DISSERTATION

GENETIC SELECTION FOR FEED INTAKE AND EFFICIENCY IN BEEF CATTLE

Submitted by

Miranda M. Culbertson

Department of Animal Sciences

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2019

Doctoral Committee:

Advisor: Scott Speidel
Co-Advisor: Mark Enns

Milt Thomas
Terry Engle
Marshall Frasier

ABSTRACT

GENETIC SELECTION FOR FEED INTAKE AND EFFICIENCY IN BEEF CATTLE

Feed costs are reported as the largest variable expense in beef production systems accounting for 50% to 70% of total production costs. Due to the large impact of feed costs on profitability, producers have become increasingly aware of the need to improve feed utilization in cattle. With the advent of technologies to measure individual feed intake in cattle, a phenotype is available for selection; however, how to implement this phenotype into a breeding objective or genetic evaluation is debatable. This dissertation examines some of these unanswered questions about feed intake and efficiency and its utility in for genetic selection. The objectives herein were 1) to evaluate the maternal genetic effects on dry matter feed intake (DMI), 2) to simulate data to examine the effects single trait selection on DMI or residual feed intake (RFI) on genetically correlated traits of weaning weight (WWT) and yearling weight (YWT), 3) to examine data generated by an ear tag accelerometer (CowManger; Agis Automatisering BV, Harmelen, Netherlands) while attached to steers located in a feedlot and on pastures to develop a proxy for measures of intake and 4) to examine the phenotypic relationship between grazing and feedlot intake.

For the examination of maternal genetic effects on intake, the American Gelbvieh Association (AGA) and the Red Angus Association of America (RAAA) provided pedigree information in addition to DMI and WWT records. Dry matter intake records were limited to animals within an age range of 240 d to 365 d to limit data to only postweaning cattle. Embryo transfer calves were removed. Gelbvieh and Red Angus data were analyzed separately.

The first analysis was a single trait model that examined the maternal genetic effects of DMI. Contemporary groups (CG) were formed using sex, pen, feed trial designation, trial length and year for both AGA and RAAA. The final data set for AGA consisted of 3,021 animals with DMI records and a 3-generation pedigree of 15,418 animals. For Red Angus, cattle with DMI records was 3,213 and a 3-generation pedigree of 13,747 animals. The heritabilities of DMI direct for Gelbvieh and Red Angus were moderate to high (0.45 ± 0.06 and 0.24 ± 0.06, respectively) but the DMI maternal heritability for Gelbvieh was 0.00 ± 0.00 and Red Angus DMI maternal heritability was very low at 0.05 ± 0.04. Resulting in little to no maternal effect for DMI.

The second analysis was a multi-trait model estimating the correlation between DMI and WWT maternal. For the multi-trait model, fixed effects for weaning weight included age, age of dam and CG. The contemporary group for WWT was sex, breeder, weaning date, and herd. For Gelbvieh, the heritabilities for DMI, WWT direct and WWT maternal were 0.45 ± 0.05, 0.36 ± 0.06 and 0.15 ± 0.05, respectively. The heritabilities for Red Angus were 0.27 ± 0.05, 0.21 ± 0.06 and 0.16 ± .07 for DMI, WWT direct and WWT maternal, respectively. The genetic correlation between DMI and WWT maternal was low at 0.12 ± 0.13 and 0.12 ± 0.24, for Gelbvieh and Red Angus, respectively. These results suggested that WWT maternal would has minimal impact on the estimation of DMI EPD using a multivariate model.

For the second objective of this dissertation, data were simulated to examine the effects on genetically correlated traits of WWT and YWT when single trait selection was conducted on DMI or RFI. Genetic parameters were established using published variance estimates weighted for the number of animals included in each estimation. Based on the weighted genetic parameters, a simulated population was established for three selection scenarios. The first and

second scenario for selection was single trait selection for DMI and RFI, respectively. The third scenario used an economic selection index as criteria for replacements. With an annual replacement rate of 20% for females and 5% for males, 10 years of offspring data was generated. Replacements were chosen based on their breeding value for the trait of interest. At the conclusion of 10 years of simulated data, the scenarios for the selection of DMI and RFI saw a decrease in DMI of 0.85 kg/year and 0.19 kg/year, respectively. Both scenarios also resulted in a decrease for YWT of 27.83 kg/year from selection of DMI and 5.13 kg/year from the selection of RFI. Selection using the economic index showed a steady increase in YWT (14.84 kg/year) but also demonstrated an increase of DMI (0.42 kg/year). The three scenarios were all examined by the amount of profitability determined from fed cattle and feed prices. Of the three scenarios, the economic index showed the greatest amount of profit due to the increase in YWT. Although DMI increased with the index, the amount of increase in yearling weights was significant enough to outweigh the increase in feed costs.

The third study of this dissertation examined a remote sensor technology as a potential proxy for DMI in addition to estimating a correlation between DMI measured in a feedlot versus grazing intake for cattle on pasture. Ninety-three steers were equipped with a CowManager ear tag accelerometer (CME) that measured the amount of time an animal spent ruminating, eating, and levels of activity. The steers were placed in a feedlot where their intakes were measured using the GrowSafe Feed Intake monitoring system. The data collected via CME and GrowSafe were analyzed to identify existing associations between the measurements. Based on the DMI measured by GrowSafe, the 15 highest and 15 lowest intake animals were identified. These low/high intake animals (LHI) were used to quantify grazing intake using the biomarker titanium dioxide (TiO_2). No association between CME and DMI measured in the feedlot were found.

Pearson's correlations for CME measurements and DMI were low and ranged from -0.11 to 0.12. A regression analysis found no significant CME variable as explanatory variables for DMI.

After a 54-d performance test, the steers were immediately transported to pasture where the steers were maintained for 43 days. Data from CME were continuously collected while cattle grazed on pasture. For the final 20 d, the LHI cattle were administered a bolus of 10 g of TiO_2 each morning. During the 6 final days of this study, rectal fecal samples were collected twice daily with collection occurring 12 h apart. Every 24 h, time of collection was advanced 2 hours to minimize effects of diurnal variation. The fecal samples were analyzed for TiO_2 concentration and based on these concentrations grazing dry matter intakes (GDMI) were estimated. The Pearson's correlation between GDMI and DMI measured in the feedlot was 0.84 ± 0.10 ($P < 0.05$) with a Spearman rank correlation of 0.99 ± 0.03 ($P < 0.05$). This result suggested a strong relationship between grazing and feedlot DMI; however, it is less than 1 indicating some change in rank between DMI and GDMI. The correlation between GDMI and CME ranged from -0.22 to 0.19 with the largest correlation (-0.22) was between GDMI and time spent eating. All of these correlation estimates were not significantly different from zero ($P > 0.05$). This study was able to show the application of remote sensor technology for monitoring cattle maintained on rangeland, but the precision of measurements from CME failed to provide an indicator for GDMI. A strong relationship between feedlot measured DMI and DMI for grazing cattle was established in this study.

ACKNOWLEDGMENTS

I would like to express my heartfelt gratitude to my advisors, Drs. Mark Enns and Scott Speidel, as well as committee member and mentor, Dr. Milt Thomas. It has been a long journey with unexpected setbacks to obtaining this degree and their guidance, assistance and encouragement throughout my program has been invaluable. I would also like to thank Drs. Terry Engle and Marshall Frasier for serving on my committee and providing valuable feedback to improve the quality of this research.

I owe a huge debt of gratitude to my family for their endless support. This would not have been possible without it. In particular, I have to thank my mother for her unwavering support both emotionally and financially throughout this endeavor.

A special thank you to my fellow graduate students. Not only was their friendship a crucial element to the graduate school experience but there was always a wealth of volunteers to help with projects when needed. Without their help, completing some these projects would not have been possible.

Finally, I owe an enormous thank you to my cousin, Dr. Avery L. Culbertson. She convinced me it was never too late to go back to school and pursue a graduate degree. I owe this Ph.D. to her for her encouragement and endless words of support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGUES

Figure Page

LIST OF EQUATIONS

Equation Page

CHAPTER I

INTRODUCTION AND OBJECTIVES

1.1 Introduction

Sustainability of any agricultural production system is dependent on the profitability of the production system. Traditionally, the beef industry has placed emphasis on improving output traits, such as growth and fertility, to improve profitability (Arthur et al., 2001; Hill, 2012) but profitability of any production system is dependent on both output and input traits. In economic environments of rising feed costs and resource constraints, producers in the beef industry have become increasingly aware of the need to improve input/cost traits, such as feed intake.

Feed costs have been attributed to the largest production expense for cow/calf producers accounting for 50% to 70% of total production costs (Anderson et al., 2005). Decreasing feed costs while maintaining animal performance could have large impacts on the profitability of a beef operation. A 10% increase in performance (i.e. gain) has been reported to contribute an 18% increase in profits, as compared to a 10% improvement of feed efficiency (increasing performance while simultaneously decrease feed intake) which is reported to increase profits by 43% (www.beefefficiency.org). A report from the Alberta Agriculture and Forestry Ministry (2006) stated that a 5% improvement in feed efficiency would have 4 times the economic effect of a 5% improvement in average daily gain. Therefore, maintaining animal performance while decreasing feed costs could have large impacts on the profitability, and therefore the sustainability, of a beef operation and could make beef more competitive against other, cheaper proteins such as pork or poultry.

In addition to the economic benefits associated with reductions in of feed intake, there are environmental benefits to consider. The world population is estimated to reach 9.7 billion by 2050 which will require an increase in agricultural output of 70% with increasing constraints on resources (FAO, 2009). To sustain animal agriculture as a protein source for this growing population, cattle will have to be produced on fewer natural resources such as feed or forage. With increases in the human population, there are also concerns of intensifying climate change and ruminants have been associated with contributing 80% of livestock greenhouse gas emissions (Gerber et al., 2013). Feed intake has been shown to be positively correlated to greenhouse gas emission from beef cattle. Herd et al. (2014) found that methane production was positively correlated with DMI (0.65 ± 0.02) and that the reduction of feed intake while maintaining production through genetic selection would contribute to a reduction in greenhouse gases. Cottle (2011) proposed a selection index using feed intake for indirect selection for the reduction of greenhouse gases. Considering that cattle finished in a feedlot spend 50 to 70% of their lifespan grazing forage prior to feedlot entry (Capper, 2011), a reduction of forage intake would likely contribute to the reduction of greenhouse gases. Furthermore, reducing forage intake while maintain stocking rates on rangelands has the potential to improve pasture quality. When cattle require less feed for production, stress placed on pastures due to drought, overgrazing, or climate change can be mitigated.

Due to the importance of feed costs to the sustainability and profitability of a beef operation, interest in selecting cattle that are more feed efficient has increased. Several breed associations have begun to publish expected progeny differences (EPD) for feed intake or traits of feed efficiency, such as residual feed intake (RFI). In order to select cattle that are more efficient, individual feed intake must be measured. Since the mid-1990's, technological advances have

allowed for a large increase in the measurement of feed intake on group housed animals (Cruz et al., 2011; Hill, 2012, Arthur et al., 2014). The collection of feed intake measures is expensive, time consuming and testing facilities have a limited capacity for the measurement of individual feed intake (Wang et al., 2006; Nielsen et al., 2013). These limitations restrict the number of animals that can be measured annually and therefore limit the amount of data generated for genetic evaluations.

Current individual feed intake measurements are obtained in feedlot environments and have a direct application to feedlot cattle. The translation and application of these feedlot-measured intakes outside of a feedlot environment is essentially unknown. This is significant given that approximately 50% of feed costs in the beef industry are attributed to the mature cow herd (Whisnant, USDA-NIFA-CRIS) which is generally maintained in much more extensive environments with forage-based diets, typically grazing on rangelands.

To select for a cowherd that is more efficient in forage utilization, individual grazing intake needs to be measured. Currently there is no technology to effectively measure feed intake on a large population of grazing cattle. Methods used for estimation of dry matter intake (DMI) for grazing animals lack precision and are often tedious, expensive and time-consuming (Undi et al., 2008). Current techniques for measuring grazing DMI typically involve digestive markers, herbage disappearance measured on group housed animals, or equations predicting DMI based on net energy requirements (NRC 2000; Meyer et al., 2008; Undi et al., 2008).

To date, there are large gaps of knowledge on feed intake as a phenotype, such as the relationship of feed intake measured in a feedlot to forage intake and the importance of maternal effects on feed intake. In addition, there is a debate as to the application of feed intake as an economically relevant trait and how it should best be used for selection decisions and for genetic

improvement. The overarching objective of this dissertation is to explore a novel approach to estimating forage intake, explore maternal effects on feed intake, and illustrate different approaches of genetic selection for feed efficiency.

1.2 Objectives

The underlying theme of this dissertation is to further examine feed intake and its application to the beef industry. This will be accomplished through 3 projects:

1. To explore and estimate parameters for the maternal genetic effects of feed intake on beef cattle.
2. To determine the effects of direct selection on DMI, residual feed intake (RFI), or an economic selection index using simulated data to illustrate which selection method would be more desirable for production.
3. To develop a method or approximation for measuring intake on grazing beef cattle through measured behavior such as eating time or rumination.

LITERATURE CITED

Alberta Agriculture and Forestry. 2006. Agri-Facts: Residual feed intake (net feed efficiency) in beef cattle. https://open.alberta.ca/dataset/91a77dec-f0a4-49c2-8c54-f172fe568e2c/resource/721e982c-b90f-4605-9de0-a3b8bb312b1f/download/2006-420-11-1.pdf. (Accessed April 30, 2017).

Anderson, R. V., R. J. Rasby, T. J. Klopfenstein, and R. T. Clark. 2005. An evaluation of production and economic efficiency of two beef systems from calving to slaughter. J. Anim. Sci. 83: 694-704.

Arthur, P. R., J. E. Pryce, and R. M. Herd. 2014. Lessons learnt from 25 years of feed efficiency research in Australia. Proc. 10th World Congr. Genet. Appl. to Livest. Prod., Vancouver, Canada.

Arthur, P. R., J. A. Archer, R. M. Herd, and G. J. Melville. 2001. Response to selection for net feed intake in beef cattle. Proc. Assoc. Advmt. Anim. Breed Genet. 14: 135-138.

Capper, J. L. 2011. The environmental impact of beef production in the United States: 1977 compared with 2007. J. Anim. Sci. 89: 4249-4261.

Cottle, D. J. 2011. Use of residual feed intake as an indirect selection trait for reduction of methane emission on grazing beef cattle. Proc. Assoc. Advmt. Anim. Breed Genet. 19: 423-425.

Cruz, G. D., J. B. Trovo, J. W. Oltjen, and R. D. Sainz. 2011. Estimating feed efficiency: Evaluation of mathematical models to predict individual intakes of steers fed in group pens. J. Anim. Sci. 89: 1640-1649.

FAO. 2009. How to feed the world in 2050. Food and Agriculture Organization of the United Nations, Rome, Italy.

Gerber, P. J., H. Steinfeld, b. Henderson, A. Mottet, C. Opio, J. Dijkman, A. Falcucci, and T. Tempio. 2013. Tackling climate change through livestock – A global assessment of emissions and mitigation opportunities. Food and Agriculture Organization of the United Nations, Rome, Italy.

Herd, R. M., P. F. Arthur, K. A. Donoghue, S. H. Bird, T. Bird-Gardiner and R. S. Hegarty. 2014. Measures of methane production and their phenotypic relationships with dry matter intake, growth, and body composition traits in beef cattle. J. Anim. Sci. 92: 5267-5274.

Hill, R. A. 2012. Feed Efficiency in the Beef Industry. Wiley-Blackwell, Oxford, UK. p. 7-18.

Meyer, A. M., M. S. Kerley, and R. L. Kaleenbach. 2008. The effect of residual feed intake classification on forage intake by grazing beef cows. J. Anim. Sci. 86: 2670-2679.

National Program for Genetic Improvement of Feed Efficiency in Beef Cattle. USDA. http://www.beefefficiency.org (Accessed April 30, 2017).

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

NRC. 2000. Nutrient requirements of beef cattle. 7th rev. ed. Natl. Acad. Press, Washington, DC.

Undi, M., C. Wilson, K. H. Ominski and K. M. Wittenburg. 2008. Comparision of techniques for estimation of forage dry matter intake by grazing beef cattle. Can. J. Anim. Sci. 88:693-701.

Wang, Z., J. D. Nkrumah, C. Li, J. A. Basarab, L. A. Goonewardene, E. K. Okine, D. H. Crews, Jr. and S. S. Moore. 2006. Test duration for growth, feed intake, and feed efficiency in beef cattle using the GrowSafe System. J. Anim. Sci. 84: 2289-2298.

Whisnant, C. USDA-NIFA-CRIS Research. http://cris.csrees.usda.gov/cgi-bin/starfinder/6244/crisassist.txt (Accessed April 30, 2017)

Chapter II

LITERATURE REVIEW

2.1 Introduction

Sustainability and profitability of agricultural systems are dependent on both outputs and inputs. In the beef industry, producers have traditionally placed emphasis on improving output traits, such as fertility and live weight, to increase profitability (Arthur et al., 2001; Hill, 2012). However, increasing outputs often lead to increased inputs, such as feed costs (Meyer et al., 2008). Feed costs are the largest variable expense in the beef industry accounting for 50 to 70% of total production costs (Anderson et. al, 2005). It has been estimated that a 10% increase in weight gain would increase profits by 18%, this is in contrast to a 10% improvement in feed efficiency which has been estimated to increase profits by 43% (www.beefefficiency.org). This suggests improvement in feed efficiency would have a larger impact on profit as compared to the same magnitude of improvement on performance. This is further supported by the Alberta Agriculture and Forestry Ministry (2006) that reported that a 5% improvement in feed efficiency would have 4 times the economic effect of a 5% improvement in ADG. Therefore, decreasing feed costs without sacrificing animal performance could have large impacts to the profitability of a beef operation.

There is debate as to the best phenotype for feed efficiency in cattle, how to incorporate it into a breeding program and what impacts selection on feed intake or efficiency would have on other performance traits (Berry and Pryce, 2014). The traditional phenotypic measures of feed efficiency in the beef industry were ratios of intake to production and generally on pens of cattle.

Also, direct selection of ratios is typically problematic due to inherent problems of selection for ratio measures (Gunsett, 1986). More recently, residual feed intake (RFI) has become the phenotype of interest for feed efficiency in livestock production because it considers production levels of the animals (Berry and Crowley, 2013).

Due to the large influence of feed costs on production profits, interest in selecting for cattle that are more feed efficient has increased. However, to select for more feed efficient cattle, individual feed intakes must be measured. Currently, individual feed intake measurements are collected in feedlot environments and the application of these measurements outside of the feedlot has yet to be quantified (Berry and Crowley, 2013). Furthermore, the relationships between these feedlot-measured intakes and the cowherd grazing on rangeland are unknown. Approximately 50% of variable feed costs and 70 to 75% of total annual energy for maintenance in the beef industry are attributed to the mature cow herd (Whisnant, USDA-NIFA-CRIS; Ferrell and Jenkins, 1985). To select for a cowherd that is more efficient for forage utilization, grazing intake needs to be measured. Currently, there are no simple technologies to measure feed intake on a population of grazing cattle. The relationship between feed intake measured in current feed trials and intake from grazing cattle are unknown.

