improve production efficiency relative to a straightbred production system using adapted germplasm. However, not all state-of-the-art technologies will necessarily improve production efficiency, as is shown herein using the example of sexed semen. The compromised pregnancy rate and less than perfect sorting technology as well as limited opportunities for cost-effective implementation of artificial insemination prevent smallholders in stressful environments from fully capitalizing on benefits from crossbreeding.

2.7 References

Anderson, P. (2012) Matching Cattle Type and Feedlot Performance. Univ. Minnesota Extension. https://www.extension.umn.edu/agriculture/beef/components/pdfs/matching_cattle_type_and_feedlot_performance.pdf

Barwick, S.A., Johnston, D.J., Burrow, H.M., Holroyd, R.G., Fordyce, G., Wolcott, M.L., Sim, W.D. and Sullivan, M.T. (2009) Genetics of Heifer Performance in Wet and Dry Seasons and their Relationships with Steer Performance in Two Tropical Beef Genotypes. Anim. Prod. Sci, 49, 367-382.

Bonsma, J.C. (1949) Breeding Cattle for Increased Adaptability to Tropical and Sub-tropical Environments. J. Agric. Sci, 39, 204-221.

Burrow, H.M. (2012) Importance of Adaptation and Genotype×Environment Interactions in Tropical Beef Breeding Systems. Animal, 6, 729-740.

Cartwright, T.C. (1970) Selection Criteria for Beef Cattle for the Future. Journal of Animal Science, 30(5), 706-711.

Corbet, N.J., Shepherd, R.K., Burrow, H.M., van der Westhuizen, J., Strydom, P.E. and Bosman, D.J. (2006) Evaluation of Bonsmara and Belmont Red Breeds in South Africa. 1. Productive performance. Aust. J. Exp. Agric, 46, 199-212.

Dadi, H., Jordaan, G.F., Schoeman, S.J. and Van der Westhuizen, J. (2002) The Effect of Charolais and Hereford Sires and Straightbred and Crossbred Dams on Pre-weaning Growth of Calves. S. Afr. J. Anim. Sci, 32, 38-43.

Delgado, C., Rosegrant, M., Steinfeld, H., Ehui, S. and Courbois, C. (1999) Livestock to 2020: The Next Food Revolution. Washington, DC, International Food Policy Research Institute, FAO, and International Livestock Research Institute.

Dickerson, G.E. (1973) Inbreeding and Heterosis in Animals. In: Proc. Anim. Breed and Genet. Symp. in honor of Dr. J.L. Lush, Champaign, Illinois. Amer. Soc. Anim. Sci, 54-77.

Ehui, S.H., Li-Pun, H.H., Mares, V. and Shapiro, B. (1998) The Role of Livestock in Food Security and Environmental Protection. Outlook on Agriculture, 27, 81-87.

Garner, D.L. and Seidel, G.E. (2008) History of Commercializing Sexed Semen for Cattle. Theriogenology, 69, 886-895.

Gregory, K.E. and Cundiff L.V. (1980) Crossbreeding in Beef Cattle: Evaluation of Systems. J. Anim. Sci, 51, 1224-1242.

Greiner, S.P. (2002) Crossbreeding in Beef Cattle. Virginia Cooperative Extension. Virginia Polytechnic Institute and State University. Publication 400–805.

Gutierrez-Gil B., Wiener, P. and Williams, J.L. (2007) Genetic Effects on Coat Colour in Cattle: Dilution of Eumelanin and Phaeomelanin Pigments in an F2-Backcross Charolais x Holstein Population. BMC Genetics 8(1), 56.

Hall, J.B. and Glaze, J.B. Jr. (2013) How Can Sexed Semen Be Used in Commercial Beef Herds? Range Beef Cow Symposium. Paper 316. http://digitalcommons.unl.edu/rangebeefcowsymp/316

Hazell, P., Poulton, C., Wiggins, S. and Dorward, A. (2007) The Future of Small Farms for Poverty Reduction and Growth. IFPRI Discussion Paper 42. http://www.ifpri.org/2020/dp/vp42.asp

Hohenboken, W.D. (1999) Applications of Sexed Semen in Cattle Production. Theri-ogenology, 52, 1421-1433.

Lardner, B. and Damiran, D. (2015) Comparison of Fixed-time Artificial Insemination vs. Natural Service in Beef Cows: Reproductive Efficiency and System Cost. http://www.wbdc.sk.ca/pdfs/fact_sheets/2015/2015-02_WBDC_Fixed-timeAI_vs_NaturalBreeding.pdf

Lin, C.Y. (1980) Relative Efficiency of Selection Methods for Improvement of Feed Efficiency. J. Dairy Sci, 63, 491-494.

Long, C. R., Cartwright, T.C. and Fitzhugh, H.A. (1975) Systems Analysis of Sources of Genetic and Environmental Variation in Efficiency of Beef Production: Cow Size and Herd Management. J. Anim. Sci, 40, 409-420.

MacNeil, M.D. (2005) Beef Cattle Management: Crossbreeding. In: Encyclopedia of Animal Science. Marcel Dekker, New York, NY.

MacNeil, M.D. and Matjuda, E. (2007) Breeding Objectives for Angus and Charolais Specialized Sire Lines for Use in the Emerging Sector of South African Beef Production. S. Afr. J. Anim. Sci, 37, 1-10.

MacNeil, M.D., Bailey, D.R., Urick, J.J., Gilbert, R.P. and Reynolds, W.L. (1991) Heritabilities and Genetic Correlations for Postweaning Growth and Feed Intake of Beef Bulls and Steers. J. Anim. Sci, 69, 3183–3189.

MacNeil, M.D., Cundiff, L.V., Dinkel, C.A. and Koch, R.M. (1984) Genetic Correlations among Sex-limited Traits in Beef Cattle. J. Anim. Sci, 58, 1171-1180.

MacNeil, M.D., Cundiff, L.V., Gregory, K.E. and Koch, R.M. (1988) Crossbreeding Systems for Beef Production. Applied Agric. Res, 3, 44-54.

MacNeil, M.D., Lopez-Villalobos, N. and Northcutt S.L. (2011) A Prototype National Cattle Evaluation for Feed Intake and Efficiency of Angus Cattle. J. Anim. Sci, 89, 3917-3923.

Maqhashu, A., Mphaphathi, M.L., Seshoka, M.M., Ramukhithi, F.V., Seolwana, F.L., Masenya, M.B., Netshirovha, R.T., Mapeka, M.H., Nengovhela, N.H., Kanuya, N.L., Muchenje, V. and Nedambale, T.L. (2016) Comparison of Oestrous Synchronization Re-sponse and Pregnancy Rate of Village Cows Following Timed Artificial Insemination in KwaZulu-Natal and Limpopo Provinces. Open J. Anim. Sci, 6, 9-15.

McLeod, A. (2011) World Livestock 2011-Livestock in Food Security. Food and Agri-culture Organization of the United Nations (FAO).

Moyo, S. (1990) Evaluation of the Productivity of Indigenous Cattle and Some Exotic Beef Breeds and their Crosses in Zimbabwe. ILCA Research Fellowship Report, ILCA, Addis Ababa, Ethiopia. 139 p.

Pienaar, L. (2014) Genetic Diversity in the Afrikaner Cattle Breed. MSc. Thesis, Univ. Free State, Bloemfontain.

Rudder, T.H., Bean, K.G., Lapworth, J.W. and Seifert, G.W. (1976) Commercial Evaluation of a Tropical Breed Developed Experimentally. Proc. Aust. Soc. Anim. Prod, 11, 89-92.

Scholtz, M.M. and Theunissen, A. (2010) The Use of Indigenous Cattle in Terminal Crossbreeding to Improve Beef Cattle Production in Sub-Saharan Africa. Anim. Genet. Res, 46, 33-39.

Scholtz, M.M., Roux, C.Z. and Lombard, P.E. (1990) Breeding Strategies for Beef Cattle in the Subtropics and Tropics: Terminal Crossbreeding. Proc. 4th World Congress on Genetics Applied to Livestock Production, Edinburgh, 23-27 July 1990, XV, 361-364.

Scholtz, M.M., Steyn, Y., Van Marle-Koester, E. and Theron, H.E. (2012) Improved Production Efficiency in Cattle to Reduce their Carbon Footprint for Beef Production. S. Afr. J. Anim. Sci, 42(suppl. 1), 450-453.

Seidel, G.E. (2003) Economics of Selecting for Sex: The Most Important Genetic Trait. Theriogenology, 59(2), 585-598.

Seifert, G.W. and Rudder, T.H. (1984) Experience with Africander and Africander Derived Genotypes in Australian Beef Herds. Proc. 2nd World Cong. Sheep Beef Cattle Breed, 671-677.

Skrypzeck, H., Schoeman, S.J., Jordaan, G.F. and Neser, F.W.C. (2000) Estimates of Crossbreeding Parameters in a Multibreed Beef Cattle Crossbreeding Project. S. Afr. J. Anim. Sci, 30(3), 193-203.

Theunissen, A. (2012) Charaterization of Breed Additive and Heterosis Effects in Beef Cattle Using Experimental Results. MSc. thesis, Univ. Free State, Bloemfontein.

Theunissen, A., Scholtz, M.M., MacNeil, M.D. and Neser, F.W.C. (2014a) Breed Additive and Heterosis Effects on Feedlot and Carcass Traits in Beef Cattle. Proc. 10th World Congress on Genetics Applied to Livestock Production. Poster No. 714.

Theunissen, A., Scholtz, M.M., Neser, F.W.C. and MacNeil, M.D. (2013) Crossbreeding to Increase Beef Production: Additive and non-Additive Effects on Weight Traits. S. Afr. J. Anim. Sci, 43(2), 143-152.

Theunissen, A., Scholtz, M.M., Neser, F.W.C. and MacNeil, M.D. (2014b) Crossbreeding to Increase Beef Production: Additive and non-Additive Effects on Fitness Traits. S. Afr. J. Anim. Sci, 44(4), 335-349.

2.8 Figures and tables

Table 2.1. Simulated values for fitness traits applicable to straightbred Afrikaner and composite dam lines with natural service and use of sexed semen via artificial insemination.

Breeding system		Fitness traits (rates)			
Method	Dam line	Pregnancy	Replacement	Calf survival	Weaning
Natural service	Afrikaner	0.809	0.191	0.953	0.771
	Composite	0.873	0.127	0.970	0.849
Sexed semen	Afrikaner	0.728	0.272	0.953	0.694
	Composite	0.786	0.214	0.970	0.762

Table 2.2. Simulated growth, feedlot and carcass values for progenies of Afrikaner and Composite cattle as straightbreds and as crosses with a Charolais terminal sire.

	Breed group			
Trait	Afrikaner	Charolais × Afrikaner	Composite	Charolais × Composite
Weaning wt., kg	184.0	217.9	228.5	230.3
Average daily gain, kg/d	0.809	1.080	1.023	1.124
Days to harvest	316	206	207	187
Daily feed intake, kg/d	6.25	7.28	7.22	7.50
Total feed intake, kg	1977	1497	1494	1400
Dressing, %	62.3	65.8	62.3	64.5
Meat yield, %	81.6	82.7	82.9	82.9
Productivity, kg (male calf)	172.5	184.4	192.7	200.0

Table 2.3. Commitments of resources and productivity per cow attributable to sections of the herd allocated to producing replacement females (RF) and market progeny (MP) for each of the simulated systems that were evaluated.

	Production system[a]									
	1		2		3		4		5	
Cow herd sections	RF	MP	RF	MP	RF	MP	RF	MP	RF	MP
Cows allocated, %	38.2	61.8	38.2	61.8	30.2	69.8	25.4	74.6	23.8	76.2
Calf sex ratio (m:f)	1:1	1:1	1:1	1:1	1:9	1:1	1:1	1:1	1:9	1:1
Productivity, kg	32.9	95.9	32.9	102.5	4.7	115.8	24.5	134.1	4.2	137.0
Total feed intake, kg	377.6	1221.7	377.6	925.4	59.3	1045.3	189.7	1044.3	35.9	1066.7
Relative merit			378		391		734		649	

a. System 1 = straightbred Afrikaner with natural service; System 2 = Use of Afrikaner as a specialized dam line, Charolais as a terminal sire, and natural service; System 3 = Use of Afrikaner as a specialized dam line, employing sexed semen in production of replacement females, and Charolais as a terminal sire; System 4 was similar to System 2, but with the Afrikaner dam line replaced by a 50% Afrikaner, 25% Hereford, and 25% Simmental composite dam line; and System 5 was similar to System 3, but also using the composite dam line.

Chapter 3. Estimation of genetic parameters and trends for growth traits in Hays Converter cattle using multiple-trait and random regression models

3.1 Abstract

Hays Converter (HC), the first registered Canadian beef breed, was developed by Harry Hays in the 1950's as a composite with contributions from Holstein, Hereford, Brown Swiss and with a later introgression of Angus. The breed is well adapted to the Western Canadian climate. Calves reach market weight early, convert feed to gain efficiently, and have excellent carcass quality and yield. The edited data consisted of 21,612 weight records taken at ages 1 to 474 days, from 8,850 animals born at Red Bow Ranching Ltd., Calgary, Alberta between 1970 and 2016. This study aimed to evaluate the genetic parameters and trends in birth, weaning and yearling weights of HC through comparison of different multiple trait models (MTM) with a random regression model (RRM). In MTM scenarios, both adjusted and unadjusted data were examined besides considering contemporary groups (GC) as fixed or random. Estimates of variance components, heritability and genetic correlations from the two approaches were not substantially different and showed similar changes along the growth trajectory. Although there was a considerable reduction in genetic trends from 2004 to 2008, due to weak sire selection, trends generally increased through 2016. Overall, both models performed similarly and fixed CG were preferred.

3.2 Introduction

Hays Converter (HC), was developed by Harry Hays in the 1950's, and was the first beef breed of Canadian origin to be registered under the Canadian Livestock Pedigree Act in 1975 (Fleming *et al*. 2016). HC combined Holstein, Hereford and Brown Swiss breeds (Fleming *et al*.

2016). It is known for rapid growth with excellent carcass grading under Canadian standards and adapts to the climatic conditions of Western Canada (Fleming 2013). Since its formation, the breeding objective for HC was to create a beef breed that excelled in growth and efficiently converted feed to gain so that the cattle reached market weight at earlier ages than other competing breeds. Sire selection always used the phenotypic records of weaning and yearling weights and a subjective evaluation of growth potential. There was no selection based on EBV or an index until 2014. However, visual inspection and use of individual phenotypes of animals that are candidates for selection may not maximize response to selection. Use of estimated breeding values (EBV) as the basis for selection is expected to produce more rapid genetic improvement (Mofakkarul Islam *et al*. 2013). Preliminary work with single-trait and bivariate models and fixed CG has produced estimates of genetic parameters and genetic trends for growth traits of HC (Fleming 2013).

Currently, most genetic evaluation programs predict EBV for growth traits in beef cattle using multiple trait models (MTM: Farquharson *et al*. 2003; Meyer 2004; Delgadillo *et al*. 2017). In this approach, weight records are collected within defined windows of time within which growth is assumed to be linear along the growth curve and standardized to 205 and 365 days of age for weaning weight and yearling weight, respectively (Beef Improvement Federation 2018). Other weights recorded outside of these windows are not used in the evaluations. This approach may lead to EBV with lower accuracies than if all available data were used (Meyer 2004; Mota *et al*. 2013).

Random regression models (RRM) facilitate use of all available weights that are recorded over time for each animal (Schaeffer 2004). Although pre-adjustment to standard ages is not

necessary, fitting RRM models is similar to MTM in computational complexity. However, RRM are sensitive to sparse data, as may occur at extremes of age-weight trajectories (Meyer 1999). Contemporary groups (CG) allow for elimination of bias caused by different environmental factors. Including CG in the genetic evaluation leads to a more accurate EBV and thus the potential for increased genetic improvement (Van Vleck 1987; Ramirez-Valverde *et al*. 2008). When they are considered as fixed effects, the bias due to non-random relationship between sires and CG is reduced. If they are considered random, then the prediction error variance is reduced because more information is used to predict the EBV (Visscher and Goddard 1993). Application of the predicted breeding values to predict the merit of progeny in currently unobservable future contemporary groups would argue for considering CG as random effects.

Therefore, the goal of this study was to evaluate alternatives for prediction of EBV in HC. Specific objectives were to assess RRM versus MTM and the use of random versus fixed CG and to estimate genetic parameters and genetic trends of birth, weaning and yearling weights.

3.3 Materials and Methods

The data used for this study was collected and provided by a commercial enterprise (Red Bow Ranching Ltd, Calgary, AB) using industry standard production practices which were generally consistent with the guidelines of the Canadian Council on Animal Care (Olfert *et al*. 1993). Raw data consisted of weights recorded at birth (BW), weaning (WW from 100 to 315 days of age), and yearling (YW from 245 to 544 days of age) (Table 3.1), from animals born between 1970 and 2016. In order to edit data for further analysis, all weaning and yearling weight records obtained at ages that deviated from the respective mean ages by more than three standard deviations were excluded from the data.

CGs were defined by concatenation of herd, year, season of birth (Jan-Mar, Apr-June, July-Sept, and Oct-Dec) and sex (male or female). CGs of less than three animals were excluded from the respective analyses where they were considered as fixed. Age of dam at calving was categorized into five classes (2 years old, 3 years old, 4 years old, 5-7 years old, and ≥8 years old). Records from calves produced through embryo transfer and also those with unidentified dams were excluded. Finally, and in contrast to Fleming (2013), all conjectural birth weights were eliminated from the data. The number of sires and dams in the pedigree were equal to 137 and 1701 in which 63% and 17% of dams were daughters of sires or dams that were used as dams, respectively (Table 3.1). Animals with unknown parents were assigned to generation zero. For animals with recorded parents, generation numbers were calculated following the approach of Brinks et al. (1961) wherein generation numbers for descendants of animals in generation 0 were calculated as the average generation number of their parents plus one. The average generation interval was estimated as the linear regression of birth year on generation number.

After this initial editing, datasets were prepared for different MTM scenarios and RRM analysis. For the RRM analysis, all weights between 1 and 474 (the maximum yearling age after ± 3SD edit) days of age were used along with 1381 records that were recorded during the period when feed intake was measured. This dataset contained 21,612 records from 8,850 animals with minimum and maximum numbers of 1 to 9 records per animal (Figure 3.1, Table 3.2). For MTM analyses, two sets of data were extracted. Dataset 1 was based on the windows of age recommended by the Beef Improvement Federation (2018) for weaning and yearling weights (i.e., ranges of +/- 45 days from the average ages at which weaning (188 d) and yearling (365 d) weights were recorded in HC) and included all available records of BW. These weights were linearly adjusted for age as follows:

$$W_{188} = BW + \frac{(WW - BW)}{Weaning\ age} \times 188$$

$$W_{365} = W_{188} + \frac{(YW - WW)}{(Yearling\ age - Weaning\ age)} \times 177$$

where: W_{188} and W_{365} represent age-adjusted weaning and yearling weights, respectively (N = 15,107 records including BW). Dataset 2 had wider windows in order to incorporate additional data making it more similar to the dataset analyzed with RRM. In the second dataset the windows for acceptable weights were expanded to +/- 65 days and age effects were estimated simultaneously in MTM models. In total, approximately 30% and 26% more records were available for the RRM analysis than in the first and second datasets used in the MTM analyses. The growth trajectory from 123 to 474 days of age is shown in Figure 3.1. Note that, except for BW with an average 40 kg, the weights recorded at ages less than 123 days were not used in either analysis due to their low frequency.

3.3.1 Models

3.3.1.1 Different MTM scenarios

First, for the analysis of age-adjusted weights, the MTM was defined as follows:

$$y_{ijlmt} = CG_i + \sum_{n=0}^{2} b_n (aod_j)^n + a_{lt} + d_{mt} + e_{ijlmt}$$

where y_{ijlmt} is the t^{th} weight record for BW, WW_{188} and YW_{365} of the animal l in the i^{th} CG and the j^{th} class of age of dam at calving (aod); b_n is a fixed regression adjusted to linear and quadratic effects for the aod as a covariate; a_{lt} is the random direct additive genetic effect of the animal l for weight t; d_{mt} is the random maternal additive genetic effect of the dam m for weight t and e_{ijlmt} is the random residual effect. To avoid the failure that occurred in the approximation of

standard errors for maternal permanent environmental parameters due to small sample size and/or over-parameterization (Meyer 2018), these effects were removed from the MTMs.

This model can be described in matrix notation as follows:

$$y = X\beta + Z_1 d + Z_2 m + \varepsilon$$

where y is the vector of weight records; β is the vector of fixed effects (CG classes and aod regressions); d is the vector of random direct additive genetic effects; m is the vector of random maternal additive genetic effects and ε is the vector of residual effects; X, Z_1 and Z_2 are the incidence matrices for the corresponding effects. The assumptions for this analysis were as follows:

$$E\begin{bmatrix} y \\ d \\ m \\ \varepsilon \end{bmatrix} = \begin{bmatrix} X\beta \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad V\begin{bmatrix} d \\ m \\ \varepsilon \end{bmatrix} = \begin{bmatrix} G_d \otimes A & 0 & 0 & 0 \\ 0 & G_m \otimes A & 0 & 0 \\ 0 & 0 & 0 & R \otimes I_n \end{bmatrix}$$

where G_d, and G_m are (co)variance matrices of random effects for direct additive genetic and maternal additive genetic effects, respectively; A is the numerator relationship matrix; I_n is the identity matrix whose order is equal to the number of records; R is a (co)variance matrix of random residual effects and ⊗ is the kronecker product operator. A parallel analysis to that just described was conducted using the weight records which had not been pre-adjusted for age and including linear covariates in the model to account for the age effects.

To evaluate consideration of CG effects as being either fixed or random, both of the datasets were analyzed similarly, but only using the data which was not pre-adjusted for age. For convenience these analyses are referred to as MTM-data1-CG-fixed, MTM-data1-CG-random, MTM-data2-CG-fixed and MTM-data2-CG-random. No maternal genetic effect was considered for YW (due to numerical errors). To test the non-linearity of age effects on WW and YW, the

second dataset was also analyzed with two additional 3-traits models, i.e. MTM-data2-CG-fixed-age quadratic and MTM-data2-CG-random-age quadratic, respectively. In each of the six models above, when considering CG as a random effect, the assumptions were as follows;

$$E\begin{bmatrix} y \\ d \\ m \\ cg \\ \varepsilon \end{bmatrix} = \begin{bmatrix} X\beta \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad V\begin{bmatrix} d \\ m \\ cg \\ \varepsilon \end{bmatrix} = \begin{bmatrix} G_d \otimes A & 0 & 0 & 0 \\ 0 & G_m \otimes A & 0 & 0 \\ 0 & 0 & G_{cg} \otimes I_g & 0 \\ 0 & 0 & 0 & R \otimes I_n \end{bmatrix}$$

where G_{cg} is the (co)variance matrix of random effects for contemporary groups and I_g is the identity matrix whose order is equal to the number of contemporary groups.

In order to estimate direct (h^2_a) or maternal (h^2_m) heritability when considering CG as fixed or random, the respective phenotypic variances (Var (p)) were calculated as follows;

$$\text{Var (p)}_{\text{CG-fixed}} = \text{Var (a)} + \text{Var (m)} + \text{Var (e)}$$

$$\text{Var (p)}_{\text{CG-random}} = \text{Var (a)} + \text{Var (m)} + \text{Var (CG)} + \text{Var (e)}$$

where Var (a), Var (m), Var (CG) and Var (e) are additive genetic, maternal genetic, contemporary group and residual variances, respectively. Therefore,

$$h^2_{\text{a/m}} = \text{Var (a/m)}_{\text{CG-fixed}} / \text{Var (p)}_{\text{CG-fixed}}$$

$$h^2_{\text{a/m}} = \text{Var (a/m)}_{\text{CG-random}} / \text{Var (p)}_{\text{CG-random}}$$

3.3.1.2 RRM analysis

In implementing the RRM analysis, quadratic Legendre polynomials were selected for the fixed regression coefficients to define changes in the population mean trend. Random regressions of different orders (k) of Legendre polynomials were modeled to describe variation in direct additive genetic (a), direct permanent environmental (p), maternal additive genetic (m) and maternal permanent environmental (c) effects, respectively. Initially, models with quadratic, cubic and quartic degrees of Legendre polynomials for the direct additive genetic and permanent environmental effects were evaluated (i.e., $k_a = k_p = 3$, 4 and 5, respectively). For the maternal additive genetic and permanent environmental effects, linear, quadratic and cubic degrees of polynomials were initially considered (i.e., $k_m = k_c = 2$, 3 and 4, respectively). Assuming heterogeneity of residual variances across the growth curve, they were categorized into four age classes as follows: 1 to 60, 61 to 205, 206 to 365 and 366 to 474 days of age, respectively. Therefore, the RRM was defined as:

$$y_{ijlmt} = CG_i + \sum_{n=0}^{2} b_n (aod_j)^n + \sum_{n=0}^{2} \beta_n \phi_n(age_t) + \sum_{n=0}^{k_a-1} \alpha_{ln} \phi_n(age_t)$$
$$+ \sum_{n=0}^{k_m-1} \gamma_{mn} \phi_n(age_t) + \sum_{n=0}^{k_p-1} \delta_{ln} \phi_n(age_t) + \sum_{n=0}^{k_c-1} \rho_{mn} \phi_n(age_t) + \varepsilon_{ijlmt}$$

$$y_{ijlmt} = CG_i + \sum_{n=0}^{2} b_n (aod_j)^n + \sum_{n=0}^{2} \beta_n \phi_n(age_t) + \sum_{n=0}^{k_a-1} \alpha_{ln} \phi_n(age_t)$$
$$+ \sum_{n=0}^{k_m-1} \gamma_{mn} \phi_n(age_t) + \sum_{n=0}^{k_p-1} \delta_{ln} \phi_n(age_t) + \sum_{n=0}^{k_c-1} \rho_{mn} \phi_n(age_t) + \varepsilon_{ijlmt}$$

where y_{ijlmt} is each of the weight records taken at age t for the animal l with the dam m, in the i^{th} CG and the j^{th} class of aod; b_n are fixed regressions relative to aod; β_n are fixed regression coefficients that model the average growth trajectory of the population; Φ_n (age_t) is the n^{th} Legendre polynomial according to age t; α_{ln}, γ_{mn}, δ_{ln} and ρ_{mn} are the n^{th} random regression

coefficients of direct additive genetic, maternal additive genetic, direct permanent environmental and maternal permanent environmental effects, respectively and ε_{ijlmt} is the random residual error associated with the age t of the l^{th} animal.

In matrix notation, the model was represented as follows:

$$y = X\beta + Z_1\alpha + Z_2\gamma + W_1\delta + W_2\rho + \varepsilon$$

where y is the vector of observations; β is the vector of fixed effects; α, γ, δ and ρ are the vectors of random regression coefficients for direct additive genetic, maternal additive genetic, direct permanent environmental and maternal permanent environmental effects, respectively; X, Z_1, Z_2, W_1 and W_2 are the incidence matrices for corresponding effects and ε is the vector of residual effects. The following assumptions were considered for the RRM:

$$E\begin{bmatrix} y \\ \alpha \\ \gamma \\ \delta \\ \rho \\ \varepsilon \end{bmatrix} = \begin{bmatrix} X\beta \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad V\begin{bmatrix} \alpha \\ \gamma \\ \delta \\ \rho \\ \varepsilon \end{bmatrix} = \begin{bmatrix} K_a \otimes A & 0 & 0 & 0 & 0 \\ 0 & K_m \otimes A & 0 & 0 & 0 \\ 0 & 0 & K_p \otimes I_{N_a} & 0 & 0 \\ 0 & 0 & 0 & K_c \otimes I_{N_m} & 0 \\ 0 & 0 & 0 & 0 & R \end{bmatrix}$$

where K_a, K_m, K_p and K_c are (co)variance matrices between random regression coefficients for direct additive genetic, maternal additive genetic, direct permanent environmental and maternal permanent environmental effects, respectively; A is the numerator relationship matrix; I is an identity matrix; N_a is the total number of individuals with records; N_m is the number of dams; R is a diagonal matrix of residual variances and ⊗ is the kronecker product operator. The covariance between direct and maternal additive genetic effects was assumed to be zero for both MTM and RRM. To find an appropriate RRM, preliminary analyses with different orders of fit for Legendre polynomials were examined from 1111 to 5343 (higher

orders produced numerical errors) where 1111 indicates the four random effect regressions being modeled as linear effects and 5343 indicated the random effects being modeled with quantic, cubic, quartic, and cubic polynomials, respectively. These analyses were initially compared using the likelihood statistics, Akaike information criterion (AIC), and Bayesian information criterion (BIC) (Table 3.3). According to these criteria, the model 5343 initially indicated improved description of variation in the data, although BIC imposed a greater penalty than AIC for the number of parameters estimated (Olori *et al*. 1999; Meyer 2001; Boligon *et al*. 2010). However, to avoid very large estimates at the boundaries of growth trajectory (Albuquerque and Meyer 2001; Nobre *et al*. 2003; Meyer 2005a), RRM models producing parameter estimates more similar to those from MTM with age-adjusted weights were compared through LRT (log-likelihood ratio test). This statistic was calculated as twice the difference in log L between complete and reduced models:

$$LRT_{ij} = 2(\log L_i - \log L_j)$$

where log L_i and log L_j were the maximum of log L for the complete model i and reduced (nested) model j, respectively. If the LRT with degrees of freedom equal to the difference between the number of parameters estimated in complete and reduced models and the significance level of $P<0.05$, was higher than a tabulated chi-square (χ^2), the complete model provided a better description of the variation (Mota *et al*. 2013). This led to selection of RRM-4333. However, in order to do an appropriate comparison with MTM-adjusted weights, an equivalent RRM without maternal permanent environmental effects, i.e. RRM-433, which was not significantly different from RRM-4333 through LRT, was used for estimating parameters.

As a residual effect in MTM is equivalent to the sum of direct permanent environmental and residual effects in RRM, when comparing both models, residual variances in RRM must be

considered as a composed variance of direct permanent environmental and residual effects (Nobre *et al*. 2003; Legarra *et al*. 2004). In RRM, the EBV of the l^{th} animal at age t was calculated as follows:

$$EBV_{lt} = \sum_{n=0}^{k_a-1} \alpha_{ln} \phi_n \left(age_t \right)$$

Coefficients of the Legendre polynomials and the resulting statistics log L, AIC, BIC, (co)variance components, genetic parameters, EBVs in MTM analysis, and random regression coefficients in RRM were calculated using the WOMBAT software implementation of restricted maximum likelihood method (REML) (Meyer 2007). Genetic trends were obtained from the EBVs by linear regression on birth year.

3.4 Results

The fixed regression curve by RRM-433 that described the changes in weight over time was nearly linear and approximately parallel to the trend in the observed weights indicating attainment of 500 kg by 463 days of age (Figure 3.2). Estimates of variance components and genetic parameters from the MTM scenarios and RRM-433 for growth traits of BW, WW and YW were presented in Table 3.4. For BW, estimates of the variance components and genetic parameters were unaffected by the analytical procedure when considering CG as fixed. For random CG, despite an increase in the Var (p) and as a result a decrease in h^2_a and h^2_m, no changes were observed in variance components. For both WW and YW, the Var (p) estimated with the RRM was greater than the corresponding estimate from the MTMs with fixed CG due primarily to the difference in the estimates of Var (e). This resulted in the estimated (h^2_a) for WW from the RRM being marginally less than that from the MTM using pre-adjusted weights and considerably less than the estimates from MTM with that incorporated simultaneous

adjustment for age due to larger estimates of additive genetic variance (Var (a)). Similarly, for YW, although the Var (a) was greater for the RRM than the MTM with pre-adjusted weights, both methods produced similar estimates of $h^2_{a.}$ The MTM in which age was accounted for simultaneously and CG were considered fixed produced a larger h^2_a due to a lower Var (p) and a higher Var (a) than the corresponding values from RRM and MTM with pre-adjusted weights, respectively. Considering CG random, h^2_a for WW was smaller than the respective values for MTM with fixed CG due to a larger Var (p). However, it was still slightly higher than RRM and MTM with pre-adjusted weights. However, in terms of YW when CG was random, the resulting h^2_a was lower than the other models due to having the largest Var (p). For all three traits the estimates of direct additive genetic variance were not detectably different whether CG were considered random or fixed. The estimates of maternal additive genetic variance (Var (m)) were small fractions of the corresponding Var (p) and except for WW in MTM-adjusted weights, the estimates of maternal heritability (h^2_m) were near or essentially zero.

Over the span of ages between 123 d and 474 d, estimates of variance components estimated using RRM fluctuated most markedly in intervals where the data were relatively sparse. This was particularly true after 365 d for all variances (Figures 3.3 and 3.4). However, both h^2_a and h^2_m reduced after yearling and weaning ages in RRM, respectively (Figure 3.5).

Estimates of the direct additive genetic correlations (r(a)) of BW and subsequent weights were similar and greater when estimated with different MTMs than with RRM, and likewise between WW and YW (Table 3.5). Estimates of maternal additive genetic correlations (r(m)) in RRM were greater between BW and WW or YW, and less between WW and YW. There were no differences in r(m) between BW and WW for all MTM scenarios. Moreover, the r(m) between BW and YW or WW and YW in MTM-adjusted weights showed the least and greatest

magnitudes, respectively. Although slightly greater, the estimates of direct permanent environmental correlations from RRM were similar to those for residual correlations in all MTM scenarios. When using CG as a random effect in MTM, estimates of CG correlations (r(CG)) between BW and WW or YW was less than WW with YW. Estimates of phenotypic correlations in both models were similar and slightly less in RRM.