2.2 Measuring Feed Intake

The ability to measure feed efficiency in cattle is dependent on the capability to measure individual feed intake (FI) and the quality of these records (Hill, 2012). Traditional methods for the collection of intakes per animal involved housing cattle individually. Research has shown an inadequacy of feed intake data for individually housed cattle for the purpose of genetic evaluations

(Hill, 2012) given that the collection of FI on individually housed animals severely limited the number of animals that could be measured and thus limited the ability for estimation of reliable genetic parameters of FI. Since the mid-1990's, advancements in technology such as radio electronic identification (RFID) of animals, have led to a large increase in the measurement of FI for group housed animals (Cruz et al., 2011; Hill, 2012, Arthur et al., 2014).

2.2.1 Equipment and Guidelines for Measuring Feed Intake

Traditional methods for measuring individual FI involved to individually housed cattle. Specifically, cattle were allocated to a pen where each animal's feed was weighed before being placed into the bunk. Feed that was not consumed throughout the day, was subsequently weighed and subtracted from the initial feed weight. The difference in feed weight was the animal's FI. This process was extremely time consuming and limited the number of animals that could be measured throughout the year. In addition, this method did not accurately reflect the behavior of group housed animals whose social behaviors may affect their FI (DeHaer and Mercks, 1992; Neilsen et al., 1995; Guiroy et al., 2001; Chapinal et al., 2007; Cruz et al., 2011)

With increased need for the collection of feed intake data, research facilities and performance-testing centers have been equipped with the technology for the capability to measure individual feed intake on group housed animals (Hill, 2012). Technology to measure feed intake must be capable of identifying individual animals, weighing rations fed to the individual animal, associate the measured feed consumed to the appropriate animal and compile the data into a useable format (Dahlke et al., 2008). Prominent technologies used for measuring feed intake are the Calan Gate System (American Calan, Inc.; Northwood, NH), GrowSafe Systems (GrowSafe

Systems, Ltd., Airdire, Alberta, Canada), Insentec (Hokofarm; Marknesses, The Netherlands) and Smartfeed (C-Lock Inc.; Rapid City, SD).

Calan Gate Systems (Figure 2.1) are a system of feed bunks that require a "key" for an animal to access the feed within their individual bunk. In this system the door of the bunk transmits an electrical signal which the animal's key recognizes and the door is unlocked. The animal pushes the door open and is able to access the feed. While an animal is consuming feed in the bunk, the narrow door to the bunk prevents other animals in the pen from accessing the feed from that specific bunk. The restriction of animals to specific bunks allows researchers to dispense individual diets to specified animals in addition to measuring individual feed intake (American Calan, Inc.; http://americancalan.com). A disadvantage of the Calan Gate System is that animals must be trained to use the system and access the bunks (Stock and Klopfenstein, 1986; Cole, 1995). Therefore, it is imperative that animals have an extended adaptation period up to 3 weeks for Calan Gates (Stock, 1986). It should also be noted that a small percentage of cattle cannot be trained to this system which limits the number of cattle for data collection.

Figure 2.1. Calan Gates are pictured on the left and animals equipped with electronic keys pictured on the bottom (http://www.animal.ufl.edu/facilities/bru/images/calang2.jpg).

In contrast, GrowSafe Systems (Figure 2.2) utilize electronic scales built into the feed bunk. Animals are equipped with an electronic ear tag that the GrowSafe system uses to identify each individual. When the ear tag enters into the GrowSafe feed bunk, the system registers the animal and weighs the feed in the bunk. When the animal removes its head from the bunk, the system weighs the feed again. The difference between the two weights is the animal's measured feed intake. The presence of the electronic ear tag initiates the feeding event and intake record. The system also measures the amount of time an animal spends at the bunk. The data is then transmitted for processing and reporting of individual feed intake (GrowSafe Systems, Ltd.; http://growsafe.com; Dahlke et al., 2008). Unlike Calan Gates, as long as the cattle are accustomed to eating from a feed bunk, the time spent to acclimate animals to GrowSafe is reduced.

Figure 2.2. Picture of GrowSafe System feed bunks at Colorado State University's Feed Intake Unit.

The Insentec monitoring system (Figure 2.3) are a series of feed bunks with barriers that are lowered to allow access the feed within the bunk. Each individual animal is equipped with an electronic ear tag that Insentec uses for animal identification. As the animal approaches, an antenna detects the electronic ear tag for the individual animal and lowers the barrier, allowing the animal access to feed inside the bunk. When the animal leaves the bin, the barrier is raised until another animal approaches the feed bin. When the animal exits the feed bunk, Insentec records the initial and finial time and feed weight to determine feeding duration and intake (Chapinal et al., 2007).

Figure 2.3. Picture of Insentec Monitoring System (http://www.hokofarmgroup.com/ric/feed-weigh.aspx).

SmartFeed (Figure 2.4) measures feed intake similar to the GrowSafe system. SmartFeed utilizes an electronic ear tag that identifies an animal when their head enters the feed bunk. The

feed within the bunk is weighed and when the animal retracts their head from the bunk, the amount of feed is measured again. The difference in feed weight is the animals FI. The primary difference between SmartFeed and GrowSafe, is the mobility of SmartFeed. SmartFeed is designed to be portable, self-contained and can be purchased as transportable trailers with 2 to 4 feed bunks per trailer. These trailers have the potential for use in pastures for measurement of supplement intake. To date, there is limited literature evaluating the use of SmartFeed to measure FI. Most literature references for the use of SmartFeed has been for supplementation and trace mineral studies; however, the application for measuring feed intake using SmartFeed is increasing (McCarthy et al., 2018; Wyfells et al., 2018).

Figure 2.4. SmartFeed feed monitoring system. On the left are the SmartFeed feed bunks in a feedlot and on the right is the SmartFeed trailer that can be transported to various locations (https://www.c-lockinc.com/shop/feed-intake-measurements/smartfeed/).

The above section described 4 technologies that are prominently used for the collection of individual feed intake in cattle. Given the increasing interest in feed intake, new technologies and methods for measuring FI will continue to develop and evolve.

2.2.2 Feeding Trials for Feed Intake

When incorporating feed intake or traits requiring feed intake measures into a breeding objective, the goal of the test is to produce the maximum number of responses for data collection and precision at the minimum cost (Archer et al., 1997). Measuring feed intake is expensive and time consuming (Wang et al., 2006; Nielsen et al., 2013), a problem which is exacerbated by limited capacity for testing facilities to measure individual feed intake (Nielsen et al., 2013). These limitations restrict the number of animals that can be measured within a year and therefore limit the amount of data generated for genetic evaluations. Optimizing testing periods for the measurement of feed intake and efficiency is crucial to increase the number of animals measured within facilities but still obtain reliable measures for intake and efficiency.

In 1997, Archer et al. published the results from a study examining the optimum test duration to standardized measurement protocols for measuring feed intake and RFI. The authors compared the genetic and environmental variances, genetic and phenotypic correlations, and the predicted correlated response to selection (selection efficiency) of 119-d test with shortened tests. The shortened test lengths ranged from 7-d to 119-d. The resulting correlations were assessed for the test length with the same traits as the 119-d test. The authors concluded that the measurement of feed intake required a minimum of a 35-d test and a 70-d test would be required to obtain RFI measures. They also concluded that although phenotypic correlations would be useful indicators of the ability to shorten test lengths, if the ultimate goal was genetic selection comparison of genetic correlations between full and shortened tests would be more relevant. Following the work by Archer et al. (1997), Archer and Bergh (1999) examined different breeds of beef cattle and concluded that testing length for residual feed intake (RFI) could be shortened to 70-d regardless

of breed. More recently, it was reported that shortening testing periods to 56-d for the calculation of RFI would lead to limited loss in information compared to a 70-d testing period (Culbertson et al., 2015). In all research examining testing length, the driving factor for test length for feed efficiency was the ability to obtain a reliable measurement of average daily gain (ADG). Retallick et al. (2015) reported a strong genetic correlation between on-test ADG and post-weaning gain (PWG) from national cattle evaluations. Based on this genetic correlation, the authors concluded that the PWG could be used as a proxy for on-test ADG, allowing producers to reduce the length of feed intake testing periods.

As a result of Archer et al. (1997), guidelines for measuring feed intake and efficiency were defined to be a 70-d test for standardize feeding trials for measuring FI. The Beef Improvement Federation (BIF) proposed a set of guidelines for the collection of feed intake and calculations of feed efficiency as well. The BIF guidelines recommend a 21-d adaptation period to allow cattle to acclimate to the testing facility. This was followed by either a 45-d or 70-d testing period depending on the phenotype of interest. A minimum of 70-d testing period was recommended to accurately measure average daily gain for the use of calculating feed efficiency and for feed intake a 45-d testing period are sufficient.

An animal's age during a testing period is related to their feed intake. Therefore, BIF recommends the age when an animal enters a feeding test be post weaning but no younger than 240-d, but the test should be completed prior to the animal reaching 390-d. Cattle should also be contemporary grouped within 60-d of age at the beginning of the testing period (BIF, 2016).

2.2.3 Measuring Grazing Intake

Currently, individual feed intake measurements are collected in feedlot environments and have a direct application to feedlot performance. The translation of these feedlot-measured intakes to the cowherd grazing on rangeland is unknown. Approximately 50% of the varied feed costs in the beef industry are attributed to the mature cow herd (Whisnant, USDA-NIFA-CRIS; accessed April 2017). To select for a cowherd that is more efficient for forage utilization, individual grazing intake needs to be measured or an indicator trait for grazing intake needs to be established. Previous methods used for estimation of FI for grazing cattle lacked precision and were often tedious, expensive and time-consuming (Undi et al., 2008). Current techniques for measuring grazing FI typically involved digestive markers, group housed animals where herbage disappearances are measured, or equations predicting FI depending on net energy requirements (Meyer et al., 2008: Undi et al., 2008). Currently, there is no technology to efficiently measure individual feed intake on a large population of grazing cattle and the relationship between feed intake measured in a feedlot and intake from grazing cattle are relatively unknown.

Studies involving group housed animals typically involve a number of animals confined to a pasture sections (referred to as paddocks) where herbage disappearance is measured and averaged for the group of animals. These studies usually utilize grazing cages (Figure 2.5), where a small section of the paddock is enclosed to keep animals from grazing and allow un-grazed forage growth. At the conclusion of the grazing trial, the difference in biomass within and outside of the cage are compared in order to calculate herbage disappearance (Burns et al., 1994; Undi et al., 2008). Grazing intake can be determined for the group of animals, but unless animals are individually housed within a paddock, determining individual grazing intake is not possible.

Figure 2.5. Picture of grazing cage (Nobel Research Institute, 2017).

The most commonly used method of forage intake estimation is the use of digestive markers such as chromium oxide, titanium dioxide and n-alkanes (Gordon 1995; Titgemeyer et al., 2001; Undi et al., 2008). Long chain n-alkanes occur naturally in plants and have been used as digestive markers for the estimation of forage FI through the recovery of alkanes in feces (Dove and Mayes, 1991). Chemical digestive markers, such as chromium oxide and titanium dioxide, can be used to estimate the FI based on the concentration level of the marker in feces of individual cattle (Undi et al, 2008; Titgemeyer et al., 2001). The implementation of digestive markers for the collection of FI phenotypes in a production setting is not practical since it is labor intensive and tedious to obtain the actual measurement.

Empirical methods of estimating FI for grazing cattle have also been tested. Prediction equations using BW and ADG have been developed to estimate FI of cattle (NRC 2000; Minson and McDonald, 1987; Undi et al., 2008). The National Research Council (2000) have published prediction equations which estimated FI as a function of dietary net energy for maintenance (NE_m) and BW. These estimations were appropriate when estimating the required FI for a group of cattle;

however, the estimations fails to account for variability in FI between individual cattle of the same BW.

2.3 Measures of Feed Intake and Efficiency

Efficiency is defined as a level of performance that uses the lowest amount of input to obtain the greatest level of output (www.investopia.com). Efficiency can be described at a production level as the saleable output per unit input, weighted according to their relative economic importance (Berry and Crowley, 2013). Feed intake and feed efficiency are important, contributing factors to the economics for production efficiency (Berry and Crowley, 2013; Hill 2012). Given that feed variable costs are the largest expense to producers, feed intake and efficiency has a direct economic relevance and is therefore considered an economically relevant trait (Neilsen et al., 2013). However, there is no definitive definition of feed efficiency in beef cattle and several definitions and calculations for feed efficiency exist. Feed efficiency in beef cattle is typically described as either a ratio or residual trait (Berry and Crowley, 2013). In the context of genetic improvement, there is often a debate as to what is the best measure for feed efficiency, how to incorporate feed efficiency into a breeding program and what the impact of selection would do to other performance traits (Berry and Pryce, 2013; Nielsen et al., 2013).

2.3.1 Feed Intake

For the purpose of this review, feed intake (FI) is the amount of feed an individual animal ingests on a dry matter basis. Factors that influence individual FI in cattle are complex and not

fully understood (NRC, 2000). Physiological factors influencing FI are body composition, frame score, physiological state and age. Sex was reported by the NRC (2000) to have limited effect on FI.

An animal's body fat composition can affect FI. When predicting FI for beef cattle, percent body fat is often considered because the relationship between FI and body fat is a management tool for feedlots to determine the appropriate slaughter condition of beef cattle based on the animals' FI (NRC, 2000). Fox et al. (1988) estimated a decrease in FI of 2.7 percent for each 1 percent increase in body fat within the range of 21.3 to 31.5 of percent body fat for growing cattle. The NRC (1987) theorized that adipose tissue may have a feedback roll that influences FI. Therefore, when assessing the FI of younger growing cattle or older mature cattle, the influence of percent body fat should be considered.

The age of an animal when placed on feed can affect that animal's FI. Yearling cattle have a higher FI when compared to weaned calves. However, this increase in FI may be attributed to the increase in body weight and size. The NRC (2000) reported that "the greater the ratio of age to body weight for yearling cattle prompts greater feed intake." This ratio of age to body weight has been related to the increase in feed intake for cattle experiencing compensatory growth. Yearling cattle have been reported to have a 10 percent increase per unit body weight in estimated FI compared to cattle started on feed as calves (NRC, 1984; Fox et al., 1988). The frame score, or body size, of an animal can also influence their intake. The larger the animal, the higher the FI. Currently, there is a lack of designed studies to further the understanding of the biological effects of age and body weight or composition on FI (NRC, 2000).

The physiological status of an animal, such as lactation or pregnancy, can significantly alter the animal's energy requirements and therefore affect FI. For lactating cattle, net energy and

protein requirements are at their highest during peak lactation (Adams et al., 1996). According to the Agricultural Research Council (ARC, 1980), FI increases by 35 to 50 percent for lactating animals when compared to non-lactating animals of the same BW and diet. Therefore, cattle bred for higher levels of milk production would be expected to have an increase in FI due to the increase in energy requirements (NRC, 2000). A cow's stage of pregnancy can also affect her FI; however, the overall physiological effect of pregnancy on individual feed intake is relatively unknown (Arthur et al., 2001). Nutritional studies have shown a decrease in FI for cows during their last stage of pregnancy. Research has shown a decrease of 1.5 to 2 percent per week of FI during the animals last month of pregnancy (Ingvartsent et al., 1992).

2.3.2 Feed Conversion Ratios

Traditional measures of feed efficiency in the beef industry were ratios of intake and production traits. Feed conversion ratio (FCR) is the most commonly used measure for feed efficiency (Berry and Pryce, 2013) and widely used in production settings (Nielsen et al., 2013). However, FCR does not account for the differences in maintenance efficiency among individual animals (Berry and Crowley, 2013). Feed conversion ratio is routinely represented as follows:

$$FCR = \frac{Average\ Daily\ Gain}{Feed\ Intake}$$

(Eq. 2.1)

Feed conversion ratios are often referred to as gain to feed (G:F) or it's reciprocal feed to gain (F:G). These measures are regularly used in feedlots using pen averages for ADG and feed intake. Animals with higher values for FCR are considered to be more efficient.

Heritabilities for FCR range from 0.06 to 0.41 (Berry and Crowley, 2013). This suggests that genetic improvement in FCR is possible but direct selection on ratios makes the expected response to selection difficult to ascertain (Gunsett, 1986; Berry and Crowley, 2013). Gunsett (1986) argued "prediction of response to selection practiced to change a trait such as feed conversion assumes that the trait has a normal distribution with some mean and variance. The fact that feed conversion is a ratio of two traits has made the ability to predict the change of the trait in future generations difficult." Placing selection pressure on the components that comprised the ratio would be more effective for improved feed efficiency (Gunsett, 1986).

Although FCR is easy to calculate and accepted for use in the U. S. beef industry, it is not an ideal measure of feed efficiency for the purpose of genetic improvement given the inherit problems with selection on ratios as the phenotype. This has led to the proposal of other traits for genetic improvement on feed efficiency, such as residual feed intake.

2.3.3 Residual Feed Intake

Koch et al. (1963) introduced the use of RFI as a measure of feed efficiency. The authors proposed that feed efficiency was not a directly measurable trait but must be calculated as a function of feed intake, increase in body weight, and time. They concluded that efficiency expressed as gain adjusted for differences in feed consumption, or the deviation from the regression of gain on feed intake, was considered the most accurate mathematical description for

phenotypic feed efficiency. Since 1963, the general definition of RFI has developed into the difference between actual feed intake and the estimated feed intake adjusted for the requirements of production (Kennedy et al., 1993). A typical RFI equation for beef cattle would be:

$$RFI = DMI - (\beta_1 ADG + \beta_2 Fat + \beta_3 WT^{0.75})$$

(Eq. 2.2)

where DMI was the individual animal's dry matter intake, ADG was an animal's average daily gain, Fat is the animal's ultrasound measurement for back fat, $WT^{0.75}$ was the metabolic mid test weight for an animal, and β_i is the regression coefficient for the corresponding predictor (i.e. ADG). Metabolic mid weight is estimated by taking the animal's weight at the midpoint of the test and raising it to the ¾ power. Including $WT^{0.75}$ in the model for RFI accounts for energy sinks associated with the energy requirements for body weight (Berry and Crowley, 2013; NRC, 2000).

Cattle with positive RFI values would be considered inefficient as their feed intake was more than what was expected for their level of performance. Negative RFI values suggest cattle who are more efficient as their feed intakes were less than what was expected given their level of performance. An attraction of RFI is its independence from variables included in the multiple regression model (i.e. ADG). Given that RFI are residual terms of the model (Eq. 2.2), they are dependent on the variables included in the model. As the complexity of the model increases, the variation in RFI will reduce but the risk of over parameterizing the model increases as well (Koch et al., 1963; Kennedy et al., 1993; Berry and Pryce, 2014).

An advantage to RFI is that it is a measure of feed efficiency (Berry and Pryce, 2014). This is compared to FI which only measures the amount of feed ingested by cattle and doesn't account

for the animal's performance. An animal's intake will depend on their level of performance, stage in life and the actual size (NRC, 2000). Larger cattle will have a higher FI while FI for smaller cattle will be lower. Selecting for cattle with lower FI may inadvertently select for smaller cattle (Nielsen et al., 2013) and therefore, a single measure of feed *efficiency* would have to account for reduced FI without sacrificing gain in body weight.