Based on the EBVs estimated by RRM, the genetic trends for direct effects on BW, WW, and YW were -3.6 ± 0.7, 116 ± 8, and 280 ± 21 g/yr, respectively (Table 3.6). The corresponding estimates of genetic trend based on MTM-adjusted weights were -7.1 ± 1, 63.3 ± 9.6, and -73 ± 16 g/yr, respectively. Although using CG as a random effect in MTMs resulted in substantially lower genetic trends for all weight traits than using fixed effects, all the six MTMs were in accordance with MTM-adjusted weights results. In total, estimates of the genetic trends from MTM were clearly less than those from RRM. Generation numbers in the recorded Hays Converter pedigree ranged from a minimum 0 for animals with unknown parents to a maximum of 5.16 over the period 1970 to 2016. The average generation interval was 5.35 ± 0.07 yr. In order to realize how sire selection over the past years has affected the genetic trends, the selection differential (SD) of HC sires were compared to the best males available that were not chosen as sires (Figure 3.6). Irrespective of how the data were analyzed, there was a consistent loss of selection pressure on the individual traits relative to the opportunities that existed in the population.

3.5 Discussion

Different orders of Legendre polynomials were evaluated with various statistical criteria to find the RRM that was used to describe the variation in body weights over time for the HC population. Baldi *et al*. (2010) suggested taking precision into account when adding random

effects (including direct additive genetic, maternal additive genetic, animal permanent environmental and maternal permanent environmental effects) to a model due to the potential for over-parameterization. Moreover, convergence problems and susceptibility to numerical errors may be avoided by excluding non-essential parameters from the model (Arango *et al*. 2004; Legarra *et al*. 2004). Although in utilizing RRM to evaluate milk test day records, Jamrozik and Schaeffer (2002) indicated that models may be assessed differently when using different statistical criteria, and which model would be the most suitable may be unclear. Similar issues were observed in this study.

The main goal of creating HC was to develop an animal that would efficiently convert feed to gain and reach the desirable market weight of 500 kg at the earliest age possible (Fleming *et al*. 2016). Considering the fixed regression curve in Figure 3.2, the attainment of 500 kg by 463 days of age was deemed consistent with the breeding goal of reaching desirable market weight at a young age. However, it only addresses that component of efficient conversion of feed to gain that results from avoiding the additional feed consumed in satisfying maintenance requirements over a longer time on feed (Nielsen *et al*. 2013).

To date, RRMs have most frequently used orthogonal (Legendre) polynomials because they flexibly model changes in variance and covariance along a continuous scale, especially at higher orders (Meyer 2005b). However, observations at the extremes are over-emphasized and this may be problematic for models that are parameterized in this way. In fact, as 'Runge's phenomenon' describes, implausible errors in variance component estimates may be observed at the ends of the growth trajectory due to small numbers of extreme observations and higher orders of polynomials (de Boor 2001; Meyer 2005b; Meyer and Kirkpatrick 2005). This is consistent with the results obtained here. Moreover, in contrast to MTM, which incorporate information

among traits only through linear covariances, RRM allow for more complex global consideration of information over the whole curve (Meyer 2005b). Similar results of unexpected estimates beyond biological reality have been observed for variance components analysis by RRM in other analyses of data from beef cattle (Boligon *et al*. 2010; Mota *et al*. 2013). In this study, B-spline RR models were not used as a panacea for RRM because they are also susceptible to the sparcity and irregularity of records distribution and choosing suitable knots and degrees of B-splines would not be convenient (Meyer 2005c). An increased frequency of data recording may not be feasible as it would increase the cost to weigh animals on a more frequent basis from birth to beyond a year of age. As a potential alternative, if the maximum degree of polynomials were established in advance, the ages at which to record weights in order to maximize the precision of random regression coefficients could be determined from statistical theory.

Usually, the number of weight records in beef cattle production is dependent on the length of growth trajectory which differs among breeds. For example, Nellore cattle typically show more data than European breeds (Albuquerque and Meyer 2001). Therefore, as mentioned above, in shorter times, getting more data points than BW, WW and YW would be possible if the relative costs are provided to weigh animals regularly. In this study, although there were fewer points available on the growth trajectory than those normally applied in RRM, the purpose was to see how the results deviated from MTM specially when working with local beef herds.

On the other hand, when there are more points available for longer trajectories, the frequency of data for those points are more important to affect the analysis than the number of points. For example, in a research study done by Meyer (2005a) regarding the use of RRM to analyze the growth curve of Australian Angus cattle, although there were more records available for the growth trajectory than this study, only 1.5% of the animals had 7-9 records and they

mostly showed four main critical points on average. Moreover, considering Boligon *et al*. (2010) and Oliveira *et al*. (2017) results, even with more points available, it was still possible to observe extreme values at the boundaries due to selecting a model with higher orders that matched the statistical criteria and not what might be reasonable with biological realities. In other words, if there are data with high frequency for the critical points of a growth trajectory, RRM will be more sensitive to the orders than the number of points and/or the length of the growth trajectory.

However, it could be argued that in the case of HC where weaning typically was at a younger age than the 205-d standard, RRM would allow more data to be used in prediction of a 205-d weight EBV than if weaning weights were edited to the 160 to 250 day window recommended by the Beef Improvement Federation (2018). This increase in the amount of data used is expected to increase accuracy of EBV for some selection candidates due to their own phenotypes being included in the analysis (Meyer 2004; Bohmanova *et al*. 2005; Mota *et al*. 2013). Furthermore, according to a research done by Bohmanova *et al*. (2005), for a specific length of growth trajectory (similar to this study), although incorporating additional records in RRM increased the accuracy compared to MTM, the change in accuracy would be small enough to conclude that both models performed similarly. Therefore, as observed in the variation results, even with fewer data points on the growth trajectory, RRM performed similar to MTM so that there may be no advantage to get more frequent data points.

Maternal effects are typically thought to be important from birth to weaning age and then gradually decreasing to the end of growth trajectory. In the present study, estimates of maternal genetic effects other than on WW were not significant. In addition, maternal permanent environmental effects accounted for negligible proportion of phenotypic variance (not shown). This may be logical in that for HC there were few calves per cow and virtually all cows produced

calves in a single herd. Therefore, as opposed to Boligon *et al*. (2010) and Mota *et al*. (2013), the exclusion of maternal permanent environmental effects not only did not affect the overall parameter estimation but also provided a better approximation for standard errors which could be considered when using RRM.

In this study, despite the usual MTM that employs adjusted weights in the analysis, other MTMs were also defined in which WW and YW records were not subjected to a priori adjustment, but rather their real ages were incorporated into the models as both linear and quadratic covariates. This increased the estimation of genetic variation and consequently estimates of direct heritability which implied that the pre-adjustments of records to defined ages in MTM and/or some standardization of ages through the process of RRM may lead to a reduction in the corresponding genetic variation.

The expanded range in ages at which weights were deemed acceptable for incorporation into the analysis (dataset 2) did not result in any specific changes of the estimated variance components and genetic parameters because the relaxed restrictions on age did not add many records to the analyses. Additionally, incorporation of ages into the model for dataset 2 as quadratic effects also did not affect the variance component estimates. This latter result may reflect the observation that weight changes across the growth trajectory were mainly linear.

Adding CG as fixed or random did not affect the overall genetic variation and other variance components except for an increase in Var (p) and consequent reduction in estimates of direct heritability. Contemporary groups are defined as a group of animals that benefit from common environmental and management factors. They are usually taken into account as fixed effects in animal models to make the results of genetic evaluations invariant and reduce biases in genetic comparisons due to the association between CG and sires (Van Vleck, 1987). However,

if they are random, the prediction error variance will be reduced due to using a larger amount of data for prediction of animal breeding values (Visscher and Goddard 1993). Today, there is no consensus about the best method of applying this effect. However, in general, CG is considered fixed in beef cattle genetic evaluations (Ramirez-Valverde *et al*. 2008). Likewise, many published literature regarding applications of RRM and MTM in beef cattle that have been referred to here, used CGs as fixed effects and hence the results of this study were more following and similar to them. Of course, some studies have shown that random CGs would be a better choice if there are numerous levels of this effect, small subclasses are predominant and limited use of AI in the population has led to a weak genetic connectedness among them (Schaeffer 2009; Vostry *et al*. 2015). Although having several herds across Canada, the HC data for this study originated from one farm so that there were not many levels or small subclass of CGs available and regarding the similarity in MTM results, it seemed reasonable to consider CGs as fixed effects.

Fleming (2013) estimated the direct heritability of BW in HC to be 0.06 through a univariate analysis. Thus, the value reported by Fleming (2013) was slightly less than the present estimate of approximately 0.10. However, Fleming (2013) also obtained greater heritability estimates for WW (0.30) and YW (0.42) using a bivariate model than were observed in this study for RRM and MTM-adjusted weights. Estimates of maternal heritability for BW, while still near zero, were greater in Fleming (2013) (0.03 vs 0.01), but greater for WW in the present study (0.13 vs 0.04). However, the MTM for the analysis of datasets 1 and 2 with fixed CG produced similar estimates of heritability for WW and YW to those of Fleming (2013). The almost zero estimation of maternal heritability for BW might be related to the low number of calves per cow available and/or simply that in HC BW was not affected by the heritable factors influencing the

uterine environment (Ferrell 1993). However, for WW, the higher maternal heritability likely reflects differences in milk production (MacNeil and Mott 2006).

Positive genetic correlations may result from a part-whole relationship between traits. For example, weaning weight makes up a part of yearling weight. However, the very large values in MTM may also arise from pre-adjustment of the data before analysis (Iwaisaki *et al*. 2005). According to Boligon et al. (2010) (although they did not report the standard errors), the similar direct and maternal genetic correlations in RRM suggest that these effects are likely controlled by the same genes and that are considered similar between different traits. However, although there was a weak maternal genetic correlation between BW with WW and YW in MTM, the resulting standard errors were high and similar to RRM which reflected the poor structure of HC data. Moreover, the moderate to high maternal genetic correlation between WW and YW in both models indicates that the maternal effect on YW is probably a carry-over effect from WW (Boligon *et al*. 2010).

With respect to RRM, BW appeared not to be very genetically correlated to WW and YW, which would be desirable from an economic point of view when selecting bulls with lower BW to facilitate ease of calving. The somewhat lower direct genetic correlation between BW and WW in RRM might be due better modeling of age in both fixed and random effects compared to MTM (Iwaisaki *et al*. 2005). In this study, direct additive genetic correlation between WW and YW in MTM was slightly greater than Fleming's result (0.81) in 2013.

Except for genetic correlations, the estimates of variance components and heritability for BW were approximately of the same magnitude in comparable MTM and RRM models. This reflected the importance of the fact that there are similarities in using MTM or RRM when enough data is available for each time point in the growth curve. Compared to BW and WW,

there was a slightly greater difference between the two models in measuring the changes over time for YW. This may originate from the low number of actual records available for 365 days of age in RRM as an end point relative to MTM. The lower magnitude of the genetic correlation between WW and YW in RRM may have resulted in higher differences between their genetic values relative to MTM. However, despite YW, EBV trends for other traits approximately followed a similar pattern in both models. Furthermore, as there is a positive genetic correlation between WW and YW, it seemed RRM estimated a more realistic increasing trend for both traits than MTM which showed a decreasing trend for YW.

Genetic trends for weight traits in HC were directly affected by the sires selected to produce calves in each year. Sire selection did not always maximize the genetic selection differentials (Figure 3.6). Sire selection always used the phenotypic records of weaning and yearling weights and there was no selection based on EBV or an index until 2014. In fact, lack of a structured management program that takes into account genetic values when selecting animals, has resulted in very little meaningful progress in genetic improvement of growth traits of a local beef breed like HC. Additionally, a large proportion of the herd was sold in 2000 leading to a meaningful reduction in the number of candidates for selection. During 2004 to 2008, little sire selection was practiced which may have also contributed to the decreasing trend in genetic values of weight traits in recent years.

3.6 Conclusion

Results of this study suggest similarities between RRM and MTM for most estimates of variance components and genetic trends of HC. This is mainly because records occur at standard points for both models. Currently, choosing MTM for HC genetic evaluation seems simpler. Increased weaning and yearling weights were the main objectives for HC from its inception.

Tighter control of when these traits are recorded and replacement of selection based on phenotypes with selection based on EBV are expected to accelerate progress toward this goal. In these data, whether CG were considered random or fixed had little effect and thus fixed CG were deemed preferable due to their being more parsimonious with other genetic evaluations for beef cattle.

3.7 References

Albuquerque, L.D., Meyer, K., 2001. Estimates of covariance functions for growth from birth to 630 days of age in Nelore cattle. J. Anim. Sci. 79, 2776-2789.

Arango, J.A., Cundiff, L.V., Van Vleck, L.D., 2004. Covariance functions and random regression models for cow weight in beef cattle. J. Anim. Sci. 82, 54-67.

Baldi, F., Albuquerque, L.G., Alencar, M.M., 2010. Random regression models on Legendre polynomials to estimate genetic parameters for weights from birth to adult age in Canchim cattle. J. Anim. Breed. Genet. 127, 289-299.

Beef Improvement Federation., 2018. Guidelines for Uniform Beef Improvement Programs. 9th ed. https://beefimprovement.org/wp-content/uploads/2018/03/BIFGuidelinesFinal_updated0318.pdf (accessed 5 January 2019).

Bohmanova, J., Misztal, I., Bertrand, J.K., 2005. Studies on multiple trait and random regression models for genetic evaluation of beef cattle for growth. J. Anim. Sci. 83, 62-67.

Boligon, A.A., Mercadante, M.E.Z., Forni, S., Lobo, R.B., Albuquerque, L.G.D., 2010. Covariance functions for body weight from birth to maturity in Nellore cows. J. Anim. Sci. 88, 849-859.

Brinks, J. S., Clark, R.T., Rice, F.J., 1961. Estimation of genetic trends in beef cattle. J. Anim. Sci. 20, 903.

de Boor, C., 2001. A practical guide to splines. Springer series in applied mathematics, vol. 27, second ed. Springer, New York.

Delgadillo, J.S., Riley, D.G., Langdon, J.M., Hanna, L.H., Herring, A.D., 2017. 178 Black Hereford genetic parameters and predictions for calf traits with alternate modelling strategies. J. Anim. Sci. 95(suppl. 4), 88.

Farquharson, R.J., Griffith, G.R., Barwick, S.A., Banks, R.G., Holmes, W.E., 2003. Estimating the returns from past investment into beef cattle genetic technologies in Australia. Econ. Res. Rep. 15, 3.

Ferrell, C.L., 1993. Factors influencing fetal growth and birth weight in cattle. Roman L. Hruska U.S. Meat Animal Res. Center 132. https://digitalcommons.unl.edu/hruskareports/132

Fleming, A., 2013. A genetic characterization of the Hays Converter. M.Sc. thesis. University of Guelph, ON, Canada.

Fleming, A.L., Hays, D.P., Moore, S., Plastow, G., Miller, S.P., 2016. History of the Hays Converter. Can. J. Anim. Sci. 96, 471-477.

Iwaisaki, H., Tsuruta, S., Misztal, I., Bertrand, J.K., 2005. Genetic parameters estimated with multitrait and linear spline-random regression models using Gelbvieh early growth data. J. Anim. Sci. 83, 757-763.

Jamrozik, J., Schaeffer, L.R., 2002. Bayesian comparison of random regression models for test-day yields in dairy cattle. Proc. 7th World Cong. Genet. Appl. Livestk. Prod. Montpellier, France. Communication (No. 01-03).

Legarra, A., Misztal, I., Bertrand, J.K., 2004. Constructing covariance functions for random regression models for growth in Gelbvieh beef cattle. J. Anim. Sci. 82, 1564-1571.

MacNeil, M.D., Mott, T.B., 2006. Genetic analysis of gain from birth to weaning, milk production, and udder conformation in Line 1 Hereford cattle. J. Anim. Sci. 84, 1639-1645.

Meyer, K., 1999. Estimates of genetic and phenotypic covariance functions for postweaning growth and mature weight of beef cows. J. Anim. Breed. Genet. 116, 181-205.

Meyer, K., 2001. Estimates of direct and maternal covariance functions for growth of Australian beef calves from birth to weaning. Genet. Sel. Evol. 33, 487-514.

Meyer, K., 2004. Scope for a random regression model in genetic evaluation of beef cattle for growth. Livestk. Prod. Sci. 86, 69-83.

Meyer, K., 2005a. Estimates of genetic covariance functions for growth of Angus cattle. J. Anim. Breed. Genet. 122, 73-85.

Meyer, K., 2005b. Advances in methodology for random regression analyses. Aust. J. Exp. Agric. 45, 847-858.

Meyer, K., 2005c. Random regression analyses using B-splines to model growth of Australian Angus cattle. Genet. Sel. Evol. 37, 473-500.

Meyer, K., 2007. WOMBAT—A tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). J. Zhejiang Uni. Science B, 8, 815-821.

Meyer, K., 2018. wombat:sterrors. The Wicked WOMBAT Wiki. http://didgeridoo.une.edu.au/womwiki/doku.php?id=wombat:sterrors&rev=1536881200 (accessed 12 January 2019).

Meyer, K., Kirkpatrick, M., 2005. Up hill, down dale: quantitative genetics of curvaceous traits. Phil. Trans. R. Soc. B. 360, 1443-1455.

Mofakkarul Islam, M., Renwick, A., Lamprinopoulou, C., Klerkx, L., 2013. Innovation in livestock genetic improvement. EuroChoices. 12, 42-47.

Mota, R.R., Marques, L., Lopes, P.S., Da Silva, L.P., Neto, F., De Resende, M., Torres, R.A., 2013. Genetic evaluation using multi-trait and random regression models in Simmental beef cattle. Genet. Mol. Res.12, 2465-2480.

Nielsen, M.K., MacNeil, M.D., Dekkers, J.C.M., Crews Jr., D.H., Rathje, T.A., Enns, R.M., Weaber, R.L., 2013. REVIEW: Life-cycle, total industry genetic improvement of feed efficiency in beef cattle: Blueprint for the Beef Improvement Federation. Professional Animal Scientist 29, 559-565.

Nobre, P.R.C., Misztal, I., Tsuruta, S., Bertrand, J.K., Silva, L.O.C., Lopes, P.S., 2003. Analyses of growth curves of Nellore cattle by multiple-trait and random regression models. J. Anim. Sci. 81, 918-926.

Olfert, E.D., Cross, B.M., McWilliam, A.A., Eds., 1993. Guide to the Care and Use of Experimental Animals, 2nd ed. Canadian Council on Animal Care, Ottawa.

Oliveira, M.R., Azevêdo, D.M., Malhado, C., Pires, L., Martins Filho, R., Sousa Júnior, S., 2017. Weight evaluation of Tabapuã cattle raised in northeastern Brazil using random-regression models. Arq. Bras. Med. Vet. Zootec. 69, 457-464.

Olori, V.E., Hill, W.G., McGuirk, B.J., Brotherstone, S., 1999. Estimating variance components for test day milk records by restricted maximum likelihood with a random regression animal model. Livestock Prod. Sci. 61, 53-63.

Ramirez-Valverde, R., Nunez-Dominguez, R., Ruiz-Flores, A., García-Muñiz, J.G., Magana-Valencia, F., 2008. Comparison of contemporary group definitions for genetic evaluation of Braunvieh cattle. Téc. Pecu. Méx. 46, 359-370.

Schaeffer, L.R., 2004. Application of random regression models in animal breeding. Livestock Prod. Sci. 86, 35-45.

Schaeffer, L.R., 2009. Contemporary Groups Are Always Random. http://animalbiosciences.uoguelph.ca/~lrs/piksLRS/ranfix.pdf (accessed 25 June 2019).

Van Vleck, L.D., 1987. Contemporary groups for genetic evaluations. J. Dairy. Sci. 70, 2456-2464.

Visscher, P.M., Goddard, M.E., 1993. Fixed and random contemporary groups. J. Dairy. Sci. 76, 1444-1454.

Vostry, L., Milerski, M., Krupa, E., Vesela, Z., Vostra-Vydrova, H., 2015. Genetic relationships among calving ease, birth weight and perinatal calf survival in Charolais cattle. Animal Science Papers and Reports. 33, 233-242.

3.8 Figures and tables

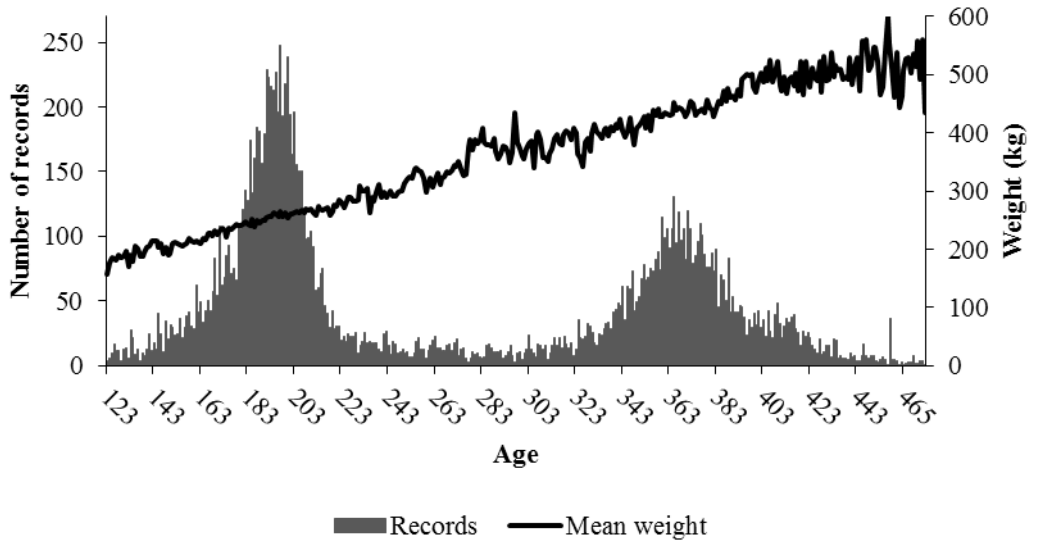

Figure 3.1. Numbers of records at each age and trajectory of average weight (kg) in the data.

Figure 3.2. Plots of the fixed regression curve of weight on age from the random regression analysis and average weight across the growth trajectory of Hays Converter cattle.

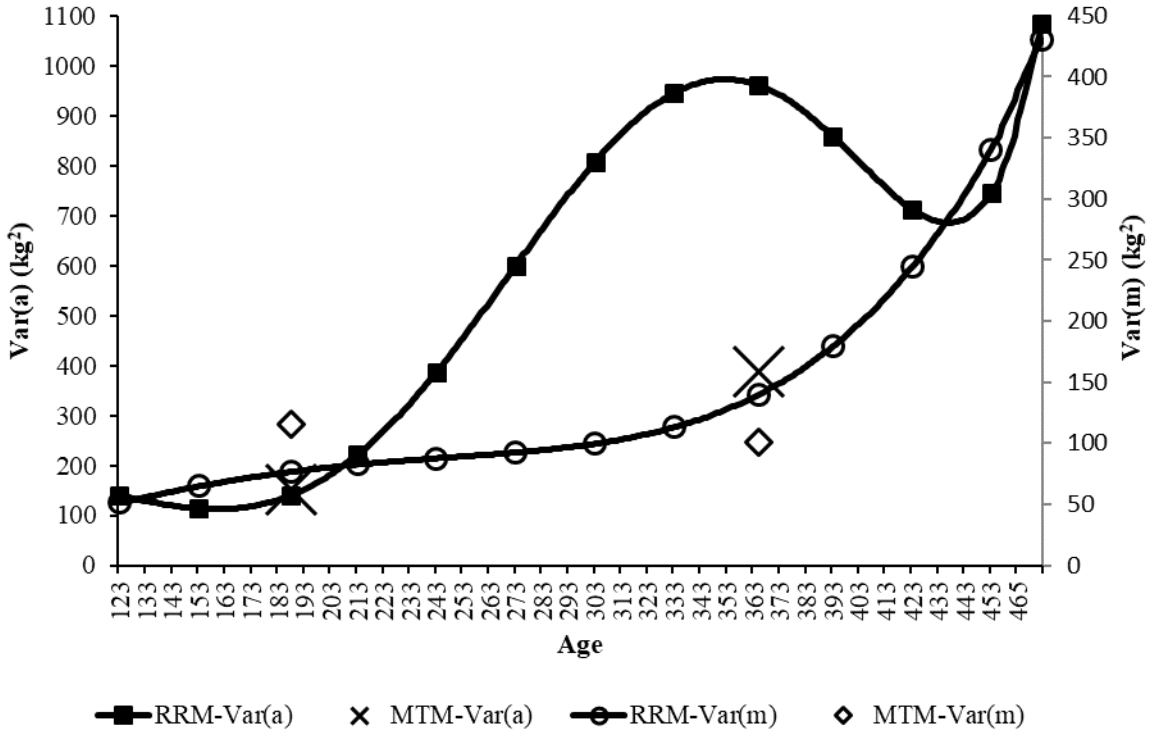

Figure 3.3. Estimates of Var (a) and Var (m) in RRM-433 together with MTM-adjusted weights along the growth trajectory.

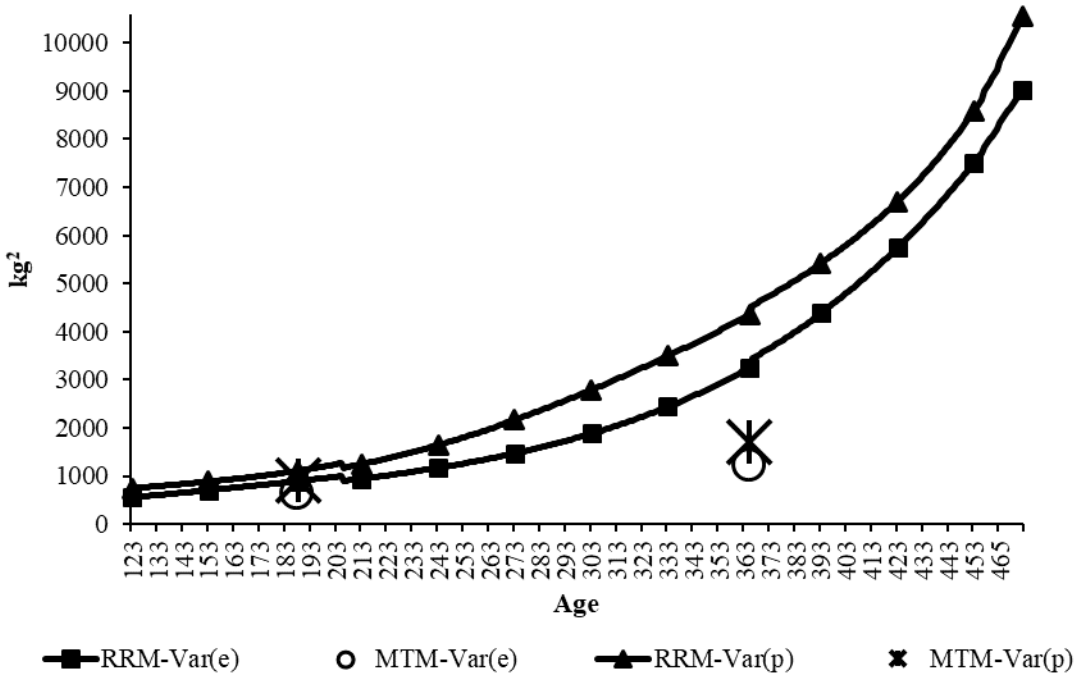

Figure 3.4. Estimates of Var (e) and Var (p) by RRM-433 and MTM-adjusted weights along the growth trajectory.

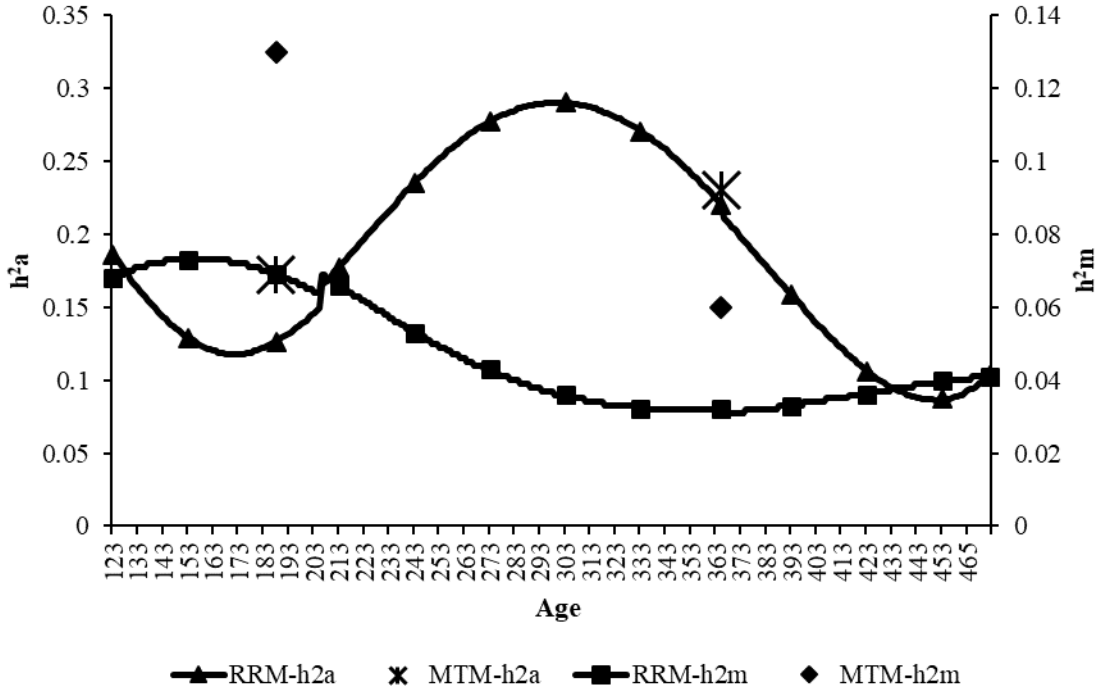

Figure 3.5. Estimates of direct and maternal heritabilities in RRM-433 and MTM-adjusted weights along the growth. trajectory.

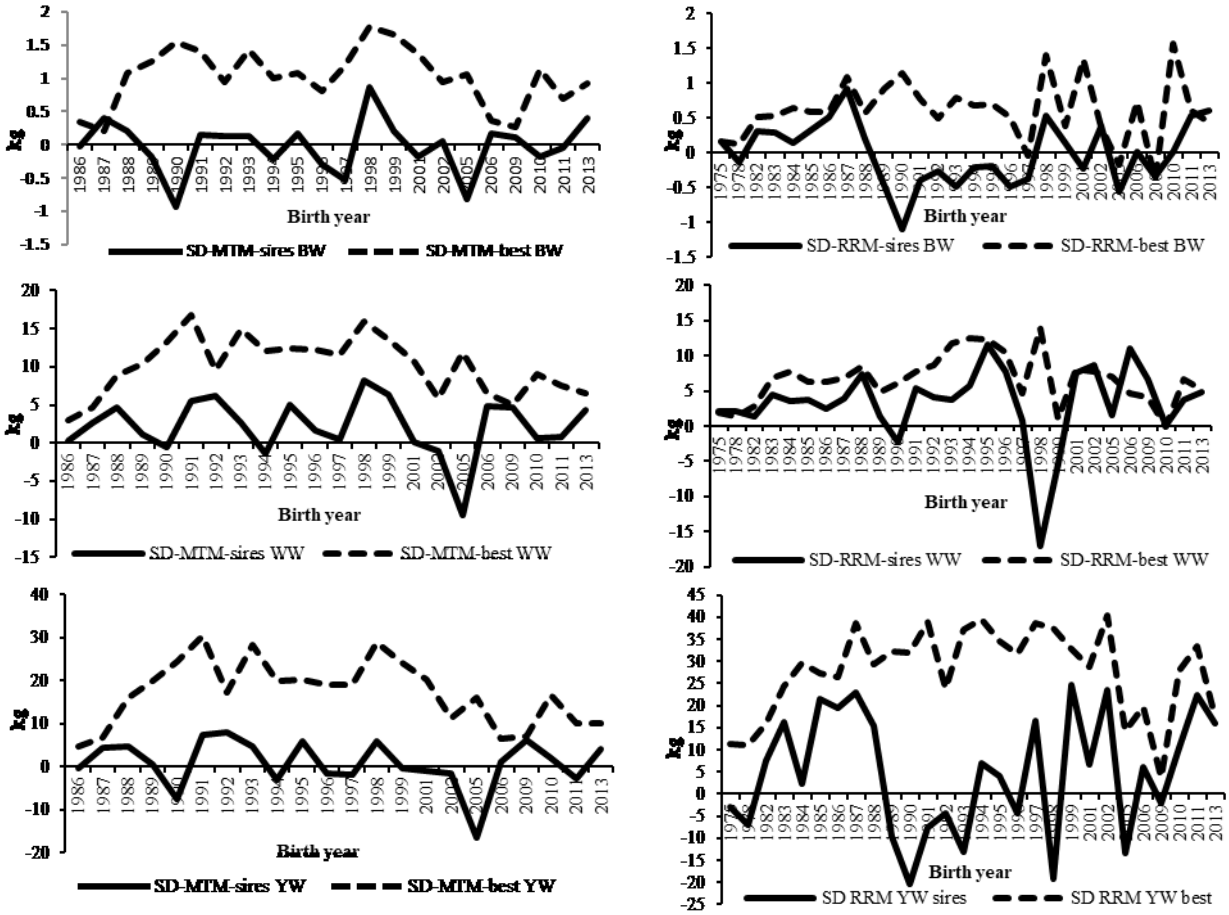

Figure 3.6. Selection differentials (SD) of HC sires and best available males for BW, WW and YW in MTM-adjusted weights and RRM-433.

Table 3.1. Descriptive statistics detailing distributions of ages at the recording of weaning and yearling weights, and the weight traits along with number of sires, dams and their daughters as dams. Statistics from the data set after editing are shown parenthetically.

	N	Mean	SD	Min	Max
Weaning age (days)	8748 (7776)	188.7	22.0	100 (123)	315 (254)
Yearling age (days)	6474 (5936)	375.9	32.8	245 (280)	544 (474)
Birth weight (kg)	7119 (6519)	39.7	5.7	13.6 (22.7)	74.8 (56.7)
Weaning weight (kg)	8585 (7776)	247.3	40.5	61.7 (126.1)	412.8 (365.1)
Yearling weight (kg)	6396 (5936)	444.0	79.3	138.8 (217.7)	739.4 (679.9)
Sires	137				
Dams	1701				
Daughters of sires as dams	1069				
Daughters of dams as dams	291				

Table 3.2. Structure of data for analysis with RRM with reference to the number of records per animal.

	No. of animals	Percentage
with records	8850	100.00%
1 record	1425	16.10%
2 records	3388	38.30%
3 records	3801	42.90%
7-9 records	236	2.70%
Records in total	21612	

Table 3.3. Different orders of Legendre polynomials in RRM along with statistical criteria of log L, AIC and BIC.

	Order of polynomials					Statistical criteria		
Model	K_a[a]	K_m[b]	K_p[c]	K_c[d]	np[e]	log L	AIC	BIC
3232	3	2	3	2	22	-73862.0	147768.0	147943.2
3233	3	2	3	3	25	-73845.2	147740.3	147939.5
3332	3	3	3	2	25	-73833.3	147716.5	147915.7
3333	3	3	3	3	28	-73826.5	147709.0	147932.1
3433	3	4	3	3	32	-73768.4	147600.8	147855.8
3343	3	3	4	3	32	-73724.2	147512.4	147767.3
3353	3	3	5	3	37	-73603.3	147280.5	147575.3
4332	4	3	3	2	29	-73723.5	147505.1	147736.1
4333	4	3	3	3	32	-73709.8	147483.6	147738.6
4432	4	4	3	2	33	-73713.1	147492.3	147755.2
4342	4	3	4	2	33	-73698.9	147463.8	147726.8
4343	4	3	4	3	36	-73691.4	147454.7	147741.6
4344	4	3	4	4	40	-73689.4	147458.9	147777.6
4443	4	4	4	3	40	-73687.9	147455.9	147774.6
5332	5	3	3	2	34	-73561.8	147191.6	147462.5
5333	5	3	3	3	37	-73554.7	147183.5	147478.3
5432	5	4	3	2	38	-73559.1	147194.2	147497
5342	5	3	4	2	38	-73529.5	147135.0	147437.8
5343	5	3	4	3	41	**-73520.9**	**147123.8**	147450.4

[a]order of fit for direct additive genetic, [b]maternal additive genetic, [c]direct permanent environmental and [d]maternal permanent environmental effects

[e]number of parameters

Table 3.4. Estimates of variance components and genetic parameters for birth weight (BW), weaning weight (WW) and yearling weight (YW) from different MTM scenarios and RRM-433 analyses.

		Variance components[a]					Genetic parameters[b]	
Trait	**Model**	Var (a)	Var (m)	Var (CG)	Var (e)[c]	Var (p)	h^2_a	h^2_m
BW	RRM	1.75±0.48	0.21±0.19	-	16.49	18.45±0.34	0.10±0.03	0.01±0.01
	MTM-adjusted weights	1.86±0.48	0.19±0.19	-	16.2±0.46	18.23±0.33	0.10±0.03	0.01±0.01
	MTM-data1-CG fixed	1.85±0.47	0.19±0.18	-	16.19±0.46	19.07±0.33	0.10±0.02	0.01±0.01
	MTM-data1-CG random	2.02±0.48	0.19±0.18	11.43±1.50	16.01±0.46	30.59±1.53	0.07±0.02	0.006±0.006
	MTM-data2-CG fixed	1.82±0.47	0.20±0.18	-	16.21±0.46	18.91±0.33	0.10±0.02	0.01±0.01
	MTM-data2-CG random	1.99±0.47	0.20±0.18	11.32±1.5	16.03±0.46	30.42±1.52	0.07±0.02	0.006±0.006
	MTM-data2-CG fixed-age quadratic	1.81±0.47	0.19±0.18	-	16.21±0.46	18.87±0.33	0.10±0.02	0.01±0.01
	MTM-data2-CG random-age quadratic	1.99±0.47	0.19±0.18	11.30±1.49	16.03±0.46	30.38±1.52	0.07±0.02	0.006±0.006
WW	RRM	140.35±24.71	76.51±14.83	-	892.97	1109.84±18.94	0.13±0.02	0.07±0.01
	MTM-adjusted weights	150.9±34.17	115.5±19.8	-	629.52±26.45	895.94±19.91	0.17±0.04	0.13±0.02
	MTM-data1-CG fixed	236.82±34.45	48.68±10.41	-	599.81±27.19	885.31±19.48	0.27±0.04	0.05±0.01
	MTM-data1-CG random	229.75±33.26	47.58±10.07	332.24±49.01	604.56±26.63	1214.1±52.17	0.19±0.03	0.04±0.008
	MTM-data2-CG fixed	238.09±33.52	44.59±10.12	-	594.04±26.31	876.73±18.9	0.27±0.03	0.05±0.01
	MTM-data2-CG random	229.31±32.25	43.99±9.8	326.27±47.43	599.05±25.72	1198.6±50.52	0.19±0.03	0.04±0.008
	MTM-data2-CG fixed-age quadratic	234.94±33.47	45.05±10.23	-	595.9±26.29	875.9±18.86	0.27±0.03	0.05±0.01
	MTM-data2-CG random-age quadratic	226±32.14	44.52±9.9	323.04±47.02	601.21±25.68	1194.8±50.12	0.19±0.03	0.04±0.008
YW	RRM	960.8±118.58	139.82±50.23	-	3259.44	4360.06±88.93	0.22±0.03	0.03±0.01
	MTM-adjusted weights	387.93±84.37	100.97±35.97	-	1215.15±62.88	1704.05±44.08	0.23±0.05	0.06±0.02
	MTM-data1-CG fixed	574.5±80.31	-	-	1112.8±65.9	1687.3±46.33	0.34±0.04	-
	MTM-data1-CG random	570.08±78.89	-	3234.3±419.14	1118.4±65.06	4922.8±420.85	0.12±0.02	-
	MTM-data2-CG fixed	622.73±79.84	-	-	1148.1±64.19	1770.8±45.7	0.35±0.04	-
	MTM-data2-CG random	610.3±78	-	3189.6±402.48	1157.7±63.20	4957.6±404.35	0.12±0.02	-
	MTM-data2-CG fixed-age quadratic	614.15±78.42	-	-	1126.9±63.04	1741.1±45.04	0.35±0.04	
	MTM-data2-CG random-age quadratic	601.74±76.58	-	3220.6±405.52	1136.3±62.04	4958.6±407.3	0.12±0.02	-

[a]Var (a) = direct additive genetic variance, Var (m) = maternal additive genetic variance, Var (CG) = contemporary group variance, Var (e) = residual variance , Var (p) = phenotypic variance

[b]h^2_a = direct heritability, h^2_m = maternal heritability

[c]For the RRM = sum of estimates of variance for permanent environmental effects due to animals and residual

Table 3.5. Estimates of direct and maternal additive genetic, contemporary group, residual (permanent environmental in RRM) and phenotypic correlations by RRM-433 and different MTM scenarios.

Traits		WW	YW	WW	YW	WW	YW	WW	YW	WW	YW
	Models	***r*(a)[a]**		***r*(m)[b]**		***r*(CG)[c]**		***r*(e)[d]**		***r*(p)[e]**	
BW	RRM	0.21±0.15	0.47±0.12	0.43±0.31	0.56±0.40	-	-	0.17±0.11	0.10±0.07	0.16±0.02	0.15±0.02
	MTM-adjusted weights	0.67±0.13	0.77±0.12	0.28±0.31	-0.11±0.47	-	-	0.14±0.02	0.11±0.03	0.21±0.01	0.20±0.02
	MTM-data1-CG fixed	0.60±0.12	0.65±0.11	0.27±0.33	-	-	-	0.14±0.02	0.12±0.03	0.21±0.01	0.21±0.02
	MTM-data1-CG random	0.53±0.12	0.62±0.11	0.31±0.34	-	0.28±0.09	0.29±0.09	0.15±0.02	0.12±0.03	0.23±0.03	0.24±0.04
	MTM-data2-CG fixed	0.59±0.12	0.60±0.11	0.23±0.33	-	-	-	0.14±0.02	0.12±0.03	0.20±0.01	0.20±0.01
	MTM-data2-CG random	0.52±0.12	0.61±0.11	0.3±0.33	-	0.26±0.10	0.20±0.09	0.15±0.02	0.12±0.03	0.22±0.03	0.19±0.04
	MTM-data2-CG fixed-age quadratic	0.59±0.12	0.60±0.11	0.22±0.33	-	-	-	0.14±0.02	0.12±0.03	0.20±0.01	0.20±0.01
	MTM-data2-CG random-age quadratic	0.52±0.12	0.62±0.11	0.29±0.34	-	0.26±0.10	0.21±0.09	0.15±0.02	0.11±0.03	0.22±0.03	0.20±0.04
WW	RRM		0.54±0.07		0.55±0.13		-		0.71±0.02		0.61±0.01
	MTM-adjusted weights		0.90±0.05		0.92±0.07		-		0.65±0.02		0.72±0.01
	MTM-data1-CG fixed		0.91±0.03		-		-		0.64±0.02		0.70±0.01
	MTM-data1-CG random		0.92±0.03		-		0.50±0.08		0.64±0.02		0.56±0.03
	MTM-data2-CG fixed		0.89±0.04		-		-		0.56±0.02		0.65±0.01
	MTM-data2-CG random		0.90±0.04		-		0.45±0.08		0.56±0.02		0.52±0.03
	MTM-data2-CG fixed-age quadratic		0.88±0.04		-		-		0.56±0.02		0.64±0.01
	MTM-data2-CG random-age quadratic		0.90±0.04		-		0.46±0.08		0.56±0.02		0.52±0.03

[a]direct additive genetic correlation

[b]maternal additive genetic correlation

[c]contemporary group correlation

[d]residual correlation in MTM and direct permanent environmental correlation in RRM

[e]phenotypic correlation

Table 3.6. Estimates of genetic trends (g/yr) for RRM-433 and different MTM scenarios.

Models	BW	WW	YW
RRM	-3.6±0.7	116±8	280±21
MTM-adjusted weights	-7.1±1	63.3±9.6	-73±16
MTM-data1-CG fixed	-5.1±1	117.7±13.5	-14±20.1
MTM-data1-CG random	-17±1	76.2±13.2	-59.45±20.5
MTM-data2-CG fixed	-6.2±1	82±13.8	-147.9±22.03
MTM-data2-CG random	-19.4±1	37.6±13.5	-194.4±21.7
MTM-data2-CG fixed-age quadratic	-6.7±1	78.9±13.7	-172.5±22
MTM-data2-CG random-age quadratic	-20.07±1	34.7±13.3	-212.8±21.7

Chapter 4. Prediction of genomic breed composition of Hays Converter with 6K and 50K SNP panels using admixture and regression methods and exploring the variability of founder proportions across each autosome

4.1 Abstract

Crossbreeding is a widely used strategy in animal breeding to benefit from complementarity and heterosis. Composite breeds are one of the products of crossbreeding in which animals maintain a stabilized combination of genetic characteristics of two or more pure breeds. Hays Converter (HC) is a Canadian composite beef breed that was reportedly founded as a mixture of primarily Hereford (HER), Holstein (HOL) and Brown Swiss (BSW) breeds. Objectives of this study were to estimate the genomic breed composition of HC based on the entirety of its genome and on the chromosome segments of 50 SNP length over 6K genomic map. Admixture and regression methods were used with both 6K and 50K SNP panels to evaluate genome-wide breed composition. The results indicated that the regression method generated similar estimates with both SNP datasets. However, the admixture method analyzed HC composition differently, depending on SNP panel size and mainly due to the ratio of number of admixed to unmixed individuals. Obviating these constraints resulted in uniformity between admixture and regression estimates. There were no important changes in HC founder breed proportions from past to date with either method. Thus the expected retained heterozygosity was almost kept constant. Presence of Angus (AN) breed could also be identified as it was occasionally used to control calving difficulty. Use of the regression-6K is recommended as it is more cost effective (for genotyping) and avoids issues with sample size that arise with the admixture method. Overall, HC genomic composition was predicted in percentage as 8±0.2 AN,

51±0.2 HER, 15±0.1 BSW and 26±0.1 HOL. Diversity of breed proportions in HC chromosomal segments relative to whole genome may be used to imply signatures of selection for specific traits from its founders.

4.2 Introduction

Crossbreeding is a widely used strategy in animal breeding that is designed to exploit complementarity and retain heterosis in future generations. In beef cattle, composite breeds are produced by continuous crossbreeding in which the genetic characteristics of two or more pure breeds are combined to eventually achieve a stabilized proportional contribution of ancestral breeds (Dickerson 1973; Cundiff 1977; Koger 1980). In farm animals, breed registries are usually used to record and maintain pedigrees of animals with specific conformational characteristics and coat color patterns (Funkhouser *et al*. 2017). In pedigreed populations, the level of crossbreeding for admixed individuals is derived from assuming strict halving of the contributions of parents across generations (Sölkner *et al*. 2010).

However, pedigrees may contain errors, be incomplete or fraudulent, or entirely missing. Furthermore, recombination, random assortment of chromosomes into gametes and selection can lead to deviation from the expected contributions based on pedigree (VanRaden and Cooper 2015; Crum et al. 2019). Therefore, molecular tools, such as single nucleotide polymorphism (SNP) panels, are required to document these events. They are more accurate than pedigrees to estimate genomic breed composition because they are capable of measuring realized parental contributions at the genomic level (Toosi *et al*. 2010; Frkonja *et al*. 2012; De Beukelaer *et al*. 2017). In addition, they can be applied to estimate the genomic architecture at the chromosome level, because specific regions of the genome may represent a specific founder breed composition that is different from the whole genome expectation. Such variation in distribution

of ancestry across chromosomes may arise from differences in natural and artificial selective processes that the cattle population has experienced (McTavish and Hillis 2014). Also, this may imply an association between differences in ancestry of chromosomal regions with significant quantitative trait loci for important traits in beef production (Bolormaa *et al*. 2011).

Through SNP genotyping, animals can be categorized into genetic clusters (founding breeds) according to their patterns of multiple-loci genotypes (or haplotypes). Thus, for animals with ancestors that originated from different populations, the genomic composition can be estimated by associating the multiple ancestries with multiple genetic clusters (Pritchard *et al*. 2000; Alexander *et al*. 2009). Alternatively, genomic breed composition can be estimated using linear regression. With linear regression, discrete random variables corresponding to counts of alleles of the reference SNPs, which are distributed across the genome, are regressed on the allele frequencies of each reference SNP in a number of known breeds (Chiang *et al*. 2010; Kuehn *et al*. 2011).

Thus, estimation of breed composition is very useful in composite individuals. For example, in animals with complex ancestries, it has utility for estimating breed complementarity and heterosis effects as well as the additive genetic merit of individuals (Crum et al. 2019). Moreover, knowing the admixture proportions will be helpful in estimating heterozygosity, understanding the breeding history of the population, sorting animals into management groups, evaluating adaptability to production environments and making management decisions (Kuehn *et al*. 2011; Pickrell and Pritchard 2012; Akanno *et al*. 2017).

Hays Converter (HC) is a Canadian composite breed with a mixture of mainly, Hereford (HER), Holstein (HOL) and Brown Swiss (BSW) breeds. According to the recorded history of Hays Converter (HC), in 1957, Harry Hays started mating eight sons of a famous HOL, Spring

Farm Fond Hope, to 700 HER females of Baker's herd to increase gainability and milk production (Figure 4.1). Unknown numbers of polled HER and Angus (AN) females were also incorporated in the foundation female herd (Fleming *et al*. 2016). From this cross and a subsequent year of mating, he selected the "best" 159 HOL-HER heifers for backcrossing to a big Certified Meat Sire; the HER bull, Silver Prince 7P. At the time of his selection, Silver Prince 7P weighed 1200 kg in breeding condition and was known to transmit size, length, bone and fleshing ability to his offspring (Fleming et al., 2016). From the resulting progeny, the top five fastest gaining bulls were selected and bred back to the 159 HOL-HER females. In 1959, four young BSW bulls, sired by a son of the well-known BSW female, Jane of Vernon (famous for perfect udders and feet, lactation persistency and high growth in progeny), were also mated to 100 HER cows of the original Baker's herd to produce BSW-HER cross females. As shown by black stars in Figure 4.1, the foundation herd, including the progeny of Silver Prince sons, the HOL-HER and BSW-HER crosses, was closed in 1963 to all outside germplasm and for the next decade, only top performing males and females were kept for breeding (Fleming *et al*. 2016). However, in order to overcome calving difficulties in first-calf heifers, AN sires were periodically used on HC (HER+HOL+BSW) heifers (especially after 2000). This resulted in some 50%-50% AN-HC heifers being retained in the population and backcrossed to HC bulls to gradually reduce the proportion of AN.

No documentation of the proportion of each of the foundation breeds was maintained beyond the initial crosses. Despite a breed composition analysis through the regression method for a few animals (Fleming 2013), no study has been done to investigate changes in contributions from the founding breeds for all HC animals over time. Therefore, the objectives of this study were; first to predict the genomic breed composition of HC with 6K and 50K SNP panels using

both admixture (Alexander et al. 2009) and regression methods (Kuehn et al. 2011) and second to dissect each autosome to explore the potential of founder breed proportions to vary across the genome.

4.3 Materials and methods

In this study, 941 HC animals from Red Bow Ranching were genotyped using different marker panels, namely, Illumina BovineLD, Illumina BovineSNP50 (Illumina Inc., San Diego, CA) and GeneSeek Genomic Profiler (GGP) BeadChips (GeneSeek, Lincoln, NE) at Delta Genomics, Edmonton, AB, Canada. The data consisted of different panels of 6K, 7K, 19K, 30K to 50K SNPs for individuals born from 1973 to 2018. Two genotype files were extracted from these panels for further analysis. The first file contained 941 individuals with 6K SNP data that were common to all animals. The second file contained a sample of 205 animals that had data for 50K SNPs. These animals were born from 1973 to 2015. Corresponding genotypes were also collected from purebred Canadian cattle data available including 2225 Angus (AN), 1027 Hereford (HER), 451 Holstein (HOL) and 109 Brown Swiss (BSW) individuals. Quality control was performed using PLINK software to eliminate SNPs with minor allele frequency less than 0.01, heterozygosity excess more than 0.15 to compare with expected heterozygosity from Hardy-Weinberg equilibrium and detect potential genotyping errors (Wiggans *et al*. 2009; Turner *et al*. 2011) and individual and SNP genotyping call rates less than 0.90 (Purcell *et al*. 2007). Only autosomal SNPs with genomic positions matched to the UMD_3.1 bovine assembly map (Zimin *et al*. 2009) were retained. This resulted in 5907 and 42099 SNPs remaining from the 6K and 50K panels, respectively.

The two datasets were then analyzed using the admixture (Alexander *et al*. 2009) and regression (Kuehn *et al*. 2011) methods to estimate genomic breed proportions of founder

populations in HC. In order to find the number of hypothetical ancestral populations (K value) with the lowest cross validation error, a 10-fold cross validation procedure was performed using the admixture program for K=1 to 7 (Alexander *et al*. 2015). Both unsupervised and supervised modes were applied in admixture analysis (Alexander and Lange 2011). The regression method was run using the following formula in R (RStudio Team 2016);

$$Y = Xb + e$$

where X = $\{f_{kj}\}$was a M×N matrix, with M and N equal to the number of SNPs and number of founder breeds (here 4), respectively and f_{kj} corresponded to the allele B frequency of SNP k in founder breed j; Y was a vector of M×1 including copies of allele B of all SNPs present in each HC animal divided by 2 (0, 0.5, 1); b was a N×1 vector of regression coefficients indicating contribution of each founder breed to the given HC animal in Y; and e was a random residual M×1 vector (Kuehn *et al*. 2011). For negative regression coefficients, it was assumed that the purported founder breed made no contribution to HC and thus, they were set to zero (Larmer *et al*. 2014). As the sum of breed proportions in admixture is set to 1, to keep homogeneity when comparing results, the sum of regression coefficients was also forced to 1.

Expected retained heterozygosity (RH) was calculated for each animal as the formula below (Dickerson 1973);

$$RH = 1 - \sum_{i=1}^{n} P_i^2$$

where n was equal to the number of founder breeds and P_i represented the estimated proportion for the i^{th} founder breed. All average breed proportion results were reported for animals that were 75% or more HC. Also, except for BSW with smaller size available, samples of 451 individuals (corresponding to HOL size) were randomly extracted from AN and HER to repeat the analyses with reference samples that were more homogeneous in size. In addition, the random sampling

from AN and HER was repeated twice as more data was available for them. To test if admixture worked differently from the regression method, all the corresponding allele B frequencies for founder breeds were extracted from admixture and put into the regression method. Considering the map position of the common 6K SNPs, each autosomal chromosome was divided into segments of 50 SNPs to explore the breed proportion diversity through regression method. Finally, breed contribution and RH trends were obtained by linear regression on birth year.

4.4 Results

4.4.1 Admixture analysis - whole genome

As shown in Figure 4.2, cross validation errors followed a similar pattern in both common 6K and 50k datasets. There was almost no difference between K=4 with 5, 6 or 7. Also, the difference between K=4 and K=3 was small in both datasets. To explore if there was potential similarity in estimating genomic breed composition among K=3 to 5 for both datasets, average ancestral contributions were calculated for all AN, HER, BSW, HOL and HC individuals (Tables 4.1 and 4.2).

According to Table 4.1, considering the common 6K dataset with K=3 and unsupervised mode, the AN and HER animals were uniquely assigned to individual clusters with probability > 90%. The HOL animals were also assigned to a single unique cluster, but with lower probability (74%). However, the BSW animals were deemed an admixture of these three founding clusters, having a probability of membership in any one cluster of < 50%. Finally, the HC were indicated as having nearly equal probabilities of membership in the clusters that uniquely identified the HER and HOL animals. When increasing K to 4 and 5, although detection of AN, HER, HOL and BSW were improved, HC animals were not dissected to these clusters and rather assigned to a unique cluster with probability around 80% (not shown). Besides, reduction of reference

sample size through random selection of 451 individuals from each AN and HER populations, resulted in almost the same breed proportions as when using all founders data.

In supervised mode with 6K SNPs and K=4 (Table 4.3), BSW showed the highest contribution in genomic composition of all HC animals (87%) and other founder breeds indicated implausibly lower proportions. However, when using a subset of 50 HC individuals, supervised mode estimated the breed proportions as 11%, 51%, 15% and 23% for AN, HER, BSW and HOL, respectively.

Looking at Table 4.2, with 50K SNP data in unsupervised mode and K=3, despite detecting HOL as a pure ancestral breed, BSW was still distinguished as a mixed population of AN, HER and HOL. In K=4, all defined HC founder breeds were almost completely detected as pure populations with relative proportions as high as 97% on average. Consequently, their contributions in HC genomic composition were estimated as 6%, 48%, 25% and 21% for AN, HER, BSW and HOL, respectively. In K=5, as opposed to 6K common SNP data, HC was not separated as the fifth subpopulation. Rather, although AN was detected as two subpopulations (not shown), all the resulting breed fractions for founder breeds and HC were approximately the same as values taken in K=4. Moreover, there was no difference in the corresponding results when using supervised mode with K=4 and 50K data (Table 4.3). Also, analyzing a subset of 50 HC individuals through supervised mode-50K resulted in approximately identical breed proportions as obtained in supervised mode-6K.

Overall, with K=4 and 50K SNP data, Pearson correlation was equal to 0.66 between recorded estimates of the percentage HC which were presumably based on pedigree and admixture-50K HC percentages. As shown in Table 4.4, with admixture-50K, no meaningful change was observed in AN percentage for HC from 1973 to 2015. HER and HOL contributions

were decreased by 0.1%/yr (P<0.01) and 0.04%/yr (P<0.05), respectively, while BSW proportion was increased by 0.15%/yr (P<0.01). In addition, the average RH was equal to 0.66±0.002 (Figure 4.3) with 0.06% increase per year (P<0.01).

4.4.2 Admixture analysis – genome dissection

In terms of the first segment of chromosome one dissected by 50 SNPs according to the 6K data map and in K=3, 4 and 5, admixture in unsupervised mode could not distinguish unique clusters attributable to only each of the AN, HER, BSW or HOL populations. In supervised mode and by using all HC animals, similar to whole genome analysis, the genomic composition of the first segment of chromosome one was detected as being predominantly BSW (51%) while the other breeds made contributions of up to 20% (Table 4.5). Applying a subset of 50 HC individuals resulted in founder proportions of 15%, 57%, 9% and 19% for AN, HER, BSW and HOL, respectively.

Using more SNPs at the equivalent location from the 50K data (324 SNPs) and with unsupervised mode, recognition of AN, HER and HOL were improved as unique clusters although the probabilities of membership in the breed-specific clusters remained less than 90%. In contrast, BSW was detected as a mixture of these founders with greater membership probability for HOL and HER, respectively (Table 4.5). This consequently made the first segment of HC genome to be detected as a composition of more likely unique clusters above, with greater probabilities for HER (50%) and HOL (36%), respectively. In supervised mode and using all HC animals available in 50K data, all founding breeds showed some membership in the first segment of the genome. In this case, HER had the largest probability (51%), BSW and HOL assigned nearly equal probabilities of membership (~22%) and AN showed the smallest contribution (6%). Moreover, using a subset of 50 HC animals in supervised mode resulted in

similarity to admixture-6K supervised breed proportions with estimation of the BSW proportion less than HOL.

4.4.3 Regression analysis – whole genome

Assuming four potential founder breeds using the regression method (Tables 4.3), there was 0.99 correlation between 6K and 50K SNP datasets in estimating average ancestral breed proportions for HC genome. Considering the 6K SNP dataset, HC was a mixture of 8% AN, 51% HER, 15% BSW and 26% HOL. The contributions for AN and HER were slightly greater than admixture-50K results while the BSW and HOL percentages were by 10% less and 5% greater, respectively. There was a high correlation of 0.93 between regression-6K and admixture-50K average breed proportions. In addition, considering a subset of 50 HC individuals, the average breed proportions were almost the same in both regression and admixture-supervised methods, in both SNP datasets and in accordance with the regression results for all HC animals, too. Pearson correlation was equal to 0.75 and 0.68 between pedigree and regression based HC percentages for 6K and 50K SNP datasets, respectively.

There were similar trends as with admixture-50K for HER and BSW proportions in HC through both regression-6K and regression-50K (Table 4.4). AN contribution showed an increase to 0.05%/yr ($P<0.05$) through regression-50K and despite an increase of 0.07%/yr ($P<0.01$) by regression-6K, HOL contribution decreased to 0.04%/yr ($P<0.05$) through admixture-50K. The RH trends based on regression-6K and regression-50K were approximately the same as admixture-50K (Table 4.4, Figure 4.3). The correlation between RH estimated through regression-6K and regression-50K or admixture-50K was equal to 0.87 while it was 0.97 between regression-50K and admixture-50K. The average RH values were equal to 0.64±0.001

and 0.64±0.002 in both regression-6K and regression-50K, respectively which were almost the same as admixture-50K.

Except for HC, allele B frequencies of founders were extracted from admixture-6K and re-analyzed by regression method. The results showed a very high correlation with breed fractions estimated in regression-6k, i.e., 99% for AN, HER, HOL and 95% for BSW, respectively. Similarly, except for BSW with 89% correlation, the extraction of allele B frequencies from admixture-50K and re-analyzing them with regression method led to a 99% correlation with regression-50K results for all founders.

4.4.4 Regression analysis – genome dissection

Contrary to admixture results for the first segment of chromosome one in HC, the average values of estimated ancestral proportions were almost uniform through regression-6K and regression-50K for all animals (Table 4.5). Compared to admixture-50K unsupervised results, there were no specific difference between AN and HER proportions in the two methods. Moreover, the summation of BSW and HOL proportions in regression-6K-50K seemed to be approximately equal to the HOL percentage in admixture-50K unsupervised. Considering 50 HC individuals, the average breed proportions in regression method were almost the same as admixture-supervised methods, in both SNP datasets and in accordance with the regression results for all HC animals.

Therefore, according to Table 4.6, genomic diversity of breed proportions across the autosomes were calculated by dividing each chromosome into segments of 50 SNP length based on 6K panel and regression method. Along with the observed diversity for all founders in each autosome, the results indicated the highest values of HER contribution in the chromosome-segments of 1-4, 6-4, 7-5 29-3 and 22-1 (from 68 to 65%), respectively. Considering HOL

proportion, the highest percentages (53 to 46%) were assigned to the chromosome-segments of 2-6, 29-1, 18-1, 11-2 and 6-2, respectively. The greatest proportions of BSW (42 to 34%) belonged to the chromosome-segments of 2-5, 8-4, 20-3, 9-5, and 20-4, respectively. In terms of AN, the highest contributions (38 to 32%) were located in the chromosome-segments of 20-2, 24-2, 26-3, 29-2 and 12-4, respectively. Conversely, regarding the bottom five lowest values, the ancestral proportions also demonstrated variation across each segment of chromosomes. They were in the range of 13 to 20% for HER located in the chromosome-segments of 6-2, 18-1, 9-3, 29-1 and 8-4, respectively. The lowest values of HOL (8 to 11%) were found in the chromosome-segments of 6-4, 2-1, 7-5, 23-1 and 15-2, respectively. BSW showed the smallest values (3 to 7%) in the chromosome-segments of 23-1, 5-4, 6-1, 27-3 and 28-2, respectively. Finally, AN indicated the smallest amounts of 3 to 4% in the chromosome-segments of 3-4, 7-2, 3-2, 4-1 and 6-6, respectively.

4.5 Discussion

To address the change in HC genomic composition from the time it was created, it was useful to have some information of the crossbreeding design and primary ancestral breed contributions (Figure 4.1). Moreover, it was important to know that except for the main founders, due to some crosses with AN, there were genotyped individuals with less than 100% HC proportion, that was a summation of HER, HOL and BSW proportions. From these animals, according to the breeder's definition, only those with 75 or more HC% were considered as HC animals for breeding. Typically, the germplasm of the cattle being predicted is important to be represented in the reference breeds (Kuehn et al. 2011). Therefore, to distinguish those individuals with less than 75% HC from all genotyped data, it was necessary to add AN to the reference population, although it was not one of the three main founding breeds. In spite of

knowing the founding breeds used to develop HC, no data were available in terms of the percentages of each founder in the pedigree file. Therefore, genomic breed composition analysis through SNP data was employed to have a better understanding of potential changes in the composite structure during time.

Although there were statistically significant changes in HC ancestral contributions and RH through both methods and datasets, they were of small magnitude and therefore negligible. This actually meant the fractions of all four founders remained fixed and there were neither backcrosses nor outcrosses with the own founders or other breeds to affect the high level of RH in HC composite from past to present. Gregory *et al*. (1982) indicated that in a three-breed or four-breed composite populations targeting contributions of 1/2A, 1/4B, 1/4C or 1/2A, 1/4B, 1/8C, 1/8D, the percentage of heterozygosity retained after *inter se* crossing in the F_3 would be equal to 62.5 and 65.6, respectively. Indeed, the main objective of HC breeding seemed to be developing a composite of three-breed combination that essentially consisted of 1/2HER, 1/4HOL and 1/4BSW. However, though contributing less than 1/8, the results demonstrated a potential contribution from AN. According to the HC history (Fleming *et al*. 2016), this could be due to contributions from the unknown number of AN females in the original females that were used to create HC. Therefore, compared to three-breed contribution in Gregory *et al*. (1982), a little increase in average RH of HC implied the effect of small contribution of AN in the HC genomic composition. The retention of initial heterosis in composite populations is proportional to retention of heterozygosity (Dickerson 1973). Thus, the approximately constant level of RH in HC might demonstrate the retention of heterosis, especially for growth traits of WW and YW as the main objectives in HC development. Considering sire selection differential results for growth traits in chapter three, it is obvious how genetic trends would improve if HC had the opportunity

to select for the best sires based on EBVs. In fact, despite weak sire selection practice, the hidden increased level of genetic trend for best sires from 1985 to 2000 might be used as an implication of potentially favorable effect of heterosis on HC growth traits. This meant if a selection index was used to improve HC performance from the inception, it might exploit the heterosis effect on growth more efficiently. Furthermore, after 2000, although the herd size greatly decreased from around 600 to 100 animals per year, it has not led to a dramatic impact on the current RH, as yet. However, due to linearity of association of retention of heterosis with retention of heterozygosity, continuous reduction in size and thus genetic variation may lead to possible dissipation of RH and appearance of inbreeding in the future (Gregory and Cundiff 1980).

The Pearson correlation between pedigree and the two methods for prediction of HC% was moderately high. This could imply the presence of possible errors in pedigree-determined HC proportion and variation attributable to chromosomal sampling (Kuehn et al. 2011).

When the estimated allele frequencies in admixture method were re-analyzed with regression method, the resulting breed fractions were very highly correlated with regression method results. Hence, admixture and regression methods likely followed a similar procedure to estimate breed fractions. Buzanskas *et al*. (2017) studied the introgression of beef breeds in Canchim cattle by admixture and regression methods and observed similarity in their behavior as the size of panels increased from 1 to 7, 15 and 32K SNPs. In addition, they concluded that the more noise observed in purebred animals were attributed to the reduction in density of SNPs.

Considering the same level of average RH, despite a little difference in BSW fraction, there was no specific difference in HC genomic composition between admixture-50K and regression-6K-50k methods. According to Shringarpur and Xing study (2014), this might be related to the sensitivity of admixture to smaller number of these animals compared to composite

individuals. They indicated that the accuracy of ancestry inference using admixture method depended on the ratio of the number of admixed individuals to unmixed individuals from each founder population. For cattle data with K=2, they observed that the accuracy of ancestry estimation was high while the ratio was less than 1 and started to drop as the ratio increased. Considering their suggestions, although no ratio was available for larger K, in order to use more homogeneous samples by size, the number of reference animals to admixed individuals should be selected with caution. Pritchard *et al*. (2000) showed that the number of loci available for analyzing with STRUCTURE had a significant effect on the recovery of individual ancestry. Likewise, Shringarpure & Xing (2014) pointed out the advantage of adding more SNPs to improve the admixture analysis and reduce the effects of sample selection bias for identical sample sizes and samples. However, this may not be always possible as observed in the case of HC; as the data set was constructed from the intersection of multiple datasets that resulted from individuals being genotyped on different platforms.