Residual feed intake is also heritable, which would mean that genetic selection for lower RFI would be possible. Numerous studies have estimated heritabilities for RFI in several populations of cattle. Estimates of RFI have been reported ranging from 0.07 to 0.62 (Berry and Pryce, 2014; Berry and Crowley, 2013). Berry and Pryce (2014) reported a pooled heritability for RFI at 0.33 from 36 previous studies which would be considered moderate. Selection on RFI compared to FI included in an index was shown to be mathematically equivalent by Kennedy et al. (1993). Selecting for RFI does have substantial benefits over selecting for FCR, since, as stated above, the selection of cattle for breeding based on ratios are problematic (Gunsett, 1986).

There are several disadvantages to RFI for application in the beef industry. One disadvantage of RFI is that it can be conceptually difficult to explain to producers (Berry and Pryce, 2014). From a technical standpoint, calculations for RFI would seem complex and difficult for some producers to understand. Residual feed intake values dependent on the group in which they were calculated (i.e. their contemporary group) and are susceptible to genetic by environment interactions (Berry and Crowley, 2013) and a direct comparison of RFI values cannot be made across contemporary groups. Since RFI is a residual from within a regression model, the phenotypic value of the RFI can only be compared to other animals that were included within the model. Actual RFI values are dependent on effects included in the model as these would change the resulting residuals. Residual feed intake is essentially a selection index (Eisen, 1997) and RFI

values are dependent on contemporary groups (CG) since phenotypic indexes do not account for environmental differences (Bourdon, 1997).

These disadvantages of RFI become problematic when trying to incorporate this variable into a genetic evaluation. Varying conditions such as different equipment between testing facilities or environmental conditions may be problematic. These varying conditions potentially affect the mean and the variance of the observations (Nielsen et al., 2013). Equations for estimating RFI would also have to be standardized. Differences in what is included in the models would change the RFI values and could result in a re-ranking of animals.

2.4 Biological Factors of Feed Efficiency

The biological functions influencing feed efficiency are not fully understood and it is to be expected that many mechanisms are associated with phenotypic feed efficiency in livestock (Herd and Arthur, 2009; NRC 2000). Intuitively, variations in feed efficiency depend on differences in biological processes (Herd and Arthur, 2009) and this variation is likely associated with at least 5 major processes: FI, feed digestion, metabolism, activity and thermoregulation (Herd et al., 2004).

Richardson and Herd (2004) described many physiological mechanisms that contribute to the variation observed in RFI for Angus steers divergently selected for low and high RFI. They concluded that metabolic heat production, body composition, and physical activity accounted for 73% of variation in RFI. Figure 2.5 illustrates the portions of variation of RFI explained by each physiological function.

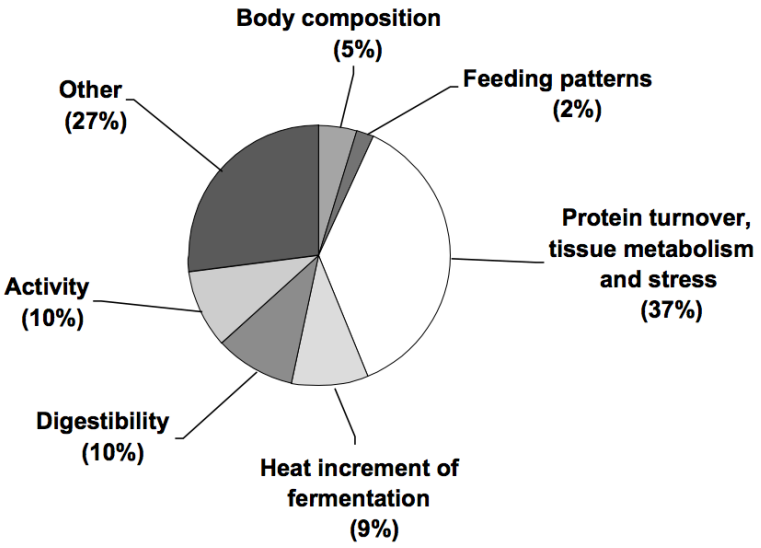

Figure 2.6. Contributions of biological mechanisms to variation in residual feed intake as determined from experiments on divergently selected cattle (Richardson and Herd, 2004).

The body composition of an animal affects the variation of feed efficiency of that animal. The deposition of fat versus lean tissue has different energy costs. There is less variation for the depositing of fat gain when compared to lean gain. Theoretical partial efficiencies of nutrient use for lean gain is estimated at 40 to 50% where the nutrient use for fat gain is estimated at 70 to 95%. The metabolism of protein has a greater amount of variation in efficiency when compared to fat metabolism. This variation in body composition and gain would influence the efficiency of nutrient utilization (Herd and Arthur, 2009).

Richardson et al. (2001) reported a correlation between chemical composition and genetic variation for RFI for Angus steers divergently selected for RFI. The authors of that report stated that steers from low-RFI parents had less whole-body fat and more whole-body protein then the steers from high-RFI parents. Also, been reported and described in literature, that animals with high RFI tend to have more fat deposition then animals with lower RFI values (Carstens et al, 2002; Nkrumah et al., 2007; Hill, 2012).

2.4.1 Mitochondrial Function

Feed efficiency is affected at many levels from the animal as a whole, from environmental effects, to differences at the cellular level (Bottje and Kong, 2013). Mitochondria are the energy production organelles of the cell, producing approximately 90% of the cell's energy and the majority of ATP (Kolath et al., 2006; Herd and Arthur, 2009). The mitochondrion have a crucial role in growth and animal development and has been linked with production efficiency (Hill, 2012). Variation in the mitochondria's ability to produce energy results in phenotypic difference in feed efficiency of an animal (Herd and Arthur, 2009). Kolath et al. (2006) found no difference in mitochondrial function between low or high RFI animals; however, the rate of mitochondrial respiration was increased in low RFI animals suggesting an efficiency of electron transfer. The authors concluded that mitochondrial function was not impaired for steers with high RFI but the flux of electrons through the electron transport chain was impaired in low RFI steers.

Hill (2012) discussed potential areas of mitochondrial inefficiencies through electron transport defects and mitochondrial defects. The electron transport chain (ETC) is a source of reactive oxygen species (ROS) and endogenous oxidative stress. Reactive oxygen species forms superoxide (O_2^-) as a result of electron leakage from the electron transport chain. Low levels of ROS are needed for cellular function and are metabolized by antioxidants. However, ROS is capable of altering gene expression and has been linked to diabetes, Alzheimer's and cancer in humans. The author suggested that elevated levels of ROS in animals that are less efficient in feed conversion could play a role in gene regulation and protein expression, and mitochondrial function with phenotypic expression of feed efficiency.

Nitric oxide (NO) is involved in the regulation of mitochondrial respiration when nitric oxide synthase produces NO near the ETC. Nitric oxide has been shown to competitively inhibit cytochrome c and regulate oxygen absorption. When NO reacts with ROS, several reactive nitrogen species are released which damages cellular structures such as complexes I and II of the ETC (Hill, 2012).

The functional integrity of mitochondria is due to a balance of mitochondrial DNA (mtDNA) and nuclear DNA (nDNA). There is a lack of protective histones for mtDNA and it has a close proximity to the ETC. As a result, mtDNA has a tendency for ROS oxidation, which leads to mitochondrial dysfunction. The damaged mtDNA has a restricted ability to encode subunits or proteins, which leads to a diminished ability of the mitochondria to achieve its respiratory ability. Complexes I, II and IV were particularly affected by this and as a result the ETC would have decreased activity. From the standpoint of animal feed efficiency, there could be a relationship between inefficient animals and a decrease in the activity of respiratory chain complexes (Hill, 2012).

Protons pumping across the inner mitochondrial membrane are used to drive ATP synthesis. Protons can also flow back into the mitochondria essentially short circuiting the coupling of ATP synthesis. This process is termed a proton leak and represents 30% and 50% of oxygen consumption in liver and muscle cells, respectively, and contributes up to 25% of total basal metabolic rate of an animal. It stands to reason that proton leak could contribute to feed efficiency in cattle. When ADP levels are high, the animal's respiration rate will be dependent on oxidative phosphorylation and the reaction that produces ATP. When ADP levels are minimal, the proton leak across the inner membrane of the mitochondria controls oxygen consumption (Hill, 2012).

The final mitochondrial inefficiency examined by Hill (2012) was the uncoupling attenuation of oxidative stress. This uncoupling is a result of cellular inefficiency but alleviates oxidative stress by reducing ROS production. When examining broilers, Ojano-Dirain et al. (2007) found uncoupling lowered ROS production in those with inferior gain to feed ratios but was not found in mitochondria from individuals with superior gain to feed ratios. It was also found that ROS production was significantly higher in broilers who were less efficient for feed utilization and as a result uncoupling would be expected. These studies indicate that higher mitochondrial ROS associated with the phenotypic expression of less efficient FCR or RFI would indicate differences in membrane characteristics that affect the proton conductance of the mitochondria (Bottje et al., 2002; Hill, 2012).

Mitochondria play a clear role in the efficiency of feed utilization in livestock animals. Proton leak and ROS production are inefficient mitochondrial activities and may be indicative of less-efficient animals. The development of genomic markers from nDNA or mtDNA to identify efficient or inefficient animals is not currently available given the complex nature of the interaction of n- and mt-encoded proteins (Hill, 2015).

2.4.2 Biological Markers of RFI

Important factors influencing the variation of feed utilization may be explained by biological differences in the growth hormone/IGF-1 axis (Hill 2012). Glucose and amino acid metabolism, protein accretion, and linear growth are all affected by endocrine actions of IGF-1 (Jones and Clemmons, 1995) suggesting a role in feed utilization and efficiency. Australian research on IGF-1 in weaned pigs reported a moderate to high positive genetic correlation between

plasma IGF-1 concentrations and FCR (Luxford et al., 1998; Hermesch et al., 2001). Bunter et al. (2010) estimated genetic correlations between RFI and IGF-1 in pigs to be 0.63 ± 0.15.

In *Bos taurus* cattle, genetic correlations between IGF-1 concentration and RFI were positive and ranged from 0.39 to 0.63 (Arthur et al., 2004; Brown et al., 2004; Moore et al., 2005; Hill, 2012). Herd and Arthur (2009) examined data from Australian Angus cattle prior to 2004 and found a genetic correlation of 0.57 between RFI and serum IGF-1 concentrations collected at or before weaning. The authors concluded that this would suggest that many of the genes responsible for greater concentrations of IGF-1 are also associated with high RFI measures. In contrast, researchers analyzed data available prior to 2007, and estimated the correlations between IGF-1 and RFI from yearling steers before harvest (feedlot) and RFI from younger cattle (postweaning). The genetic correlation between postweaning RFI and feedlot RFI with IGF-1 were 0.17± 0.11 and -0.22 ± 0.16, respectively. These lower correlations suggested that the effect of IGF-1 for RFI is lower than initially expected and the effect of many genes on RFI differ from postweaning and feedlot cattle (Hill, 2012). In addition, as cattle become more physiologically mature, the relationship between IGF-1 and RFI decreased or becames negative. Lancaster et al. (2008) suggested that body composition be considered when examining the relationship between serum IGF-1 concertation and RFI.

Limited research involving RFI has been conducted in *Bos indicus* cattle. Wolcott et al. (2006) examined correlations between serum concentrations of IGF-1 and RFI in *Bos indicus* and *Bos taurus* crosses. Resulting genetic correlations were -0.12 for Brahman cattle and -0.80 for tropical composites. These results would suggest that genotype (i.e. species) would affect the relationship between IGF-1 and RFI.

The relationship between leptin and RFI have also been examined. Leptin is commonly associated with fatness in cattle. Nkrumah et al. (2007) reported a significant phenotypic correlation of 0.31 for serum leptin concentrations and RFI, whereas, Brown et al. (2004) did not results in a phenotypic correlation between leptin and RFI.

2.5. Genetic Parameters of Feed Efficiency and Intake

2.5.1 Heritability

Heritability is defined as the measure of strength of the relationship between breeding values (genetics) and performance and is an indication of what we observe is due to inheritance (Bourdon, 1997). Heritability is expressed as a ratio of variances:

$$h^2 = \frac{\sigma_{BV}^2}{\sigma_P^2}$$

(Eq. 2.3)

where σ_{BV}^2 was the variance of breeding values and σ_P^2 was the phenotypic variance. For selection and genetic improvement, heritability is crucial for polygenetic traits, such as feed efficiency. Increases in heritability leads to an increase in response to selection. The following section is a summarization of heritability estimates for FI and feed efficiency.

2.5.1.1 Heritabilities for Feed Efficiency and Intake

There is a large range of heritabilities reported for feed efficiency (RFI and FCR) and FI which is not surprising considering the differences in analysis that estimated the heritabilities and the various methods used to collect phenotypic measures. Tables 2.1, 2.2 and 2.3 present the range of heritabilities for RFI, FI and FCR for growing and mature cattle, respectively (adapted from Berry and Crowley, 2013). Berry and Pryce (2013) presented an abbreviated version of table 2.1 and 2.2 with pooled heritabilities for all the traits (Table 2.4).

Table 2.1. Heritability estimates (SE) for feed intake (FI), residual feed intake (RFI) and feed conversion ratio (FCR) for a range of studies prior to 2000 on growing animals across different breeds.

FI	RFI	FCR	Breeds[1]	Animals, n	Reference
0.62 (0.12)	0.62 (0.14)	0.42 (0.13)	AN, HE, Polled HE, SH	760	Archer et al. (1997)
0.59 (0.07)	0.44 (0.07)	0.31 (0.09)	AN, HE, SH	966	Arthur et al. (1997)
0.37 (0.19)		0.19 (0.16)	Bonsmara	298	Bergh et al. (1992)
0.06 (0.12)		0.46 (0.20)	FR x HE	327	Bishop et al. (1992)
		0.14 (0.07)	AN	393	Brown et al. (1988)
		0.13 (0.08)	HE	340	Brown et al. (1988)
0.27 (0.15)	0.23 (0.12)	0.35 (0.22)	AN	263	Fan et al. (1995)
0.18 (0.10)	0.07 (0.13)	0.08 (0.09)	HE	271	Fan et al. (1995)
0.58 (na)		0.16 (na)	BB	1,442	Gengler et al. (1995)
0.24 (na)		0.14 (na)	HE	486	Glaze and Schalles (1995)
0.26 (na)		0.27 (na)	HO, BS	630	Jensen et al. (1991)
	0.28 (0.11)		HO, BS	650	Jensen et al. (1992)
0.64 (0.12)	0.28 (0.11)		AN, HE, SH	1,324	Koch et al. (1963)
0.56 (0.11)	0.22 (0.11)	0.18 (0.08)	HO, HO x Dutch	417	Korver et al. (1991)
		0.33 (0.10)	HE	452	Mrode et al. (1990)
0.46 (0.14)	0.19 (0.12)	0.37 (0.14)	HO	360	Van Arendonk et al. (1991)

[1]AN = Angus; BA = Brangus; BB = Belgium Blue; BS = Brown Swiss; CH = Charolais; CHX = Charolais crossbred; FR x HE = Freisian and Hereford crossbreds; HE = Hereford; HO = Holstein; LI = Limousin; SH = Shorthorn; SI = Simmental.

Table 2.2. Heritability estimates (SE) for feed intake (FI), residual feed intake (RFI) and feed conversion ratio (FCR) for a range of studies after 2000 on growing animals across different breeds.

FI	RFI	FCR	Breeds[1]	Animals, n	Reference
0.39 (0.03)	0.39 (0.03)	0.29 (0.04)	AN	1,180	Arthur et al. (2001a)
0.48 (0.04)	0.39 (0.04)	0.46 (0.04)	CH	792	Arthur et al. (2001c)
0.48 (0.06)	0.43 (0.04)	0.31 (0.06)	CH	397	Arthur et al. (2001c)
0.49 (0.15)	0.24 (0.11)		Brahman	700	Barwick et al. (2009)
0.51 (0.14)	0.38 (0.12)		Tropical Composite	787	Barwick et al. (2009)
0.30 (0.08)	0.26 (0.10)	0.30 (0.12)	BA	678	Bouquet et al. (2010)
0.48 (0.14)	0.45 (0.18)	0.23 (0.15)	LI	708	Bouquet et al. (2010)
	0.30 (0.06)		CH sired steers	281	Crews et al. (2003)
	0.26 (0.07)		CH sired steers	274	Crews et al. (2003)
0.49 (0.07)	0.45 (0.07)	0.30 (0.06)	AH, CH, HE, SI, LI	2,605	Crowley et al. (2010)
0.30 (0.15)	0.19 (0.12)	0.07 (0.09)	ANX, CHX	402	Durunna et al. (2011b)
0.43 (0.14)	0.36 (0.13)	0.26 (0.12)	ANX, CHX	419	Durunna et al. (2011b)
0.21 (0.07)	0.14 (0.06)	0.18 (0.07)	AN, BR, BA	1,129	Elzo et al. (2010)
0.31 (0.08)	0.16 (0.08)	0.17 (0.09)	HE	540	Herd and Bishop (2000)
0.34 (0.11)	0.24 (0.11)	0.15 (0.04)	Wagyu	740	Hoque et al. (2006)
0.36 (0.09)	0.49 (0.09)	0.38 (0.07)	Wagyu	514	Hoque et al. (2009)
0.70 (0.11)	0.22 (0.07)	0.11 (0.05)	Wagyu	863	Inoue et al. (2011)
0.48 (0.14)	0.47 (0.13)	0.29 (0.12)	Brangus	468	Lancaster et al. (2009)
0.36 (0.05)			AN	4,215 to 18,169	MacNeil et al. (2011)
	0.18 (0.14)		AN, CH, composite	387	Muijibi et al. (2010)
0.41 (0.12)	0.29 (0.12)		AN, CH, composite	721	Muijibi et al. (2011)
0.54 (0.15)	0.21 (0.12)	0.41 (0.15)	AN, CH, composite	464	Nkumah et al. (2007b)
0.33	0.29	0.14	Wagyu	1,304	Okanishi et al. (2006)
0.27 (0.06)	0.18 (0.06)	0.06 (0.04)	Tropically adapted	1,481	Robinson and Oddy (2004)
0.14	0.14		AN	698	Rolf et al. (2010)
0.40 (0.02)	0.52 (0.14)	0.27 (0.10)	Multibreed	1,141	Rolf et al. (2011)
	0.5		AH, CH, HE, SI, LI	3,724	Savietto et al. (2014)
0.44 (0.06)	0.38 (0.07)	0.37 (0.06)	AN, CH, LI, SI, HE, BA	2,284	Schenkel et al. (2004)
	0.31	0.34	Bonsmara	6,738	Van der Westhuizen et al. (2004)
0.17 (0.10)	0.27 (0.12)		HO	903	Williams et al. (2011)

[1]AN = Angus; BA = Brangus; BB = Belgium Blue; BS = Brown Swiss; CH = Charolais; CHX = Charolais crossbred; FR x HE = Freisian and hereford crossbreds; HE = Hereford; HO = Holstein; LI = Limousin; SH = Shorthorn; SI = Simmental.

Table 2.3. Heritability estimates (SE) for feed intake (FI), residual feed intake (RFI) and feed conversion ratio (FCR) for a lactating or dry mature animals across different breeds adapted from Berry and Crowley (2013).

FI	RFI	FCR	Breeds[2]	Animals, n	Reference
0.28	0.23	0.26	AN, HE, SH	751	Archer et al. (2002)
0.02 (0.02)	0.23 (0.11)	0.18 (0.15)	HE	295	Fan et al. (1996a)
0.11 (0.02)	0.03 (0.01)	0.11 (0.06)	HE, multibreed	1,174	Fan et al. (1996a)
0.03 (0.01)	0.03 (0.02)	0.11 (0.10)	HE	206	Fan et al. (1996b)
0.16 (0.02)	0.22 (0.04)	0.05 (0.01)	HE, multibreed	729	Fan et al. (1996b)
	0.05 to 0.38		Holstein-Friesian	463	Lopez-Vallalobos et al. (2008)
	0.16		HO	247	Ngwerume and mao (1992)
0.18 to 0.63		0.32 (0.13)	HO	402	Spurlock et al. (2012)
	0.04 (0.05)		Norwegian	353	Svendsen et al. (1993)
	0.38		HO	204	Veerkamp et al. (1995)
0.16 (0.09)	0.01 (0.05)	0.14 (0.06)	HO	970	Vallimont et al. (2011)

[1]na = not available

[2]AN = Angus; HE = Hereford; HO = Holstein; SH = Shorthorn.

Table 2.4. Number of studies (N), pooled heritabilities (pooled), minimum (min) and maximum (max) heritability estimates for average daily gain (ADG), weight (WT), dry matter feed intake (FI), residual feed intake (RFI), feed conversion ratio (FCR) adapted from Berry and Pryce (2013).

	ADG	WT	FI	RFI	FCR
N	35	25	37	36	34
Pooled (S.E.)	0.31 (0.01)	0.39 (0.01)	0.40 (0.01)	0.33 (0.01)	0.23 (0.01)
Min	0.06	0.30	0.06	0.07	0.06
Max	0.65	0.88	0.70	0.62	0.46

Minimum heritability estimates for FI, RFI and FCR were 0.02, 0.01 and 0.05, respectively, with respective maximum estimates of 0.70, 0.62 and 0.46. Heritability estimates for mature cows were less than estimates for growing cattle, however, the number of studies was considerably less and indicative of the lack of knowledge for mature cattle (Berry and Crowley, 2013). When studies presented heritabilities for both FI and RFI, the heritabilities for FI were higher than RFI with the exception of two studies (Rolfe et al., 2011 and Williams et al., 2011).

2.5.2 Genetic Correlations

Genetic correlation is defined as the relationship between breeding values of one trait and the breeding values of another trait (Bourdon, 1997). When traits are genetically correlated, selection on one trait can influence the expression of another trait. In addition, performance in one trait can be used to predict the performance of a genetically correlated trait.

In the context of feed efficiency and intake, understanding the relationship between genetically correlated traits is crucial for two reasons. First, there is a lack of knowledge surrounding the effect of selection for feed efficiency on other performance traits, especially survival traits (i.e. health and reproduction). Second, measuring feed intake is expensive and time consuming in addition to a limitation to the availability of facilities to measure intake. The use of an indicator trait could help to mitigate the expense of measurement and enable the ability to measure a greater number of animals. The following section presents genetic correlation estimates reported in literature.

2.5.2.1 Genetic Correlations of Feed Efficiency and Intake

The opportunity to improve production efficiency is dependent on the magnitude of genetic correlations between FI or feed efficiency and other production traits (Herd et al., 2003). In general, estimation of precise genetic correlations of feed efficiency or intake is generally not achievable due to relatively small datasets. In addition, genetic correlations with lowly heritable traits tend to reduce the ability to estimate precise genetic correlations (Berry and Crowley,

2013; Berry and Pryce, 2013). Tables 2.5, 2.6 and 2.7 present genetic correlations of feed efficiency and feed intake to performance, carcass, mature cow and reproductive traits.

Table 2.5. Genetic correlations between feed efficiency traits

[1]Traits	[2]FI	FCR	References
FCR	0.31 (0.07)		Arthur et al. (2001a)
FCR	-0.49 (0.22)		Robinson and Oddy (2004)
FCR	0.64 (0.07)		Arthur et al. (2001b)
FCR	-0.60 (0.18)		Rolfe et al. (2011)
RFI	0.69 (0.03)	0.66 (0.05)	Arthur et al. (2001a)
RFI	0.64 (0.16)	0.70 (0.22)	Herd and Bishop (2000)
RFI	0.43 (0.15)	0.41 (0.32)	Robinson and Oddy (2004)
RFI	0.79 (0.04)	0.85 (0.05)	Arthur et al. (2001b)
RFI	0.66 (0.12)		Rolfe et al. (2011)

[1]FCR = Feed conversion ratio; RFI = Residual feed intake

[2]FI = Feed intake

Table 2.6. Genetic correlations (S.E.) for feed efficiency and postweaning traits.

[1]Trait	[2]FI	[3]FCR	[4]RFI	Reference
ADG	0.54 (0.06)	-0.62 (0.06)	-0.04 (0.08)	Arthur et al. (2001a)
ADG			0.09 (0.29)*	Herd and Bishop (2000)
ADG	0.87 (0.05)	-0.86 (0.10)	0.09 (0.20)*	Robinson and Oddy (2004)
ADG	0.39 (0.08)	-0.46 (0.08)	0.32 (0.10)	Arthur et al. (2001b)
ADG	0.56 (0.16)			Rolfe et al. (2011)
MMWT	0.65 (0.03)	-0.01 (0.07)*	-0.06 (0.06)*	Arthur et al. (2001a)
MMWT			0.22 (0.29)*	Herd and Bishop (2000)
MMWT	0.76 (0.07)	-0.62 (0.18)	-0.20 (0.16)	Robinson and Oddy (2004)
MMWT	0.71 (0.11)			Rolfe et al. (2011)
LWT	0.83 (0.04)	0.24 (0.09)	-0.10 (0.13)*	Arthur et al. (2001b)
200d WT	0.28 (0.15)	-0.21 (0.20)	-0.45 (0.17)	Arthur et al. (2001a)
200d WT			0.34 (0.34)*	Herd and Bishop (2000)
400d WT	0.56 (0.09)	-0.09 (0.15)	-0.26 (0.13)	Arthur et al. (2001a)
400d WT			0.15 (0.28)*	Herd and Bishop (2000)

[1] WT = weight; ADG = Average daily gain; MMWT = metabolic mid-weight ($WT^{0.75}$); LWT = Live weight;

[2]FI = Feed intake

[3]FCR = Feed conversion ratio of feed to gain

[4]RFI = Residual feed intake

* indicates that the estimate is not statistically different from zero as stated by the authors of the respective study.

Table 2.7. Genetic correlations between feed efficiency and carcass traits.

[1]Trait	[2]FI	[3]FCR	[4]RFI	Reference
RIB	0.27 (0.05)	0.03 (0.06)	0.17 (0.05)	Arthur et al. (2001a)
RIB	0.61 (0.11)	0.38 (0.32)	0.48 (0.12)	Robinson and Oddy (2004)
P8	0.14 (0.05)	-0.04 (0.07)*	0.06 (0.06)*	Arthur et al. (2001a)
P8	0.59 (0.10)	0.40 (0.33)	0.72 (0.17)	Robinson and Oddy (2004)
LMA	0.43 (0.07)	-0.12 (0.11)	0.09 (0.09)*	Arthur et al. (2001a)
LMA	0.23 (0.16)	0.20 (0.42)*	-0.24 (0.26)*	Robinson and Oddy (2004)
IMF%	0.39 (0.14)	0.08 (0.28)*	0.22 (0.17)	Robinson and Oddy (2004)
LCC			-0.43 (0.23)	Herd and Biship (2000)
CC	-0.07 (0.16)*	0.12 (0.19)*	-0.06 (0.17)	Crowley et al. (2011)
CF	0.02 (0.21)*	0.15 (0.23)*	0.19 (0.21)*	Crowley et al. (2011)
CWT	0.16 (0.17)*	0.29 (0.19)	-0.01 (0.17)*	Crowley et al. (2011)

[1]RIB = 12/13th rib fat depth; P8 = rump P8 fat depth; LMA = longissimus muscle area; LCC = lean carcass content; IMF% = intramuscular fat; CC = carcass conformation; CF = carcass confirmation; CWT = carcass weight.

[2]FI = Feed intake

[3]FCR = Feed conversion ratio

[4]RFI = Residual feed intake

* indicates that the estimate is not statistically different from zero as stated by the authors of the respective study.

Genetic correlations between RFI and postweaning traits suggest that selection for lower RFI has the potential to reduce feed intake without a reduction in cattle size and therefore improving the efficiency of cattle (Herd et al., 2003). The association between RFI and reproduction traits are in general unfavorable but have large standard errors. Crowley et al. (2011a) found unfavorable genetic correlations for feed efficiency measures with age at first calving (-0.55 $\pm$ 0.14 for FCR and -0.29 $\pm$ 0.14 for RFI) but correlations for calving interval and calving to first service were not different from zero. Basarab et al. (2011) analyzed the significance of calving rates when ultrasound back fat is included or removed from the calculation of RFI. They showed when ultrasound back fat was excluded from the calculation of RFI, a significant difference in calving rates was found for low RFI heifers (72.63%) compared to high RFI heifers

(84.21%). However, when ultrasound back fat was included, there was no difference in the calving rates. When compared to other production traits, few studies have examined the impact of selection for feed efficiency on reproductive performance. Arthur et al. (2005) reported, from a divergent RFI selection experiment, a 5-d later calving date in cows with low RFI compared to high RFI cows. The effect of selection on feed efficiency and the correlated to response to fertility and survival traits merit further research (Berry and Crowley, 2013). However, the association of selection for reduced RFI and forage intake for mature, grazing cattle is unknown (Herd et al., 2003).

Genetic correlations between RFI and carcass traits lacked consistency across studies and resulted in large standard errors. That said, cattle with higher RFI values had improved carcass conformation (Berry and Crowley, 2013). There was a tendency for RFI to be negatively correlated with carcass traits in beef cattle suggesting that as RFI decreased (improves), carcass merit also improved. There was a tendency for FI, FCR and RFI to be positively correlated with body fat in beef cattle, 0.28 ± 0.04, 0.08 ± 0.05 and 0.20 ± 0.04 respectively (pooled estimates; Berry and Crowley, 2013). There are few studies that have examined the genetic correlation between feed efficiency and meat quality. Genetic correlations between both RFI and FCR and meat fatty acid composition were examined in Japanese Black steers. The correlations with RFI were zero and with FCR was -0.38 (C18:0 and C18:1) to 0.43 (C14:1) with standard errors of approximately 0.20 (Inoue et al., 2011; Berry and Crowley, 2013).

2.5.3 Genetic Improvement of Feed Efficiency and Intake

There is an ongoing debate on whether to include feed intake or RFI in a breeding objective (Berry and Pryce, 2013). However, reducing feed intake only should not be the goal of the breeding objective. Selection pressure should also be placed on production or output traits while attempting to reduce feed intake (Nielsen et al., 2013). From the perspective of genetic improvement for feed efficiency, selection for feed efficiency can be accomplished through selection using an index for output (i.e. body weight) and input traits (i.e. FI). Appropriately weighting the index traits for output as positive and input traits as negative, feed efficiency would not have to be explicitly be calculated (MacNeil et al, 2013; Neilsen et al., 2013). Berry and Crowley (2013) stated "achieving improved efficiency of production through genetic selection can be best achieved through a balanced breeding goal, selecting on all traits influencing profitability simultaneously rather than selection on individual traits." Kennedy et al. (1993) showed the equivalence of RFI to the selection index when economic weights were calculated correctly.

Efficiency measures, such as RFI, that are calculated from feed intake and performance traits, provide no additional information beyond the traits used to calculate the efficiency measures (Kennedy et al., 1993). Van der Werf (2004) stated that in growing animals, including all component traits of RFI in a breeding objective or selection index is mathematically equivalent to including RFI assuming no fixed effects were included in the genetic evaluation model. Therefore, there was no additional benefit of including RFI (or any feed efficiency measure) in a breeding goal or selection index that already included the individual feed and production traits. In the context of national cattle evaluation (NCE), it was recommended to analyze feed intake. In order to increase accuracy of feed intake, known genetically correlated traits that were more easily or

cheaply measured (i.e. post weaning weight) were suggested to be included in NCE for feed intake (Neilsen et al., 2013). Selection decision for genetic improvement of feed efficiency should be based on genetic prediction from multi-trait genetic evaluations of feed intake (MacNeil et al., 2013; Neilsen et al., 2013).

Estimating higher accuracies for selection of feed efficiency and intake has been a substantial obstacle for the implementation of feed efficiency traits in breeding objectives (Berry and Pryce, 2014). Given the expense of measuring feed intake on cattle, producers tend to test their more elite animals. As a result, databases of recorded FI may not be truly representative of the cattle population which has led to use of densely recorded traits, such as weaning weight, in multi-trait analyses for the publication of EPD.

2.6 Conclusion

There is a large body of research on the subject of feed intake and feed efficiency; however there are gaps in knowledge pertaining to feed intake and efficiency. There are few studies examining feed efficiency and the correlation to cow performance or health traits. The robustness of feed reduction and how it affects feed efficient cattle is largely unknown. In addition, the relationship of feed intake and grazing intake is unknown. In order to select for a cowherd that is more efficient for forage utilization, grazing intake needs to be measured. Currently there is no technology to measure feed intake on a population of grazing cattle and the relationship between feed intake measured in a feedlot and intake from grazing cattle are unknown.

LITERATURE CITED

Adams, D.C., R.T. Clark, T.J. Klopfenstein, and J.D. Volesky. 1996. Matching the cow with forage resources. Rangelands. 18: 57-62.

Agricultural Research Council. 1980. The nutrient requirements of ruminant livestock. Technical Review. Farnham Royal, U.K.: Commonwealth Ag. Bureaux.

Alberta Agriculture and Forestry. 2006. Agri-Facts: Residual feed intake (net feed efficiency) in beef cattle. https://open.alberta.ca/dataset/91a77dec-f0a4-49c2-8c54-f172fe568e2c/resource/721e982c-b90f-4605-9de0-a3b8bb312b1f/download/2006-420-11-1.pdf. (Accessed April 30, 2017).

Anderson, R. V., R. J. Rasby, T. J. Klopfenstein, and R. T. Clark. 2005. An evaluation of production and economic efficiency of two beef systems from calving to slaughter. J. Anim. Sci. 83: 694-704.

Archer, J. A., P. F. Arthur, R. M. Herd, P. F. Parnell, and W. S. Pitchford. 1997. Optimum postweaning test for measurement of growing rate, feed intake, and feed efficiency in British breed cattle. J. Anim. Sci. 75: 2024-2032.

Archer, J. A. and L. Bergh. 1999. Duration of performance tests for growth rate, feed intake and feed efficiency in four biological types of beef cattle. Livest. Prod. Sci. 65: 47-55.

Arthur, P. F., J. A. Archer, R. M. Herd, E. C. Richardson, S. C. Exton, J. H. Wright, K. C. P. Dibley, and D. A. Burton. 1997. Genetic and phenotypic variation in feed intake, feed efficiency and growth in beef cattle. Proc. 12th Assoc. Adv. Anim. Breed. Genet. 12:234–237.

Arthur, P. F., J. A. Archer, D. J. Johnson, R. M. Herd, E. C. Richardson, and P. F. Parnell. 2001a. Genetic and phenotypic variance and covariance components for feed intake, feed efficiency and other postweaning traits in Angus cattle. J. Anim. Sci. 79:2805–2811

Arthur, P. R., J. A. Archer, R. M. Herd, and G. J. Melville. 2001b. Response to selection for net feed intake in beef cattle. Proc. Assoc. Advmt. Anim. Breed Genet. 14: 135-138.

Arthur, P. F., R. M. Herd, D. J. Johnson, K. Moore, and H. U. Graser. 2004. Incorporating information on plasma IGF-I concentration in genetic improvement strategies for feed efficiency in beef cattle. In: Canadian Society of Animal Science Annual Meeting (Abstr.). p. 31.

Arthur, P. F., R. M. Herd, J. F. Wilkins and J. A. Archer. 2005. Maternal productivity of Angus cows divergently selected for post-weaning residual feed intake. Anim. Prod. Sci. 45:985-993.

Arthur, P. F., J. A. Archer, R. M. Herd and G. J. Melville. 2011. Response to selection for net feed intake in beef cattle. Proc. Assoc. Advmt. Anim. Breed. Genet. 14:135-138.

Arthur, P. R., J. E. Pryce, and R. M. Herd. 2014. Lessons learnt from 25 years of feed efficiency research in Australia. Proc. 10th World Congr. Genet. Appl. to Livest. Prod., Vancouver, Canada.

Barwick, S. A., M. L. Wolcott, D. J. Johnston, H. M. Burrow and M. T. Sullivan. 2009. Genetics of steer daily and residual feed intake in two tropical beef genotypes, and relationships among intake, body composition, growth and other post-weaning mea- sures. Anim. Prod. Sci. 49:351–366.

Basarab, J. A., M. G. Colazo, D. J. Ambrose, S. Novak, D. McCartney, and V. S. Baron. 2011. Residual feed intake adjusted for backfat thickness and feed frequency is independent of fertility in beef heifers. Can. J. Anim. Sci. 91:573–584.

Bergh, L., M. M. Scholtz, and G. J. Erasmus. 1992. The identifica- tion and assessment of the best animals. The Kleiber ratio as a selection criterion for beef cattle. Proc. 7th Aust. Assoc. Anim. Breed. Genet. 10:338–340.

Berry, D. P. and J. J. Crowley. 2013. Cell Biology Symposium: Genetics of feed efficiency in dairy and beer cattle. J. Anim. Sci. 91: 1594-1613.

Berry, D. P. and J. E. Pryce. 2013. Feed efficiency in growing and mature animals. Proc. 10th World Congr. Genet. Appl. to Livest. Prod., Vancouver, Canada.

BIF. 2016. Guidelines for uniform beef improvement programs. 9th ed. Beef Improvement Federation. p. 25-28.

Bishop, S. C., J. S. Broadbent, R. M. Kay. I. Rigby, and A. V. Fisher. 1992. The performance of Hereford × Friesian offspring of bulls selected for lean growth rate and lean food conversion efficiency. Anim. Prod. 54:23–30.

Bottje, W. G. and G. E. Carstens. 2009. Association of mitochondrial function and feed efficiency in poultry and livestock species. J. Anim. Sci. 87: E48-63.

Bottje, W. G. and B. W. Kong. 2013. Cell Biology Symposium: Mitochondrial function to global gene expression. J. Anim. Sci. 91: 1582-1593.

Bouquet, A., M.-N. Fouilloux, G. Renand, and F. Phocas. 2010. Genetic parameters for growth, muscularity, feed efficiency and carcass traits of young beef bulls. Livest. Sci. 129:38–48.

Bourdon, R.M. 1997. Understanding Animal Breeding. Prentice-Hall. p 185-239.

Brown, A. H., Jr., Z. B. Johnson, J. J. Chewning, and C. J. Brown. 1988. Relationships among absolute growth rate, relative growth rate and feed conversion during postweaning feedlot performance tests. J. Anim. Sci. 66:2524–2529.

Brown, E. G., G. E. Carstens, J. T. Fox and K. O. Jr. Curley, T. M. Bryan. L J. Slay, T. H. Welsh, Jr., R. D. Randel, J. W. Holloway and D. H. Keisler, 2004. Physiological indicators of performance and efficiency of gain traits in growing steers and bulls. Beef Cattle Research in Texas. Texas A&M University, Agricultural Experiment Station.