Patterson et al. (2006) mentioned that the number of markers needed to resolve populations was inversely proportional to the genetic distance (Fst) between the populations. Moreover, Shringarpure & Xing (2014) suggested that Fst had an effect on sample selection bias so that well-differentiated populations were easy to separate even with biased sampling. In this study, there was a smaller Fst between BSW and HOL than HER or AN. Hence, in addition to small number of SNPs (324 on 50K map) and higher sample ratio of HC to BSW (205/109=1.88), admixture might have mostly detected BSW as HOL in the first segment of chromosome one results (Table 4.5). Therefore, as opposed to regression method, it was probable that the potential contribution of BSW in HC was hidden in HOL% in admixture-50K unsupervised.

Alexander and Lange (2011) also provided a supervised mode in admixture method in which ancestry estimates could be more accurate due to less uncertainty in allele frequencies. This mode provided a straightforward way to compare the admixture and regression methods by assuming similar knowledge of the founding breeds. It also facilitated faster analysis owing to reduction of dimensionality of parameter space. Although the interpretation of results would be simplified, it was only suitable when the ancestral populations were known and fairly homogeneous. Otherwise, they suggested unsupervised mode as the standard option for exploratory analyses. Therefore, in the case of HC, with known story about the founders, the analyses were repeated in supervised mode for the whole genome. As observed in Table 4.3, when considering all individuals, admixture-6K detected HC mostly as BSW which was discrepant from regression-6K results. Moreover, using 50K dataset produced no meaningful difference to unsupervised mode results. This might be related to the elaborate likelihood formulation in admixture method and its assumptions about linkage equilibrium between markers and/or the underlying distribution of genetic markers both of which are relaxed in regression method (Boerner and Wittenburg 2018). However, as proposed by Shringarpur and Xing (2014), due to smaller number of BSW individuals in this study (109), 50 HC animals were separated from its population to have a ratio of less than one with all reference populations. This made the admixture-supervised be able to analyze the HC composition similar to the regression method. The procedure worked for both the whole genome and its first segment (Table 4.5) in both SNP datasets and the proportions were in accordance with the regression results for all HC animals, too. The results were also similar for a second sample of 50 HC animals (not shown). Hence, it looked possible to repeat the admixture-supervised for all HC animals through groups of 50 individuals and then get the overall average. As a recommendation, it seems there should

always be a regression analysis beside admixture to support results, especially with smaller SNP data and the ratio of higher than one for admixed to unmixed individuals. Moreover, although subdividing admixed population to several groups and repeating the admixture analysis will provide similar results as regression method, depending on the sample size of populations, it could spend longer computational time than regression method.

It was obvious that estimation of genomic composition for each animal was more accurate if 50K dataset were used relative to 6K common SNPs. However, there was uniformity in average estimated breed contributions between regression-6K and regression-50K, either for all or each segments of autosomes. Therefore, in order to be cost-effective for future animal genotyping and avoid possible discrepancy in admixture due to small number of animals or SNPs, the regression-6K is proposed as the preferred method. Fleming (2013) calculated the breed composition for 125 HC animals using 50K SNP data and regression method. Although she selected the potential founders from 13 beef and dairy breeds, the largest components on average were still found to be relative to HER, HOL and BSW at approximately 43.6%, 17.8% and 7%, respectively. Thus, if there is sufficient information about the founder breeds and mating design of a composite, adding more breeds to admixture or regression analysis may lead to high computational demand and biased interpretation of results. As there was no specific difference between K=5,6,7 with K=4 in cross validation errors for both 6K and 50K data, it looked reasonable to represent HC as a composite of at most four breeds.

This study used founder populations only containing individuals which were 100% purebred. To detect fractions of founders accurately by smaller density SNP panels, it was recommended to have pure populations indicating ≥85% assignment to their breed of registration (Crum *et al*. 2019). In admixture-6K, when K=5, BSW was approved as a pure-bred population

with the self-ancestral assignment more than 90% (not shown). However, the own composite was not decomposed to its constructing founders and rather considered as the fifth sub-population (not shown). In admixture-50K, using K=5 did not change the structure of HC as a composite, though AN was sub-divided to two clusters (not shown). Adding putative clusters of founders in the analysis of the 50K data (not shown) again assigned HC to a unique cluster. This is circular reasoning to have HC as an ancestral population of HC and it is not true. Therefore, depending on the founder sample size, its purity and SNP panel density, it was speculated that increasing K from the expected number of founders might affect the interpretation of clusters that were revealed by the admixture method. Paim *et al*. (2020) performed admixture analysis for all autosomes and each chromosome using Brangus individuals with 700K SNPs and AN and Brahman as founder populations. In K=2, they ascertained some bias in AN proportion from the theoretical expectation (70.4% vs 62.5%) and suggested breeders to be aware of this situation if they want to maintain the Brahman component. However, considering the own Brangus as the third cluster (K=3), they regressed the proportional cluster assignments on generation number and inferred the increase in Brangus as the formation of a new breed which it is actually a circular reasoning and Brangus cannot be a founder of itself. Similarly, Blackburn *et al*. (2014) analyzed genomic composition of Braford using STRUCTURE (Pritchard *et al*. 2000) with smaller set of SNPs (60K) and founder populations of HER, Nelore and Brahman. Although the three progenitor breeds were found to be distinct from each other through principle component analysis, due to founders sample selection bias, estimating the construction of Braford was suspected. Moreover, with K=4, although there was a very mild increase in Braford cluster assignment (10% on average) over time, they recognized it as the formation of the new breed. Taking such results suggests that incorporation of the own composite as one of the K clusters

may result in biased interpretation of the effects of selection and genetic drift over time on its distinction from founders. To avoid diminishing complementarity from the defined expectations, the ancestral proportions must be assessed with K equal to the number of constructing founders. In this way, the disproportionate moving of composite toward one of its ancestral clusters can be monitored which will then help manage the genetic diversity of composite population effectively.

As shown in Table 4.6, on a within-chromosome basis, the proportion contributed by each founder varied substantially. This may provide new insights into the formation of composite breeds. A search in the AnimalQTL database (https://www.animalgenome.org/cgi-bin/QTLdb/index) identified that chromosomes showing larger proportion of a founder, have been associated to traits that were mostly studied on that founder, too. For example, chromosomes 1, 6 and 7 (with the highest HER%) showed many QTL-association studies in HER population that were related to traits of average daily gain and body weight (BW, YW). For chromosome 2 (with the highest HOL% in HC), there were many association studies in HOL population that were predominantly related to traits including reproduction, milk chemical contents, milk yield, milk protein and fat percentage and length of productive life. In BSW population, there were multiple QTL-association studies for milk fat% on chromosome 20, or in AN population, for meat and carcass quality on chromosome 24. Interestingly, both chromosomes showed high contributions of BSW and AN in HC, respectively. Therefore, chromosomal segments showing higher percentage of founder origin may be considered as the relative founder enriched regions (Goszczynski *et al*. 2017). These examples suggest that selection and complementarity may work on favorable alleles of each founder for the traits of

interest in various segments of the genome and persist as the new breed continues to develop its own signature (Paim *et al*. 2020).

4.6 Conclusion

The results demonstrated similarity in predicting average genomic breed composition in HC through admixture and regression methods. However, to avoid potential problems with sample selection bias and SNP density in admixture program, and to be cost effective, regression method with 6K dataset is recommended to estimate average genomic breed composition for all the genome and each chromosome, too. Although no AN was targeted in HC development as a founder, its presence was confirmed. Overall, HC genomic composition was predicted as 8% AN, 51% HER, 15% BSW and 26% HOL. Variability in founder contributions at the chromosome level may reveal signatures of selection for traits from those founders in the genomic structure of HC as a composite.

4.7 References

Akanno, E.C., Chen, L., Abo-Ismail, M.K., Crowley, J.J., Wang, Z., Li, C., Basarab, J.A., MacNeil, M.D. and Plastow, G., 2017. Genomic prediction of breed composition and heterosis effects in Angus, Charolais, and Hereford crosses using 50K genotypes. Canadian Journal of Animal Science, 97(3), pp.431-438.

Alexander, D.H. and Lange, K., 2011. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. BMC bioinformatics, 12(1), p.246.

Alexander, D.H., Novembre, J. and Lange, K., 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome research, 19(9), pp.1655-1664.

Alexander, D.H., Shringarpure, S.S., Novembre, J. and Lange, K., 2015. Admixture 1.3 software manual. Los Angeles: UCLA Human Genetics Software Distribution.

Blackburn HD, Paiva, SR, Sollero, BP, Biegelmeyer, P, Caetano, AR, Cardoso, FF (2014) A Dedicated SNP Panel for Evaluating Genetic Diversity in a Composite Cattle Breed. In Proceedings of the World Congress on Genetics Applied to Livestock Production, 048. World Congress on Genetics Applied to Livestock Production.

Boerner, V. and Wittenburg, D., 2018. On estimation of genome composition in genetically admixed individuals using constrained genomic regression. Frontiers in Genetics, 9, p.185.

Bolormaa, S., Hayes, B.J., Hawken, R.J., Zhang, Y., Reverter, A. and Goddard, M.E., 2011. Detection of chromosome segments of zebu and taurine origin and their effect on beef production and growth. Journal of animal science, 89(7), pp.2050-2060.

Buzanskas, M.E., Ventura, R.V., Chud, T.C.S., Bernardes, P.A., de Abreu Santos, D.J., de Almeida Regitano, L.C., de Alencar, M.M., de Alvarenga Mudadu, M., Zanella, R., da

Silva, M.V.G.B. and Li, C., 2017. Study on the introgression of beef breeds in Canchim cattle using single nucleotide polymorphism markers. PloS one, 12(2).

Chiang, C.W., Gajdos, Z.K., Korn, J.M., Kuruvilla, F.G., Butler, J.L., Hackett, R., Guiducci, C., Nguyen, T.T., Wilks, R., Forrester, T. and Haiman, C.A., 2010. Rapid assessment of genetic ancestry in populations of unknown origin by genome-wide genotyping of pooled samples. PLoS genetics, 6(3).

Crum, T.E., Schnabel, R.D., Decker, J.E., Regitano, L.C. and Taylor, J.F., 2019. CRUMBLER: a tool for the prediction of ancestry in cattle. PloS one, 14(8).

Cundiff, L.V., 1977. Foundations for animal breeding research. Journal of Animal Science, 44(2), pp.311-319.

De Beukelaer, H., Badke, Y., Fack, V. and De Meyer, G., 2017. Moving beyond managing realized genomic relationship in long-term genomic selection. Genetics, 206(2), pp.1127-1138.

Dickerson, G. E. 1973. Inbreeding and heterosis in animals. Proc. Anim. Breed. Genet. Symp. Honor of Dr. Jay L. Lush, Amer. Soc. Anim. Sci., Champaign, IL, pp. 54-77

Dickerson, G.E., 1973. Inbreeding and heterosis in animals. Journal of Animal Science, 1973(Symposium), pp.54-77.

Fleming, A., 2013. A genetic characterization of the Hays Converter. M.Sc. thesis. University of Guelph, ON, Canada.

Fleming, A.L., Hays, D.P., Moore, S., Plastow, G., Miller, S.P., 2016. History of the Hays Converter. Can. J. Anim. Sci. 96, 471-477.

Frkonja, A., Gredler, B., Schnyder, U., Curik, I. and Soelkner, J., 2012. Prediction of breed composition in an admixed cattle population. Animal genetics, 43(6), pp.696-703.

Funkhouser, S.A., Bates, R.O., Ernst, C.W., Newcom, D. and Steibel, J.P., 2017. Estimation of genome-wide and locus-specific breed composition in pigs. Translational Animal Science, 1(1), pp.36-44.

Goszczynski, D.E., Corbi-Botto, C.M., Durand, H.M., Rogberg-Muñoz, A., Munilla, S., Peral-Garcia, P., Cantet, R.J.C. and Giovambattista, G., 2018. Evidence of positive selection towards Zebuine haplotypes in the BoLA region of Brangus cattle. animal, 12(2), pp.215-223.

Gregory, K.E. and Cundiff, L., 1980. Crossbreeding in beef cattle: evaluation of systems. Journal of animal science, 51(5), pp.1224-1242.

Gregory, K.E., Trail, J.C.M., Koch, R.M. and Cundiff, L.V., 1982. Heterosis, crossbreeding and composite breed utilization in the tropics. In Proceedings of the 2nd World Congress on Genetics Applied to Livestock Production (p. 279).

Koger, M., 1980. Efective Crossbreeding Systems Utilizing Zebu Cattle. Journal of Animal Science, 50(6), pp.1215-1220.

Kuehn, L.A., Keele, J.W., Bennett, G.L., McDaneld, T.G., Smith, T.P.L., Snelling, W.M., Sonstegard, T.S. and Thallman, R.M., 2011. Predicting breed composition using breed frequencies of 50,000 markers from the US Meat Animal Research Center 2,000 Bull Project. Journal of animal science, 89(6), pp.1742-1750.

Larmer, S., Ventura, R., Buzanskas, M.E., Sargolzaei, M. and Schenkel, F.S., 2014. Assessing admixture by quantifying breed composition to gain historical perspective on dairy cattle in Canada. In Vancouver, Canada: 10th World Congress on Genetics Applied to Livestock Production.

McTavish, E.J. and Hillis, D.M., 2014. A genomic approach for distinguishing between recent and ancient admixture as applied to cattle. Journal of Heredity, 105(4), pp.445-456.

Paim, T.D.P., Hay, E.H.A., Wilson, C., Thomas, M.G., Kuehn, L.A., Paiva, S.R., McManus, C. and Blackburn, H.D., 2020. Dynamics of genomic architecture during composite breed development in cattle. Animal genetics, 51(2), pp.224-234.

Patterson, N., Price, A.L. and Reich, D., 2006. Population structure and eigenanalysis. PLoS genetics, 2(12).

Pickrell, J. and Pritchard, J., 2012. Inference of population splits and mixtures from genome-wide allele frequency data. Nature Precedings, pp.1-1.

Pritchard, J.K., Stephens, M. and Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics, 155(2), pp.945-959.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., De Bakker, P.I., Daly, M.J. and Sham, P.C., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. The American journal of human genetics, 81(3), pp.559-575.

RStudio Team (2016). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA. Retrieved from http://www.rstudio.com/.

Shringarpure, S. and Xing, E.P., 2014. Effects of sample selection bias on the accuracy of population structure and ancestry inference. G3: Genes, Genomes, Genetics, 4(5), pp.901-911.

Sölkner, J., Frkonja, A., Raadsma, H.W., Jonas, E., Thaller, G., Gootwine, E., Seroussi, E., Fuerst, C., Egger-Danner, C. and Gredler, B., 2010. Estimation of individual levels of

admixture in crossbred populations from SNP chip data: examples with sheep and cattle populations. Interbull Bulletin, (42), pp.62-62.

Toosi, A., Fernando, R.L. and Dekkers, J.C.M., 2010. Genomic selection in admixed and crossbred populations. Journal of Animal Science, 88(1), pp.32-46.

Turner, S., Armstrong, L.L., Bradford, Y., Carlson, C.S., Crawford, D.C., Crenshaw, A.T., De Andrade, M., Doheny, K.F., Haines, J.L., Hayes, G. and Jarvik, G., 2011. Quality control procedures for genome-wide association studies. Current protocols in human genetics, 68(1), pp.1-19.

VanRaden, P.M. and Cooper, T.A., 2015. Genomic evaluations and breed composition for crossbred US dairy cattle. Interbull Bulletin, (49), pp.19-23.

Wiggans, G.R., Sonstegard, T.S., VanRaden, P.M., Matukumalli, L.K., Schnabel, R.D., Taylor, J.F., Schenkel, F.S. and Van Tassell, C.P., 2009. Selection of single-nucleotide polymorphisms and quality of genotypes used in genomic evaluation of dairy cattle in the United States and Canada. Journal of dairy science, 92(7), pp.3431-3436.

Zimin, A.V., Delcher, A.L., Florea, L., Kelley, D.R., Schatz, M.C., Puiu, D., Hanrahan, F., Pertea, G., Van Tassell, C.P., Sonstegard, T.S. and Marçais, G., 2009. A whole-genome assembly of the domestic cow, Bos taurus. Genome biology, 10(4), p.R42.

4.8 Figures and tables

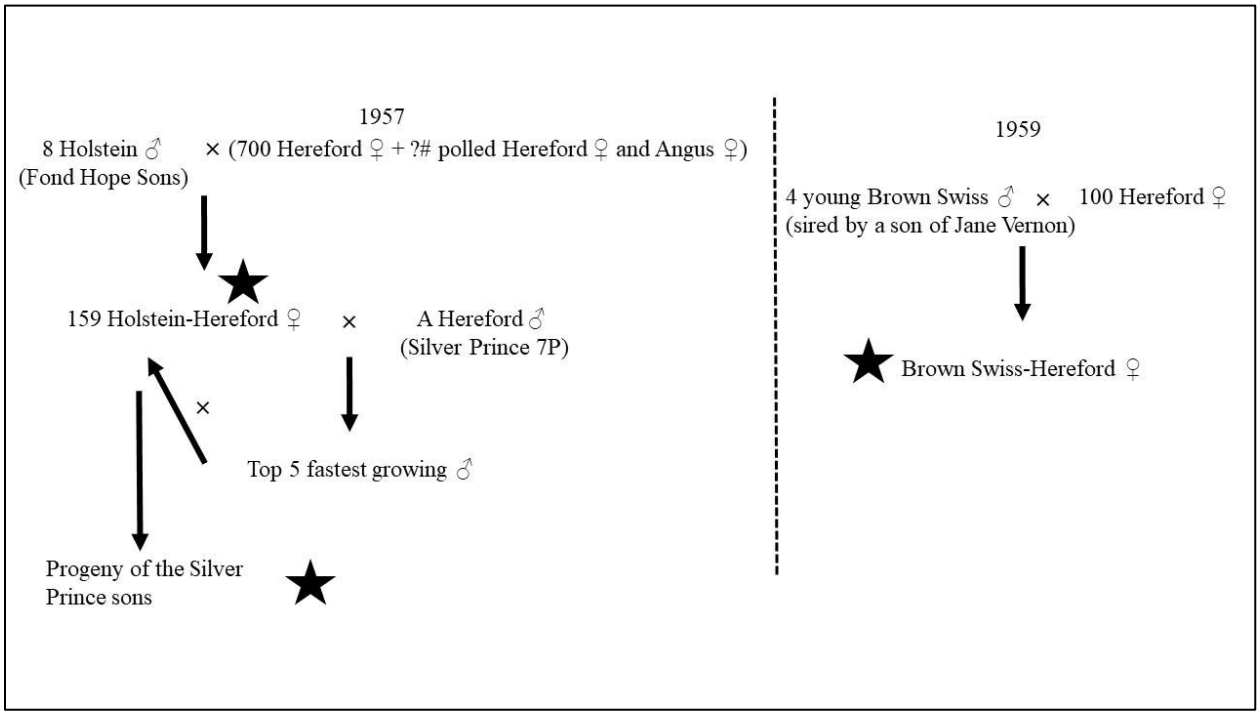

Figure 4.1. Schematic shape of the HC breed foundation.

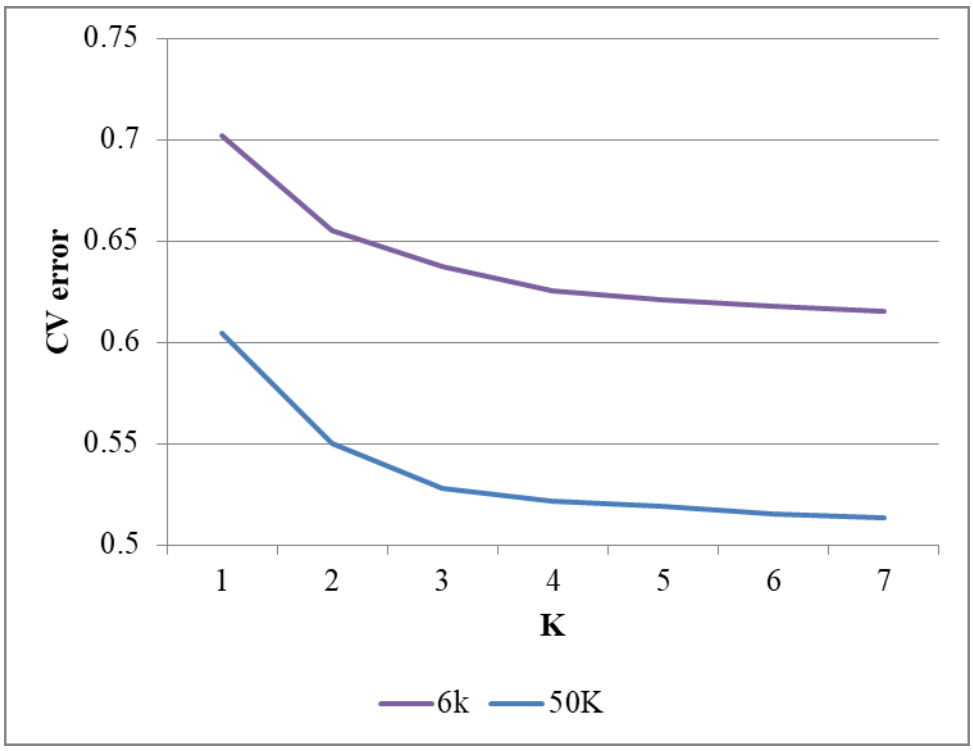

Figure 4.2. Cross validation (CV) errors for K=1 to 7 clusters and two datasets of 6K and 50K SNPs through admixture method.

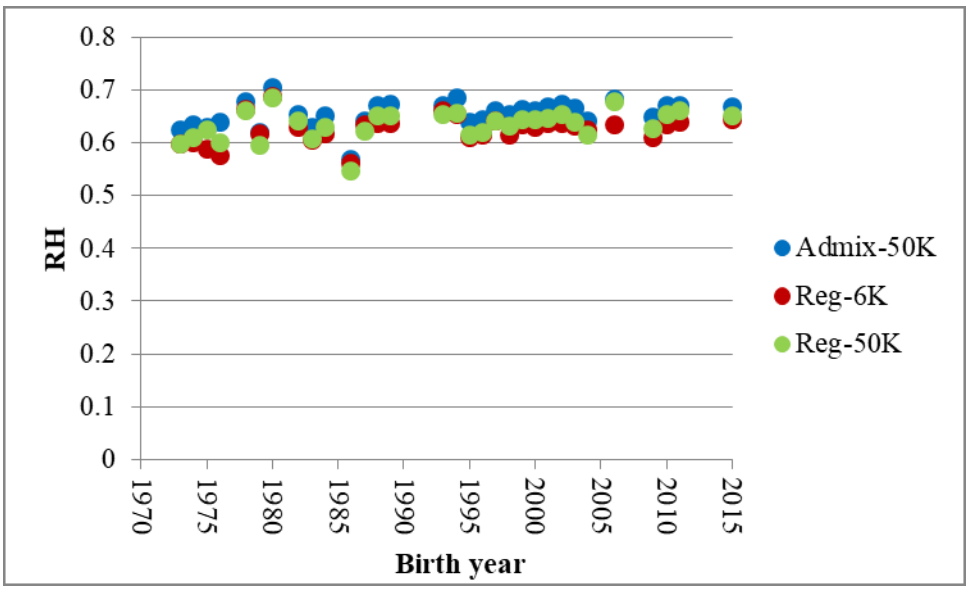

Figure 4.3. Retained heterozygosity (RH) trend in HC from 1973 to 2015 through admixture method with 50K SNP data, K = 4 and regression method with both 6K and 50K SNP data.

Table 4.1. Average genomic breed composition of founder breeds and HC through admixture method with unsupervised mode considering cluster (K) = 3 with 6K SNP data.

K	Population	AN	HER	BSW	HOL
3	AN	0.94±0.0009	0.04±0.0006	-	0.02±0.0005
	HER	0.04±0.001	0.94±0.001	-	0.02±0.0005
	BSW	0.33±0.002	0.20±0.002	-	0.47±0.002
	HOL	0.25±0.001	0.01±0.001	-	0.74±0.001
	HC	0.04±0.002	0.46±0.001	-	0.50±0.002

Table 4.2. Average genomic breed composition of founder breeds and HC through admixture method with unsupervised mode considering clusters (K) = 3 to 5 with 50K SNP data.

K	Population	AN	HER	BSW	HOL
3	AN	0.97±0.0008	0.01±0.0005	-	0.02±0.0004
	HER	0.02±0.0007	0.97±0.001	-	0.01±0.0004
	BSW	0.26±0.0008	0.22±0.001	-	0.52±0.001
	HOL	0.015±0.001	0.005±0.0004	-	0.98±0.001
	HC	0.12±0.003	0.54±0.003	-	0.34±0.002
4	AN	0.97±0.0009	0.01±0.0005	0.01±0.0004	0.01±0.0002
	HER	0.01±0.0006	0.96±0.001	0.02±0.0006	0.004±0.0002
	BSW	0.009±0.001	0.006±0.001	0.97±0.004	0.015±0.002
	HOL	0.01±0.0008	0.002±0.0003	0.018±0.0009	0.97±0.002
	HC	0.06±0.003	0.48±0.003	0.25±0.002	0.21±0.002
5	AN	0.98±0.0005	0.01±0.0003	0.006±0.0002	0.004±0.0002
	HER	0.02±0.0008	0.96±0.001	0.017±0.0005	0.003±0.0002
	BSW	0.01±0.001	0.005±0.001	0.97±0.004	0.015±0.002
	HOL	0.01±0.0009	0.003±0.0003	0.017±0.0008	0.97±0.002
	HC	0.08±0.003	0.48±0.003	0.24±0.002	0.20±0.002

Table 4.3. Comparison of average genomic breed composition of HC through admixture method in both unsupervised and supervised modes with regression method, when K = 4, for all and a subset of 50 individuals (>75% HC) using 6K and 50K SNP datasets.

Method	N	AN	HER	BSW	HOL
Admix-6k-supervised	861	0.05±0.005	0.07±0.003	0.87±0.006	0.01±0.001
Regression-6k	861	0.08±0.002	0.51±0.002	0.15±0.001	0.26±0.001
Admix-50k-unsupervised	196	0.06±0.003	0.48±0.003	0.25±0.002	0.21±0.002
Admix-50k-supervised	196	0.05±0.003	0.50±0.003	0.24±0.003	0.21±0.002
Regression-50k	196	0.11±0.003	0.51±0.003	0.13±0.002	0.24±0.002
Admix-6k-supervised	50	0.11±0.007	0.51±0.007	0.15±0.005	0.23±0.005
Regression-6k	50	0.08±0.008	0.51±0.007	0.15±0.005	0.26±0.006
Admix-50k-supervised	50	0.10±0.007	0.52±0.007	0.15±0.005	0.23±0.004
Regression-50k	50	0.11±0.007	0.52±0.007	0.13±0.005	0.24±0.004

Table 4.4. Trends of retained heterozygosity (RH) and founder breeds proportions in HC composition through admixture method using 50K SNP data with K = 4 and regression method using both 6K and 50K SNP data.

	Admixture method-50K	Regression method-6K	Regression method-50K
RH	0.0006**±0.0001	0.0007**±0.0002	0.0008**±0.0002
AN	0.0002±0.0002	-0.0004±0.0003	0.0005*±0.0002
HER	-0.001**±0.0003	-0.001**±0.0002	-0.001**±0.0003
BSW	0.0015**±0.0002	0.0011**±0.0002	0.0006**±0.0002
HOL	-0.0004*±0.0002	0.0007**±0.0002	-0.00002±0.0002

***P<0.05**
****P<0.01**

Table 4.5. Average genomic breed composition of founder breeds and HC for the first segment of chromosome one, through admixture method in both unsupervised and supervised modes and regression method, when K = 4, for all and a subset of 50 HC individuals (>75% HC) using 6K and 50K SNP datasets.

Method	N	Population	AN	HER	BSW	HOL
Admixture-6K-supervised	861	HC	0.11±0.007	0.20±0.008	0.51±0.01	0.18±0.008
Admixture-50K-unsupervised	2225	AN	0.79±0.003	0.08±0.002	-	0.13±0.002
	1027	HER	0.07±0.003	0.87±0.004	-	0.06±0.002
	109	BSW	0.19±0.007	0.38±0.008	-	0.43±0.007
	451	HOL	0.06±0.003	0.09±0.004	-	0.85±0.005
	196	HC	0.14±0.008	0.50±0.01	-	0.36±0.01
Admixture-50K-supervised	196	HC	0.06±0.006	0.51±0.02	0.21±0.02	0.22±0.01
Regression-6K	861	HC	0.19±0.006	0.50±0.009	0.12±0.006	0.19±0.007
Admixture-6K-supervised	50	HC	0.15±0.02	0.57±0.04	0.09±0.03	0.19±0.03
Regression-6K	50	HC	0.16±0.02	0.51±0.04	0.09±0.02	0.24±0.03
Regression-50K	196	HC	0.12±0.01	0.49±0.02	0.14±0.01	0.25±0.01
Admixture-50K-supervised	50	HC	0.08±0.02	0.56±0.04	0.13±0.03	0.23±0.03
Regression-50K	50	HC	0.11±0.01	0.54±0.04	0.11±0.02	0.24±0.03

Table 4.6. Average genomic breed composition of founder breeds in 50-SNP length segments of HC autosomes through regression-6K method (bold and underlined for the largest and smallest, respectively).

Chromosome-segment	AN	HER	BSW	HOL
1-1	0.19±0.006	0.50±0.009	0.12±0.006	0.19±0.007
1-2	0.11±0.005	0.56±0.007	0.22±0.006	0.11±0.005
1-3	0.14±0.007	0.48±0.009	0.20±0.006	0.17±0.006
1-4	0.13±0.006	**0.68±0.008**	0.08±0.004	0.11±0.005
1-5	0.20±0.007	0.41±0.008	0.25±0.009	0.14±0.006
1-6	0.06±0.004	0.37±0.009	0.17±0.007	0.40±0.009
1-7	0.22±0.007	0.32±0.008	0.23±0.007	0.23±0.008
2-1	0.16±0.005	0.49±0.01	0.25±0.009	0.10±0.006
2-2	0.13±0.005	0.50±0.01	0.11±0.004	0.26±0.008
2-3	0.17±0.007	0.36±0.008	0.16±0.007	0.31±0.007
2-4	0.21±0.007	0.40±0.009	0.24±0.007	0.15±0.005
2-5	0.10±0.005	0.24±0.007	**0.42±0.008**	0.24±0.008
2-6	0.09±0.004	0.23±0.006	0.15±0.007	**0.53±0.01**
3-1	0.09±0.005	0.61±0.007	0.15±0.006	0.15±0.006
3-2	0.04±0.003	0.60±0.008	0.17±0.005	0.19±0.007
3-3	0.09±0.005	0.57±0.008	0.21±0.005	0.12±0.005
3-4	0.03±0.003	0.41±0.01	0.32±0.007	0.24±0.008
3-5	0.08±0.004	0.49±0.007	0.23±0.006	0.20±0.006
3-6	0.10±0.006	0.47±0.009	0.11±0.005	0.32±0.008
4-1	0.04±0.003	0.52±0.007	0.15±0.005	0.29±0.008
4-2	0.10±0.005	0.57±0.008	0.15±0.006	0.18±0.006
4-3	0.11±0.005	0.33±0.01	0.25±0.007	0.31±0.008
4-4	0.07±0.004	0.51±0.009	0.22±0.008	0.20±0.005
4-5	0.07±0.005	0.43±0.008	0.21±0.006	0.29±0.008
4-6	0.10±0.006	0.49±0.01	0.23±0.008	0.18±0.008
5-1	0.23±0.008	0.33±0.008	0.25±0.008	0.19±0.008
5-2	0.20±0.007	0.26±0.008	0.27±0.009	0.27±0.008
5-3	0.18±0.006	0.45±0.01	0.12±0.005	0.25±0.007
5-4	0.17±0.007	0.42±0.01	0.05±0.003	0.36±0.008
5-5	0.23±0.008	0.26±0.009	0.19±0.005	0.32±0.009
5-6	0.18±0.008	0.27±0.009	0.18±0.007	0.37±0.01
6-1	0.23±0.006	0.39±0.008	0.06±0.004	0.32±0.006
6-2	0.18±0.008	0.13±0.006	0.23±0.007	**0.46±0.01**
6-3	0.05±0.004	0.37±0.008	0.14±0.005	0.44±0.008

6-4	0.13±0.005	**0.68±0.006**	0.11±0.004	0.08±0.004
6-5	0.07±0.005	0.51±0.01	0.21±0.006	0.21±0.007
6-6	0.04±0.003	0.45±0.01	0.29±0.01	0.22±0.009
7-1	0.22±0.007	0.46±0.008	0.09±0.004	0.23±0.007
7-2	0.03±0.003	0.64±0.007	0.11±0.005	0.22±0.005
7-3	0.15±0.008	0.36±0.01	0.32±0.008	0.17±0.006
7-4	0.13±0.007	0.39±0.008	0.17±0.006	0.31±0.008
7-5	0.14±0.005	**0.67±0.008**	0.09±0.005	0.10±0.005
8-1	0.13±0.006	0.31±0.008	0.14±0.006	0.42±0.009
8-2	0.27±0.009	0.32±0.008	0.19±0.007	0.22±0.007
8-3	0.15±0.006	0.44±0.008	0.16±0.006	0.25±0.007
8-4	0.17±0.006	0.20±0.008	**0.38±0.008**	0.25±0.007
8-5	0.24±0.006	0.38±0.007	0.18±0.006	0.20±0.006
8-6	0.17±0.008	0.29±0.01	0.17±0.007	0.37±0.01
9-1	0.24±0.008	0.22±0.007	0.20±0.005	0.34±0.009
9-2	0.26±0.009	0.38±0.008	0.18±0.006	0.18±0.007
9-3	0.16±0.006	0.19±0.007	0.29±0.008	0.36±0.009
9-4	0.23±0.007	0.33±0.008	0.11±0.004	0.33±0.007
9-5	0.07±0.004	0.26±0.009	**0.35±0.007**	0.32±0.008
10-1	0.17±0.007	0.48±0.008	0.08±0.005	0.27±0.007
10-2	0.09±0.006	0.48±0.01	0.19±0.007	0.24±0.008
10-3	0.22±0.009	0.37±0.01	0.29±0.007	0.12±0.005
10-4	0.11±0.006	0.42±0.008	0.20±0.005	0.27±0.007
10-5	0.10±0.006	0.53±0.007	0.17±0.007	0.20±0.006
11-1	0.10±0.005	0.40±0.008	0.22±0.006	0.28±0.006
11-2	0.10±0.006	0.31±0.007	0.12±0.004	**0.47±0.008**
11-3	0.26±0.006	0.26±0.008	0.19±0.006	0.29±0.007
11-4	0.27±0.006	0.40±0.007	0.15±0.005	0.18±0.007
11-5	0.17±0.008	0.51±0.009	0.12±0.004	0.20±0.007
12-1	0.14±0.006	0.50±0.009	0.19±0.007	0.17±0.007
12-2	0.13±0.006	0.37±0.01	0.17±0.007	0.33±0.009
12-3	0.10±0.006	0.28±0.008	0.22±0.007	0.40±0.009
12-4	**0.32±0.008**	0.25±0.008	0.19±0.006	0.24±0.007
13-1	0.13±0.006	0.44±0.007	0.20±0.006	0.23±0.006
13-2	0.12±0.006	0.44±0.008	0.23±0.006	0.21±0.008

13-3	0.18±0.006	0.41±0.01	0.15±0.006	0.26±0.007
13-4	0.28±0.007	0.24±0.006	0.32±0.007	0.16±0.006
14-1	0.09±0.005	0.39±0.008	0.23±0.008	0.29±0.008
14-2	0.25±0.007	0.26±0.007	0.22±0.007	0.27±0.008
14-3	0.13±0.006	0.47±0.009	0.16±0.006	0.24±0.008
14-4	0.20±0.007	0.30±0.007	0.19±0.006	0.31±0.009
15-1	0.17±0.006	0.47±0.01	0.16±0.006	0.20±0.006
15-2	0.26±0.007	0.39±0.01	0.24±0.007	0.11±0.006
15-3	0.15±0.006	0.34±0.008	0.21±0.007	0.30±0.008
15-4	0.13±0.005	0.43±0.01	0.27±0.008	0.17±0.008
16-1	0.13±0.006	0.56±0.01	0.12±0.004	0.19±0.006
16-2	0.15±0.005	0.49±0.01	0.10±0.004	0.26±0.009
16-3	0.13±0.005	0.47±0.008	0.11±0.005	0.29±0.008
16-4	0.27±0.008	0.21±0.008	0.08±0.005	0.44±0.008
17-1	0.10±0.005	0.33±0.009	0.14±0.005	0.43±0.009
17-2	0.14±0.005	0.55±0.009	0.10±0.005	0.21±0.007
17-3	0.17±0.007	0.46±0.009	0.21±0.007	0.16±0.008
17-4	0.30±0.01	0.38±0.01	0.10±0.005	0.22±0.008
18-1	0.22±0.008	0.18±0.007	0.12±0.005	**0.48±0.01**
18-2	0.18±0.007	0.43±0.008	0.18±0.007	0.21±0.008
18-3	0.17±0.006	0.41±0.009	0.12±0.006	0.30±0.007
19-1	0.11±0.006	0.36±0.01	0.10±0.004	0.43±0.01
19-2	0.10±0.005	0.34±0.008	0.19±0.005	0.37±0.008
19-3	0.14±0.006	0.50±0.008	0.11±0.004	0.25±0.007
20-1	0.09±0.005	0.50±0.007	0.18±0.006	0.23±0.006
20-2	**0.38±0.008**	0.34±0.009	0.09±0.005	0.19±0.005
20-3	0.07±0.004	0.41±0.007	**0.36±0.008**	0.16±0.005
20-4	0.05±0.004	0.47±0.01	**0.34±0.01**	0.14±0.005
21-1	0.09±0.005	0.60±0.01	0.14±0.005	0.17±0.007
21-2	0.09±0.005	0.61±0.01	0.14±0.005	0.16±0.007
21-3	0.22±0.007	0.53±0.007	0.12±0.004	0.13±0.006
22-1	0.06±0.004	**0.65±0.009**	0.09±0.004	0.20±0.006
22-2	0.14±0.005	0.53±0.009	0.20±0.008	0.13±0.005

22-3	0.22±0.007	0.24±0.008	0.11±0.004	0.43±0.01
23-1	0.30±0.006	0.57±0.007	0.03±0.003	0.10±0.005
23-2	0.20±0.007	0.36±0.007	0.23±0.006	0.21±0.008
23-3	0.17±0.007	0.47±0.01	0.11±0.005	0.25±0.008
24-1	0.10±0.005	0.21±0.008	0.25±0.007	0.44±0.01
24-2	**0.38±0.008**	0.23±0.007	0.12±0.006	0.27±0.005
24-3	0.09±0.005	0.43±0.008	0.16±0.006	0.32±0.008
25-1	0.14±0.006	0.28±0.008	0.17±0.006	0.41±0.009
25-2	0.15±0.006	0.28±0.008	0.18±0.006	0.39±0.008
25-3	0.25±0.007	0.28±0.008	0.22±0.007	0.25±0.009
26-1	0.10±0.006	0.44±0.009	0.29±0.008	0.17±0.006
26-2	0.21±0.007	0.46±0.006	0.14±0.006	0.19±0.006
26-3	**0.34±0.008**	0.42±0.008	0.09±0.005	0.15±0.006
27-1	0.22±0.008	0.46±0.009	0.11±0.005	0.21±0.007
27-2	0.09±0.007	0.47±0.007	0.09±0.005	0.35±0.007
27-3	0.08±0.006	0.65±0.01	0.07±0.004	0.20±0.009
28-1	0.20±0.008	0.37±0.007	0.08±0.005	0.35±0.008
28-2	0.18±0.007	0.34±0.009	0.07±0.003	0.41±0.008
29-1	0.08±0.004	0.19±0.007	0.23±0.007	**0.50±0.009**
29-2	**0.34±0.007**	0.37±0.007	0.10±0.004	0.19±0.007
29-3	0.09±0.004	**0.67±0.007**	0.08±0.005	0.16±0.006

Chapter 5. Exploring genome-wide signatures of selection associated with the founding breeds of Hays Converter composite cattle through Fst and runs of homozygosity

5.1 Abstract

Hays Converter (HC) is a Canadian composite breed formed in the late 1950's and early 1960's by crossing Hereford (HER), Holstein (HOL) and Brown Swiss (BSW). These breeds were selected mainly to capture benefits from fertility and carcass traits (HER), milk production and growth potential (HOL) and strong feet and udders (BSW). However, a small percentage of Angus (AN) has also been found to be present in the genome due to its occasional usage to control calving difficulty in first calf heifers. Except for body weight records, phenotypic information characterizing the HC breed is scarce. Therefore, the aim of this study was to explore indicators of selection across the genome using Fst and runs of homozygosity (ROH). Subsequently, the AnimalQTL database was used to quantitative trait loci (QTL) that are co-located with the selection signatures. Twenty eight chromosomal segments showing over-representation of ancestral breeds relative to the entire genome were identified using Grubbs' test. For each chromosomal segment, average Fst between HC and the ancestral breeds were evaluated. Also, to measure the pressure of selection, trends in ROH length per chromosomal segment were analyzed through regression on birth year. Within each meaningful segment, for each SNP, Fst values were ranked (1 to 4) between HC and the four ancestral breeds from smallest to largest indicating the degree of similarity relative to each of the founding breeds. Then, considering the over-represented breed in each segment, predominant genomic areas were detected through regression analysis for SNP sequences showing continuous ranks of 1 and/or 2

for that breed. Although the average level of genomic inbreeding was observed greater than the pedigree estimate (7.5%), the frequency of SNPs in ROH was not alarming. In addition, only three chromosomal segments showed a positive trend in ROH. They included several numbers of autozygous fragments originated from HOL, BSW or AN which overlapped to QTLs associated to traits of body weight and milk production. The pressure of selection of HOL and BSW in HC was a bit higher in BTA-segment 6-2 than 29-1 due to having higher than average Fst between HER and dairy populations. Although the selection of AN fragments in BTA-segment 20-2 was not still strong due to lower than average Fst between HER/AN populations, if the trend continues to increase, it will be at the expense of specially HOL for milk traits. Overall, the effects of such trends were not meaningful because they were mainly as a result of reduction in herd size after the year 2000 and not selection. Moreover, the lack of a genetic selection program has probably made a weak selection of ancestral breeds' haplotypes in over-represented areas so that almost all the 28 chromosomal segments showed a lower Fst with HER. As HER comprised the highest percentage across the composite genome, too, these findings might imply its sustainable role for weight traits, body features, milk production, fertility and carcass standards. Accordingly, due to having the higher genetic distance with BSW, no fragments were identified in over-represented areas and co-located with QTL for traits that were the basis for adding BSW to HC, i.e., lactation persistency and eye area pigmentation. Although AN represented the smallest fraction of the whole genome, in order to control the use of breed, the increasing ROH trend observed for its over-representation should be monitored for the future development of HC. The use of Fst, ROH, the analysis of breed proportions and the AnimalQTL database help to interpret signatures of selection of breeds contributing to the HC composite.

5.2 Introduction

The main advantages of developing composite breeds are to maintain heterosis over time and to exploit complementarity among breeds to achieve an optimum additive genetic composition. As inbreeding influences retention of heterosis, it is worth to highlight that it should be avoided in order to maintain high levels of heterozygosity in composite breeds (Geregory et al. 1999). Today, practices such as intense sire selection, artificial insemination and embryo transfer have led to a reduction in effective population size and genetic diversity. This might consequently affect the levels of heterozygosity across the genome and increase the levels of homozygosity and thus, inbreeding within breeds (Purfield et al. 2012).

Runs of homozygosity (ROH) are defined as contiguous lengths of homozygous genotypes across the genome of an animal. They are thought to be inherited by transmitting identical haplotypes from parents to offspring. Thus, their extent and frequency may inform on the ancestry of a population and its evolution over time (Purfield *et al*. 2012). With the widespread use of high density SNP panels, an increasing interest has been aroused in identifying autozygous segments from molecular information based on ROH. This information helps to disclose the genetic relationships among individuals and to be a more accurate estimator for detection of effects of inbreeding (Ferenčaković et al. 2013). In addition, it can reveal selection within populations, because selection tends to cause homozygous stretches on the genome and generate shared autozygous areas with reduced genetic diversity, which might consequently harbor targets of positive selection (Pemberton *et al*. 2012; Kim *et al*. 2013).

A classic approach to the detection of selection signatures is based on using the fixation index or Fst which was first defined by Wright (1949). It is the degree of genetic differentiation between populations that is quantified by differences in allele frequencies. Indeed, it is able to

provide information on genomic variation at each SNP locus and be used as an evidence of selection, i.e. high Fst values indicate local positive adaptation while low values specify neutral selection (Zhao *et al*. 2015).

Hays Converter (HC) is a composite animal that mainly developed as a mixture of Hereford (HER), Holstein (HOL) and Brown Swiss (BSW). Data documenting the specific matings among the founding animals are unavailable. A general description of the foundation follows. The HC breed was started by selecting 700 commercial HER cows that were smaller than average, fecund and very tough cattle adapted to the western Canada ranges and Alberta winters (Fleming 2013). They were mated to 8 sons of a big HOL bull (Spring Farm Fond Hope), whose progeny were known for their large size, strong constitution, excellent feet and outstanding udders. Also, his daughters were popular in the dairy industry, having produced over 4% butterfat on record of performance tests with an average milk production of around 5443 kg. By adding HOL blood, the breeder (Harry Hays) expected to obtain more growth potential and greater milk production in the composite and thus produce fast growing calves (Fleming *et al*. 2016). The resulting crossbred HOLxHER heifers were backcrossed to the famous big Canadian HER bull (Silver Prince 7P) in order to add more size, length, bone, ruggedness, rapid maturity and improved carcass characteristics. From this mating, the top five fastest gaining bulls were kept and bred back to the HOL×HER females. The final breed addition occurred as a cross between four young BSW bulls with 100 HER females. The bulls were grandsons of Jane of Vernon, a BSW cow well-known for her perfect udder, and high milk and butterfat production (Fleming 2013). With the BSW genetics, the thought was to get strong udders, lactation persistency, additional growth, excellent feet and legs, and also pigmented skin in order to reduce problems with eyes and udder chapping. By 1963, the best females of this cross along

with the progeny of Silver Price sons and HOL×HER crosses were placed in the foundation herd (Fleming *et al*. 2016). At this time, the herd was closed and top performing males and females were retained and mated over the next decade. They were selected intensely for the traits most important to the breeder; weaning weight, yearling weight and udder conformation for the replacement heifers. In terms of fertility traits, once cows were in the brood herd, remaining depended on their ability to produce offspring and survive. In fact, females had to calve between 23 to 25 months of age and wean a calf every year after (CDA Livestock Division 1976). Moreover, they had to be pregnant on the bull's first service and calves had to be born unassisted in the open pasture. The breeder mainly used yearling bulls of 500 kg (the HC goal) to qualify for service and only a few were retained as herd sires for several years (Fleming et al. 2016).

Though not being considered as one of the main ancestral breeds in developing HC, there were an unknown number of AN females in the foundation herd. In addition, there was an evidence of occasional use of AN bulls to control calving difficulty in first-calf heifers. Therefore, the breed proportion of AN was 8% across whole genome as estimated through the 6k-regression method analysis that was presented in chapter four.

The aim of this study was to assess genome-wide signatures of selection in HC through Fst, ROH and breed proportion analyses to firstly identify potential over-represented chromosomal segments that originated from HER, HOL, BSW or AN. Secondarily, in addition to using ROH beside pedigree for estimation of inbreeding, ROH trends were analyzed for each over-represented segment to identify pressure of selection. Finally, segments with increasing trends were analyzed through Fst and associated with phenotypic traits through comparison with the QTL found in the AnimalQTL database.

5.3 Materials and methods

5.3.1 Inbreeding

In this study, the inbreeding coefficient of a HC individual was estimated through two approaches: 1) using pedigree-based analysis (F_{PED}) to calculate the probability of identity by descent (IBD) (Wright 1922; Keller and Waller 2002) through CFC software package (Sargolzaei *et al*. 2006) and 2) measuring autozygosity across the genome based on ROH to estimate genomic inbreeding (F_{ROH}) (McQuillan *et al*. 2008).

In order to find ROH, 205 HC animals were genotyped with a 50K SNP panel. The SNP genotypes on 29 autosomes were quality controlled using PLINK (Purcell *et al*. 2007). Both animals and SNPs with call rates <90% were eliminated. Also, SNPs with MAF <1% and those deviating from Hardy Weinberg equilibrium (P <0.0001) were discarded (Purfield *et al*. 2012). This resulted in 40578 SNP remaining in the data. Next, to identify ROH segments across the HC genome, a sliding window of 50 SNPs [--homozyg-window-snp 50] was determined using PLINK in which up to 5 SNPs with missing genotypes [--homozyg-window-missing 5] and only one heterozygous SNP [--homozyg-window-het 1] were allowed (Purfield *et al*. 2012). The scanning window hit rate was set to 0.05 [--homozyg-window-threshold 0.05]. Then, the following parameters were used to define ROH segments; a minimum number of 30 consecutive homozygous SNPs [--homozyg-snp 30], a minimum density of one SNP per 500 kb inside an ROH [--homozyg-density 500], a maximum gap of 500 kb between consecutive homozygous SNPs [--homozyg-gap 500] and a maximal amount of one heterozygous SNP in the final ROH segment [--homozyg-het 1] (Sumreddee *et al*. 2019). Also, a minimum ROH length of 1000 kb [--homozyg-kb 1000] was set to exclude possible short and common ROH segments less than one Mb that occurred due to LD (Ferenčaković *et al*. 2013).

The resulting ROH segments were classified into five groups; 1-2 Mb, 2-4 Mb, 4-8 Mb, 8-16 Mb and >16 Mb (Kirin *et al*. 2010). For each genotyped HC animal, F_{ROH} was calculated as the total length of ROH segments divided by the total length of the autosomal genome covered by SNPs (McQuillan *et al*. 2008);

$$F_{ROH} = \frac{\sum L_{ROH}}{L_{TOTAL}}$$

where ΣL_{ROH} was the sum of ROH segments length per genotyped individual and L_{TOTAL} was the total length of the autosomal genome of 2,512,189 Kb based on the consensus map (Sumreddee *et al*. 2019). To find the pressure of selection, in addition to all chromosomes, trends of sum of lengths of ROH were analyzed per over-represented chromosomal segment through linear regression on birth year. Pairwise correlations were computed between inbreeding estimates obtained from the pedigree and ROH, as well.

5.3.2. Detection of signatures of selection

It was hypothesized that the proportional representation of each of the founding breeds in the segments of the HC genome was equal to the proportions across the entire genome. To test this hypothesis, the ancestral breed proportions were calculated for the 123 chromosomal segments that were described in chapter four (Table 4.6). The hypothesis was then tested using Grubbs' test (Grubbs 1950). This test is a standard procedure for detection of outliers (Stefansky 1972). With the outliers deemed to be indicative of over- or under- representation of a breed in the chromosomal segment compared to the entire genome. Based on the normal distribution, the test compares the values of two standardized variables computed from test data with tabulated values of the relevant Grubbs' critical parameter (Barbato *et al*. 2011). In terms of ordered data, two standardized variables of G_l and G_u were defined as follows;

$$G_l = \frac{\bar{x} - x_{min}}{s}$$

$$G_u = \frac{x_{max} - \bar{x}}{s}$$

where $\bar{x}$ and s were the genome-wide mean and standard deviation among animals estimated for each of four ancestral breed proportions as estimated in chapter 4 through the regression method using the 6K SNP data. The x_{min} and x_{max} were the smallest and largest breed-specific proportions calculated for each of the 123 segments in Table 4.6. For a two-tailed test, the hypothesis of no outliers was rejected if either G_l or G_u exceeded the critical value (G_{crit}) given by the following formula;

$$G_{crit} = \frac{n-1}{\sqrt{n}} \sqrt{\frac{t^2_{(\alpha/2n,n-2)}}{n-2+t^2_{(\alpha/2n,n-2)}}}$$

where n denoted the number of data points and t(α/2n,n-2) was the critical value of the t-distribution with (n-2) degrees of freedom and a significance level of (α/2n) (α was set to 0.05). Then, from detected outliers indicating over-represented proportions of HER, HOL, BSW and AN, 28 chromosomal segments were analyzed further for signatures of selection.

To measure the degree of genetic differentiation at SNP levels between HC and each of the four constructive breeds, Wright's fixation index (Fst) (Wright, 1965; Nei, 1977) was calculated as described by Frankham *et al*. (2002):

$$F_{ST} = \frac{H_T - H_S}{H_T}$$

$$H_T = 1 - \sum(\bar{p}^2 + \bar{q}^2)$$

$$H_S = \frac{\sum_{i=1}^{n} H_{exp_i} \times n_i}{N_T}$$

where H_T, $\bar{p}$ and $\bar{q}$ denote the expected heterozygosity, the frequency of allele A and B over the total population, respectively. H_S stands for the expected heterozygosity averaged across all sub-populations in which H_{expi} and n_i denote the expected heterozygosity and sample size for the i^{th} sub-population and N_T equals the total number of individuals in all sub-populations. For the 28 selected chromosomal segments, corresponding SNPs from 50K panel in chapter four were used to calculate Fst values between HC and each of the four constructing breeds.

The Fst value is a measure of genetic divergence among populations and ranges from 0 to 1. It represents the most genetic similarity (identity) between populations at 0 while complete genetic differentiation at 1 (Qanbari *et al*. 2011). As there were four Fst values at each SNP locus, they were first ranked on 1 to 4 from the smallest to largest to allow distinguish the amount of HC similarity in genetic material relative to each of the four founding breeds. Considering the sequence of SNPs in chromosomal segments with significant ROH trends, the most consecutive (unbroken) sequences of ranks of 1, 2 or 1 and 2 in Fst values were identified for the HC versus over-represented ancestral breed and the over-representation (predominance) was approved by the regression method used in chapter four. For example, as shown in some part of the segment 2 of chromosome 20 (Table 5.1), the four different Fst values were first ranked from 1 to 4. Then, considering AN as the over-represented ancestral breed in this segment (38%), HC-AN Fst values were colored according to their ranks. The approach was based on an assumption that if an ancestral breed was over-represented in one segment of a composite animal chromosome, there was more probability to find SNP sequences originating from that breed showing smaller Fst ranks (1 and 2). For confirmed sequences, a search was done in the AnimalQTL database (Hu *et al*. 2019) to identify their potential overlapping location with previously published QTLs, SNP and/or gene association data on bovine species

(https://www.animalgenome.org/cgi-bin/QTLdb/BT/search). The HC animals were counted if they had common ROH segments matching to these regions.

To detect positive selection in HC population for the SNPs located at the confirmed regions above, only those significant Fst values (the ones positioned at the extreme 2.5% of the empirical distribution ($\mu+2\sigma$)) were considered in which the SNP expected heterozygosity in HC was lower than the compared ancestral breed.

5.4 Results and discussion

5.4.1 Inbreeding

This chapter followed usual parameters that have been used for detection of ROH segments through 50K SNP data (Purfield *et al*. 2012, Ferenčaković *et al*. 2013). The descriptive statistic results of ROH analysis for 205 HC animals were summarized in Table 5.2. On average, there were approximately 126 SNPs per ROH segment (SD ≈ 105) with minimum and maximum numbers ranging from 30 to 1158. The average density (the mean distance between two consecutive SNPs in one ROH segment) was found to be 61 kb (SD ≈ 12 kb) with a range from 36 to 129 kb and roughly corresponded to the average density for most livestock and pet species in medium density SNP arrays (50 kb) (Meyermans *et al*. 2020). Each animal's autosomal genome contained approximately 25 ROH segments (SD ≈ 9) with minimum and maximum numbers of 2 to 52, respectively. The average length of a ROH was 7.5 Mb (SD ≈ 6.3) with 1.39 and 70.02 Mb as the shortest and longest segments, respectively. Finally, the total length of all ROH segments across the genome was nearly 190 Mb (SD ≈ 107.5 Mb) with the minimum and maximum sum of 7.8 and 624 Mb, respectively. Although the mean length for the ROH segments was a bit larger in HC compared to the Line 1 Hereford cattle population (6.83 Mb), there were similarities in minimum (1.36) and maximum length (64.86) (Sumreddee *et al*. 2019).

On the other hand, the average number of ROH segments observed in HC animals was much smaller than in the Line 1 Hereford population (82.92) (Sumreddee *et al*. 2019), Italian Brown (94.6) and Holstein (81.7) (Marras *et al*. 2015). Thus based on ROH, HC was deemed less inbred than these populations (F_{ROH} = 0.075 in HC, 0.23 in Line 1 Hereford, 0.145 in Italian Brown and 0.116 in Holstein). This could be due to the populations above were being older relative to HC and have been under more intensive selection. Moreover, the high rate of consanguineous mating through artificial insemination using a small number of high genetic merit sires would expect to increase the levels of IBD areas in such breeds over time (Mc Parland *et al*. 2007; MacNeil 2009).

Figure 5.1 depicts the average inbreeding coefficients for genotyped animals per year calculated through pedigree (F_{PED}) and ROH data (F_{ROH}) from 1973 to 2015. Except for two years, the F_{PED} values were generally lower than F_{ROH} (0.038 vs 0.075 on average, respectively). After 1989, both trends started to follow a similar pattern and the difference between them became reduced during the last years. The increase in both trends was relatively low with an annual rate of 0.001±0.0003 ($P < 0.01$). Moreover, the Pearson correlation of the two estimates was equal to 0.60. The quality of a pedigree file influences the estimation of inbreeding coefficients from it. As the depth and completeness of the pedigree increases, it allows to more accurately calculating the inbreeding coefficients. (Ferenčaković *et al*. 2013; Pryce *et al*. 2014). In the case of HC population, the pedigree-derived inbreeding coefficients were underestimated due to the incomplete recording of parentage (Fleming 2013) and thus potential contributions of unknown ancestors could not be quantified (Cassell *et al*. 2003). However, as the recorded pedigree became more complete, the estimated inbreeding coefficients showed less difference between the two methods. The correlation between estimates of inbreeding was in line with the

studied range in different dairy and beef breeds (0.51 to 0.70) (Pryce *et al*. 2014; Marras *et al*. 2015; Sumreddee *et al*. 2019). The average genomic inbreeding coefficient in HC was similar to estimates from most beef and dairy cattle breed societies with extremely large and deep pedigrees where the majority of animals had an inbreeding coefficient of less than 10% (Wiggans et al. 1995; Gengler et al. 1998; Cleveland et al. 2005). Also, though significant, the increase in the level of inbreeding was observed to be negligible. This might be because of the usage of expanded number of sires and dams in breeding until the year 2000. However, after that, due to a considerable reduction in HC population size, it was expected to see some increase in the amount of inbreeding coefficients as a result of unavoidable mating between relatives.

Figure 5.2 showed the frequency distribution of ROH segments per autosomal chromosome into five groups according to their length. Except for 9% of chromosome 12, there was almost negligible occurrence of 1% in the very short ROH segments (1-2 Mb). Small differences were observed among autosomes, in terms of short (2-4Mb), medium (4-8Mb) and long (8-16Mb) ROH segments, with the averages of 30%, 40% and 20% respectively. It seemed medium to long segments prevailed in the HC genome accounting for 60% of all ROH detected per chromosome. Also, except for chromosome 9, small differences were observed in the frequency of very long ROH segments (>16Mb) which on average accounted for 9% of each autosome. The abundance of relatively long ROH segments can be interpreted as indicating relatively recent inbreeding. This is consistent with observations for the recently developed composite Montana Tropical beef cattle (Peripolli *et al*. 2020). Furthermore, these results differ from those reported in pure beef and dairy cattle in which the total length of ROH was composed of an abundant number of shorter segments (Ferenčaković et al., 2013; Marras et al. 2015; Peripolli et al. 2018) which indicate inbreeding that occurred further back in time. Indeed, in

recently formed and young populations like HC and other composite populations, ROH segments are expected to be longer. This is because not enough generations have passed for recombination to break up those IBD areas. Ancient populations including pure dairy and beef breeds tend to reflect shorter ROH because the segments have been broken down by repeated meiosis over many generations (Kirin *et al*. 2010). In HC, there was also a significant increasing trend in the frequency of long and very long ROH segments together (0.26%/yr, P <0.01) and a corresponding significant decrease in the frequency of 2-4Mb ROH group (0.26%/yr, P <0.01). Although the frequency change per year was not quite remarkable, it might be as a result of the reduction occurred in the population size after the year 2000 and should be considered as some caution regarding the control of level of inbreeding in the future.

Partitioning the ROH segments to chromosomes revealed that chromosomes 2, 6, 1, 7 and 3 had on average the greatest ROH length (intercept, Table 5.3) and numbers of F_{ROH} (Figure 5.3), respectively. Also, all other autosomes were significantly involved in inbreeding (intercept, Table 5.3), though the contribution typically depended on the length of the chromosome (Zimin *et al*. 2009). Looking at each chromosome length separately, BTA 6, 21, 27, 2, and 7 had the largest ROH proportions, respectively (Figure 5.3). In addition, summarized in Table 5.3 are annual changes for sum of ROH per chromosome from 1973 to 2015. The trends showed that only chromosomes 2, 3, 6, 11 and 20 had significantly increasing lengths of autozygosity over years. Therefore, most of the chromosomes seemed to indicate the maintenance of a constant level of inbreeding in recent years. This might be the result of balancing selection so that selection for increased fitness increased heterozygosity at a rate sufficient to offset other trait-specific selection that would be manifest as increased autozygosity. This could have stabilized the length of ROH over time. In addition, based on Table 5.3, BTA 21 and 27 didn't make any

meaningful change in the level of inbreeding over time though they had upper ranks relative to their own length. Considering the concept of F_{ROH} to describe inbreeding (ROH proportion to the whole genome length), looking at the own chromosome length for regional inbreeding might be a bit tricky and lead to wrong ranks, especially when the chromosome length is very long (1 and 3) or short (21 and 27).

From all the number of SNPs applied in ROH, 76.5% was repeated from 1 to 20 times, 19.1% from 21 to 30, 3.8% from 31 to 40 and only 0.6% from 41 to 52 (i.e., the number of animals out of 205). Therefore, it seemed the level of inbreeding should not be alarming specific to all QTLs affecting each trait of interest.

5.4.2 Signatures of selection of HC founding breeds

Table 5.4 shows the 28 (out of 123) chromosomal segments that were selected by Grubbs' test for over-representation of HER, HOL, BSW and AN. For each of these chromosomal segments, it also indicates the genetic distance of HC based on average Fst values relative to the four ancestral breeds. Furthermore, trends in ROH length were added for each segment, separately.

In terms of HER over-representation, there were only four chromosomal segments which showed the smallest Fst with HER, too. For almost all other cases, HOL, BSW and AN over-representation were observed at the expense of HER under-representation. Here, although the difference among Fst values were not large, the results indicated that the segments were in general closer to AN and HER and became genetically farther from HOL to BSW. Considering trends in ROH length from 1973 to 2015, only three chromosomal segments of 6-2 (location: 21.55-45.90 Mb), 20-2 (location: 16.24-38.76 Mb) and 29-1 (location: 0.63-21.64 Mb) showed significant increase of less than one Mb, annually from which two were in line with the trends

observed for chromosomes 6 and 20 (Table 5.3). In order to find potential HOL, BSW and AN origin regions under selection in these three segments, they were partitioned by Fst ranks and originated by regression method (as described in Table 5.1). The found fragments were included in Tables 5.5 to 5.7. They were sequences of 5 to 16 SNPs, originally predominant for one of the three ancestral breeds above and all recognized as shared ROH areas with some frequencies out of 205 genotyped animals. Across each of these regions, all associated traits studied in HC ancestral populations were included along with their AnimalQTLdb publication IDs. If any SNPs of each fragment showed significant Fst between HC and its ancestral breeds and also indicated positive selection in HC were added in the respective tables. Finally, regarding the location of each fragment, the respective Fst between HER and breed origin (HOL, BSW or AN) populations were added to compare the selection pressure in HC.

5.4.2.1 Over-representation of HOL and BSW

Considering Table 5.5, there were 8 and 5 found fragments, respectively originated from HOL and BSW that mainly showed the lowest Fst with their breed origin populations, (HC vs breed column). Also, depending originated from HOL or BSW, they showed the highest Fst with BSW or HOL, reversely which meant they were specifically under selection from each dairy breeds. Here, except for BSW, several QTL studies were observed for HOL, HER and AN populations in AnimalQTL database which were associated to milk and body weight traits.

The main objectives of adding HOL and BSW to HER were to obtain more growth potential and milk production. This idea was due to the fact of genetic differentiation between HER and HOL/BSW populations that led to diversity in the performance of animals (Lawson 1982). Thus, based on this fact, for each found fragment, the average differentiation (Fst)

between HER and HOL or BSW populations were compared to the respective value between HER and HC.

As shown in Table 5.5, the Fst between HC and HER was exactly the same as HER and HOL/BSW populations for HOL fragments 1, 4, 5, 7 and BSW fragment 3 (green). This implied that HOL/BSW selection in HC for these genomic regions made a similar effect on body weight or milk traits as average differentiation observed between HER and dairy breeds. On the other hand, the Fst between HC and HER was larger than HER and HOL/BSW populations for HOL fragments 2, 3 and BSW fragment 1, 2, 4 and 5 (red). In this case, there should be a specific selection of HOL/BSW haplotypes in HC that made a larger than average differentiation between HER and dairy breeds, specifically effective for weight traits. In other words, the increasing trend in ROH length of chromosomal segment 6-2 might be essentially related to the selection of these genomic areas. In terms of HOL fragments 6 and 8, the Fst between HC and HER was smaller than the average differentiation between HER and dairy breeds (purple). This meant there should be a weak selection from HOL/BSW haplotypes that made a lower than average differentiation between HER and dairy populations. Thus, it was more probable that HC acted similar to HER than dairy breeds for these regions. Similarly, in terms of Table 5.7 for chromosomal segment 29-1, in spite of over-representation, HOL selection in HC was still not strong enough to make equal Fst as or higher than average differentiation between HER and HOL populations for milk protein traits (a few BSW fragments were found but not reported due to no available QTL data). Moreover, since the whole chromosome 29 did not show a trend in ROH (Table 5.3), the trend in segment 1 should have started recently after the year 2000.

It is noteworthy to mention that the increase in ROH trend was more probably as a result of reduction in population size and not selection. This could be why the deviation of HC from Fst

between HER and dairy populations were not meaningfully large for all found fragments. Moreover, the number of chromosomal segments showing positive ROH trend was actually much smaller than all over-represented segments for HOL and BSW (Table 5.4). This accordingly made a smaller number of HOL/BSW origin fragments to have higher Fst than average differentiation from HER. Considering all QTLs affecting traits of interest, the effects of individual SNPs from these fragments will be negligible to create meaningful differentiation in HC weights relative to HER.