Bunter, K. L., W. Cai, D. J. Johnston and J. C. M. Dekkers. 2010. Selection to reduce residual feed intake in pigs proudces a correlated response in juvenile insulin-like growth factor-I concentration. J. Anim. Sci. 88:1973-1981.

Burns, J. C., K. R. Pond, and D. S. Fisher. 1994. Measurement of forage intake. Page 494-514 in Forage quality, evaluation, and utilization. G. C. Fahey, M. Collins, D. R. Mertens, and L. E. Moser, ed. ASA, CSSA, SSSA, Madison, WI.

Capper, J. L. 2011. The environmental impact of beef production in the United States: 1977 compared with 2007. J. Anim. Sci. 89: 4249-4261.

Carstens, G. E., C. M. Theis, M. B. White, T. H. Welsh, Jr., B. G. Warrington, R. D. Randel, T. D. A. Forbes, H. Lippke, L. W. Greene, and D. K. Lunt. 2002. Residual feed intake in beef steers: I. Correlations with performance traits and ultrasound measures of body composition. Proc. West. Sec. Meet. Am. Soc. Anim. Sci. 53:552-555.

Chapinal, N., D. M. Veira, D. M. Weary, and M. A. G. von Keyserlingk. 2007. Technical Note: Validation of a system for monitoring individual feeding and drinking behavior and intake on group-housed cattle. J. Dairy Sci. 90:5732-5736.

Cole, N. A. 1995. Intake control systems. In: F. N. Owens, D. R. Gill, K. S. Lusby, and F. T. McCollum (Ed.) Symposium: Intake by Feedlot Cattle. Okla. Agric. Exp. Sta. P-942. Okla. State Univ., Stillwater. p 156-161.

Cottle, D. J. 2011. Use of residual feed intake as an indirect selection trait for reduction of methane emission on grazing beef cattle. Proc. Assoc. Advmt. Anim. Breed Genet. 19: 423-425.

Crews, D. H., Jr., N. H. Shannon, B. M. A. Genswein, R. E. Crews, C. M. Johnson, and B. A. Kendrick. 2003. Genetic parameters for net feed efficiency of beef cattle measured during postweaning growing versus finishing periods. Proc. West. Sect. Am. Soc. Anim. Sci. 54:125–128.

Crowley, J. J., M. McGee, D. A. Kenny, D. H. Crews Jr., R. D. Evans, and D. P. Berry. 2010. Phenotypic and genetic parameters for different measures of feed efficiency in different breeds of Irish performance tested beef bulls. J. Anim. Sci. 88:885–894.

Crowley, J. J., R. D. Evans, N. Mc Hugh, D. A. Kenny, M. McGee, D. H. Crews Jr., and D. P. Berry. 2011a. Genetic relationships between feed efficiency in growing males and beef cow performance. J. Anim. Sci. 89:3372–3381.

Crowley, J. J., R. D. Evans, N. Mc Hugh, T. Pabiou, D. A. Kenny, M. McGee, D. H. Crews Jr., and D. P. Berry. 2011b. Genetic associations between feed efficiency measured in a performance test station and performance of growing cattle in commercial beef herds. J. Anim. Sci. 89:3382–3393.

Cruz, G. D., J. B. Trovo, J. W. Oltjen, and R. D. Sainz. 2011. Estimating feed efficiency: Evaluation of mathematical models to predict individual intakes of steers fed in group pens. J. Anim. Sci. 89: 1640-1649.

Culbertson, M. M., S. E. Speidel, R. K. Peel, R. R. Cockrum, M. G. Thomas, and R. M. Enns. 2015. Optimum measurement period for evaluating feed intake traits in beef cattle. J. Anim. Sci. 93: 2482-2487.

Dahlke, G., D. R. Strohbehn, C. Ingle, and P. Beedle. 2008. A feed intake monitoring system for cattle. ISU Anim. Industry Report: AS 654. http://lib.dr.iastate.edu/ans_air/vol654/iss1/28.

De Haer, L. C. M., and J. W. M. Mercks. 1992. Patterns of daily food intake in growing pigs. Anim. Prod. 54: 95-104.

Donoghue, K. A., P. F. Arthur, J. F. Wilkins and R. M. Herd. 2010. Onset of puberty and early-life reproduction in Angus females divergently selected for post-weaning residual feed intake. Anim. Prod. Sci. 51:183-190.

Dove, H. and R. W. Mayes. 1991. The use of plant wax alkanes as marker substances in studies of the nutrition of herbivores. Aust. J. Agric. Res. 42:913-952.

Durunna, O. N., M. G. Colazo, D. J. Ambrose, D. McCartney, V. S. Baron, and J. A. Basarab. 2012. Evidence of residual feed in- take reranking in crossbred replacement heifers. J. Anim. Sci. 90:734–741.

Durunna, O. N., G. Plastow, F. D. N. Mujibi, J. Grant, J. Mah, J. A. Basarab, E. K. Okine, S. S. Moore, and Z. Wang. 2011. Genetic parameters and genotype × environment interaction for feed efficiency traits in steers fed grower and finisher diets. J. Anim. Sci. 89:158–167.

Eisen, E. J. 1997. Restricted selection index: an approach to selecting for feed efficiency. J. Anim. Sci. 44:958-972.

Elzo, M. A., D. D. Johnson, G. C. Lamb, T. D. Maddock, R. O. Myer, D. G. Riley, G. R. Hansen, J. G. Wasdin, and J. D. Driver. 2010. Heritabilities and genetic correlations between postweaning feed intake, growth, and ultrasound traits in a multibreed Angus-Brahman cattle population in the subtropics. In: Proc. 9th World Congr. Genet. Appl. Livest. Prod., Leipzig, Germany. p. 203–207.

Fan, L. Q., D. R. C. Bailey, and N. H. Shannon. 1995. Genetic pa- rameter estimation of postweaning gain, feed intake and feed efficiency for Hereford and Angus bulls fed two different diets. J. Anim. Sci. 73:365–372.

FAO. 2009. How to feed the world in 2050. Food and Agriculture Organization of the United Nations, Rome, Italy.

Ferrell, C. L. and T. G. Jenkins. 1985. Cow type and the nutritional environment: Nutritional aspects. J. Anim. Sci. 3: 725-741.

Fox, D. G., C. J. Sniffen, and J. D. O'Connor. 1988. Adjusting nutrient requirements of beef cattle for animal and environmental variations. J. Anim. Sci. 66:1475–1495.

Gengler, N., C. Seutin, F. Boonen, and L. D. Van Vleck. 1995. Estimation of genetic parameters for growth, feed consumption, and conformation traits for double-muscled Belgian blue bulls performance-tested in Belgium. J. Anim. Sci. 73:3269–3273.

Glaze, J. B., Jr., and R. R. Schalles. 1995. Selection response and genetic parameter estimation for feed intake, gain and feed conversion. http://krex.k-state.edu/dspace/handle/2097/4876. (Accessed December 21, 2017).

Gordon, I. J. 1995. Animal-based techniques for grazing ecology research. Small Ruminant Res. 16:203-214.

Guiroy, P. J., D. G. Fox, L. O. Tedeshchi, M. J. Baker, and M. D. Cravey. 2001. Predicting individual feed requirements of cattle fed in groups. J. Anim. Sci. 79: 1983-1995.

Gunsett, F. C. 1986. Problems associated with selection for traits defined as a ratio of tow component traits. Proc. 3rd World Congr. Genet. Appl. Livest. Prod. XI:435.

Herd, R. M., and S. C. Bishop. 2000. Genetic variation in residual feed intake and its association with other production traits in British Hereford cattle. Livest. Prod. Sci. 63:111–119.

Herd, R. M., V. H. Oddy, and E. C. Richardson. 2004. Biological basis for variation in residual feed intake in beef cattle. 1. Review of potential mechanisms. Aust. J. Exp. Agric. 44:423-430.

Herd, R. M. and P. F. Arthur. 2009. Physiological basis for residual feed intake. J. Anim. Sci. 87:E64-E71.

Herd, R. M., P. F. Arthur, K. A. Donoghue, S. H. Bird, T. Bird-Gardiner and R. S. Hegarty. 2014. Measures of methane production and their phenotypic relationships with dry matter intake, growth, and body composition traits in beef cattle. J. Anim. Sci. 92: 5267-5274.

Hermesch, S., K. L. Bunter and G. Luxford. 2001. Estimates of genetic correlations between IGF-1 recorded at 4 weeks of age and individual piglet weights at birth and 14 days, along with lifetime growth rate and backfat. Proc. 14th Conf. Assoc. for Adv. Anim. Breed. Gen, Queenstown, New Zealand. 227-230.

Hill, R. A. 2012. Feed Efficiency in the Beef Industry. Wiley-Blackwell, Oxford, UK. p. 7-18, 21-28, 175-198, 225-274.

Hoque, M. A., P. F. Arthur, K. Hiramoto, and T. Oikawa. 2006. Genetic relationship between different measures of feed ef- ficiency and its component traits in Japanese Black (Wagyu) bulls. Livest. Sci. 99:111–118.

Hoque, M. A., M. Hosono, T. Oikawa, and K. Suzuki. 2009. Genetic parameters for measures of energetic efficiency of bulls and their relationships with carcass traits of field progeny in Japanese Black cattle. J. Anim. Sci. 87:99–106.

Ingvartsen, K. L., H. R. Andersen, and J. Foldager. 1992. Effect of sex and pregnancy on feed intake capacity of growing cattle. Acta. Agr. Scand. (Sect. A). 42:40-46.

Inoue, K., M. Kobayashi, N. Shoji, and K. Kato. 2011. Genetic parameters for fatty acid composition and feed efficiency traits in Japanese Black cattle. Animal 5:987–994.

Jensen, J., I. L. Mao, B. B. Andersen, and P. Madsen. 1991. Genetic parameters of growth, feed intake, feed conversion and carcass composition of dual-purpose bulls in performance testing. J. Anim. Sci. 69:931–939.

Jensen, J., I. L. Mao, B. B. Andersen, and P. Madsen. 1992. Phenotypic and genetic relationships between residual energy intake and growth, feed intake, and carcass traits of young bulls. J. Anim. Sci. 70:386–395.

Jones, J. I. and D. R. Clemmons. 1995. Insulin-like growth factors and their binding proteins: biological actions. Endocr. Rev. 16:3-34.

Kennedy, B. W., J. H. J. van der Werf and T. H. E. Meuwissen. 1993. Genetic and statistical properties of residual feed intake. J. Anim. Sci. 71:3239-3250.

Koch, R. M., L. A. Swiger, D. Chambers and K. E. Gregory. 1963. Efficiency of feed use in beef cattle. J. Anim. Sci. 22:486-494.

Korver, S., E. A. M. van Eekelen, H. Vos, G. J. Nieuwhof, and J. A. M. van Arendonk. 1991. Genetic parameters for feed intake and feed efficiency on growing dairy heifers. Livest. Prod. Sci. 29:49–59.

Lancaster, P. A., G. E. Casterns, F. R. B. Ribeiro, M. E. Davis, J. G. Lyons and T. H. Welsh, Jr. 2008. Effects of divergent selection for serum insulin-like growth facto-I concentration on performance, feed efficiency and ultrasound measures of carcass composition traits in Angus bulls and heifers. J. Anim. Sci. 86:2862-2871.

Lancaster, P. A., G. E. Carstens, D. H. Crews Jr., T. H. Welsh Jr., T. D. A. Forbes, D. W. Forrest, L. O. Tedeschi, R. D. Randel, and F. M. Rouquette. 2009. Phenotypic and genetic relationships of residual feed intake with performance and ultrasound carcass traits in Brangus heifers. J. Anim. Sci. 87:3887–3896.

Luxford, B. G., K. L. Bunter, P. C. Owens and R. G. Campbel. 1998. Genetic relationship between insulin-like growth factor-1 and pig performance. J. Anim. Sci. 44:189-194.

MacNeil, M. D., N. Lopez-Villalobos, and S. L. Northcutt. 2011. A prototype national cattle evaluation for feed intake and efficien- cy of Angus cattle. J. Anim. Sci. 89:3917–3923.

MacNeil, M. D., M. M. Scholtz, and A. Maiwashe. 2013. Estimates of variance components for postweaning feed intake and growth in Bonsmara bulls and evaluation of alternative measures of feed efficiency. S. Afr. J. Anim. Sci. 43:18–24.

McCarthy, K. L., M. Undi, and C. R. Dahlan. 2018. Mineral intake, feeding behavior, and growth performance of cow-calf pairs grazing native range. J. Anim. Sci. 96:S83-S84.

Meyer, A. M., M. S. Kerley, and R. L. Kaleenbach. 2008. The effect of residual feed intake classification on forage intake by grazing beef cows. J. Anim. Sci. 86: 2670-2679.

Moore, K. L., D. J. Johnson, H. U. Graser, and R. Herd. 2005. Genetic and phenotypic relationships between insulin-like growth factor-I (IGF-I) and net feed intake, fat, and growth traits in Angus beef cattle. Aust. J. Agric. Res. 56:211-218.

Mrode, R. A., C. Smith, and R. Thompson. 1990. Selection for rate and efficiency of lean gain in Hereford cattle. 1. Selection ap- plied and direct responses. Anim. Prod. 51:23–34.

Mujibi, F. D. N., S. S. Moore, D. J. Nkrumah, Z. Wang and J. A. Basarab. 2010. Season of testing and its effect on feed intake and efficiency in growing beef cattle. J. Anim. Sci. 88:3789- 3799.

Mujibi, F. D. N. , J. D. Nkrumah, O. N. Durunna, P. Stothard, J. Mah, Z. Wang, J. Basarab, G. Plastow, D. H. Crews, Jr. and S. S. Moore. 2011. Accuracy of genomic breeding values for residual feed intake in crossbred beef cattle. J. Anim. Sci. 89:3353-3361.

Nielsen, B. L., A. B. Lawrence, and C. T. Whittemore, 1995. Effect of individual housing on the feeding behavior of previously group housed pigs. Appl. Anim. Behav. Sci. 47: 149-161.

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

Nkrumah, J. D., J. A. Basarab, Z. Wang, C. Li, M. A. Price, E. K. Okine, D. H. Crews, Jr., and S. S. Moore. 2007. Genetic and phenotypic relationships of feed intake and measures of efficiency with growth and carcass merit of beef cattle. J. Anim. Sci. 85:2711-2720.

Nobel Research Institute. 2017. https://www.noble.org/news/publications/ag-news-and-views/2017/january/use-exclusion-cages-to-manage-stocking-rates/ (Accessed January 2, 2018).

NRC. 1984. Nutrient requirements of beef cattle. 6th rev. ed. Natl. Acad. Press, Washington, DC.

NRC. 1987. Predicting Feed Intake of Food-Producing Animals. Natl. Acad. Press, Washington, D. C.

NRC. 2000. Nutrient requirements of beef cattle. 7th rev. ed. Natl. Acad. Press, Washington, DC.

Ojano-Dirain, C., N. B. Tinsley, T. Wing, M. Cooper, and W. G. Bottje. 2007. Membrane potential and H_2O_2 production in duodenal mitochondria from broiler chickesn (*Gallus gallus domesticus*) with low and high feed efficiency. Comp. Biochem. Phys. (A) 147: 934-941.

Okanishi, T., S. Masayuki, K. Tomohiro, O. Kenji, and F. Mukai. 2008. Genetic analysis of residual feed intakes and other performance test traits of Japanese Black cattle from revised protocol. Anim. Sci. J. 79:291–296.

Retallick, K. J., J. M. Bormann, R. L. Weber and M. D. MacNeil. 2015. Genetic variance and covariance components for feed intake, average daily gain, and postweaning gain in growing beef cattle. Kansas Ag. Exper. Station Research Reports. Vol 1: Article 9. http://newprairiepress.org/cgi/viewcontent.cgi?article=1023&context=kaesrr.

Richardson, E. C., R. M. Herd, V. H. Oddy, J. M. Thompson, J. A. Archer, and P. F. Arthur. 2001. Body composition and implications for heat production of Angus steer progeny of parents selected for and against residual feed intake. Aust. J. Exp. Agric. 41: 1065-1072.

Richardson, E. C. and R. M. Herd. 2004. Biological basis for variation in residual feed intake. 2. Synthesis of results following divergent selection. Aust. J. Exp. Agric. 44:431-440.

Richardson, E. C., R. M. Herd, J. A. Archer, and P. F. Arthur. 2004. Metabolic differences in Angus steers divergently selected for residual feed intake. Aust. J. Exp. Agric. 44:441-452.

Rolf, M. M., J. F. Taylor, R. D. Schnabel, S. D. McKay, M. C. McClure, S. L. Northcutt, M. S. Kerley and R. L. Weaber. 2010. Impact of reduced marker set estimation of genomic relationship matrices on genomic selection for feed efficiency in Angus cattle. BMC Genetics. 11:24.

Rolfe, K. M., W. M. Snelling, M. K. Nielsen, H. C. Freetly, C. L. Ferrell, and T. G. Jenkins. 2011. Genetic and phenotypic parame- ter estimates for feed intake and other traits in growing beef cattle, and opportunities for selection. J. Anim. Sci. 89:3452–3459.

Schenkel, F. S., S. P. Miller, and J. W. Wilton. 2004. Genetic parameters and breed differences for feed efficiency, growth and body com- position traits of young beef bulls. Can. J. Anim. Sci. 84:177–185.

Stock, R. A. and T. J. Klopfenstein. 1986. Calan gates for intake measurement. Proc. 1986 Symposium on Feed Intake of Cattle. 321-324.

Titgemeyer, E. C., C. K. Armendariz, D. J. Bindel, R. H. Greenwood and C. A. Loest. 2001. Evaluation of titanium dioxide as a digestibility marker for cattle. J. Anim. Sci. 79: 1059-1063.

Undi, M., C. Wilson, K. H. Ominski and K. M. Wittenburg. 2008. Comparision of techniques for estimation of forage dry matter intake by grazing beef cattle. Can. J. Anim. Sci. 88:693-701.

Van Arendonk, J. A. M., G. J. Nieuwhof, H. Vos, and S. Korver. 1991. Genetic aspects of feed intake and efficiency in lactating dairy heifers. Livest. Prod. Sci. 29:263–275.

Van der Werf, J. H. J. 2004. Is it useful to define residual feed intake as a trait in animal breeding programs? Aust. J. Exp. Agric. 44:405–409.

Van der Westhuizen, R. R., J. Van der Westhuizen, and S. J. Schoeman. 2004. Genetic variance components for residual feed intake and feed conversion ratio and their correlations with other production traits in beef bulls. South Afr. J. Anim. Sci. 54:257–264.

Wang, Z., J. D. Nkrumah, C. Li, J. A. Basarab, L. A. Goonewardene, E. K. Okine, D. H. Crews, Jr. and S. S. Moore. 2006. Test duration for growth, feed intake, and feed efficiency in beef cattle using the GrowSafe System. J. Anim. Sci. 84: 2289-2298.

Wang, Z., M. G. Colazo, J. A. Basaarab, L. A. Goonewardene, D. J. Abrose, E. Marques, G. Plastow, S. P. Miller and S. S. Moore. 2012. Impact of selection for residual feed intake on breeding soundness and reproductive performance of bulls on pasture-based multisire mating. J. Anim. Sci. 90:2963-2696.