Exploring AnimalQTL database, among autosomal chromosomes affecting BW, WW, YW and body weight gain, BTA 6 comprised tremendous number of studied SNPs and/or QTL areas and thus was substantially important (Snelling *et al*. 2010). Comparing phenotypic averages for weight records between HER and HOL or BSW populations demonstrated that HER is generally lighter in weight for BW, WW and YW (MacNeil *et al*. 2000; Coffey *et al*. 2006; Chin-Colli *et al*. 2016). Before the year 2000, HC phenotypically looked as heavy as HOL/BSW populations for BW and WW and even more for YW. In terms of BW, occupation of key genomic regions by HOL and BSW might bring to mind the potential reason of calf size at birth for HC first-calf heifers and why they had problem in calving. After this time, the remaining composite animals showed considerable decrease in BW similar to HER average again which might be as a result of more usage of AN. Furthermore, in spite of small reduction in WW and YW, they were still higher than HER averages. These comparisons along with the lower differentiation from HER in chromosome 6 (Table 5.4) demonstrated the fact that though having HOL/BSW origin fragments in over-represented areas of this chromosome (all not shown), if no systematic selection is applied (as shown in chapter three), the genetic performance of HC weights might gradually return to be close to that of HER.

Chromosome 6 also plays a fundamental role in milk production and its components (Snelling *et al*. 2010). It has been shown that milk components account for a significant proportion of variation in weaning weight (Butson *et al*. 1980). Although there were no phenotypic records or direct selection on such traits in HC, some signatures of HOL and BSW were found. This might refer to the point that selection on weight traits have probably had an indirectly correlated effect on selection of QTLs associated to milk yield and its components. In terms of areas under positive ROH trend in chromosomal-segment 29-1, when trends of ROH are not as a result of direct or purposive selection, it might be more probable to observe QTL areas that are not associated to key traits of selection.

5.4.2.2 Over-representation of AN and HER

Regarding Table 5.6, although AN had on average the lowest percentage across the HC genome (8%), it was over-represented in some chromosomal segments and was under specific selection through increased ROH trend in the chromosomal segment 20-2 (Table 5.4). In this region, 7 AN origin fragments were found by Fst ranking method which also indicated the lowest Fst with AN population. Moreover, they typically had the highest differentiation with dairy breeds. Considering body weight traits, only fragment 3 showed a meaningful substitution of HER by AN in which HC showed higher than average Fst between HER and AN populations. In terms of association to calving ease, fragment 4 showed a larger than average differentiation between AN and HER and/or HOL (not shown). For milk traits, except for the fragment 4, other fragments showed a lower than average Fst between AN and HER. In addition, all the four fragments had higher than average differentiation between AN and HOL populations (not shown). Chromosome 20 has been found to have QTLs affecting body weight and calving ease (Snelling *et al*. 2010; Cole *et al*. 2011). It is also among the most important regions in dairy

cattle that affect milk production and components (Meredith *et al*. 2012). Despite small number of AN fragments with SNPs having negligible effects, the ROH trend occurred in these regions might imply promising candidates of AN introduction into HC for effects on birth weight and calving difficulty. However, its importation at the expense of HOL QTLs for milk production traits should be monitored for the future of HC development.

In terms of HER over-represented chromosomal segments, although no ROH trend was observed, it seemed 6-4 was the most differentiated segment for HER selection (Table 5.4). Searching in AnimalQTL database substantially demonstrated the presence of abundant number of QTLs in this segment effective for milk production and protein components (Buitenhuis *et al*. 2016). This might imply the fact that HER still has the main role of milk production in HC (Zimmerman 1980; Lawson 1981). Moreover, there were considerable numbers of loci/QTLs that studied eye area and facial pigmentation in Fleckvieh cattle (Mészáros *et al*. 2015). Breeds with white heads like Fleckvieh and HER are more susceptible to eye cancer due to a lack of ambilateral circumocular pigmentation (Pausch *et al*. 2012). On the contrary, BSW has a dark eye pigmentation which helps the breed to resist extreme solar radiation. Eye cancer is the most prevalent tumour affecting cattle and causes substantial economic losses (Pausch *et al*. 2012). Therefore, the hypothesis was that adding BSW to the HC composition could reduce problems with HER eye pigmentation such as eye cancer (Pausch *et al*. 2012). There is a polygenic inheritance pattern of pigmentation in cattle studies which shows selection of animals with eye area pigmentation rapidly reduces the incidence of cancer. As such QTLs explain almost half of the phenotypic variation of animals with pigmented eyes, they provide a basis for future effective genomic selection against eye cancer disease (Pausch *et al*. 2012). In this case, searching on associated areas across the HC genome found no BSW origin fragments to support the

hypothesis. Indeed, there was an increased frequency observed for HER origin alleles in these regions. Grosz and MacNeil (1999) also found a QTL on this chromosomal segment that was linked to HER piebald patterns. Thus, it is probable to mention why HC is more like HER animals with white face. Although coat color inheritance was never an important issue for the HC breeder (Fleming *et al*. 2016), the present composite appearance looks more like 'black baldies' which are the results of mating between Hereford and Angus, progenies with the coat color of Angus while the white markings of Hereford. Exploring AnimalQTL database demonstrated chromosome 5 as the location of majority of QTLs associated to coat color (Mészáros *et al*. 2015) which included around 20% AN proportion in HC. Looking at Fst ranking of the respective positions showed AN predominant relative to HOL. Thus, this might strengthen the idea that adding Angus females in the HC foundation herd might have had a meaningful effect on the coat color of animals.

In terms of over-representation of HER on BTA 1, 7 and 29, some fragments were found to overlap with QTL regions that are associated to fertility and carcass traits (Allais *et al*. 2014; Doran *et al*. 2014; Saatchi *et al*. 2014). This might indicate the importance of HER selection in HC for such traits.

In total, as observed in Table 5.4, despite having HOL, BSW and AN over-representation across the HC genome, chromosomal segments made a lower differentiation with HER due to lack of genetic selection. Moreover, for all other chromosomal segments than these 28, HER percentage corresponded to the average proportion across the HC genome (i.e., 51%). Therefore, it was more probable for this breed to have the priority of effect on all the traits of interest in founders, especially relative to BSW that had the highest differentiation from HC and contained a low percentage across the HC genome.

In this study, it was mainly attempted to focus on over-represented areas relative to whole genome averages in order to have more chance in finding potential SNP sequences originating from a specific ancestral breed. Totally, breed proportions across chromosomes seem to act like a zero-sum game. This means the breed proportion for one breed cannot be low/high without the breed proportion for other breeds being high/low and this consequently indicates selection against/in favor of that breed relative to others.

A meaningful increase in the sum of ROH per chromosome over time might imply some sort of intense selection that made not only an increase in the frequency of some specific autozygous regions, but also the fraction of the autosomal genome under ROH. Since homozygous stretches printed on the genome may have emerged by artificial selection, autozygosity based on ROH can strongly unfold the conception of genetic selection (Marras *et al*. 2015). Therefore, trends of sum of ROH per chromosome could be helpful to monitor severity of selection from ancestral breeds.

In total, it seems the approach of looking at a sequence of Fst ranks of 1 and 2 across the genome properly works to identify its ancestral breed origin. However, to confirm the origin, doing a regression analysis is also necessary. This is because there are some genomic areas that show similar Fst ranks of 1 and 2 between dairy and/or beef breeds which means either selection will have similar impact on the composite performance. Besides, the occurrence of such fragments on shared ROH regions might imply the selection of these sequences from ancestral breeds. Therefore, incorporation of Fst, ROH and breed proportion analysis through regression method will be efficient to trace signatures of selection of ancestral breeds of a composite.

5.5 Conclusion

Considering average genomic inbreeding coefficient along with a few chromosomes showing annually small positive trend for sum of ROH and also, the frequency range of SNPs in ROH, it seems the levels of inbreeding in HC is still not alarming for different traits. HC data was lack of phenotypic information for many traits of importance in founding breeds; however, the use of Fst ranks, ROH, breed proportion analysis through regression method and AnimalQTL database helps to find genomic areas that might imply signatures of selection of those traits from constructing breeds. Despite over-representation of HOL and BSW haplotypes across the HC chromosomal segments, they mainly showed lower differentiation with HER. This was as a result of lack of genetic selection programs which could make HC to act more like the HER population for all the traits of interest in ancestral breeds. Along with selection, due to having averagely lower BSW proportion across the HC genome, no fragments were found on relative over-represented regions associated to lactation persistency and eye area pigmentation, the two traits that HC breeder intended to get from BSW. Although AN had the lowest percentage across the genome, an increasing trend of ROH was observed for its over-representation, especially at the expense of HOL origin for association to milk production which meant it should be monitored for the future development of HC.

5.6 References

Allais, S., Levéziel, H., Hocquette, J.F., Rousset, S., Denoyelle, C., Journaux, L. and Renand, G., 2014. Fine mapping of quantitative trait loci underlying sensory meat quality traits in three French beef cattle breeds. Journal of Animal Science, 92(10), pp.4329-4341.

Barbato, G., Barini, E.M., Genta, G. and Levi, R., 2011. Features and performance of some outlier detection methods. Journal of Applied Statistics, 38(10), pp.2133-2149.

Buitenhuis, B., Poulsen, N.A., Gebreyesus, G. and Larsen, L.B., 2016. Estimation of genetic parameters and detection of chromosomal regions affecting the major milk proteins and their post translational modifications in Danish Holstein and Danish Jersey cattle. BMC genetics, 17(1), pp.1-12.

Butson, S., Berg, R.T. and Hardin, R.T., 1980. Factors influencing weaning weights of range beef and dairy-beef calves. Canadian Journal of Animal Science, 60(3), pp.727-742.

Cassell, B.G., Adamec, V. and Pearson, R.E., 2003. Effect of incomplete pedigrees on estimates of inbreeding and inbreeding depression for days to first service and summit milk yield in Holsteins and Jerseys. Journal of Dairy Science, 86(9), pp.2967-2976.

CDA Livestock Division. 1976. The Hays Converter. Agriculture Canada, Canadex: 420.32.

Chin-Colli, R.D.C., Estrada-León, R., Magaña-Monforte, J., Segura-Correa, J. and Núñez-Domínguez, R., 2016. Genetic parameters for growth and reproductive traits of Brown Swiss cattle from Mexico. Ecosistemas y recursos agropecuarios, 3(7), pp.11-20.

Cleveland, M.A., Blackburn, H.D., Enns, R.M. and Garrick, D.J., 2005. Changes in inbreeding of US Herefords during the twentieth century. Journal of animal science, 83(5), pp.992-1001.

Coffey, M.P., Hickey, J. and Brotherstone, S., 2006. Genetic aspects of growth of Holstein-Friesian dairy cows from birth to maturity. Journal of Dairy Science, 89(1), pp.322-329.

Cole, J.B., Wiggans, G.R., Ma, L., Sonstegard, T.S., Lawlor, T.J., Crooker, B.A., Van Tassell, C.P., Yang, J., Wang, S., Matukumalli, L.K. and Da, Y., 2011. Genome-wide association analysis of thirty one production, health, reproduction and body conformation traits in contemporary US Holstein cows. BMC genomics, 12(1), pp.1-17.

Doran, A.G., Berry, D.P. and Creevey, C.J., 2014. Whole genome association study identifies regions of the bovine genome and biological pathways involved in carcass trait performance in Holstein-Friesian cattle. BMC genomics, 15(1), pp.1-12.

Ferenčaković, M., Hamzić, E., Gredler, B., Solberg, T.R., Klemetsdal, G., Curik, I. and Sölkner, J., 2013. Estimates of autozygosity derived from runs of homozygosity: empirical evidence from selected cattle populations. Journal of Animal Breeding and Genetics, 130(4), pp.286-293.

Fleming, A., 2013. A genetic characterization of the Hays Converter. M.Sc. thesis. University of Guelph, ON, Canada.

Fleming, A.L., Hays, D.P., Moore, S., Plastow, G., Miller, S.P., 2016. History of the Hays Converter. Can. J. Anim. Sci. 96, 471-477.

Frankham, R., Ballou, S.E.J.D., Briscoe, D.A. and Ballou, J.D., 2002. Introduction to conservation genetics. Cambridge university press.

Gengler, N., Misztal, I., Bertrand, J.K. and Culbertson, M.S., 1998. Estimation of the dominance variance for postweaning gain in the US Limousin population. Journal of animal science, 76(10), pp.2515-2520.

Gregory, K. E., Cundiff, L. V., & Koch, R. M. (1999). Composite breeds to use heterosis and breed differences to improve efficiency of beef production. U.S. Department of Agriculture Technical Bulletin No. 1875, 81 pp.

Grosz, M.D. and MacNeil, M.D., 1999. The" spotted" locus maps to bovine chromosome 6 in a Hereford-Cross population. The Journal of heredity, 90(1), pp.233-236.

Grubbs, F.E., 1950. Sample criteria for testing outlying observations. The Annals of Mathematical Statistics, 21(1), pp.27-58.

Hu, Z.L., Park, C.A. and Reecy, J.M., 2019. Building a livestock genetic and genomic information knowledgebase through integrative developments of Animal QTLdb and CorrDB. Nucleic acids research, 47(D1), pp.D701-D710.

Keller, L.F. and Waller, D.M., 2002. Inbreeding effects in wild populations. Trends in ecology & evolution, 17(5), pp.230-241.

Kim, E.S., Cole, J.B., Huson, H., Wiggans, G.R., Van Tassell, C.P., Crooker, B.A., Liu, G., Da, Y. and Sonstegard, T.S., 2013. Effect of artificial selection on runs of homozygosity in US Holstein cattle. PloS one, 8(11), p.e80813.

Kirin, M., McQuillan, R., Franklin, C.S., Campbell, H., McKeigue, P.M. and Wilson, J.F., 2010. Genomic runs of homozygosity record population history and consanguinity. PloS one, 5(11), p.e13996.

Lawson, J.E., 1981. Milk production and weight of cows of the Hereford and Highland breeds and their reciprocal crosses. Canadian Journal of Animal Science, 61(4), pp.861-869.

Lawson, J.E., 1982. Performance of dairy crosses with the HEREFORD in a range environment. Canadian Journal of Animal Science, 62(3), pp.717-724.

MacNeil, M.D., 2009. Invited review: Research contributions from seventy-five years of breeding Line 1 Hereford cattle at Miles City, Montana. Journal of animal science, 87(8), pp.2489-2501.

MacNeil, M.D., Urick, J.J. and Decoudu, G., 2000. Characteristics of Line 1 Hereford females resulting from selection by independent culling levels for below-average birth weight and high yearling weight or by mass selection for high yearling weight. Journal of animal science, 78(9), pp.2292-2298.

Marras, G., Gaspa, G., Sorbolini, S., Dimauro, C., Ajmone-Marsan, P., Valentini, A., Williams, J.L. and Macciotta, N.P., 2015. Analysis of runs of homozygosity and their relationship with inbreeding in five cattle breeds farmed in Italy. Animal genetics, 46(2), pp.110-121.

McQuillan, R., Leutenegger, A.L., Abdel-Rahman, R., Franklin, C.S., Pericic, M., Barac-Lauc, L., Smolej-Narancic, N., Janicijevic, B., Polasek, O., Tenesa, A. and MacLeod, A.K., 2008. Runs of homozygosity in European populations. The American Journal of Human Genetics, 83(3), pp.359-372.

Meredith, B.K., Kearney, F.J., Finlay, E.K., Bradley, D.G., Fahey, A.G., Berry, D.P. and Lynn, D.J., 2012. Genome-wide associations for milk production and somatic cell score in Holstein-Friesian cattle in Ireland. BMC genetics, 13(1), pp.1-11.

Mészáros, G., Petautschnig, E., Schwarzenbacher, H. and Sölkner, J., 2015. Genomic regions influencing coat color saturation and facial markings in Fleckvieh cattle. Animal genetics, 46(1), pp.65-68.

Meyermans, R., Gorssen, W., Buys, N. and Janssens, S., 2020. How to study runs of homozygosity using PLINK? A guide for analyzing medium density SNP data in livestock and pet species. BMC genomics, 21(1), pp.1-14.

Nei, M., 1977. F-statistics and analysis of gene diversity in subdivided populations. Annals of human genetics, 41(2), pp.225-233.

Pausch, H., Wang, X., Jung, S., Krogmeier, D., Edel, C., Emmerling, R., Götz, K.U. and Fries, R., 2012. Identification of QTL for UV-protective eye area pigmentation in cattle by progeny phenotyping and genome-wide association analysis. PloS one, 7(5), p.e36346.

Pemberton, T.J., Absher, D., Feldman, M.W., Myers, R.M., Rosenberg, N.A. and Li, J.Z., 2012. Genomic patterns of homozygosity in worldwide human populations. The American Journal of Human Genetics, 91(2), pp.275-292.

Peripolli, E., Metzger, J., De Lemos, M.V.A., Stafuzza, N.B., Kluska, S., Olivieri, B.F., Feitosa, F.L.B., Berton, M.P., Lopes, F.B., Munari, D.P. and Lôbo, R.B., 2018. Autozygosity islands and ROH patterns in Nellore lineages: Evidence of selection for functionally important traits. BMC genomics, 19(1), pp.1-14.

Peripolli, E., Stafuzza, N.B., Amorim, S.T., de Lemos, M.V.A., Grigoletto, L., Kluska, S., Ferraz, J.B.S., Eler, J.P., Mattos, E.C. and Baldi, F., 2020. Genome-wide scan for runs of homozygosity in the composite Montana Tropical® beef cattle. Journal of Animal Breeding and Genetics, 137(2), pp.155-165.

Pryce, J.E., Haile-Mariam, M., Goddard, M.E. and Hayes, B.J., 2014. Identification of genomic regions associated with inbreeding depression in Holstein and Jersey dairy cattle. Genetics Selection Evolution, 46(1), pp.1-14.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., De Bakker, P.I., Daly, M.J. and Sham, P.C., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. The American journal of human genetics, 81(3), pp.559-575.

Purfield, D.C., Berry, D.P., McParland, S. and Bradley, D.G., 2012. Runs of homozygosity and population history in cattle. BMC genetics, 13(1), p.70.

Qanbari, S., Gianola, D., Hayes, B., Schenkel, F., Miller, S., Moore, S., Thaller, G. and Simianer, H., 2011. Application of site and haplotype-frequency based approaches for detecting selection signatures in cattle. BMC genomics, 12(1), pp.1-12.

Saatchi, M., Schnabel, R.D., Taylor, J.F. and Garrick, D.J., 2014. Large-effect pleiotropic or closely linked QTL segregate within and across ten US cattle breeds. BMC genomics, 15(1), pp.1-17.

Sargolzaei, M., Iwaisaki, H. and Colleau, J.J., 2006. CFC: a tool for monitoring genetic diversity. Proc. 8th World Congr. Genet. Appl. Livest. Prod., CD-ROM Communication, (27-28), pp.13-18.

Snelling, W.M., Allan, M.F., Keele, J.W., Kuehn, L.A., McDaneld, T., Smith, T.P.L., Sonstegard, T.S., Thallman, R.M. and Bennett, G.L., 2010. Genome-wide association study of growth in crossbred beef cattle. Journal of animal science, 88(3), pp.837-848.

Stefansky, W., 1972. Rejecting outliers in factorial designs. Technometrics, 14(2), pp.469-479.

Sumreddee, P., Toghiani, S., Hay, E.H., Roberts, A., Agrrey, S.E. and Rekaya, R., 2019. Inbreeding depression in line 1 Hereford cattle population using pedigree and genomic information. Journal of animal science, 97(1), pp.1-18.

Wiggans, G.R., VanRaden, P.M. and Zuurbier, J., 1995. Calculation and use of inbreeding coefficients for genetic evaluation of United States dairy cattle. Journal of dairy science, 78(7), pp.1584-1590.

Wright, S., 1922. Coefficients of inbreeding and relationship. The American Naturalist, 56(645), pp.330-338.

Wright, S., 1949. The genetical structure of populations. Annals of eugenics, 15(1), pp.323-354.

Wright, S., 1965. The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution, pp.395-420.

Zhao, F., McParland, S., Kearney, F., Du, L. and Berry, D.P., 2015. Detection of selection signatures in dairy and beef cattle using high-density genomic information. Genetics Selection Evolution, 47(1), pp.1-12.

Zimin, A.V., Delcher, A.L., Florea, L., Kelley, D.R., Schatz, M.C., Puiu, D., Hanrahan, F., Pertea, G., Van Tassell, C.P., Sonstegard, T.S. and Marçais, G., 2009. A whole-genome assembly of the domestic cow, Bos taurus. Genome biology, 10(4), pp.1-10.

Zimmerman, W.E., 1980. Lactation and weaning weight relationships in Hereford and Simmental-Hereford cows in Southern Utah.

5.7 Figures and tables

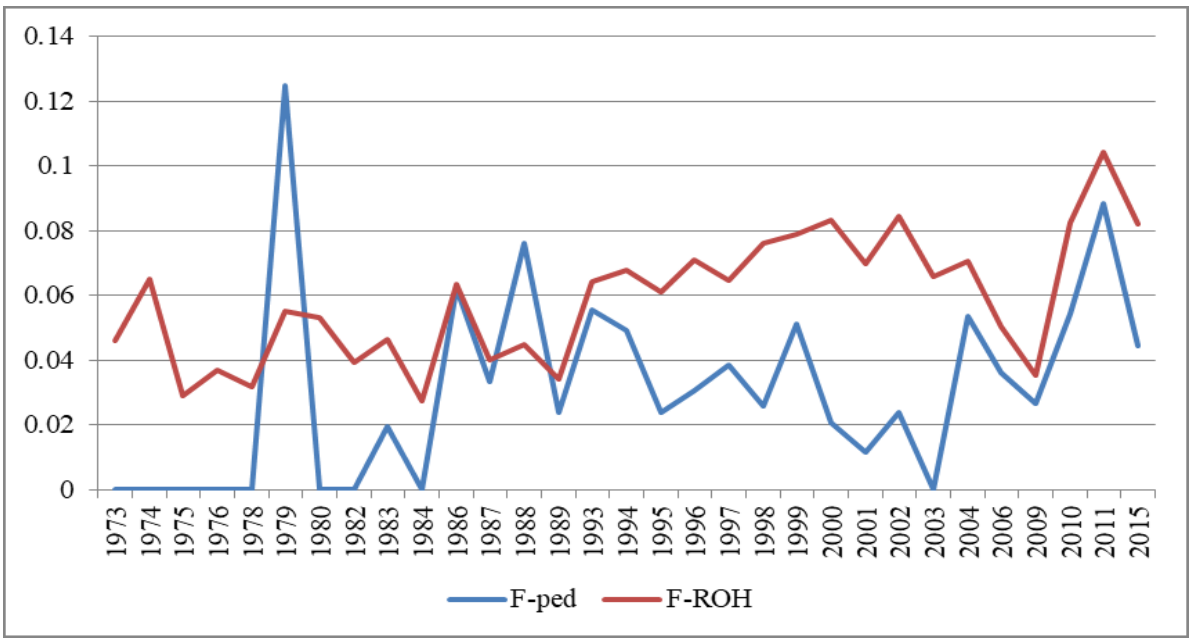

Figure 5.1. The trends of average inbreeding coefficients calculated through pedigree and ROH data from 1973 to 2015.

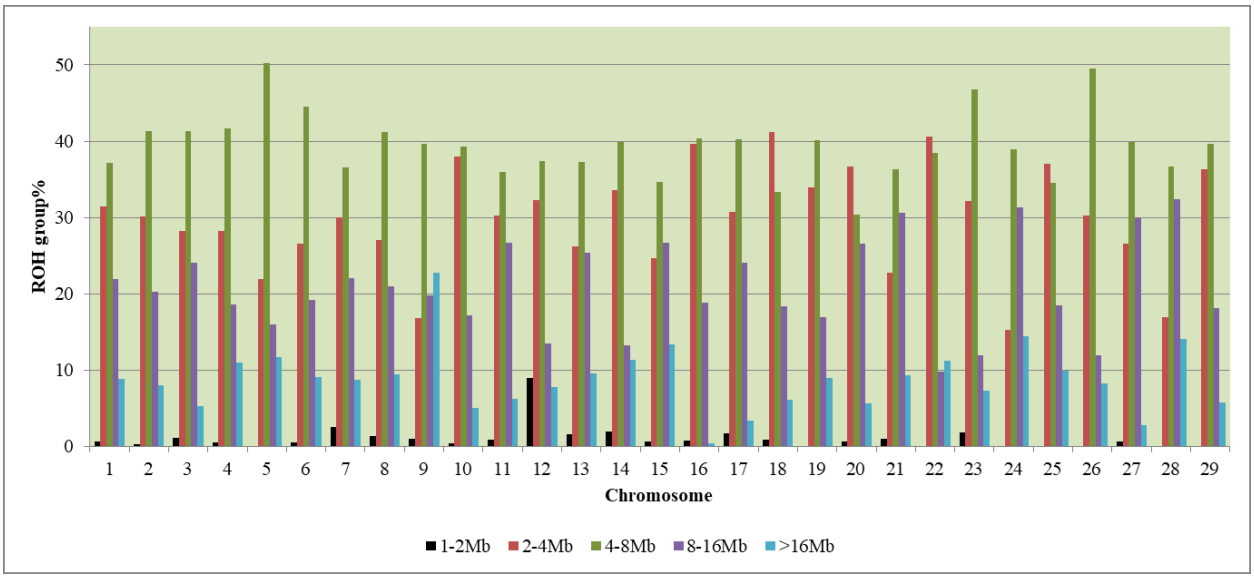

Figure 5.2. The frequency distribution of ROH segments groups according to their length per chromosome.

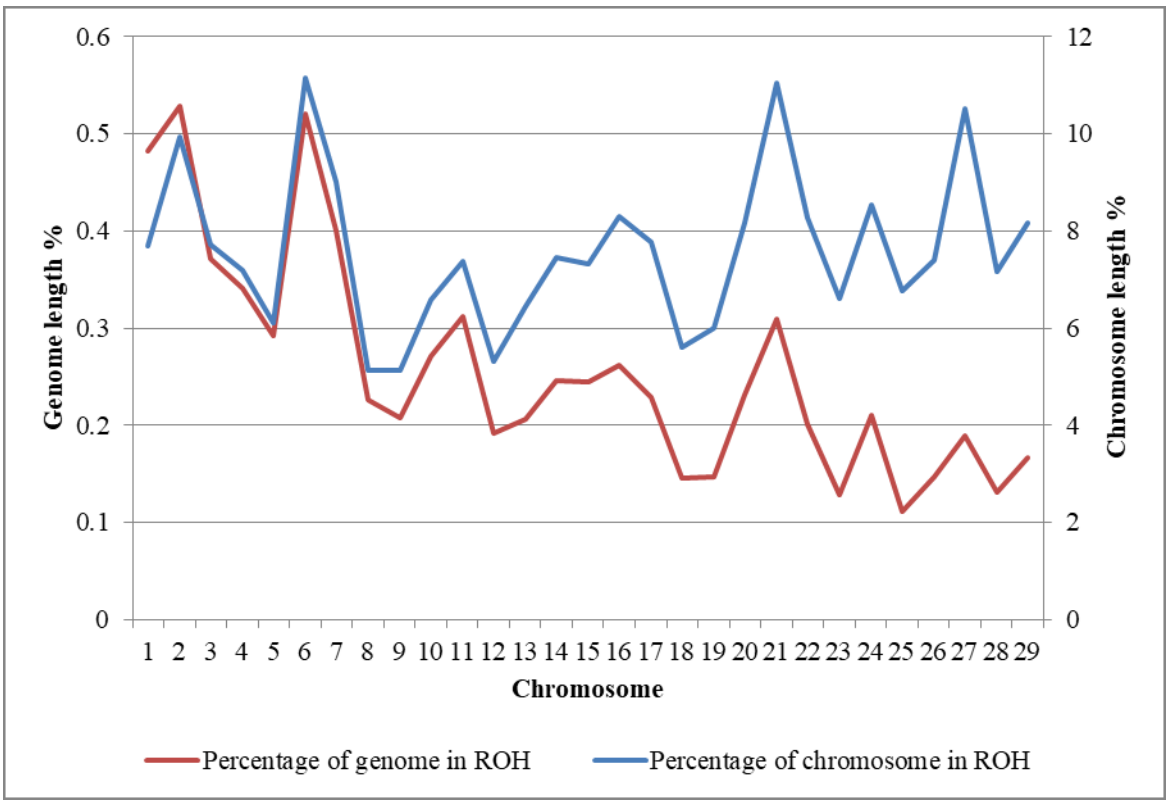

Figure 5.3. Partition of ROH segments to chromosomes to estimate proportion relative to the whole genome and each chromosome lengths, respectively.

Table 5.1. Ranking of Fst values (multiplied by 1000) in some part of segment 2 of chromosome 20, over-represented for AN (yellow and blue colors for HC-AN Fst values as ranks 1 and 2, respectively).

SNP	location	Fst				Fst rank			
Order	(Mb)	HC-AN	HC-HER	HC-BSW	HC-HOL	HC-AN	HC-HER	HC-BSW	HC-HOL
33080	24.74	12.61	48.27	68.25	211.57	1	2	3	4
33081	24.77	0.48	27.65	40.75	67.41	1	2	3	4
33082	24.8	12.2	27.1	13.57	21.23	1	4	2	3
33083	24.83	4.13	56.11	0.01	119.33	2	3	1	4
33084	24.86	4.75	27.74	5.69	72.7	1	3	2	4
33085	24.88	32.65	102.8	112.44	29.86	2	3	4	1
33086	24.97	0.38	23.36	225.11	17.99	1	3	4	2
33087	25.19	0.42	1.98	215.86	57.99	1	2	4	3
33088	25.26	1.74	23.28	44.99	42.04	1	2	4	3
33089	25.29	0.03	0.74	33.35	178.16	1	2	3	4

Table 5.2. Descriptive statistics of ROH results through PLINK.

	Mean	SD	Min	Max
ROH_{nSNPs}[1]	125.82	104.54	30	1158
ROH_D[2](Mb)	0.061	0.012	0.036	0.129
ROH_N[3]	25.15	9.41	2	52
ROH_L[4](Mb)	7.54	6.29	1.39	70.02
ROH_T[5](Mb)	189.66	107.45	7.81	624.06

[1]number of SNPs per ROH segment
[2]density of ROH: The mean distance between two consecutive SNPs in a ROH segment
[3]number of ROH segments per animal
[4]one ROH segment length
[5]sum of ROH segments length per animal

Table 5.3. Sum of ROH trends and averages (Mb) per chromosome from 1973 to 2015.

Chromosome	b (trend)	a (intercept)
Chr1	0.10±0.09	12.19**±0.94
Chr2	0.25*±0.10	13.42**±1.11
Chr3	0.16*±0.07	9.41**±0.79
Chr4	0.007±0.07	8.63**±0.79
Chr5	0.02±0.07	7.39**±0.80
Chr6	0.29**±0.09	13.22**±1.01
Chr7	0.15±0.08	10.16**±0.90
Chr8	0.11±0.06	5.73**±0.65
Chr9	0.10±0.07	5.26**±0.73
Chr10	-0.05±0.06	6.82**±0.67
Chr11	0.18**±0.07	7.92**±0.77
Chr12	0.09±0.06	4.86**±0.66
Chr13	0.10±0.05	5.24**±0.61
Chr14	0.11±0.06	6.25**±0.68
Chr15	0.06±0.07	6.20**±0.75
Chr16	0.03±0.05	6.60**±0.58
Chr17	0.06±0.05	5.78**±0.60
Chr18	0.03±0.04	3.67**±0.43
Chr19	0.05±0.04	3.71**±0.43
Chr20	0.13*±0.05	5.89**±0.60
Chr21	-0.13±0.07	7.76**±0.78
Chr22	0.03±0.05	5.09**±0.54
Chr23	0.02±0.04	3.25**±0.41
Chr24	0.11±0.06	5.33**±0.66
Chr25	0.06±0.04	2.82**±0.39
Chr26	0.09±0.05	3.74**±0.52
Chr27	0.05±0.05	4.78**±0.50
Chr28	0.04±0.04	3.33**±0.45
Chr29	0.07±0.05	4.22**±0.50

*(P<0.05)
**(P<0.01)

Table 5.4. The 28 chromosomal segments selected by Grubbs test showing over (green) or under (yellow) representation of AN, HER, BSW and HOL proportions relative to the entire genome averages with trends in ROH and Fst.

Chr region	Breed proportions				Trend in ROH over time (Mb)	Genetic distance, Fst (HC vs breed)			
	AN	HER	BSW	HOL		AN	HER	BSW	HOL
1,4	0.13±0.006	0.68±0.008	0.08±0.004	0.11±0.005	0.04±0.03	0.03	0.01	0.07	0.07
1,6	0.06±0.004	0.37±0.009	0.17±0.007	0.4±0.009	0.03±0.03	0.03	0.03	0.06	0.05
2,5	0.1±0.005	0.24±0.007	0.42±0.008	0.24±0.008	0.06±0.03	0.02	0.03	0.04	0.06
2,6	0.09±0.004	0.23±0.006	0.15±0.007	0.53±0.01	0.03±0.02	0.02	0.05	0.05	0.03
6,2	0.18±0.008	0.13±0.006	0.23±0.007	0.46±0.01	**0.09*±0.04**	0.02	0.03	0.08	0.03
6,3	0.05±0.004	0.37±0.008	0.14±0.005	0.44±0.008	0.04±0.03	0.02	0.02	0.08	0.03
6,4	0.13±0.005	0.68±0.006	0.11±0.004	0.08±0.004	0.07±0.04	0.06	0.02	0.10	0.10
7,5	0.14±0.005	0.67±0.008	0.09±0.005	0.1±0.005	0.05±0.03	0.02	0.01	0.06	0.05
8,1	0.13±0.006	0.31±0.008	0.14±0.006	0.42±0.009	-0.003±0.02	0.02	0.02	0.04	0.03
8,4	0.17±0.006	0.2±0.008	0.38±0.008	0.25±0.007	0.006±0.02	0.02	0.03	0.04	0.04
9,5	0.07±0.004	0.26±0.009	0.35±0.007	0.32±0.008	0.01±0.01	0.02	0.02	0.04	0.04
12,3	0.1±0.006	0.28±0.008	0.22±0.007	0.4±0.009	0.02±0.03	0.02	0.03	0.05	0.03
12,4	0.32±0.008	0.25±0.008	0.19±0.006	0.24±0.007	0.02±0.01	0.02	0.02	0.07	0.04
16,4	0.27±0.008	0.21±0.008	0.08±0.005	0.44±0.008	-0.003±0.01	0.02	0.02	0.04	0.04
17,1	0.1±0.005	0.33±0.009	0.14±0.005	0.43±0.009	0.02±0.02	0.02	0.02	0.05	0.04
18,1	0.22±0.008	0.18±0.007	0.12±0.005	0.48±0.01	0.001±0.02	0.02	0.04	0.07	0.03
19,1	0.11±0.006	0.36±0.01	0.1±0.004	0.43±0.01	0.01±0.02	0.03	0.02	0.05	0.03
20,2	0.38±0.008	0.34±0.009	0.09±0.005	0.19±0.005	**0.07**±0.02**	0.02	0.03	0.05	0.07
20,3	0.07±0.004	0.41±0.007	0.36±0.008	0.16±0.005	0.03±0.02	0.03	0.02	0.05	0.07
20,4	0.05±0.004	0.47±0.01	0.34±0.01	0.14±0.005	-0.01±0.01	0.03	0.02	0.05	0.06
22,3	0.22±0.007	0.24±0.008	0.11±0.004	0.43±0.01	0.03±0.02	0.02	0.02	0.06	0.04
24,1	0.1±0.005	0.21±0.008	0.25±0.007	0.44±0.01	0.04±0.03	0.02	0.04	0.04	0.03
24,2	0.38±0.008	0.23±0.007	0.12±0.006	0.27±0.005	0.02±0.02	0.02	0.04	0.05	0.05
25,1	0.14±0.006	0.28±0.008	0.17±0.006	0.41±0.009	0.02±0.02	0.02	0.03	0.06	0.04
25,2	0.15±0.006	0.28±0.008	0.18±0.006	0.39±0.008	0.02±0.02	0.02	0.02	0.05	0.03
28,2	0.18±0.007	0.34±0.009	0.07±0.003	0.41±0.008	0.04±0.03	0.01	0.02	0.06	0.03
29,1	0.08±0.004	0.19±0.007	0.23±0.007	0.5±0.009	**0.04*±0.02**	0.01	0.03	0.05	0.02
29,3	0.09±0.004	0.67±0.007	0.08±0.005	0.16±0.006	0.003±0.009	0.02	0.02	0.07	0.06

*(P<0.05)
**(P<0.01)

Table 5.5. Signatures of selection of HOL and BSW in chromosome 6, segments 2 (relative results underlined if breed origin was equivalent to the studied breed).

order	chr,seg	#SNPs	from-to (Mb)	Predominant breed origin in HC through regression method	PUBMED_ID	trait	studied breed	Fst (HC vs breed)				common ROH frequency	Positive selection through Fst with	Fst between HER and HOL/BSW populations
								AN	HER	BSW	HOL			
1	6,2	6	22.96-23.35	HOL	31138106, 31139206, **19966163**	Milk yield, Milk fat and protein yield, Milk protein percentage, **Body weight (birth)**	HOL **HER AN**	0.01	**0.09**	0.08	0.007	28/205		0.09
2	6,2	6	23.43-23.61	HOL	20412936, **19966163**	Lactation persistency, **Body weight (birth, yearling), Body weight gain**	HOL **HER AN**	0.01	**0.05**	0.14	0.01	28/205	BSW	0.03
3	6,2	7	25.47-25.66	HOL	19966163	Body weight (birth, yearling), Body weight gain	HER AN	0.05	**0.05**	0.05	0.02	34/205	HER	0.03
4	6,2	9	39.21-39.59	HOL	19966163, **28521758**	Body weight (birth, weaning, yearling), Body weight gain, **Metabolic body weight**	HER **AN**	0.01	**0.04**	0.17	0.01	30/205		0.04
5	6,2	13	42.12-42.65	HOL	19966163	Body weight (birth, weaning, yearling), Body weight gain	HER AN	0.02	**0.02**	0.05	0.02	24/205		0.02
6	6,2	8	44.33-44.67	HOL	27485317	Milk protein percentage	HOL	0.01	**0.05**	0.08	0.004	24/205		0.07
7	6,2	11	44.96-45.32	HOL	22497262, 22486504, 25511820, **19966163**	Bone weight, Body weight and length (birth), Milk yield, Milk protein yield, Milk fat and protein percentage, **Body weight (birth, yearling), Body weight gain**	HOL **AN HER**	0.02	**0.03**	0.08	0.007	25/205		0.03
8	6,2	5	45.56-45.9	HOL	**19966163,** 20412936	**Body weight (birth, yearling), Body weight gain,** Lactation persistency	**HER AN** HOL	0.07	**0.02**	0.07	0.03	25/205	AN	0.04
1	6,2	5	26.2-26.55	BSW	19966163	Body weight (birth, weaning, yearling), Body weight gain	HER AN	0.03	**0.03**	0.01	0.10	23/205		0.005
2	6,2	8	31.01-31.46	BSW	19966163	Body weight (birth, yearling),	HER AN	0.01	**0.04**	0.02	0.05	25/205		0.02

						Body weight gain								
3	6,2	8	40.71-40.98	BSW	19966163	Body weight (birth, yearling)	HER AN	0.01	**0.01**	0.008	0.04	26/205		0.01
4	6,2	7	42.68-43.03	BSW	19966163	Body weight (birth, weaning, yearling), Body weight gain	HER AN	0.03	**0.01**	0.008	0.03	24/205		0.008
5	6,2	7	44.69-44.89	BSW	**19966163,** 24796806, 28711251	**Body weight (birth), Body weight gain,** Milk yield, Milk protein yield, Milk protein percentage, Milk fat percentage, Milk casein percentage	**HER AN** HOL	0.01	**0.02**	0.004	0.05	25/205		0.01

Table 5.6. Signatures of selection of AN in chromosome 20, segments 2 (relative results underlined if breed origin was equivalent to the studied breed).

order	chr,seg	#SNPs	from-to (Mb)	Predominant breed origin in HC through regression method	PUBMED_ID	trait	studied breed	Fst (HC vs breed)				common ROH frequency	Positive selection through Fst with	Fst between HER and AN populations
								AN	HER	BSW	HOL			
1	20,2	12	21.6-22.42	AN	19966163	Body weight gain	AN HER	0.005	**0.03**	0.11	0.02	11/205	BSW	0.06
2	20,2	16	22.69-23.86	AN	19966163	Body weight (birth)	AN HER	0.004	**0.03**	0.06	0.07	11/205		0.03
3	20,2	10	24.74-25.29	AN	22497295	Average daily gain	AN	0.007	**0.03**	0.08	0.08	12/205	BSW HOL	0.01
4	20,2	5	28.14-28.32	AN	21831322, 27287773, 22449276	Calving ease (maternal), Milk protein percentage, Milk yield	HOL	0.006	**0.1**	0.05	0.05	9/205		0.07
5	20,2	6	28.44-28.88	AN	27287773	Milk protein and fat percentage	HOL	0.02	**0.04**	0.05	0.11	8/205		0.05
6	20,2	6	36.71-37.06	AN	27287773, 22449276	Milk protein and fat percentage, Milk, fat and protein yield	HOL	0.005	**0.04**	0.05	0.07		HER	0.07
7	20,2	10	38.22-38.76	AN	20630249, 22449276	Milk protein and fat percentage, Milk, fat and protein yield	HOL	0.02	**0.04**	0.06	0.09	15/205	HER HOL BSW	0.05

Table 5.7. Signatures of selection of HOL in chromosomes 29, segment 1 (relative results underlined if breed origin was equivalent to the studied breed).

order	chr,seg	#SNPs	from-to (Mb)	Predominant breed origin in HC through regression method	PUBMED_ID	trait	studied breed	Fst (HC vs breed)				common ROH frequency	Positive selection through Fst with	Fst between HER and HOL populations
								AN	HER	BSW	HOL			
1	29,1	6	11.15-11.61	HOL	27485317	Milk kappa-casein percentage	HOL	0.01	**0.05**	0.008	0.007	19/205		0.10
2	29,1	6	12.11-12.35	HOL	27485317	Milk kappa-casein percentage	HOL	0.01	**0.04**	0.05	0.006	14/205		0.05
3	29,1	5	19.61-19.92	HOL	27485317	Milk glycosylated kappa-casein percentage	HOL	0.01	**0.02**	0.009	0.002	13/205		0.03
4	29,1	6	20.02-20.49	HOL	**28521758,** 27485317	**Average daily gain,** Milk glycosylated kappa-casein percentage	**HER** HOL	0.02	**0.03**	0.10	0.01	17/205	HER	0.07

Chapter 6. General conclusion and recommendations for further research

The present work was first an attempt to investigate key concepts of developing composite cattle with a special insight into their efficiency in productivity under (sub) tropical environments. A simulation study was designed with an interest in the indigenous Afrikaner cattle, a specialized dam line known for limited calving difficulties and improved performance of progeny when crossing with exotic terminal sires. The aim of study was to evaluate opportunities of use of technologies of crossbreeding as a multi-breed composite dam and sexed semen for potential improvement of production efficiency in South Africa.

Following that, the research focused on a Canadian composite breed named Hays Converter (HC) by gathering available pedigree data, phenotypic and genomic records to study their current genetic and genomic characterisations. Since the breeding objective for HC was to create a beef breed that excelled in growth, selection mainly emphasized weaning and yearling weights. However, there was no selection index by genetic values from its inception. Therefore, one of the purposes was to evaluate genetic trends and genetic parameters of body weight measures to observe the current situation of HC.

Due to incompleteness in the HC pedigree file and selling large part of the main herd after the year 2000, it looked necessary to have a more accurate estimation of inbreeding by genomic data. Also, as there was no information on each HC founder proportion available, the next purpose was set to evaluate genomic breed composition of animals and their potential change over time. Finally, because of lack of records of most traits that founders selected for, signatures of their selection were explored through over/under representation of the relative breed proportion on several segments of each HC chromosome.

Looking at chapter two, the simulation study could be considered a demonstration of the utility of systems analysis techniques to explore the effects of many more factors simultaneously when it is not feasible to address all experimentally. This resulted in the suggestion of application of crossbreeding for improvement of production efficiency of indigenous breeds adapted to (sub)tropical climatic conditions. In terms of HC, there has been some interest from the Australian market, with breeders hoping to cross this composite with Brahman cattle to improve crossbred production. It has just been introduced to Australia in 2016 through an embryo transfer project by a commercial farmer, Stewart Murray who started by a small herd of 50 calves. He hopes HC will play a role in increasing the efficiency and sustainability of Australia's cattle industry. His objective is to develop a nucleus composite herd with incorporation of important traits from indigenous breeds like tick resistance and heat tolerance to make them more suitable for the Northern Territory. However, as Dr Stephen Moore, director of the Centre of Animal Science at the University of Queensland believed, establishing a new breed in another country is not easy. Due to being only one generation available, measuring of these animals might not be a true measure of HC or at least a second generation will be required. Also, due to some noticeable limitation in the number of live animals being produced in Canadian HC herds, the importation by Murray or Australian industry seems to be currently not feasible to provide a sufficient supply for further analysis. Moreover, as the Australian industry is largely based on pasture fed beef, it is more difficult to measure the efficiency of HC animals for food utilization rather than feedlot. Therefore, at the moment, it makes the industry hardly convinced that HC is worthwhile to be used in Australia.

Considering chapter three, similarities were found in the EBV for birth, weaning and yearling weight produced from genetic evaluations conducted using random regression and

multiple trait models. This was due to the fact that collected data of weight records mainly occurred at standard points for both models. Therefore, currently choosing MTM for future HC genetic evaluation seems simpler. Increased weaning and yearling weights have always been the main objectives for the composite productivity from the HC inception. However, due to focus on selection by phenotypes, little meaningful progress has occurred in genetic improvement of growth traits over time. In fact, as genetic trends for weight traits in progeny were directly affected by sires, as shown by the sire selection differential results, if there was a structured management program that took into account genetic values when selecting for best sires, it was expected to see a dramatic change in genetic trends for weight traits.

In chapter four, in order to avoid potential problems arising from admixture analysis due to sample selection bias occurred by number of animals and/or SNP density along with elaborate likelihood functions with specific assumptions, regression method was selected as the appropriate approach in analyzing genomic breed composition of the composite. Also, to report average founder breeds proportions of the population, 6K SNP panel was consistent with the historical record of the founding of HC. In total, HC was detected as a mixture of 51% HER, 26% HOL, 15% BSW and 8% AN without specific change in the proportions over years. The reason for presence of some percentage AN was as a result of importation of several AN females in the foundation herd and periodic application of this breed in controlling calving difficulty of some first-calf heifers.

The division of whole genome to small segments of identical length and analyzing the respective breed proportions was able to identify chromosomal regions with possible over-representation of each of the constructing founders. As explained in chapter five, this actually meant signatures of selection of that over-represented ancestral breed which were explorable

across the genome by finding SNP sequences showing lower Fst between HC and that breed. This approach helped to find genomic areas across the genome that most probably originated from the ancestral breeds. Unless for weight, there were no phenotypic data available regarding traits Harry Hays hoped to implement in HC. Therefore, the found sequences were searched for potential association to traits of interest in Animal QTL database. Although the results in general demonstrated the importation of HOL on key chromosomal segments affecting body weight growth, the selection was weak due to having lower Fst with HER population. Similarly, chromosomal segments showed to be genetically farther from BSW relative to HER. This along with lower proportion on the HC genome even meant a weaker selection on BSW so that no fragments were found in its over-represented areas for association to lactation persistency and eye area pigmentation. In terms of HER, though only four over-represented chromosomal segments remained across the genome, the breed still showed large number of fragments associable to milk yield and milk components traits. This might imply the fact that despite importation of HOL to improve these traits, HER was still under meaningful selection. Regarding AN, the breed showed the lowest contribution across the genome; however, its positive selection on over-represented areas should be considered as some caution. Overall, despite mainly HOL/BSW over-representation, lower Fst with HER shows that there was a weak selection of haplotypes in HC. This along with HER as the main breed proportion for all other chromosomal segments implies that HC should act more like HER populations for many the traits of interest.

In terms of inbreeding, the genomic level average through ROH analysis was still in accordance with an amount to 10% for most beef and dairy cattle breed societies (7.5%). A little increase that happened after the year 2000 was actually as a result of selling large part of the

main herd that led to an unavoidable mating between relatives. However, this has not resulted in a meaningful reduction in the heterozygosity level of the genome and the significant increase observed in the frequency of long and very long ROH group was very small. Overall, although the size of population has started to decrease, it is still soon to have any conclusion about inbreeding. Moreover, since this study was mainly done on one segment of the HC population, the estimated inbreeding cannot be generalized.

HC possesses a strong historical background and has the potential to play a role in Canadian beef production industry and even excel in certain environmental production systems. However, the future of the breed depends on a continuous co-operation among the association members, animal science researchers and industry operators. They will need to define a well-structured program in which besides increasing the total size of Canadian herds, useful genetic and genomic tools are applied to gradually improve its situation from a local breed to an industrial position. Therefore, to advance knowledge about HC and boost its position, further acts, research and/or suggestions are recommended as follows:

- A regular control of recording traits

  It is recommended beside tighter control of recording weight traits in all herds, other traits targeted from HC founders are also considered such as fertility, milking ability, conformation, feet and udders features, eye area pigmentation and especially carcass quality which was the ultimate mission of developing this composite. Phenotypes are the main components of genetic analyses for production improvement and this is why high-throughput phenotyping technologies are growing in livestock systems due to their ability to generate real-time, non-invasive, and accurate animal-level information. For example, new techniques are now suggested for more reliable estimation of live cattle birth weight

such as Weigh tape and Schaeffer's formula for WW and YW. Also, near-infrared spectrometry techniques are being developed as tools to predict complex dairy and beef phenotypes, such as milk composition, feed efficiency, fertility, health status, and meat quality traits.

- Comparing HC performance with its founders populations

Except for weight records, there are no research documents to approve claims about the composite capabilities in converting feed to gain efficiently, milk production, fertility, resistance to climatic conditions and qualified carcass. One potential action can be comparing the performance of HC relative to its founders or other beef and dairy breeds to realize how much these claims are true or the composite is different.

- Defining a selection index for HC

For genetic improvement of the HC, it looks like mandatory for the association to provide a well-defined selection index in which traits important for the breed are included. This could help to affect the future profitability of the composite and build a competitive opportunity relative to popular beef breed associations. As the fastest genetic progress is made with sire selection, it is recommended to divide bulls into two groups of terminal and maternal sires and define separate selection indices for each group. As terminal bulls are selected to make progeny that will not enter the reproductive system, it looks sufficient to include growth and carcass traits in the index. However, maternal bulls should be used to produce replacement females and therefore, rather maternal traits including stayability (health), fertility, calving ease, body construction (feet and udders) and milk production should be included in the index. The common point for two indices should be selection for low birth weight EBV bulls.

- Evaluation of genomic breed proportions of all HC

Except for the main herd developed by Dan Hays, there is no information of founder breed proportions from other herds available. Since there have been crosses with other breeds in a number of herds, it looks necessary for the association to update their genomic knowledge of breed proportions for all animals. To genotype animals, providing 50K SNP data for at least all parents and main ancestors of HC descendants will be beneficial for future analyses like imputation, genomic inbreeding, signatures of selection and genomic selection. Furthermore, it will help to develop selection indices in which breed proportions are taken into account. Thus, there will be potential to control birth weight by gradually exclude the importation of AN and back to a HC which is 25% BSW.

- Estimation of inbreeding of all available HC animals

As there is now limitation in creating new HC animals from founding breeds, to control the inbreeding level of the main herd maintaining at U of A, it is recommended the association provides some information about the inbreeding level of other herds by both pedigree and genomic data. This will help to structure mating designs in which the level of autozygosity can be monitored by switching animals among herds.

- And studying potential characterization of HC × indigenous breeds

Since the breed associations relative to HC founders, i.e. HER, HOL and BSW have been established in Australia for many years, they might have access to genetic evaluation data for different traits of interest which consequently could reflect a kind of adapted information for Australian climatic conditions. Therefore, looking at chapter two as a practical example will help to establish scenarios in which it is possible to define an artificial HC of 50% HER, 25% HOL and 25% BSW. This will also let to evaluate its incorporation into the straightbred Brahman or

Afrikaner system through measuring potential production efficiency. In addition, it might help to evaluate the trade-off between costs of local development versus importation.

Bibliography

Akanno, E.C., Chen, L., Abo-Ismail, M.K., Crowley, J.J., Wang, Z., Li, C., Basarab, J.A., MacNeil, M.D. and Plastow, G., 2017. Genomic prediction of breed composition and heterosis effects in Angus, Charolais, and Hereford crosses using 50K genotypes. Canadian Journal of Animal Science, 97(3), pp.431-438.

Albuquerque, L.D., Meyer, K., 2001. Estimates of covariance functions for growth from birth to 630 days of age in Nelore cattle. J. Anim. Sci. 79, 2776-2789.

Alexander, D.H. and Lange, K., 2011. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. BMC bioinformatics, 12(1), p.246.

Alexander, D.H., Novembre, J. and Lange, K., 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome research, 19(9), pp.1655-1664.

Alexander, D.H., Shringarpure, S.S., Novembre, J. and Lange, K., 2015. Admixture 1.3 software manual. Los Angeles: UCLA Human Genetics Software Distribution.

Allaire, F.R. and Henderson, C.R., 1965. Inbreeding within an artificially bred dairy cattle population. Journal of dairy science, 48(10), pp.1366-1371.

Allais, S., Levéziel, H., Hocquette, J.F., Rousset, S., Denoyelle, C., Journaux, L. and Renand, G., 2014. Fine mapping of quantitative trait loci underlying sensory meat quality traits in three French beef cattle breeds. Journal of Animal Science, 92(10), pp.4329-4341.

Anderson, P. (2012) Matching Cattle Type and Feedlot Performance. Univ. Minnesota Extension. https://www.extension.umn.edu/agriculture/beef/components/pdfs/matching_cattle_type_and_feedlot_performance.pdf

Arango, J.A., Cundiff, L.V., Van Vleck, L.D., 2004. Covariance functions and random regression models for cow weight in beef cattle. J. Anim. Sci. 82, 54-67.

Baldi, F., Albuquerque, L.G. and Alencar, M.M., 2010. Random regression models on Legendre polynomials to estimate genetic parameters for weights from birth to adult age in Canchim cattle. Journal of Animal Breeding and Genetics, 127(4), pp.289-299.

Barbato, G., Barini, E.M., Genta, G. and Levi, R., 2011. Features and performance of some outlier detection methods. Journal of Applied Statistics, 38(10), pp.2133-2149.

Barwick, S.A., Johnston, D.J., Burrow, H.M., Holroyd, R.G., Fordyce, G., Wolcott, M.L., Sim, W.D. and Sullivan, M.T. (2009) Genetics of Heifer Performance in Wet and Dry Seasons and their Relationships with Steer Performance in Two Tropical Beef Genotypes. Anim. Prod. Sci, 49, 367-382.

Beef Improvement Federation., 2018. Guidelines for Uniform Beef Improvement Programs. 9th ed. https://beefimprovement.org/wp-content/uploads/2018/03/BIFGuidelinesFinal_updated0318.pdf (accessed 5 January 2019).

Bjelland, D.W., Weigel, K.A., Vukasinovic, N. and Nkrumah, J.D., 2013. Evaluation of inbreeding depression in Holstein cattle using whole-genome SNP markers and alternative measures of genomic inbreeding. Journal of dairy science, 96(7), pp.4697-4706.

Blackburn HD, Paiva, SR, Sollero, BP, Biegelmeyer, P, Caetano, AR, Cardoso, FF (2014) A Dedicated SNP Panel for Evaluating Genetic Diversity in a Composite Cattle Breed. In Proceedings of the World Congress on Genetics Applied to Livestock Production, 048. World Congress on Genetics Applied to Livestock Production.

Boerner, V. and Wittenburg, D., 2018. On estimation of genome composition in genetically admixed individuals using constrained genomic regression. Frontiers in Genetics, 9, p.185.

Bohmanova, J., Misztal, I., Bertrand, J.K., 2005. Studies on multiple trait and random regression models for genetic evaluation of beef cattle for growth. J. Anim. Sci. 83, 62-67.

Boichard, D., Ducrocq, V., Croiseau, P. and Fritz, S., 2016. Genomic selection in domestic animals: principles, applications and perspectives. Comptes rendus biologies, 339(7-8), pp.274-277.

Boligon, A.A., Mercadante, M.E.Z., Forni, S., Lobo, R.B., Albuquerque, L.G.D., 2010. Covariance functions for body weight from birth to maturity in Nellore cows. J. Anim. Sci. 88, 849-859.

Bolormaa, S., Hayes, B.J., Hawken, R.J., Zhang, Y., Reverter, A. and Goddard, M.E., 2011. Detection of chromosome segments of zebu and taurine origin and their effect on beef production and growth. Journal of animal science, 89(7), pp.2050-2060.

Bonsma, J.C. (1949) Breeding Cattle for Increased Adaptability to Tropical and Sub-tropical Environments. J. Agric. Sci, 39, 204-221.

Bovine HapMap Consortium, 2009. Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. Science, 324(5926), pp.528-532.

Brinks, J. S., Clark, R.T., Rice, F.J., 1961. Estimation of genetic trends in beef cattle. J. Anim. Sci. 20, 903.

Buitenhuis, B., Poulsen, N.A., Gebreyesus, G. and Larsen, L.B., 2016. Estimation of genetic parameters and detection of chromosomal regions affecting the major milk proteins and

their post translational modifications in Danish Holstein and Danish Jersey cattle. BMC genetics, 17(1), pp.1-12.

Burrow, H.M. (2012) Importance of Adaptation and Genotype×Environment Interactions in Tropical Beef Breeding Systems. Animal, 6, 729-740.

Burrow, H.W., 2006. Utilization of diverse breed resources for tropical beef production. Proceedings of the 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte, Brasil, August 13–18

Butson, S., Berg, R.T. and Hardin, R.T., 1980. Factors influencing weaning weights of range beef and dairy-beef calves. Canadian Journal of Animal Science, 60(3), pp.727-742.

Buzanskas, M.E., Ventura, R.V., Chud, T.C.S., Bernardes, P.A., de Abreu Santos, D.J., de Almeida Regitano, L.C., de Alencar, M.M., de Alvarenga Mudadu, M., Zanella, R., da Silva, M.V.G.B. and Li, C., 2017. Study on the introgression of beef breeds in Canchim cattle using single nucleotide polymorphism markers. PloS one, 12(2).

Caetano, S.L., Savegnago, R.P., Boligon, A.A., Ramos, S.B., Chud, T.C.S., Lôbo, R.B. and Munari, D.P., 2013. Estimates of genetic parameters for carcass, growth and reproductive traits in Nellore cattle. Livestock Science, 155(1), pp.1-7.

Cardoso, F.F. and Templeman, R.J., 2004. The value of hierarchical Bayes models on genetic evaluation of multiple-breed beef cattle populations. J. Anim. Sci, 83, pp.62-72.

Cartwright, T.C. (1970) Selection Criteria for Beef Cattle for the Future. Journal of Animal Science, 30(5), 706-711.

Cartwright, T.C., 1974. Net effects of genetic variability on beef production systems. Genetics, 78(1), p.541.

Cassell, B.G., Adamec, V. and Pearson, R.E., 2003. Effect of incomplete pedigrees on estimates of inbreeding and inbreeding depression for days to first service and summit milk yield in Holsteins and Jerseys. Journal of Dairy Science, 86(9), pp.2967-2976.

CDA Livestock Division. 1976. The Hays Converter. Agriculture Canada, Canadex: 420.32.

Chiang, C.W., Gajdos, Z.K., Korn, J.M., Kuruvilla, F.G., Butler, J.L., Hackett, R., Guiducci, C., Nguyen, T.T., Wilks, R., Forrester, T. and Haiman, C.A., 2010. Rapid assessment of genetic ancestry in populations of unknown origin by genome-wide genotyping of pooled samples. PLoS genetics, 6(3).

Chin-Colli, R.D.C., Estrada-León, R., Magaña-Monforte, J., Segura-Correa, J. and Núñez-Domínguez, R., 2016. Genetic parameters for growth and reproductive traits of Brown Swiss cattle from Mexico. Ecosistemas y recursos agropecuarios, 3(7), pp.11-20.

Cleveland, M.A., Blackburn, H.D., Enns, R.M. and Garrick, D.J., 2005. Changes in inbreeding of US Herefords during the twentieth century. Journal of animal science, 83(5), pp.992-1001.

Coffey, M.P., Hickey, J. and Brotherstone, S., 2006. Genetic aspects of growth of Holstein-Friesian dairy cows from birth to maturity. Journal of Dairy Science, 89(1), pp.322-329.

Cole, J.B., VanRaden, P.M., O'Connell, J.R., Van Tassell, C.P., Sonstegard, T.S., Schnabel, R.D., Taylor, J.F. and Wiggans, G.R., 2009. Distribution and location of genetic effects for dairy traits. Journal of dairy science, 92(6), pp.2931-2946.

Cole, J.B., Wiggans, G.R., Ma, L., Sonstegard, T.S., Lawlor, T.J., Crooker, B.A., Van Tassell, C.P., Yang, J., Wang, S., Matukumalli, L.K. and Da, Y., 2011. Genome-wide association analysis of thirty one production, health, reproduction and body conformation traits in contemporary US Holstein cows. BMC genomics, 12(1), pp.1-17.

Colleau, JJ, Fritz, S., Guillaume, F., Baur, A., Dupassieux, D., Boscher, MY, Journaux, L., Eggen, A. and Boichard, D., 2015. Simulation of the potential of genomic selection in dairy cattle. Animal productions , 28 (3), pp. 251-258.

Corbet, N.J., Shepherd, R.K., Burrow, H.M., van der Westhuizen, J., Strydom, P.E. and Bosman, D.J. (2006) Evaluation of Bonsmara and Belmont Red Breeds in South Africa. 1. Productive performance. Aust. J. Exp. Agric, 46, 199-212.

Crum, T.E., Schnabel, R.D., Decker, J.E., Regitano, L.C. and Taylor, J.F., 2019. CRUMBLER: a tool for the prediction of ancestry in cattle. PloS one, 14(8).

Cundiff, L.V. and Gregory, K.E., 1999. What Is Systematic Crossbreeding? BEEF Cow/Calf Management 1999. Intertec Publ. Corp. Overland Park, KS, pp.8-16.

Cundiff, L.V., 1977. Foundations for animal breeding research. Journal of Animal Science, 44(2), pp.311-319.

Dadi, H., Jordaan, G.F., Schoeman, S.J. and Van der Westhuizen, J. (2002) The Effect of Charolais and Hereford Sires and Straightbred and Crossbred Dams on Pre-weaning Growth of Calves. S. Afr. J. Anim. Sci, 32, 38-43.

De Beukelaer, H., Badke, Y., Fack, V. and De Meyer, G., 2017. Moving beyond managing realized genomic relationship in long-term genomic selection. Genetics, 206(2), pp.1127-1138.

de Boor, C., 2001. A practical guide to splines. Springer series in applied mathematics, vol. 27, second ed. Springer, New York.

Delgadillo, J.S., Riley, D.G., Langdon, J.M., Hanna, L.H., Herring, A.D., 2017. 178 Black Hereford genetic parameters and predictions for calf traits with alternate modelling strategies. J. Anim. Sci. 95(suppl. 4), 88.

Delgado, C., Rosegrant, M., Steinfeld, H., Ehui, S. and Courbois, C. (1999) Livestock to 2020: The Next Food Revolution. Washington, DC, International Food Policy Research Institute, FAO, and International Livestock Research Institute.

Dickerson, G. E. 1973. Inbreeding and heterosis in animals. Proc. Anim. Breed. Genet. Symp. Honor of Dr. Jay L. Lush, Amer. Soc. Anim. Sci., Champaign, IL, pp. 54-77

Doran, A.G., Berry, D.P. and Creevey, C.J., 2014. Whole genome association study identifies regions of the bovine genome and biological pathways involved in carcass trait performance in Holstein-Friesian cattle. BMC genomics, 15(1), pp.1-12.

Ehui, S.H., Li-Pun, H.H., Mares, V. and Shapiro, B. (1998) The Role of Livestock in Food Security and Environmental Protection. Outlook on Agriculture, 27, 81-87.

Elsik, C.G., Tellam, R.L. and Worley, K.C., 2009. The genome sequence of taurine cattle: a window to ruminant biology and evolution. Science, 324(5926), pp.522-528.

Farquharson, R.J., Griffith, G.R., Barwick, S.A., Banks, R.G., Holmes, W.E., 2003. Estimating the returns from past investment into beef cattle genetic technologies in Australia. Econ. Res. Rep. 15, 3.

Ferenčaković, M., Hamzić, E., Gredler, B., Solberg, T.R., Klemetsdal, G., Curik, I. and Sölkner, J., 2013. Estimates of autozygosity derived from runs of homozygosity: empirical evidence from selected cattle populations. Journal of Animal Breeding and Genetics, 130(4), pp.286-293.

Ferrell, C.L., 1993. Factors influencing fetal growth and birth weight in cattle. Roman L. Hruska U.S. Meat Animal Res. Center 132. https://digitalcommons.unl.edu/hruskareports/132

Fleming, A. 2013. A genetic characterization of the Hays Converter. M.Sc. thesis. University of Guelph, ON, Canada. 129 pp.

Fleming, A.L., Hays, D.P., Moore, S., Plastow, G., and Miller, S.P. 2016. History of the Hays Converter. Can. J. Anim. Sci. 96: 471-477.

Frankham, R., Ballou, S.E.J.D., Briscoe, D.A. and Ballou, J.D., 2002. Introduction to conservation genetics. Cambridge university press.

Frisch, J.E. and O'neill, C.J., 1998a. Comparative evaluation of beef cattle breeds of African, European and Indian origins. 1. Live weights and heterosis at birth, weaning and 18 months. Animal Science, 67(1), pp.27-38.

Frisch, J.E. and O'Neill, C.J., 1998b. Comparative evaluation of beef cattle breeds of African, European and Indian origins. 2. Resistance to cattle ticks and gastrointestinal nematodes. Animal Science, 67(1), pp.39-48.

Frkonja, A., Gredler, B., Schnyder, U., Curik, I. and Soelkner, J., 2012. Prediction of breed composition in an admixed cattle population. Animal genetics, 43(6), pp.696-703.

Funkhouser, S.A., Bates, R.O., Ernst, C.W., Newcom, D. and Steibel, J.P., 2017. Estimation of genome-wide and locus-specific breed composition in pigs. Translational Animal Science, 1(1), pp.36-44.

Garcia, M.D., Chapman, C.K. and Despain, D., 2019. Applying Principles of Crossbreeding to Maximize Hybrid Vigor. https://digitalcommons.usu.edu/cgi/viewcontent.cgi?article=2994&context=extension_curall (accessed, October, 4, 2019).

Garner, D.L. and Seidel, G.E. (2008) History of Commercializing Sexed Semen for Cattle. Theriogenology, 69, 886-895.

Gengler, N., Misztal, I., Bertrand, J.K. and Culbertson, M.S., 1998. Estimation of the dominance variance for postweaning gain in the US Limousin population. Journal of animal science, 76(10), pp.2515-2520.

Gobena, M., Elzo, M.A. and Mateescu, R.G., 2018. Population structure and genomic breed composition in an angus–brahman crossbred cattle population. Frontiers in Genetics, 9, p.90.

Goddard M. E., Meuwissen T. H. E. 2005. The use of linkage disequilibrium to map quantitative trait loci. Australian Journal of Experimental Agriculture 45, 837-845.

Golden, B.L., Garrick, D.J. and Benyshek, L.L., 2009. Milestones in beef cattle genetic evaluation. Journal of animal science, 87(suppl_14), pp.E3-E10.

González-Recio, O., De Maturana, E.L. and Gutiérrez, J.P., 2007. Inbreeding depression on female fertility and calving ease in Spanish dairy cattle. Journal of dairy science, 90(12), pp.5744-5752.