Whisnant, C. USDA-NIFA-CRIS Research. http://cris.csrees.usda.gov/cgi-bin/starfinder/6244/crisassist.txt (Accessed April 30, 2017)

Williams, Y. J., J. E. Pryce, C. Grainger, W. J. Wales, N. Linden, M. Porker, and B. J. Hayes. 2011. Variation in residual feed intake in Holstein Friesian dairy heifers in southern Australia. J. Dairy Sci. 94:4715–4726.

Wolcott, M. L., D. J. Johnston, S. A. Barwick and H. M. Burrow. 2006. Genetic correlations of steer growth, fatness and IGF-1 with feed intake and efficiency in tow tropically adapted genotypes. Proc. 9th World Congr. Genet. Appl. to Livest. Prod., Belo Horizonte, Brazil.

Wolfger, B., E. Timsti. E. A. Pajor, N. Cook, H. W. Barkema, and K. Orel. 2015. Technical note: Accuracy of an ear tag-attached accelerometer to monitor rumination and feeding behavior in feedlot cattle. J. Anim. Sci. 93: 3164-3168.

Wyfells, S. A., A. R. Williams, C. T. Parsons, J. M. Dafoe, D. L. Boss, T. DelCurto, N. G. Davis and J. P. Bowman. 2018. The influence of age and environmental conditions on supplement intake and behavior of winter grazing cattle on mixed-grass rangeland. Transl. Anim. Sci. 2:S89-S92.

Chapter III

ESTIMATION OF VARIANCE COMPONENTS DUE TO DIRECT AND MATERNAL EFFECTS FOR FEED INTAKE FOR RED ANGUS AND GELBVIEH CATTLE

3.1 Introduction

Feed costs are the largest variable expense in the beef industry, accounting for 50 to 70% of total production costs (Anderson et al., 2005). Genetic improvement of feed utilization has the potential to decrease production costs and increase profitability. As a result, interest in selecting for cattle that are more feed efficient has increased, and several beef cattle breed associations are currently publishing expected progeny differences (EPD) for feed intake or for a feed efficiency trait, such as residual feed intake (RFI) and residual average daily gain (RADG). The accurate calculations of these EPD are dependent on reliable estimation of variance components for the traits of interest.

Through the environment provided by the dam, maternal effects are known to influence performance of some traits, such as weaning weight. The phenotype of progeny is influenced by the dam's contribution of half of direct, additive genes for that trait, in addition to her genetic potential for the environment she provides for her off-spring. As a result of the environmental influence of the dam, the maternal effect can be an important source of variation. For traits that are influenced by maternal effects, the exclusion of these effects when estimating variance components can substantially inflate direct heritability estimates (Meyer, 1992). Studies have shown that maternal effects have an important influence on traits before and after weaning for up

to a year (Mavrogenis et al., 1978; Meyer, 1992; Lee et al., 2000), but the magnitude of the influences tends to dissipate for traits measured after weaning (Meyer, 1992). Beef Improvement Federation (BIF) guidelines for the measurement of DMI proposed that feed intake be measured on cattle postweaning and no older than 390 d at the conclusion of data collection (BIF, 2010). To date, few studies have examined maternal effects on DMI and the significance of its inclusion for the estimation of variance components should be explored as maternal effects are not currently included in genetic evaluations for feed intake. In addition, some beef cattle breed associations are incorporating weaning weight in a multivariate estimation for DMI EPD as a method to account for selection bias for animals with DMI records, but little is known for the relationship between DMI and weaning weight maternal. If a substantial correlation between DMI and weaning weight maternal exists, the genetic correlation between DMI and weaning weight direct could be inflated if maternal effects are not included. Therefore, the objectives of this study were to 1) estimate the variance components for direct and maternal effects on feed intake and 2) to estimate the genetic correlation between DMI and weaning weight maternal on two populations of cattle, Red Angus and Gelbvieh.

3.2 Materials and Methods

Data. The data used was provided from preexisting databases used for genetic evaluations. Therefore, animal care and use committee approval was not needed for this study.

To address the first objective, estimates of variance components were obtained from a single trait animal model for dry matter feed intake (DMI) where direct and maternal effects were included. The second objective was met using a multivariate model with DMI as the first

trait and weaning weight included as the second with maternal genetic and permanent environmental effects included for weaning weight only.

Two data sets containing records for dry matter feed intake (DMI), weaning weight (WWT), and pedigree information were provided by the American Gelbvieh Association (AGA) and the Red Angus Association of American (RAAA). The weaning weight data provided were not pre-adjusted for age and DMI records were available for both males and females. The contemporary group (CG) for DMI established by both AGA and RAAA were formed using sex, pen, feed trial designation, trial length and year. Age at the start of test (AGE) was calculated by subtracting the animal's birth date from the start date of the test. At the beginning of a test, animals younger than 240 d (BIF, 2016) and older than 365 d were removed from the analysis to restrict the data to postweaning cattle less than a year of age. Considering maternal effects have been shown to influence postweaning traits up to a year of age, the data was truncated to only include postweaning cattle up to a year of age to focus on the age category where maternal effects are most likely to be found. For traits measured on animals older than a year of age, the maternal effects are often considered to be negligible (Meyer, 1992) and were therefore removed. For the purpose of this study, embryo transfer calves were also removed since pedigree and breed information for surrogate dams were unknown and therefore the ability to account for the maternal environment provided by the surrogate dams was not possible (Shaeffer and Kennedy, 1989). For the single trait analysis of DMI, Gelbvieh had a final data set which consisted of 3,021 animals with DMI records in 95 unique CG with an average size of 32 and Red Angus data consisted of 3,213 animals in 104 unique CG with the average size of 31. Starting with animals with records, a 3-generation pedigree was formed. The pedigree for Gelbvieh included 15,418 animals with 3,027 unique sires and 9,494 unique dams. For Red

Angus, the pedigree was comprised of 13,747 animals with 2,476 and 8,117 unique sire and dams, respectively.

A multi-trait analysis was also conducted to examine the potential genetic correlation of DMI to maternal weaning weight. Due to limited capacity of testing facilities to measure DMI and the relative expense of collecting intake data, only select animals were measured. To account for this selection bias, multi-trait models including DMI and correlated traits, such as weaning weight, are often used. This was accomplished by including all the animals recorded in the correlated trait's CG from which animals with DMI records were also recorded (Neilsen et al., 2013). Weaning weight CG was defined as weaning date, breeder, herd, and sex. To ensure complete CG for both DMI and WWT, all animals within the WWT CG for animals with DMI records were included. As a result, animals were included with no DMI records, but their WWT records were included due to being in their weaning weight contemporary group of animals with DMI observations. As with the DMI records, embryo transfer calves were removed. Following Beef Improvement Federation Guidelines (BIF, 2016), age of dam (AOD) were assigned for animals with observations from dams aged 2, 3, 4, 10, 11 and 12 years old as their respective age in years. Dams 5 to 9 years of age received an AOD classification of 5 years of age and dams greater than or equal to 13 years of age were classified as 13. Age at weaning was determined by subtracting the animal's birth date from the weaning date. The final data set for the multi-trait analysis consisted of 7,792 and 4,342 records for Gelbvieh and Red Angus, respectively. A 3-generation pedigree was built for Gelbvieh with 26,412 animals and 16,676 animals for Red Angus. For weaning weight, Gelbvieh had 525 unique contemporary groups for weaning weight with an average size of 15 animals and Red Angus had 284 unique contemporary groups, also averaging in size of 15 animals.

Statistical Analysis. Data for Gelbvieh and Red Angus were analyzed separately. Genetic and residual (co)variance parameters were estimated using ASREML 3.0 software package (Gilmour et al., 2009). For the single trait analysis, a linear animal mixed model including both genetic and environmental effects was performed using the following model:

$$\boldsymbol{y} = \boldsymbol{Xb} + \boldsymbol{Z_d u_d} + \boldsymbol{Z_m u_m} + \boldsymbol{Z_{pe} u_{pe}} + \boldsymbol{e}$$

(Eq. 3.1)

where $\boldsymbol{y}$ was the vector of observations for DMI, $\boldsymbol{b}$ was a vector of unknown fixed effect solutions, $\boldsymbol{u_d}$ was a vector of additive genetic effects, $\boldsymbol{u_m}$ was a vector of maternal genetic effects, $\boldsymbol{u_{pe}}$ was a vector of permanent environmental effects of the dam, $\boldsymbol{e}$ was a vector of residual effects, and $\boldsymbol{X}$, $\boldsymbol{Z_d}$, $\boldsymbol{Z_m}$, $\boldsymbol{Z_{pe}}$ were known incidence matrices relating effects in $\boldsymbol{b}$, $\boldsymbol{u_d}$, $\boldsymbol{u_m}$, and $\boldsymbol{u_{pe}}$ to observations in y, respectively. Fixed effects for both Gelbvieh and Red Angus were AGE as a covariate and DMI CG.

Variances and means included in the model were assumed to be the following:

$$E(\boldsymbol{y}) = \boldsymbol{Xb}$$

$$E(\boldsymbol{e}) = E(\boldsymbol{u}) = 0$$

$$var\begin{bmatrix} \boldsymbol{u_d} \\ \boldsymbol{u_m} \\ \boldsymbol{u_{pe}} \\ \boldsymbol{e} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}\sigma_d^2 & \boldsymbol{A}\sigma_{dm} & 0 & 0 \\ \boldsymbol{A}\sigma_{dm} & \boldsymbol{A}\sigma_m^2 & 0 & 0 \\ 0 & 0 & \boldsymbol{I_{pe}}\sigma_{pe}^2 & 0 \\ 0 & 0 & 0 & \boldsymbol{I_e}\sigma_e^2 \end{bmatrix}$$

(Eq. 3.2)

where $\boldsymbol{A}$ was the numerator relationship matrix, $\boldsymbol{I_{pe}}$ was an identity matrix with an order of the number dams, $\boldsymbol{I_e}$ was an identity matrix with an order of the number of animals with

observations, $\boldsymbol{\sigma_d^2}$ was the direct additive genetic variance, $\boldsymbol{\sigma_{dm}}$ was the direct-maternal genetic covariance, $\boldsymbol{\sigma_m^2}$ was the maternal additive genetic variance, $\boldsymbol{\sigma_{pe}^2}$ was the maternal permanent environmental variance, and $\boldsymbol{\sigma_e^2}$ was the residual variance.

For the multi-trait model, the following animal model was used:

$$\begin{bmatrix} \boldsymbol{y_1} \\ \boldsymbol{y_2} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X_1} & 0 \\ 0 & \boldsymbol{X_2} \end{bmatrix} \begin{bmatrix} \boldsymbol{b_1} \\ \boldsymbol{b_2} \end{bmatrix} + \begin{bmatrix} \boldsymbol{Z_{d1}} & 0 \\ 0 & \boldsymbol{Z_{d2}} \end{bmatrix} \begin{bmatrix} \boldsymbol{u_{d1}} \\ \boldsymbol{u_{d2}} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & \boldsymbol{Z_{m2}} \end{bmatrix} \begin{bmatrix} 0 \\ \boldsymbol{u_{m2}} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & \boldsymbol{Z_{pe2}} \end{bmatrix} \begin{bmatrix} 0 \\ \boldsymbol{u_{pe2}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{e_1} \\ \boldsymbol{e_2} \end{bmatrix}$$

(Eq. 3.3)

Where $\boldsymbol{y_i}$ was the vector of observations for the i^{th} trait (1=DMI, 2= weaning weight), $\boldsymbol{b_i}$ was the vector of fixed effects, $\boldsymbol{u_{di}}$ was the vector of random direct additive genetic effects and $\boldsymbol{u_{mi}}$ was also a vector for maternal additive genetic effects, $\boldsymbol{u_{pei}}$ was a vector of maternal permanent environment effects, $\boldsymbol{e_x}$ was a vector of residual effects and $\boldsymbol{X_i}$, $\boldsymbol{Z_{di}}$, $\boldsymbol{Z_{mi}}$, $\boldsymbol{Z_{pei}}$ were the incidence matrices corresponding to $\boldsymbol{b_i}$, $\boldsymbol{u_{di}}$, $\boldsymbol{u_{mi}}$, and $\boldsymbol{u_{pei}}$, respectively. The fixed effects for DMI remained the same as the single trait analysis. For weaning weight, age, AOD and weaning CG were included as fixed effects. The weaning CG for both Gelbvieh and Red Angus included sex, breeder, breed percentage, weaning working group and weaning date. Breed percentages were partitioned into four groups: 25% to 43.75%, 43.76% to 62.5%, 62.3% to 87.5%, and 87.6% to 100%.

The variances and means of the model are as follows:

$$E(\boldsymbol{y_i}) = \boldsymbol{X_i b_i}$$

$$E(\boldsymbol{e_i}) = E(\boldsymbol{u_i}) = 0$$

$$var\begin{bmatrix} \boldsymbol{u_{d1}} \\ \boldsymbol{u_{d2}} \\ \boldsymbol{u_{m2}} \\ \boldsymbol{u_{pe2}} \\ \boldsymbol{e_1} \\ \boldsymbol{e_2} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}\sigma^2_{d1} & \boldsymbol{A}\sigma_{d1,d2} & \boldsymbol{A}\sigma_{d1,m2} & 0 & 0 & 0 \\ \boldsymbol{A}\sigma_{d1,d2} & \boldsymbol{A}\sigma^2_{d2} & \boldsymbol{A}\sigma_{d2,m2} & 0 & 0 & 0 \\ \boldsymbol{A}\sigma_{d1,m2} & \boldsymbol{A}\sigma_{d2,m2} & \boldsymbol{A}\sigma^2_{m2} & 0 & 0 & 0 \\ 0 & 0 & 0 & I\sigma^2_{pe2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \boldsymbol{I}\sigma^2_{e1} & \boldsymbol{I}\sigma_{e1,e2} \\ 0 & 0 & 0 & 0 & \boldsymbol{I}\sigma_{e1,e2} & \boldsymbol{I}\sigma^2_{e2} \end{bmatrix}$$

(Eq. 3.4)

where $\boldsymbol{A}$ was the numerator relationship matrix, $\boldsymbol{I_{pei}}$ was an identity matrix with an order of the number dams for trait *i*, $\boldsymbol{I_{ei}}$ was an identity matrix with an order of the number of animals with observations of the i^{th} trait, $\boldsymbol{\sigma^2_{di}}$ was the direct additive genetic variance for trait *i*, $\boldsymbol{\sigma_{di,di'}}$ was the direct genetic covariance for between the traits, $\boldsymbol{\sigma_{di,mi}}$ was the direct-maternal genetic covariance for the i^{th} trait, $\boldsymbol{\sigma^2_m}$ was the maternal additive genetic variance for WWT, $\boldsymbol{\sigma^2_{pe}}$ was the maternal permanent environmental variance for WWT, $\boldsymbol{\sigma^2_{ei}}$ was the residual variance for trait *i*, and $\boldsymbol{\sigma_{ei,ei'}}$ was the residual covariance between traits. The convergence criteria for the model was a REML log-likelihood change of less than 0.002 and the parameter estimates changed by less than 1% (Gilmour et al., 2009).

3.3 Results and Discussion

Summary statistics for DMI observations are presented in Table 3.1 On average, Gelbvieh cattle had trait values higher compared to Red Angus. The variation for DMI observations for Gelbvieh cattle was approximately half of the variation for Red Angus cattle. As for age, Gelbvieh cattle were on average 6 days younger than Red Angus cattle but the standard deviations for both data sets in regard to AGE were similar.

Table 3.1. Summary statistics for cattle with feed intake records for Gelbvieh and Red Angus cattle.

	Gelbvieh			Red Angus		
	[1]DMI	[2]Age	[3]Test	[1]DMI	[2]Age	[3]Test
n		3038			3213	
mean	11.08	289.14	71.84	10.58	295.77	52.17
SD	1.71	34.71	9.44	2.14	33.95	12.20
minimum	4.60	240.00	51.00	2.07	240.00	35.00
maximum	24.10	365.00	133.00	20.07	365.00	90.00

[1]DMI = Dry matter feed intake in kg

[2]Age = Age of cattle in days at the start of feed test

[3]Test = Test length in days for the collection DMI.

Table 3.2 presents the summary statistics for WWT and age at weaning for Gelbvieh and Red Angus cattle. On average, the age at weaning and distribution for age were similar among breeds. For WWT, Gelbvieh cattle were on average heavier at weaning compared to Red Angus cattle but the standard deviation for both breeds were similar.

Table 3.2. Weaning weight summary statistics for cattle with feed intake records for Gelbvieh and Red Angus cattle.

	Gelbvieh		Red Angus	
	[1]WWT	[2]Age	[1]WWT	[2]Age
n	7,792		4,342	
mean	270.86	200.05	253.56	202.42
SD	44.62	21.70	41.32	21.33
minimum	129.27	160.00	106.59	160.00
maximum	464.93	249.00	423.66	249.00

[1]WWT = weaning weight in kg

[2]Age = Age of cattle in days at weaning

Direct and maternal heritabilities from the single trait analysis are shown in table 3.3. The heritabilities for DMI direct for Red Angus and Gelbvieh cattle were moderate to high (0.24 ± 0.06 and 0.45 ± 0.06, respectively) and were both within the range of previously reported heritabilities for DMI (0.14 to 0.70; Berry and Crowley 2013). The DMI maternal heritability for Gelbvieh was zero and Red Angus DMI maternal heritability was very low at 0.05 ± 0.04. The correlation between DMI direct and maternal was $0.70 \times 10^{-3} \pm 0.10 \times 10^{-3}$ and -0.77 ± 0.20 for Gelbvieh and Red Angus cattle, respectively. The difference in genetic variances estimated for Gelbvieh and Red Angus would be expected given the different population structures, both genetic and non-genetic, between the two breeds.

There are few studies that have examined maternal effects of DMI in cattle. Hoque et al. (2007) examined the maternal effects for 740 Japanese Black cattle. In that study, no important influence of maternal effects on feed intake or feed efficiency was identified and Hoque et al. (2007) concluded the inclusion of maternal effects for feed intake in a genetic evaluation was not warranted. However, in a study conducted by Crowley et al. (2010), non-zero maternal heritability of 0.10 for DMI was estimated for 2,605 Irish beef bulls with an average age of 309 days. Crowley et al. (2010) found the inclusion of maternal genetic effects accounted for variability in DMI that was previously attributed to the direct genetic effect and that the inclusion of maternal effects decreased heritabilities for DMI from 0.49 to 0.38. The results obtained for Gelbvieh in the current study would support the findings of Hoque et al. (2007). Unlike Gelbvieh, the DMI maternal heritability for Red Angus was not zero, but it was lower than the estimates obtained by Crowley et al. (2010). The genetic variance for DMI maternal for Red Angus was 0.10 which only accounted for 6% of the phenotypic variance for DMI; therefore, the effect of the inclusion of a maternal effect for DMI in a genetic evaluation would be minimal.

Table 3.3. Direct and maternal variance components of dry matter feed intake for Gelbvieh and Red Angus cattle.