Gosey, J. A. 1991. Crossbreeding systems and the theory behind composite breeds. In: Proceedings of the Range Beef Cow Symposium XII, Ft. Collins, CO. p 33-35.

Goszczynski, D.E., Corbi-Botto, C.M., Durand, H.M., Rogberg-Muñoz, A., Munilla, S., Peral-Garcia, P., Cantet, R.J.C. and Giovambattista, G., 2018. Evidence of positive selection towards Zebuine haplotypes in the BoLA region of Brangus cattle. animal, 12(2), pp.215-223.

Gouveia, J.J.D.S., Silva, M.V.G.B.D., Paiva, S.R. and Oliveira, S.M.P.D., 2014. Identification of selection signatures in livestock species. Genetics and molecular biology, 37(2), pp.330-342.

Gregory, K. E., Cundiff, L. V., & Koch, R. M. (1999). Composite breeds to use heterosis and breed differences to improve efficiency of beef production. U.S. Department of Agriculture Technical Bulletin No. 1875, 81 pp.

Gregory, K.E. and Cundiff, L., 1980. Crossbreeding in beef cattle: evaluation of systems. Journal of animal science, 51(5), pp.1224-1242.

Gregory, K.E., Swiger, L.A., Koch, R.M., Sumption, L.J., Rowden, W.W. and Ingalls, J.E., 1965. Heterosis in preweaning traits of beef cattle. Journal of Animal Science, 24(1), pp.21-28.

Gregory, K.E., Trail, J.C.M., Koch, R.M. and Cundiff, L.V., 1982. Heterosis, crossbreeding and composite breed utilisation in the tropics. In Proceedings of the 2nd World Congress on Genetics Applied to Livestock Production (p. 279).

Greiner, S.P. (2002) Crossbreeding in Beef Cattle. Virginia Cooperative Extension. Virginia Polytechnic Institute and State University. Publication 400–805.

Grisart, B., Coppieters, W., Farnir, F., Karim, L., Ford, C., Berzi, P., Cambisano, N., Mni, M., Reid, S., Simon, P. and Spelman, R., 2002. Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. Genome research, 12(2), pp.222-231.

Grosz, M.D. and MacNeil, M.D., 1999. The" spotted" locus maps to bovine chromosome 6 in a Hereford-Cross population. The Journal of heredity, 90(1), pp.233-236.

Grubbs, F.E., 1950. Sample criteria for testing outlying observations. The Annals of Mathematical Statistics, 21(1), pp.27-58.

Gutierrez-Gil B., Wiener, P. and Williams, J.L. (2007) Genetic Effects on Coat Colour in Cattle: Dilution of Eumelanin and Phaeomelanin Pigments in an F2-Backcross Charolais x Holstein Population. BMC Genetics 8(1), 56.

Hall, J.B. and Glaze Jr, J.B., 2014. System application of sexed semen in beef cattle. The Professional Animal Scientist, 30(3), pp.279-284.

Hall, J.B. and Glaze, J.B. Jr. (2013) How Can Sexed Semen Be Used in Commercial Beef Herds? Range Beef Cow Symposium. Paper 316. http://digitalcommons.unl.edu/rangebeefcowsymp/316

Hayes, B.J., Bowman, P.J., Chamberlain, A.J. and Goddard, M.E., 2009. Invited review: Genomic selection in dairy cattle: Progress and challenges. Journal of dairy science, 92(2), pp.433-443.

Hazell, P., Poulton, C., Wiggins, S. and Dorward, A. (2007) The Future of Small Farms for Poverty Reduction and Growth. IFPRI Discussion Paper 42. http://www.ifpri.org/2020/dp/vp42.asp

He, J., Guo, Y., Xu, J., Li, H., Fuller, A., Tait, R.G., Wu, X.L. and Bauck, S., 2018. Comparing SNP panels and statistical methods for estimating genomic breed composition of individual animals in ten cattle breeds. BMC genetics, 19(1), p.56.

Henderson, C.R., 1973. Sire evaluation and genetic trends. Journal of Animal Science, 1973(Symposium), pp.10-41.

Henderson, C.R., 1975. Best linear unbiased estimation and prediction under a selection model. Biometrics, pp.423-447.

Hohenboken, W.D. (1999) Applications of Sexed Semen in Cattle Production. Theri-ogenology, 52, 1421-1433.

Hohenboken, W.D., 1999. Applications of sexed semen in cattle production. Theriogenology, 52(8), pp.1421-1433.

Hosokawa, D., Ishii, A., Yamaji, K., Sasazaki, S., Oyama, K. and Mannen, H., 2012. Identification of divergently selected regions between Japanese Black and Holstein cattle using bovine 50k SNP array. Animal science journal, 83(1), pp.7-13.

Hu, Z.L., Park, C.A. and Reecy, J.M., 2019. Building a livestock genetic and genomic information knowledgebase through integrative developments of Animal QTLdb and CorrDB. Nucleic acids research, 47(D1), pp.D701-D710.

Iwaisaki, H., Tsuruta, S., Misztal, I., Bertrand, J.K., 2005. Genetic parameters estimated with multitrait and linear spline-random regression models using Gelbvieh early growth data. J. Anim. Sci. 83, 757-763.

Jamrozik, J., Schaeffer, L.R., 2002. Bayesian comparison of random regression models for test-day yields in dairy cattle. Proc. 7th World Cong. Genet. Appl. Livestk. Prod. Montpellier, France. Communication (No. 01-03).

Keller, L.F. and Waller, D.M., 2002. Inbreeding effects in wild populations. Trends in ecology & evolution, 17(5), pp.230-241.

Keller, M.C., Visscher, P.M. and Goddard, M.E., 2011. Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. Genetics, 189(1), pp.237-249.

Kemper, K.E., Reich, C.M., Bowman, P.J., Vander Jagt, C.J., Chamberlain, A.J., Mason, B.A., Hayes, B.J. and Goddard, M.E., 2015. Improved precision of QTL mapping using a nonlinear Bayesian method in a multi-breed population leads to greater accuracy of across-breed genomic predictions. Genetics Selection Evolution, 47(1), p.29.

Kim, E.S., Cole, J.B., Huson, H., Wiggans, G.R., Van Tassell, C.P., Crooker, B.A., Liu, G., Da, Y. and Sonstegard, T.S., 2013. Effect of artificial selection on runs of homozygosity in US Holstein cattle. PloS one, 8(11).

Kinghorn, B., 1982. Genetic effects in crossbreeding: II. Multibreed selection indices. Zeitschrift für Tierzüchtung und Züchtungsbiologie, 99(1-4), pp.315-320.

Kinghorn, B.P., Shepherd, R.K. and Banks, R.G., 1989. A note on the formation of optimal composite populations. Theoretical and applied genetics, 78(3), pp.318-320.

Kirin, M., McQuillan, R., Franklin, C.S., Campbell, H., McKeigue, P.M. and Wilson, J.F., 2010. Genomic runs of homozygosity record population history and consanguinity. PloS one, 5(11), p.e13996.

Koger, M., 1980. Efective Crossbreeding Systems Utilizing Zebu Cattle. Journal of Animal Science, 50(6), pp.1215-1220.

Koger, M., Peacock, F.M., Kirk, W.G. and Crockett, J.R., 1975. Heterosis effects on weaning performance of Brahman-Shorthorn calves. Journal of Animal Science, 40(5), pp.826-833.

Kuehn, L.A., Keele, J.W., Bennett, G.L., McDaneld, T.G., Smith, T.P.L., Snelling, W.M., Sonstegard, T.S. and Thallman, R.M., 2011. Predicting breed composition using breed frequencies of 50,000 markers from the US Meat Animal Research Center 2,000 Bull Project. Journal of animal science, 89(6), pp.1742-1750.

Lardner, B. and Damiran, D. (2015) Comparison of Fixed-time Artificial Insemination vs. Natural Service in Beef Cows: Reproductive Efficiency and System Cost. http://www.wbdc.sk.ca/pdfs/fact_sheets/2015/2015-02_WBDC_Fixed-timeAI_vs_NaturalBreeding.pdf

Larmer, S., Ventura, R., Buzanskas, M.E., Sargolzaei, M. and Schenkel, F.S., 2014. Assessing admixture by quantifying breed composition to gain historical perspective on dairy cattle in Canada. In Vancouver, Canada: 10th World Congress on Genetics Applied to Livestock Production.

Lawson, J.E., 1981. Milk production and weight of cows of the Hereford and Highland breeds and their reciprocal crosses. Canadian Journal of Animal Science, 61(4), pp.861-869.

Lawson, J.E., 1982. Performance of dairy crosses with the HEREFORD in a range environment. Canadian Journal of Animal Science, 62(3), pp.717-724.

Legarra, A., Misztal, I., Bertrand, J.K., 2004. Constructing covariance functions for random regression models for growth in Gelbvieh beef cattle. J. Anim. Sci. 82, 1564-1571.

Lewontin, R.C. and Krakauer, J., 1973. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. Genetics, 74(1), pp.175-195.

Lin, C.Y. (1980) Relative Efficiency of Selection Methods for Improvement of Feed Efficiency. J. Dairy Sci, 63, 491-494.

Long, C. R., Cartwright, T.C. and Fitzhugh, H.A. (1975) Systems Analysis of Sources of Genetic and Environmental Variation in Efficiency of Beef Production: Cow Size and Herd Management. J. Anim. Sci, 40, 409-420.

Lund, M.S., de Roos, A.P., de Vries, A.G., Druet, T., Ducrocq, V., Fritz, S., Guillaume, F., Guldbrandtsen, B., Liu, Z., Reents, R. and Schrooten, C., 2011. A common reference population from four European Holstein populations increases reliability of genomic predictions. Genetics Selection Evolution, 43(1), p.43.

MacNeil, M.D. (2005) Beef Cattle Management: Crossbreeding. In: Encyclopedia of Animal Science. Marcel Dekker, New York, NY.

MacNeil, M.D. and Matjuda, E. (2007) Breeding Objectives for Angus and Charolais Specialized Sire Lines for Use in the Emerging Sector of South African Beef Production. S. Afr. J. Anim. Sci, 37, 1-10.

MacNeil, M.D., 1987. Formation of optimal composite populations. Theoretical and applied genetics, 74(6), pp.837-840.

MacNeil, M.D., 2009. Invited review: Research contributions from seventy-five years of breeding Line 1 Hereford cattle at Miles City, Montana. Journal of animal science, 87(8), pp.2489-2501.

MacNeil, M.D., Bailey, D.R., Urick, J.J., Gilbert, R.P. and Reynolds, W.L. (1991) Heritabilities and Genetic Correlations for Postweaning Growth and Feed Intake of Beef Bulls and Steers. J. Anim. Sci, 69, 3183–3189.

MacNeil, M.D., Cundiff, L.V., Dinkel, C.A. and Koch, R.M. (1984) Genetic Correlations among Sex-limited Traits in Beef Cattle. J. Anim. Sci, 58, 1171-1180.

MacNeil, M.D., Cundiff, L.V., Gregory, K.E. and Koch, R.M. (1988) Crossbreeding Systems for Beef Production. Applied Agric. Res, 3, 44-54.

MacNeil, M.D., Lopez-Villalobos, N. and Northcutt S.L. (2011) A Prototype National Cattle Evaluation for Feed Intake and Efficiency of Angus Cattle. J. Anim. Sci, 89, 3917-3923.

MacNeil, M.D., Mott, T.B., 2006. Genetic analysis of gain from birth to weaning, milk production, and udder conformation in Line 1 Hereford cattle. J. Anim. Sci. 84, 1639-1645.

MacNeil, M.D., Urick, J.J. and Decoudu, G., 2000. Characteristics of Line 1 Hereford females resulting from selection by independent culling levels for below-average birth weight and

high yearling weight or by mass selection for high yearling weight. Journal of animal science, 78(9), pp.2292-2298.

Maqhashu, A., Mphaphathi, M.L., Seshoka, M.M., Ramukhithi, F.V., Seolwana, F.L., Masenya, M.B., Netshirovha, R.T., Mapeka, M.H., Nengovhela, N.H., Kanuya, N.L., Muchenje, V. and Nedambale, T.L. (2016) Comparison of Oestrous Synchronization Re-sponse and Pregnancy Rate of Village Cows Following Timed Artificial Insemination in KwaZulu-Natal and Limpopo Provinces. Open J. Anim. Sci, 6, 9-15.

Marras, G., Gaspa, G., Sorbolini, S., Dimauro, C., Ajmone-Marsan, P., Valentini, A., Williams, J.L. and Macciotta, N.P., 2015. Analysis of runs of homozygosity and their relationship with inbreeding in five cattle breeds farmed in Italy. Animal genetics, 46(2), pp.110-121.

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

McLeod, A. (2011) World Livestock 2011-Livestock in Food Security. Food and Agri-culture Organization of the United Nations (FAO).

McQuillan, R., Leutenegger, A.L., Abdel-Rahman, R., Franklin, C.S., Pericic, M., Barac-Lauc, L., Smolej-Narancic, N., Janicijevic, B., Polasek, O., Tenesa, A. and MacLeod, A.K., 2008. Runs of homozygosity in European populations. The American Journal of Human Genetics, 83(3), pp.359-372.

McTavish, E.J. and Hillis, D.M., 2014. A genomic approach for distinguishing between recent and ancient admixture as applied to cattle. Journal of Heredity, 105(4), pp.445-456.

Meredith, B.K., Kearney, F.J., Finlay, E.K., Bradley, D.G., Fahey, A.G., Berry, D.P. and Lynn, D.J., 2012. Genome-wide associations for milk production and somatic cell score in Holstein-Friesian cattle in Ireland. BMC genetics, 13(1), pp.1-11.

Mészáros, G., Petautschnig, E., Schwarzenbacher, H. and Sölkner, J., 2015. Genomic regions influencing coat color saturation and facial markings in Fleckvieh cattle. Animal genetics, 46(1), pp.65-68.

Meuwissen, T., Hayes, B. and Goddard, M., 2016. Genomic selection: A paradigm shift in animal breeding. Animal frontiers, 6(1), pp.6-14.

Meuwissen, T.H., Hayes, B.J. and Goddard, M.E., 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics, 157(4), pp.1819-1829.

Meyer, K., 1999. Estimates of genetic and phenotypic covariance functions for postweaning growth and mature weight of beef cows. J. Anim. Breed. Genet. 116, 181-205.

Meyer, K., 2001. Estimates of direct and maternal covariance functions for growth of Australian beef calves from birth to weaning. Genet. Sel. Evol. 33, 487-514.

Meyer, K., 2004. Scope for a random regression model in genetic evaluation of beef cattle for growth. Livestk. Prod. Sci. 86, 69-83.

Meyer, K., 2005a. Estimates of genetic covariance functions for growth of Angus cattle. J. Anim. Breed. Genet. 122, 73-85.

Meyer, K., 2005b. Advances in methodology for random regression analyses. Aust. J. Exp. Agric. 45, 847-858.

Meyer, K., 2005c. Random regression analyses using B-splines to model growth of Australian Angus cattle. Genet. Sel. Evol. 37, 473-500.

Meyer, K., 2007. WOMBAT—A tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). J. Zhejiang Uni. Science B, 8, 815-821.

Meyer, K., 2018. wombat:sterrors. The Wicked WOMBAT Wiki. http://didgeridoo.une.edu.au/womwiki/doku.php?id=wombat:sterrors&rev=1536881200 (accessed 12 January 2019).

Meyer, K., Kirkpatrick, M., 2005. Up hill, down dale: quantitative genetics of curvaceous traits. Phil. Trans. R. Soc. B. 360, 1443-1455.

Meyermans, R., Gorssen, W., Buys, N. and Janssens, S., 2020. How to study runs of homozygosity using PLINK? A guide for analyzing medium density SNP data in livestock and pet species. BMC genomics, 21(1), pp.1-14.

Miller, S., 2010. Genetic improvement of beef cattle through opportunities in genomics. Revista Brasileira de Zootecnia, 39, pp.247-255.

Mofakkarul Islam, M., Renwick, A., Lamprinopoulou, C., Klerkx, L., 2013. Innovation in livestock genetic improvement. EuroChoices. 12, 42-47.

Mota, R.R., Marques, L., Lopes, P.S., Da Silva, L.P., Neto, F., De Resende, M., Torres, R.A., 2013. Genetic evaluation using multi-trait and random regression models in Simmental beef cattle. Genet. Mol. Res.12, 2465-2480.

Moyo, S. (1990) Evaluation of the Productivity of Indigenous Cattle and Some Exotic Beef Breeds and their Crosses in Zimbabwe. ILCA Research Fellowship Report, ILCA, Addis Ababa, Ethiopia. 139 p.

Nei, M., 1977. F-statistics and analysis of gene diversity in subdivided populations. Annals of human genetics, 41(2), pp.225-233.

Nielsen, M.K., MacNeil, M.D., Dekkers, J.C.M., Crews Jr., D.H., Rathje, T.A., Enns, R.M., Weaber, R.L., 2013. REVIEW: Life-cycle, total industry genetic improvement of feed efficiency in beef cattle: Blueprint for the Beef Improvement Federation. Professional Animal Scientist 29, 559-565.

Nobre, P.R.C., Misztal, I., Tsuruta, S., Bertrand, J.K., Silva, L.O.C., Lopes, P.S., 2003. Analyses of growth curves of Nellore cattle by multiple-trait and random regression models. J. Anim. Sci. 81, 918-926.

Olfert, E.D., Cross, B.M., McWilliam, A.A., Eds., 1993. Guide to the Care and Use of Experimental Animals, 2nd ed. Canadian Council on Animal Care, Ottawa.

Oliveira, M.R., Azevêdo, D.M., Malhado, C., Pires, L., Martins Filho, R., Sousa Júnior, S., 2017. Weight evaluation of Tabapuã cattle raised in northeastern Brazil using random-regression models. Arq. Bras. Med. Vet. Zootec. 69, 457-464.

Olori, V.E., Hill, W.G., McGuirk, B.J., Brotherstone, S., 1999. Estimating variance components for test day milk records by restricted maximum likelihood with a random regression animal model. Livestock Prod. Sci. 61, 53-63.

Paim, T.D.P., Hay, E.H.A., Wilson, C., Thomas, M.G., Kuehn, L.A., Paiva, S.R., McManus, C. and Blackburn, H.D., 2020. Dynamics of genomic architecture during composite breed development in cattle. Animal genetics, 51(2), pp.224-234.

Patterson, N., Price, A.L. and Reich, D., 2006. Population structure and eigenanalysis. PLoS genetics, 2(12).

Pausch, H., Wang, X., Jung, S., Krogmeier, D., Edel, C., Emmerling, R., Götz, K.U. and Fries, R., 2012. Identification of QTL for UV-protective eye area pigmentation in cattle by progeny phenotyping and genome-wide association analysis. PloS one, 7(5), p.e36346.

Pemberton, T.J., Absher, D., Feldman, M.W., Myers, R.M., Rosenberg, N.A. and Li, J.Z., 2012. Genomic patterns of homozygosity in worldwide human populations. The American Journal of Human Genetics, 91(2), pp.275-292.

Peripolli, E., Metzger, J., De Lemos, M.V.A., Stafuzza, N.B., Kluska, S., Olivieri, B.F., Feitosa, F.L.B., Berton, M.P., Lopes, F.B., Munari, D.P. and Lôbo, R.B., 2018. Autozygosity islands and ROH patterns in Nellore lineages: Evidence of selection for functionally important traits. BMC genomics, 19(1), pp.1-14.

Peripolli, E., Stafuzza, N.B., Amorim, S.T., de Lemos, M.V.A., Grigoletto, L., Kluska, S., Ferraz, J.B.S., Eler, J.P., Mattos, E.C. and Baldi, F., 2020. Genome-wide scan for runs of homozygosity in the composite Montana Tropical® beef cattle. Journal of Animal Breeding and Genetics, 137(2), pp.155-165.

Pickrell, J. and Pritchard, J., 2012. Inference of population splits and mixtures from genome-wide allele frequency data. Nature Precedings, pp.1-1.

Pienaar, L. (2014) Genetic Diversity in the Afrikaner Cattle Breed. MSc. Thesis, Univ. Free State, Bloemfontain.

Prayaga, K.C., 2003a. Evaluation of beef cattle genotypes and estimation of direct and maternal genetic effects in a tropical environment. 1. Growth traits. Australian Journal of Agricultural Research, 54(10), pp.1013-1025.

Prayaga, K.C., 2003b. Evaluation of beef cattle genotypes and estimation of direct and maternal genetic effects in a tropical environment. 2. Adaptive and temperament traits. Australian Journal of Agricultural Research, 54(10), pp.1027-1038.

Pritchard, J.K., Stephens, M. and Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics, 155(2), pp.945-959.

Pryce, J. and Hayes, B., 2012. A review of how dairy farmers can use and profit from genomic technologies. Animal Production Science, 52(3), pp.180-184.

Pryce, J.E., Haile-Mariam, M., Goddard, M.E. and Hayes, B.J., 2014. Identification of genomic regions associated with inbreeding depression in Holstein and Jersey dairy cattle. Genetics Selection Evolution, 46(1), pp.1-14.

Pryce, J.E., Hayes, B.J. and Goddard, M.E., 2012. Novel strategies to minimize progeny inbreeding while maximizing genetic gain using genomic information. Journal of dairy science, 95(1), pp.377-388.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., De Bakker, P.I., Daly, M.J. and Sham, P.C., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. The American journal of human genetics, 81(3), pp.559-575.

Purfield, D.C., Berry, D.P., McParland, S. and Bradley, D.G., 2012. Runs of homozygosity and population history in cattle. BMC genetics, 13(1), p.70.

Qanbari, S. and Simianer, H., 2014. Mapping signatures of positive selection in the genome of livestock. Livestock Science, 166, pp.133-143.

Qanbari, S., Gianola, D., Hayes, B., Schenkel, F., Miller, S., Moore, S., Thaller, G. and Simianer, H., 2011. Application of site and haplotype-frequency based approaches for detecting selection signatures in cattle. BMC genomics, 12(1), pp.1-12.

Ramirez-Valverde, R., Nunez-Dominguez, R., Ruiz-Flores, A., García-Muñiz, J.G., Magana-Valencia, F., 2008. Comparison of contemporary group definitions for genetic evaluation of Braunvieh cattle. Téc. Pecu. Méx. 46, 359-370.

Rincon, E.J., Schermerhorn, E.C., Mc Dowell, R.E. and Mc Daniel, B.T., 1982. Estimation of genetic effects on milk yield and constituent traits in crossbred dairy cattle. Journal of Dairy Science, 65(5), pp.848-856.

RStudio Team (2016). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA. Retrieved from http://www.rstudio.com/.

Rudder, T.H., Bean, K.G., Lapworth, J.W. and Seifert, G.W. (1976) Commercial Evaluation of a Tropical Breed Developed Experimentally. Proc. Aust. Soc. Anim. Prod, 11, 89-92.

Saatchi, M., Schnabel, R.D., Taylor, J.F. and Garrick, D.J., 2014. Large-effect pleiotropic or closely linked QTL segregate within and across ten US cattle breeds. BMC genomics, 15(1), pp.1-17.

Sabeti, P.C., Reich, D.E., Higgins, J.M., Levine, H.Z., Richter, D.J., Schaffner, S.F., Gabriel, S.B., Platko, J.V., Patterson, N.J., McDonald, G.J. and Ackerman, H.C., 2002. Detecting recent positive selection in the human genome from haplotype structure. Nature, 419(6909), pp.832-837.

Sargolzaei, M., Iwaisaki, H. and Colleau, J.J., 2006. CFC: a tool for monitoring genetic diversity. Proc. 8th World Congr. Genet. Appl. Livest. Prod., CD-ROM Communication, (27-28), pp.13-18.

Schaeffer, L.R., 2004. Application of random regression models in animal breeding. Livestock Prod. Sci. 86, 35-45.

Schaeffer, L.R., 2009. Contemporary Groups Are Always Random. http://animalbiosciences.uoguelph.ca/~lrs/piksLRS/ranfix.pdf (accessed 25 June 2019).

Schenkel, F.S., Miller, S.P. and Wilton, J.W., 2004. Genetic parameters and breed differences for feed efficiency, growth, and body composition traits of young beef bulls. Canadian Journal of Animal Science, 84(2), pp.177-185.

Scholtz, M.M. and Theunissen, A. (2010) The Use of Indigenous Cattle in Terminal Crossbreeding to Improve Beef Cattle Production in Sub-Saharan Africa. Anim. Genet. Res, 46, 33-39.

Scholtz, M.M., Roux, C.Z. and Lombard, P.E. (1990) Breeding Strategies for Beef Cattle in the Subtropics and Tropics: Terminal Crossbreeding. Proc. 4th World Congress on Genetics Applied to Livestock Production, Edinburgh, 23-27 July 1990, XV, 361-364.

Scholtz, M.M., Steyn, Y., Van Marle-Koester, E. and Theron, H.E. (2012) Improved Production Efficiency in Cattle to Reduce their Carbon Footprint for Beef Production. S. Afr. J. Anim. Sci, 42(suppl. 1), 450-453.

Seidel, G.E. (2003) Economics of Selecting for Sex: The Most Important Genetic Trait. Theriogenology, 59(2), 585-598.

Seifert, G.W. and Rudder, T.H. (1984) Experience with Africander and Africander Derived Genotypes in Australian Beef Herds. Proc. 2nd World Cong. Sheep Beef Cattle Breed, 671-677.

Shringarpure, S. and Xing, E.P., 2014. Effects of sample selection bias on the accuracy of population structure and ancestry inference. G3: Genes, Genomes, Genetics, 4(5), pp.901-911.

Skrypzeck, H., Schoeman, S.J., Jordaan, G.F. and Neser, F.W.C. (2000) Estimates of Crossbreeding Parameters in a Multibreed Beef Cattle Crossbreeding Project. S. Afr. J. Anim. Sci, 30(3), 193-203.

Smith, L.A., Cassell, B.G. and Pearson, R.E., 1998. The effects of inbreeding on the lifetime performance of dairy cattle. Journal of dairy science, 81(10), pp.2729-2737.

Snelling, W.M., Allan, M.F., Keele, J.W., Kuehn, L.A., McDaneld, T., Smith, T.P.L., Sonstegard, T.S., Thallman, R.M. and Bennett, G.L., 2010. Genome-wide association study of growth in crossbred beef cattle. Journal of animal science, 88(3), pp.837-848.

Sölkner, J., Frkonja, A., Raadsma, H.W., Jonas, E., Thaller, G., Gootwine, E., Seroussi, E., Fuerst, C., Egger-Danner, C. and Gredler, B., 2010. Estimation of individual levels of admixture in crossbred populations from SNP chip data: examples with sheep and cattle populations. Interbull Bulletin, (42), pp.62-62.

Stefansky, W., 1972. Rejecting outliers in factorial designs. Technometrics, 14(2), pp.469-479.

Sumreddee, P., Toghiani, S., Hay, E.H., Roberts, A., Agrrey, S.E. and Rekaya, R., 2019. Inbreeding depression in line 1 Hereford cattle population using pedigree and genomic information. Journal of animal science, 97(1), pp.1-18.

Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics, 123(3), pp.585-595.

Theunissen, A. (2012) Charaterization of Breed Additive and Heterosis Effects in Beef Cattle Using Experimental Results. MSc. thesis, Univ. Free State, Bloemfontein.

Theunissen, A., Scholtz, M.M., MacNeil, M.D. and Neser, F.W.C. (2014a) Breed Additive and Heterosis Effects on Feedlot and Carcass Traits in Beef Cattle. Proc. 10th World Congress on Genetics Applied to Livestock Production. Poster No. 714.

Theunissen, A., Scholtz, M.M., Neser, F.W.C. and MacNeil, M.D. (2013) Crossbreeding to Increase Beef Production: Additive and non-Additive Effects on Weight Traits. S. Afr. J. Anim. Sci, 43(2), 143-152.

Theunissen, A., Scholtz, M.M., Neser, F.W.C. and MacNeil, M.D. (2014b) Crossbreeding to Increase Beef Production: Additive and non-Additive Effects on Fitness Traits. S. Afr. J. Anim. Sci, 44(4), 335-349.

Theunissen, A., Scholtz, M.M., Neser, F.W.C. and MacNeil, M.D., 2013. Crossbreeding to increase beef production: additive and non-additive effects on weight traits. South African Journal of Animal Science, 43(2), pp.143-152.

Tier, B. and Meyer, K., 2004. Approximating prediction error covariances among additive genetic effects within animals in multiple-trait and random regression models. Journal of Animal Breeding and Genetics, 121(2), pp.77-89.

Toosi, A., Fernando, R.L. and Dekkers, J.C.M., 2010. Genomic selection in admixed and crossbred populations. Journal of Animal Science, 88(1), pp.32-46.

Turner, S., Armstrong, L.L., Bradford, Y., Carlson, C.S., Crawford, D.C., Crenshaw, A.T., De Andrade, M., Doheny, K.F., Haines, J.L., Hayes, G. and Jarvik, G., 2011. Quality control procedures for genome-wide association studies. Current protocols in human genetics, 68(1), pp.1-19.

Upperman, L.R., Kinghorn, B.P., MacNeil, M.D. and Van Eenennaam, A.L., 2019. Management of lethal recessive alleles in beef cattle through the use of mate selection software. Genetics Selection Evolution, 51(1), p.36.

Utsunomiya, Y.T., Milanesi, M., Utsunomiya, A.T.H., Torrecilha, R.B.P., Kim, E.S., Costa, M.S., Aguiar, T.S., Schroeder, S., Do Carmo, A.S., Carvalheiro, R. and Neves, H.H.R., 2017. A PLAG1 mutation contributed to stature recovery in modern cattle. Scientific reports, 7(1), pp.1-15.

Van Eenennaam, A.L., Weigel, K.A., Young, A.E., Cleveland, M.A. and Dekkers, J.C., 2014. Applied animal genomics: results from the field. Annu. Rev. Anim. Biosci., 2(1), pp.105-139.

Van Vleck, L.D., 1987. Contemporary groups for genetic evaluations. J. Dairy. Sci. 70, 2456-2464.

VanRaden, P.M. and Cooper, T.A., 2015. Genomic evaluations and breed composition for crossbred US dairy cattle. Interbull Bulletin, (49), pp.19-23.

VanRaden, P.M., Olson, K.M., Wiggans, G.R., Cole, J.B. and Tooker, M.E., 2011. Genomic inbreeding and relationships among Holsteins, Jerseys, and Brown Swiss. Journal of Dairy Science, 94(11), pp.5673-5682.

Veneroni, G.B., Meirelles, S.L., Grossi, D.A., Gasparin, G., Ibelli, A.M.G., Tizioto, P.C., Oliveira, H.N.D., Alencar, M.M. and Regitano, L.C.A., 2010. Prospecting candidate SNPs for backfat in Canchim beef cattle. Genetics and Molecular Research, pp.1997-2003.

Visscher, P.M., Goddard, M.E., 1993. Fixed and random contemporary groups. J. Dairy. Sci. 76, 1444-1454.

Voight, B.F., Kudaravalli, S., Wen, X. and Pritchard, J.K., 2006. A map of recent positive selection in the human genome. PLoS biology, 4(3).

Vostry, L., Milerski, M., Krupa, E., Vesela, Z., Vostra-Vydrova, H., 2015. Genetic relationships among calving ease, birth weight and perinatal calf survival in Charolais cattle. Animal Science Papers and Reports. 33, 233-242.

Weaber, Bob, 2008. Crossbreeding for Commercial Beef Production. Beef Sire Selection Manual. National Beef Cattle Evaluation Consortium.

http://www.nbcec.org/producers/sire_selection/chapter5.pdf. pp. 50-57. Accessed October 10, 2019.

Weikard, R., Widmann, P., Buitkamp, J., Emmerling, R. and Kuehn, C., 2012. Revisiting the quantitative trait loci for milk production traits on BTA6. Animal genetics, 43(3), pp.318-323.

Wiggans, G.R., Cole, J.B., Hubbard, S.M. and Sonstegard, T.S., 2017. Genomic selection in dairy cattle: the USDA experience. Annual review of animal biosciences, 5, pp.309-327.

Wiggans, G.R., Sonstegard, T.S., VanRaden, P.M., Matukumalli, L.K., Schnabel, R.D., Taylor, J.F., Schenkel, F.S. and Van Tassell, C.P., 2009. Selection of single-nucleotide polymorphisms and quality of genotypes used in genomic evaluation of dairy cattle in the United States and Canada. Journal of dairy science, 92(7), pp.3431-3436.

Wiggans, G.R., VanRaden, P.M. and Cooper, T.A., 2011. The genomic evaluation system in the United States: Past, present, future. Journal of dairy science, 94(6), pp.3202-3211.

Wiggans, G.R., VanRaden, P.M. and Zuurbier, J., 1995. Calculation and use of inbreeding coefficients for genetic evaluation of United States dairy cattle. Journal of dairy science, 78(7), pp.1584-1590.

Wright, S., 1922. Coefficients of inbreeding and relationship. The American Naturalist, 56(645), pp.330-338.

Wright, S., 1949. The genetical structure of populations. Annals of eugenics, 15(1), pp.323-354.

Wright, S., 1965. The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution, pp.395-420.

Zhao, F., McParland, S., Kearney, F., Du, L. and Berry, D.P., 2015. Detection of selection signatures in dairy and beef cattle using high-density genomic information. Genetics Selection Evolution, 47(1), p.49.

Zimin, A.V., Delcher, A.L., Florea, L., Kelley, D.R., Schatz, M.C., Puiu, D., Hanrahan, F., Pertea, G., Van Tassell, C.P., Sonstegard, T.S. and Marçais, G., 2009. A whole-genome assembly of the domestic cow, Bos taurus. Genome biology, 10(4), pp.1-10.

Zimmerman, W.E., 1980. Lactation and weaning weight relationships in Hereford and Simmental-Hereford cows in Southern Utah.

Zwambag, A.W., 2007. Genetic improvement of beef tenderness in a multi-breed population utilizing genetic markers. M.Sc. Thesis – University of Guelph, Guelph, Ontario.