	Gelbvieh	Red Angus
σ^2_d	0.52 ± 0.07	0.46 ± 0.13
σ^2_m	0.16 x 10^{-7} ± 0.15 x 10^{-8}	0.10 ± 0.08
σ^2_e	0.64 ± 0.06	1.26 ± 0.09
σ^2_{pe}	0.75 x 10^{-7} ± 0.69 x 10^{-8}	0.09 ± 0.06
h^2_d	0.45 ± 0.06	0.26 ± 0.07
h^2_m	0.00 ± 0.00	0.06 ± 0.05

σ^2_d = direct additive genetic variance; σ^2_m = maternal additive genetic variance; σ^2_e = residual variance; σ^2_{pe} = permanent environment variance

The heritabilities and genetic correlations from the multi-trait analysis of DMI and WWT for Gelbvieh and Red Angus cattle are presented in table 3.4. Given the physiological relationship between DMI and body size (NRC, 2000), a moderate to high genetic correlation for DMI and WWT direct was anticipated. For Red Angus cattle, the estimated genetic correlation between DMI and WWT direct was moderate (0.54 ± 0.17) but the same genetic correlation for Gelbvieh was lower than expected (0.11 ± 0.13).

There is a large body of literature describing genetic correlations between DMI and performance traits such as ADG or post weaning traits. However, only three studies (Arthur et al. 2001; Bouquet et al. 2010; Crowley et al., 2011) examined DMI and weaning weight. Arthur et al. (2001) reported a genetic correlation between DMI and weaning weight direct of 0.28 ± 0.15 for Angus cattle with an average age of 268 days. Bouquet et al. (2010) used records from Blonde d'Aquitaine and Limousin bulls 7 to 9 months in age and reported higher correlations of 0.91 ± 0.09 and 0.62 ± 0.12. The current study's estimated genetic correlation for Red Angus

cattle would fall within the correlations reported by Arthur et al. (2001) and Bouquet et al. (2010) but the DMI heritability estimated for Red Angus cattle (0.27 ± 0.05) was lower than the heritabilities reported by Arthur et al. (2001) and Bouquet et al. (2010) of 0.39 and 0.30, respectively. These previously reported heritabilities were similar to our heritability estimated for Gelbvieh cattle; however, the genetic correlation for DMI and WWT direct estimated for Gelbvieh cattle (0.11 ± 0.13) was below the same correlation reported by Arthur et al. (2001) and Bouquet et al. (2010). Crowley et al. (2011) did not report a genetic correlation with DMI and WWT direct.

Of the above-mentioned studies, only Arthur et al. (2001) and Crowley et al. (2011) included weaning weight maternal in their analysis. The Bouquet et al. (2010) study did not include maternal genetic effect to overcome calculation limitations. Arthur et al. (2001) and Crowley et al. (2011) estimated genetic correlations between DMI and WWT maternal of 0.45 ± 0.16 and 0.32 ± 0.22, respectively. Both reported correlations were larger than the genetic correlations found in the current analysis. In this current study, low genetic correlations of weaning weight maternal to DMI of 0.12 ± 0.24 resulted for Red Angus cattle and 0.12 ± 0.13 for Gelbvieh cattle with large standard errors for both analyses.

Table 3.4. Estimated heritabilities (± S.E.) on the diagonal, above the diagonal is the residual correlation, and genetic correlations (± S.E.) below the diagonal for dry matter intake (DMI), weaning weight direct (WWT_D) and weaning weight maternal (WWT_M) for Gelbvieh and Red Angus cattle.

Breed	Gelbvieh			Red Angus		
[1]Trait	DMI	WWT_D	WWT_M	DMI	WWT_D	WWT_M
DMI	**0.45 ± 0.05**	0.37 ±0.06	-	**0.27 ± 0.05**	0.37 ± 0.05	-
WWT_D	0.11 ± 0.13	**0.36 ± 0.06**	-	0.54 ± 0.17	**0.21 ± 0.06**	-
WWT_M	0.12 ± 0.13	-0.68 ± 0.12	**0.15 ± 0.05**	0.12 ± 0.24	-0.21 ± 0.32	**0.16 ± 0.07**

[1]Traits are measured in kilograms

3.4 Conclusion

The accurate estimation of EPD for feed intake traits are dependent on the reliable estimation of variance components. The exclusion of maternal effects when maternal effects can influence the phenotype of a trait, can significantly inflate heritability estimates for traits. This is significant given the increasing interest on selecting cattle for improved feed utilization and the increasing number of breed associations publishing EPD for feed intake and efficiency. The inclusion of maternal effects for the estimation of variance components for feed intake is zero for Gelbveih cattle and only 6% percent of the variability observed in Red Angus cattle. The estimates from this study indicate that the inclusion of maternal effects in genetic evaluations for feed intake would not be warranted.

LITERATURE CITED

Anderson, R. V., R. J. Rasby, T. J. Klopfenstein, and R. T. Clark. 2005. An evaluation of production and economic efficiency of two beef systems from calving to slaughter. J. Anim. Sci. 83: 694-704.

Arthur, P. F., J. A. Archer, D. J. Johnson, R. M. Herd, E. C. Richardson, and P. F. Parnell. 2001a. Genetic and phenotypic variance and covariance components for feed intake, feed efficiency and other postweaning traits in Angus cattle. J. Anim. Sci. 79:2805–2811.

Berry, D. P. and J. J. Crowley. 2013. Cell Biology Symposium: Genetics of feed efficiency in dairy and beef cattle. J. Anim. Sci. 91: 1594-1613.

BIF. 2016. Guidelines for uniform beef improvement programs. 9th ed. Beef Improvement Federation. p. 25-28.

Bouquet, A., M.-N. Fouilloux, G. Renand, and F. Phocas. 2010. Genetic parameters for growth, muscularity, feed efficiency and carcass traits of young beef bulls. Livest. Sci. 129:38–48.

Crowley, J. J., M. McGee, D. A. Kenny, D. H. Crews Jr., R. D. Evans, and D. P. Berry. 2010. Phenotypic and genetic parameters for different measures of feed efficiency in different breeds of Irish performance-tested beef bulls. J. Anim. Sci. 88:885-894.

Crowley, J. J., R. D. Evans, N. Mc Hugh, D. A. Kenny, M. McGee, D. H. Crews Jr., and D. P. Berry. 2011. Genetic relationships between feed efficiency in growing males and beef cow performance. J. Anim. Sci. 89:3372–3381.

Gilmour, A. R., B. Goegel, B. Cullis, R. Thompson, and D. Butler. 2009. ASReml user guide release 3.0. BSN International Ltd, Hemel Hepstead, UK.

Hoque, M. A., P. F. Arthur, K. Hiramoto, A. R. Gilmour, and T. Oikawa. 2007. Variance components due to direct genetic, maternal genetic and permanent environmental effect for growth and feed-efficiency traits in young male Japanese Black cattle. J. Anim. Breed. Genet. 124: 102-107.

Lee, J. W., S. B. Choi, Y. H. Jung, J. F. Keown, and L. D. Van Vleck. 2000. Parameter estimates for direct and maternal genetic effects on yearling, eighteen-month, and slaughter weights of Korean native cattle. J. Anim. Sci. 78:1414-1421.

Mavrogenis, A. P., E. U. Dillard, and O. W. Robinson. 1978. Genetic analysis of postweaning performance of Hereford bulls. J. Anim. Sci. 47:1004-1013.

Meyer, K. 1992. Variance components due to direct and maternal effects for growth traits of Australian Beef Cattle. Livest. Prod. Sci. 31:179-204

Nielsen, M. K., M. D. MacNeil, J. C. M. Dekkers, D. H. Crews Jr., T. A. Rathje, R. M. Enns, and R. L. Weaber. 2013. Review: Lifecycle, total-industry genetic improvement of feed efficiency in beef cattle: Blueprint for the Beef Improvement Federation. Prof. Anim. Sci. 29:559-565.

NRC. 2000. Nutrient requirements of beef cattle. 7th rev. ed. Natl. Acad. Press, Washington, DC.

Schaeffer, L. R. and B. W. Kennedy. 1989. Effects of embryo transfer in beef cattle on genetic evaluation methodology. J. Anim. Sci. 67:2536-2543.

Chapter IV

A SIMULATION STUDY EXAMINING GENETIC SELECTION FOR FEED INTAKE AND RESIDUAL FEED INTAKE ON CORRELATED PERFORMANCE TRAITS IN BEEF CATTLE

4.1 Introduction

Feed costs are reported as the largest variable expense to producers in the beef industry, accounting for 50 to 70% of total production costs (Anderson et. al, 2005). Due to the large impact of feed cost on profitability, beef cattle producers have become increasingly aware of the need to improve feed utilization. However, there is a debate as to what is the best phenotype for feed efficiency in cattle, how to incorporate it into a breeding program or genetic evaluation, and how the selection for feed efficiency impacts other performance traits (Berry and Pryce, 2013).

With advancements in technology, dry matter feed intake (DMI) has become a trait of interest and is readily accepted and understood by beef cattle producers since it is a direct measure of feed consumption. Given that DMI is only a measure of the amount of feed consumed by an animal, it gives no indication of an animal's performance for other production traits. As such, DMI is not a measure of feed efficiency but is a significant component for the determination of that measure (Berry and Pryce, 2013).

Feed efficiency is not a directly measurable trait but must be calculated as a function both inputs and outputs, which in this case are feed intake and production (Koch et al., 1963). Residual feed intake (RFI) has become a prominent, yet contested, trait of interest of feed

efficiency (Lu et al., 2015) and is defined as the difference between the actual feed intake and the estimated feed intake adjusted for the requirements of production (Kennedy et al., 1993). The attraction of RFI is that on a phenotypic level, it is independent from the production traits included in its calculation, such as ADG. However, on a genetic level, RFI has been found to be negatively correlated with other productions traits, such as yearling weight. In addition, RFI is conceptually difficult to explain to producers as the calculations for RFI seems complex and difficult to understand (Berry and Pryce, 2013).

In the context of genetic improvement of feed efficiency, reducing feed intake should not be the sole goal of a breeding program. Selection pressure should be placed on increasing production traits relevant to a producer's marketing scheme while simultaneously reducing feed intake (Nielsen et al., 2013). Currently, the effects of selection for RFI or DMI on correlated performance traits, such as weaning or yearling weight, is generally unknown. The objective of this study is to use simulated data to examine the effects of single trait selection on either RFI or DMI on genetically correlated traits of weaning and yearling weight. A third simulation also examines the same traits when a selection is based on the highest-ranking animals from an economic selection index weighted for intake and weight traits. After generating 10 years of simulated data for single trait selection of DMI or RFI, or selecting animals from an index, the effects of the genetically correlated traits of weaning weight and yearling weight were compared.

4.2 Materials and Methods

For this study, the performance traits of interest were weaning weight (WWT) and yearling weight (YWT). Three scenarios were simulated using R program (R Core Team, 2018).

The first scenario examined single trait selection for average daily dry matter intake (DDMI), the second examined single trait selection for RFI and the final scenario selected animals based on an economic selection index (ESI). In order to generate simulated populations of cattle, phenotypic means were required in addition to genetic and residual (co)variance matrices which were obtained from previously published genetic parameters. Once simulated populations were established, the top 5% of bulls and 20% of heifers for the specific scenario's trait of interest were chosen as replacements. A 90% conception rate was assumed with cows older than 16 years of age removed and bulls were replaced after 2 years.

4.2.1 Estimation of Weighted Means of Phenotypic Averages and Genetic (Co)variances

Given that the objective of this study was the effect of single trait selection for DDMI or RFI on genetically correlated traits, the estimation of the (co)variance components for DDMI, RFI, WWT and YWT are crucial. The estimation of breeding values (EBV) using best linear unbiased prediction (BLUP) requires that variance components were known and as such, the estimation of variance components are important for genetic evaluations. These (co)variances were the dispersion parameters that described the random blocks (i.e. genetic and residual effects) for Henderson's mixed model equations. The genetic (co)variances described the genetic variation within a trait through direct genetic variance and covariances between. Residual (co)variance describes the within and between trait environmental influences. Within a typical multi-trait BLUP model, each trait included in the model lends information to other traits in the model when genetic and residual covariances exist. The general matrix form of these equations were as follows:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

(Eq. 4.1)

where y_i was a vector of observations for the trait, β_i was a vector of fixed effects, $u_i$was a vector of additive genetic effects, $e_i$was a vector for residual effects, and X_i and Z_i were incidence matrices corresponding to β_i, and u_i, respectively. Variances included in the model were as follows:

$$var \begin{bmatrix} u_1 \\ u_2 \\ e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} G_{1,1} & G_{1,2} & 0 & 0 \\ G_{2,1} & G_{2,2} & 0 & 0 \\ 0 & 0 & R_{1,1} & R_{1,2} \\ 0 & 0 & R_{2,1} & R_{2,2} \end{bmatrix} = \begin{bmatrix} \sigma^2_{u_1} A & \sigma_{u_1 u_2} A & 0 & 0 \\ \sigma_{u_1 u_2} A & \sigma^2_{u_2} A & 0 & 0 \\ 0 & 0 & I_1 \sigma^2_{e_1} & I \sigma_{e_1 e_2} \\ 0 & 0 & I \sigma_{e_1 e_2} & I_2 \sigma^2_{e_2} \end{bmatrix}$$

(Eq. 4.2)

where *A* was the numerator relationship matrix, I was the identity matrix with an order of the number of animals with observations, $\sigma^2_{u_i}$was the direct additive genetic variance, $\sigma_{u_i u_{i'}}$was the genetic covariance between traits, $\sigma^2_{e_i}$was the residual variance, and $\sigma_{e_i e_{i'}}$ was the residual covariance. To accurately model the response of genetically correlated traits from selection on DDMI or RFI, both genetic and residual (co)variances needed to be estimated which were the elements of two respective matrices ***G*** and ***R.***

In order to simulate data that mimic real herd situations, weighted averages of (co)variances components were compiled from phenotypic, genetic and residual estimates previously published in literature using procedures outlined by Koots et al. (1994a). Weighted phenotypic averages were estimated using equation 4.3:

$$\bar{x}_w = \frac{\sum_{i=1}^{k} n_i \bar{x}_i}{\sum_{i=1}^{k} n_i}$$

(Eq. 4.3)

where $\bar{x}_w$ was the weighted phenotypic mean of the trait of interest, $\bar{x}_i$ was the trait mean from the *i*th study, n_i was the number of records in the *i*th study, and *k* was the number of studies. The weighted phenotypic averages for DDMI, RFI, WWT and YWT were presented in table 4.1.

Table 4.1. Weighted averages calculated from literature for residual feed intake (RFI), daily dry matter intake (DDMI), weaning weight (WWT), and yearling weight (YWT).

Trait	Est. [a]	n	Source [b]
RFI	-0.01	15	1, 2, 3, 4, 5, 8, 10, 11, 12, 13, 17, 18, 19, 22, 24
DDMI (kg)	10.01	15	1, 2, 3, 4, 5, 8, 10, 11, 12, 13, 16, 17, 19, 22, 24
WWT (kg)	258.91	8	2, 5, 7, 11, 14, 15, 20, 23
YWT (kg)	479.99	7	2, 5, 7, 10, 11, 14, 23

[a]Est. = Weighted averages calculated from literature.
[b] 1 = Archer et al. (1997), 2 = Arthur et al. (2001a), 3 = Arthur et al. (2001c), 4 = Barwick et al. (2009), 5 = Bouquet et al., 2010, 6 = Ceacero et al. (2016), 7 = Costa et al. (2011), 8 = Crowley et al. (2010), 9 = Durunna et al. (2011b), 10 = Fan et al. (1995), 11 = Herd and Bishop (2000), 12 = Hoque et al. (2006), 13 = Hoque et al. (2009), 14 = Iwaisaki et al. (2005a), 15 = Iwaisaki et al. (2005b), 16 = Korver et al. (1991), 17 = Lancaster et al. (2009), 18 = Mujibi et al. (2011), 19 = Nkrumah et al. (2007b), 20 = Phocas et al. (2004), 19 = Robinson and Oddy (2004), 22 = Rolf et al. (2011), 23 = Roughsedge et al. (2005), 24 = Schenkel et al. (2004), 25 = Williams et al. (2011).

Heritability estimates obtained from a literature review were averaged (Eq. 4.4: Koots et al., 1994a) using a weighting factor of the inverse of the sampling variance for each estimate (Eq. 4.5; Koots et al., 1994a).

$$h_w^2 = \frac{\sum_{i=1}^{n} h_i^2 \Big/ \left(SE_{h_i^2}\right)^2}{\sum_{i=1}^{n} 1 \Big/ \left(SE_{h_i^2}\right)^2}$$

(Eq. 4.4)

where h_w^2 is the weighed mean for heritability, h_i^2 is the heritability estimated from the *i*th published study, and $SE_{h_i^2}$ is the standard error estimate corresponding to the heritability.

$$weight\ of\ h_i^2 = \frac{1}{\left(SE_{h_i^2}\right)^2}$$

(Eq. 4.5)

The standard errors for heritability weighted means (SE_w) were calculated by taking the square root of the summation of weighted factors (Eq. 4.6; Koots et al., 1994a).

$$SE_w = \sqrt{\frac{1}{\sum_{i=1}^{n}\left(1/SE_{h_i^2}\right)}}$$

(Eq. 4.6)

All phenotypic and genetic correlations (*r*) were transformed to approximate a normal scale using a Fisher's Z transformation (Eq. 4.7; Steele and Torrie, 1980; Koots et al., 1994b).

$$Z = 0.5log\frac{(1+r)}{(1-r)}$$

(Eq. 4.7)

With standard errors:

$$SE_Z = \sqrt{1/(n-3)}$$

(Eq. 4.8)

where *r* was the correlation (phenotypic and genetic) from literature and *n* was the number of animals for a phenotypic correlation and number sires for a genetic correlation. Over all studies, the value of Z was pooled with weighting by the inverse of the sampling variance (Eq. 4.9; Koots et al., 1994b):

$$Z_{pooled} = \frac{\sum_{i=1}^{n} Z_i / \left(SE_{Z_i}\right)^2}{\sum_{i=1}^{n} 1 / \left(SE_{Z_i}\right)^2}$$

(Eq. 4.9)

The pooled Z values was transformed back to a correlation (Eq. 4.10; Koots et al., 1994b):

$$r = \frac{(e^{2Z} - 1)}{(e^{2Z} + 1)}$$

(Eq. 4.10)

Table 4.2 are the weighted genetic variances, correlations (genetic and residual) and heritabilities for DDMI, RFI, WWT, and YWT.

Table 4.2. Number of estimates (Est.) from literature (n) and weighted estimates with standard error in parenthesis for genetic variance (diagonal), genetic correlation (above diagonal), residual correlation (below diagonal), and heritability (h^2) for residual feed intake (RFI), average daily dry matter intake (DDMI), weaning weight (WWT) and yearling weight (YWT) with sources cited.

Trait	RFI			DDMI			WWT			YWT		
	Est.	n	Source [a]	Est.	n	Source [a]	Est.	n	Source [a]	Est.	n	Source [a]
RFI	**0.18 (0.01)**	16	1, 2, 3, 4, 5, 8, 10, 11, 12, 13, 17, 18, 19, 22, 24, 25	0.73 (0.08)	12	2, 3, 4, 6, 8, 9, 12, 17, 18, 19, 21, 22	-0.21 (0.23)	2	2, 5	-0.21 (0.23)	2	2, 5
DDMI	0.57 (0.05)	12	2, 3, 4, 6, 8, 9, 12, 17, 18, 19, 21, 22	**0.69 (0.02)**	16	1, 2, 3, 4, 5, 8, 10, 11, 12, 13, 17, 18, 19, 22, 24, 25	0.79 (0.20)	2	2, 5	0.93 (0.13)	2	2, 5
WWT	0.02 (0.14)	2	2, 5	0.02 (0.14)	1	5	**230.77 (0.20)**	8	2, 5, 7, 11, 14, 15, 20, 23	0.88 (0.08)	4	2, 7, 15, 23
YWT	0.03 (0.14)	2	2, 5	0.53 (0.14)	1	5	0.59 (0.08)	2	7, 15	**598.33 (0.003)**	7	2, 5, 7, 10, 11, 15, 23
h^2	0.37 (0.07)	16	1, 2, 3, 4, 5, 8, 10, 11, 12, 13, 17, 18, 19, 22, 24, 25	0.41 (0.06)	16	1, 2, 3, 4, 5, 8, 10, 11, 12, 13, 17, 18, 19, 22, 24, 25	0.29 (0.04)	8	2, 5, 7, 11, 14, 15, 20, 23	0.35 (0.11)	7	2, 5, 7, 10, 11, 15, 23

[a] 1 = Archer et al. (1997), 2 = Arthur et al. (2001a), 3 = Arthur et al. (2001b), 4 = Barwick et al. (2009), 5 = Bouquet et al., 2010, 6 = Ceacero et al. (2016), 7 = Costa et al. (2011), 8 = Crowley et al. (2010), 9 = Durunna et al. (2011), 10 = Fan et al. (1995), 11 = Herd and Bishop (2000), 12 = Hoque et al. (2006), 13 = Hoque et al. (2009), 14 = Iwaisaki et al. (2005a), 15 = Iwaisaki et al. (2005b), 16 = Korver et al. (1991), 17 = Lancaster et al. (2009), 18 = Mujibi et al. (2011), 19 = Nkrumah et al. (2007b), 20 = Phocas et al. (2004), 19 = Robinson and Oddy (2004), 22 = Rolf et al. (2011), 23 = Roughsedge et al. (2005), 24 = Schenkel et al. (2004), 25 = Williams et al. (2011).

4.2.2 Simulated Data

In order to simulate a population of beef cattle that would represent industry herds, a base population of 10,000 cows was established to obtain an age distribution that was similar to what occurs in industry. No selection pressure was placed on the base population. From this base population, true breeding values (TBV) were estimated using the weighted heritabilities and (co)variances presented in table 4.1. Selection for decreased DMI, RFI or using an economic selection index (ESI) began using the progeny from the base population as replacements. Replacements were selected dependent on their true breeding values (TBV) for the trait of interest depending on the selection scenario. The simulation continued until 10 years of progeny was generated. The method used to develop the base population and the subsequent selection for decreased DMI and RFI are detailed in the following sections.

Female Base Population. A base population of females was established using the R statistical software package (R Core Team, 2018) and the *R* code written by Larry Schaeffer (http://www.aps.uoguelph.ca/~lrs/Summer2012Full/MTiter.R), this simulated base population began with 10,000 females randomly assigned to 50 herds. A 20% replacement rate was assigned for the establishment of this female base population. In addition, no selection pressure was placed on the population, and 20% of the females were randomly removed for 16 years. Once 16 years of a population was established, females from the initial year were removed (females greater than 16 years of age). With the removal of females greater than 16 years of age, the population of 10,000 females reached an equilibrium for the distribution of age.

Table 4.3 was the number and percentage of females expected for each age group within the base population. The generation interval for the age distribution of the female base population was determined (Eq. 4.11).

$$L = \frac{\sum(age * n_{age})}{n_{total}}$$

(Eq. 4.11)

Where *age* is the age of the females (i.e. 2 to 16) from the base population, n_{age} is the number of females for the specific age, and n_{total} was the total number of females within the population, in this case 10,000. The distribution of female age resulted in a generation interval of 5.52 years. This population of simulated females was used as the base population of females for the following selection scenarios for DDMI, RFI or ESI.

Table 4.3. Number of females for each age group used as a base population for a simulation study (n=10,000).

[1]Females	Age	Percentage
0	17	0%
137	16	1%
118	15	1%
148	14	1%
170	13	2%
207	12	2%
287	11	3%
377	10	4%
417	9	4%
530	8	5%
701	7	7%
824	6	8%
1048	5	10%
1372	4	14%
1664	3	17%
2000	2	20%

[1]Females are the number of females for each age.

Simulation of Data. Four hundred fifty bulls were simulated to establish a 1:22 bull to cow ratio and were randomly assigned to the 50 herds established during the development of the female base population. The females from the established base population were randomly mated to the bulls with an assumed 90% conception rate for females, resulting in 9,000 progeny born year 1 of the simulation (P1). Progeny were classified by the year they are born as Pn, where n represents the year of birth. For example, progeny born during year 1, 2 and 3 are designated as P1, P2, and P3, respectively. Cows that did not produce offspring (open) were removed from the simulation resulting in 9,000 females from the base population that were retained to year 2. Open cows were designated by the simulation randomly, regardless of their TBV. Selection for the trait of interest only occurred on replacement heifers and bulls based on their TBV for that specific trait. Since cattle are 2 years of age when their first offspring are born, there is a delay before replacement heifers and bulls are contributing progeny to the population. As a result, selection for the traits of interest did not occur until progeny from the base population were old enough to be used as replacements. The progeny born year 1 of the simulation (P1) contributed offspring to the population year 3, and progeny born year 2 (P2) contributed offspring to year 4, and so forth until year 10 (Figure 1).

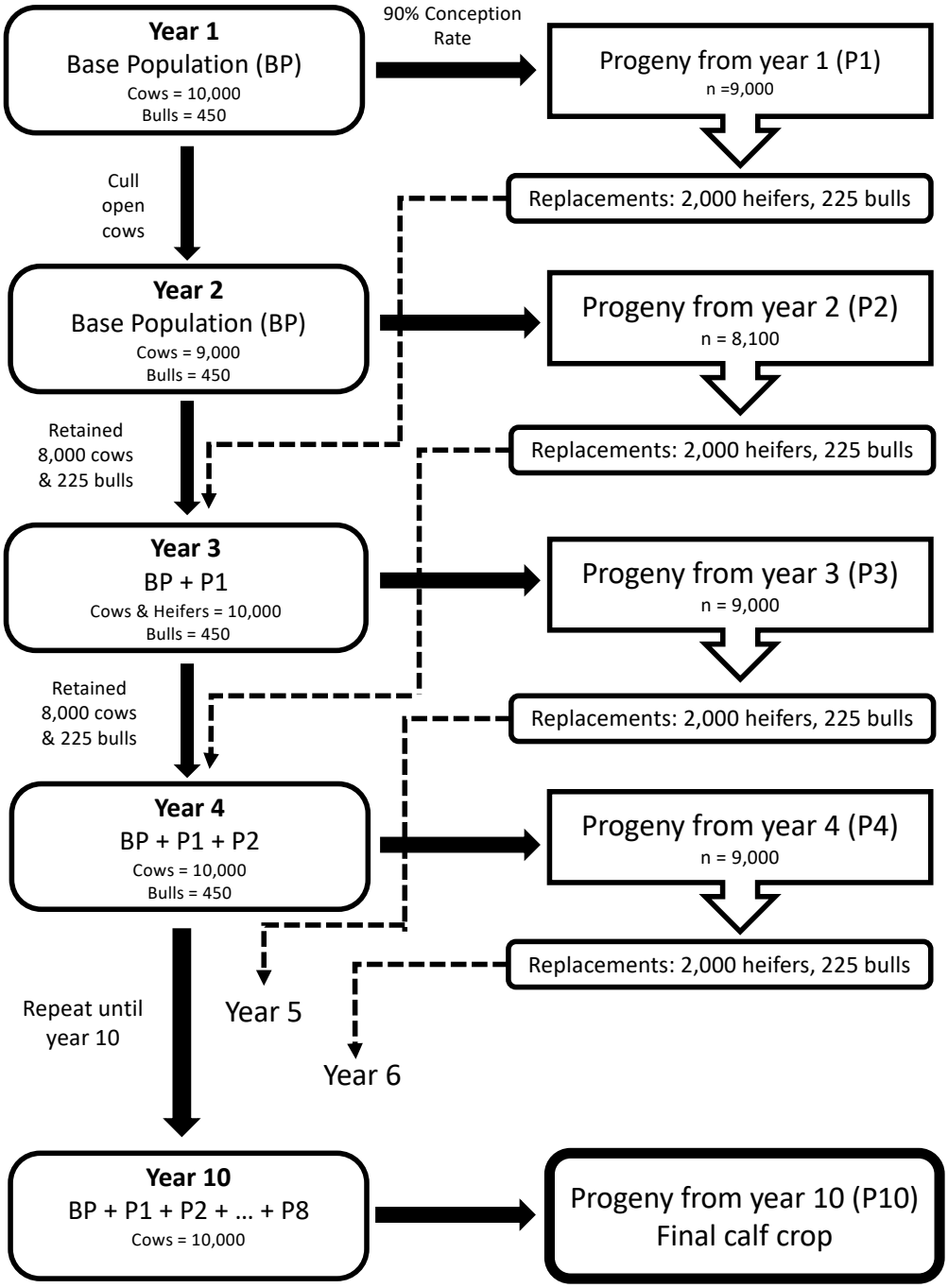

Figure 4.1. Schematic representation of simulated data used to generate populations of cattle for single trait selection for either daily dry matter intake, residual feed intake, or an economic selection index.

For year 2, only animals from the base population were used for the simulation of P2 since replacement heifers and bulls would be too young to produce offspring until year 3. At a conception rate of 90%, the number of females from the base population, as well as the number of progeny, would reduce from 9,000 to 8,100 for year 3. Year 2 is the only year were the calf crop is below 9,000, each subsequent year, the calf crop would remain at 9,000. Beginning during year 2 and subsequent years, females greater than 16 years of age were removed. From year 3 to year 10, only 8,000 cows would be retained for the next year. Once cows were culled due to reproductive failure or age, the simulation randomly selected 8,000 retained cows to produce progeny for the next year. Progeny from replacement heifers and bulls born year 1 (P1) were introduced into the breeding population during year 3. Progeny from each year were ranked by their TBV with the top 2,000 heifers and 225 bulls chosen as replacements each year. For progeny from year 3 to year 10, yearling and 2-year-old bulls were used resulting in a generation interval of 2.5 years for males. Animals that failed to gain more than 40 kg from weaning to yearling weight, were removed as potential replacements regardless of their TBV. Selection for the trait of interest began with the replacement heifers and bulls producing calves during year 3. From year 3 to year 10, all 3-year-old bulls were replaced each year and 20% of the cows were culled and replaced with 2,000 heifers reestablishing a female population of 10,000 each year.

Based on the *R* code written by Larry Schaeffer (http://www.aps.uoguelph.ca/~lrs/Summer2012Full/MTiter.R), true breeding values (TBV) and observations were created using the R program (R Core Team, 2018). Genetic (G) and residual (R) covariance matrices were constructed using the weighted heritabilities, genetic variances and correlations mentioned previously (Table 4.1). No observations were simulated for the base

population, but true breeding values were estimated for the initial base population using a Cholesky decomposition as a function of the G and R matrices.

The Cholesky decomposition factors a positive-definite matrix into the product of a lower triangular matrix and its conjugate transpose (Meyer, 1996; Mrode, 2005). For the Cholesky decomposition of matrix A (Eq. 4.12):

$$\mathrm{A} = \mathrm{LL}'$$

(Eq. 4.12)

where L was the lower triangular matrix (Cholesky factor) with elements l_{ij} $\left(l_{ij} = 0 \text{ for } j > i\right)$ and L′was the inverse of L. The elements of L can be derived by the following (Eq. 4.13 and 4.14)

$$\mathrm{L}_{j,j} = \sqrt{A_{j,j} - \sum_{k=1}^{j-1} \mathrm{L}_{j,k}^2}$$

(Eq. 4.13)

$$\mathrm{L}_{i,j} = \frac{1}{\mathrm{L}_{j,j}}\left(A_{i,j} - \sum_{k=1}^{j-1} \mathrm{L}_{i,k}\mathrm{L}_{j,k}\right) \quad for\ i > j$$

(Eq. 4.14)

True breeding values for progeny were simulated using the average of the parents' TBV and a mendelian sampling effect. The mendelian sampling effect was estimated by multiplying a vector of randomly generated numbers with an average of zero and a standard deviation of 1 by the Cholesky factor of the G matrix. For each year's progeny, observations were estimated as the trait average, fixed effects (sex and herd), and TBV (Eq. 4.15):

$$y_{ij} = \mu_j + h_i + tbv_{ij}$$

(Eq. 4.15)

Where y_{ij} was the simulated observation for the *i*th animal for the *j*th trait, μ_j was the average of the *j*th trait, h_i was the fixed effects of sex and herd, and tbv_{ij} was the true breeding value. The fixed effect for herd was estimated by using randomly generated, normally distributed vector of numbers with a mean of zero and a non-specified standard deviation. Previously published sex effects for weaning and yearling weight (Van Vleck and Cundiff, 1998) were used. Initial observations for progeny born year 1 used weighted phenotypic averages calculated using equation 4.1 (Table 4.1; Koots et al., 1994a). For each year, new population phenotypic averages would be recalculated using the simulated observations. It was anticipated that the phenotypic observations for each trait would change as genetic selection occurred.

Selection Scenarios. Three selection scenarios were used for this study. The first scenario (S_{DDMI}) was the single trait selection for DDMI, the second scenario (S_{RFI}) was the single trait selection for RFI, and the last scenario (S_{ESI}) was selection using an economic selection index (ESI). Selection on DDMI and RFI followed the same procedures. Selection on the index followed a similar procedure except for the calculation of the ESI and selection which was based on the index values (Table 4.5) instead of DDMI or RFI. All three scenarios used simulated data from the female base population.

For selection scenarios S_{DDMI} and S_{RFI}, true breeding values for DDMI and RFI were used for the selection of replacement bulls and heifers. For the third selection scenario (S_{ESI}), TBV were weighted for DDMI, average daily gain, weaning and yearling weight to estimate

economic selection index values. Animals were ranked based on those index values and chosen from the top 5% for bulls and top 20% for heifers to be used as replacements.

Economic Selection Index. An economic selection index is a method of multi-trait selection by applying economic weights to the breeding values for relevant traits to predict an aggregate breeding value for an individual and is represented as follows (Eq. 4.16; Bourdon, 1997):

$$H = v_1 BV_1 + v_2 BV_2 + \cdots + v_t BV_t \quad \text{(Eq. 4.16)}$$

Where H was the index value, v_i was the economic weight for the trait, BV_i was the breeding values for the trait, and t is the total number of traits incorporated in the index.

The economic selection index for this study was developed to account for yearling cattle prices as a revenue source and feed prices as an expense for the profitability of the simulated cattle populations. Other factors such as cow costs, labor and herd health were not considered for this index. Cattle prices were estimated using monthly futures market prices for United States fed cattle (www.investing.com/commodities/live-cattle-historical-data), averaged over 10 years resulting in a fed cattle price of $2.64/kg. A finishing ration of 80% corn, 10% hay, 5% distillers' grain and 5% supplement were assumed for the calculation of feed costs. The cost of each feedstuff within the ration were calculated using a 10-year average of commodity prices with the exception of the supplement which was considered as a constant. Table 4.3 is the average price for fed cattle, the cost of each feed ingredient per kilogram and the total ration cost per kilogram.

Table 4.4. Ten-year average for cattle and commodity prices.

	[1]Average per kg	[2]Source
Cattle	$2.64	www.investing.com/commodities/live-cattle-historical-data
Corn	$0.18	www.nass.usda.gov/Charts_and_Maps/graphics/data/pricecn.txt
[3]DDGS	$0.18	www.marketnews.usda.gov/mnp/ls-report-config
Hay	$0.17	www.asi.k-state.edu/about/newsletters/focus-on-feedlots/monthly-reports.html
Supplement	$0.29	www.iowabeefcenter.org/calculators.html
Total Ration Cost	**$0.18**	

[1]Ten year average per kilogram
[2]Reference for cattle and commodity prices
[3]Distillers' grain

To estimate total feed consumed for finished yearling cattle in a feedlot, a 10-year average for days on feed was calculated (Focus on Feedlots Monthly Reports, https://www.asi.k-state.edu/about/newsletters/focus-on-feedlots/monthly-reports.html) resulting in 157 d as the average days to finish. Using the methods previously described, phenotypes for DDMI, ADG, WWT and YWT as well as TBV for a population of cattle was simulated using the G and R matrices from the DDMI simulation. Using this population of cattle, the profit for each animal was calculated as follows (Eq. 4.17):

$$Profit = \$2.64 * YWT + 157 * \$0.18 * DDMI \quad \text{(Eq. 4.17)}$$

where YWT and DDMI are the phenotypes for yearling weight and average daily dry matter intake, respectively, and 157 was the average days on feed.

To estimate an economic index, profit was regressed on the TBV of the simulated population of cattle (Eq. 4.18):

$$Profit = TBV_{DDMI}\, v_1 + TBV_{ADG} v_2 + TBV_{WWT} v_3 + TBV_{YWT} v_4 + e$$

(Eq. 4.18)

Regression coefficients and corresponding p-values are presented in table 4.4. Average daily gain and weaning weight were not important ($p > 0.05$) in the model. Average daily dry matter intake had the largest, negative coefficient which was expected given its significant cost to producers. With the establishment of the index (Eq. 4.19), data was simulated using the same genetic parameters as the DDMI simulation with the selection of replacement bulls and heifers dependent on their index value instead of the TBV.

Table 4.5. Regression coefficients for the selection index for dry matter intake.

Trait	Coefficient	P-value
[1]DDMI	-29.669	<0.0001
[2]ADG	21.994	0.555
[3]WWT	0.097	0.554
[4]YWT	2.551	<0.0001

[1] Average daily dry matter Intake breeding value
[2] Average daily gain breeding value
[3] Weaning weight breeding value
[4] Yearling weight breeding value

$$H = -29.669 BV_{DDMI} + 21.994 BV_{ADG} + 0.097 BV_{WWT} + 2.551 BV_{YWT}$$

(Eq. 4.19)

4.3 Results and Discussion

Figure 4.2 illustrated the change in DDMI and YWT over the ten years of simulated data and the averages for DDMI, WWT and YWT for all 3 simulations are presented in table 4.5. In the first simulation, for the selection on DDMI demonstrated the largest decline in all traits included in the simulation. There was a yearly average decrease of 0.85 kg/year in DDMI and a decrease of 7.09 kg for progeny averages of P1 compared to P10. This simulation also demonstrated a drastic decline for both genetically correlated traits, weaning and yearling weights. There was an annual average loss of 14.03 kg/year and 27.83 kg/year for weaning and yearling weight, respectively, resulting in decreased progeny averages of 119.14 kg for weaning weight and 240.18 kg for yearling weights over the specified time frame. In addition to the loss in growth traits, there was also a decrease in the number of progeny. Since the simulation assigned observations (phenotypes) based on a distribution and accounted for environmental effects such as herd and sex, the rapid decrease in DDMI resulted in some progeny having a negative DDMI. Since this was not biologically plausible, simulated cattle with DDMI less than 1 were removed resulting in a reduction in population size for this simulation (table 4.6).

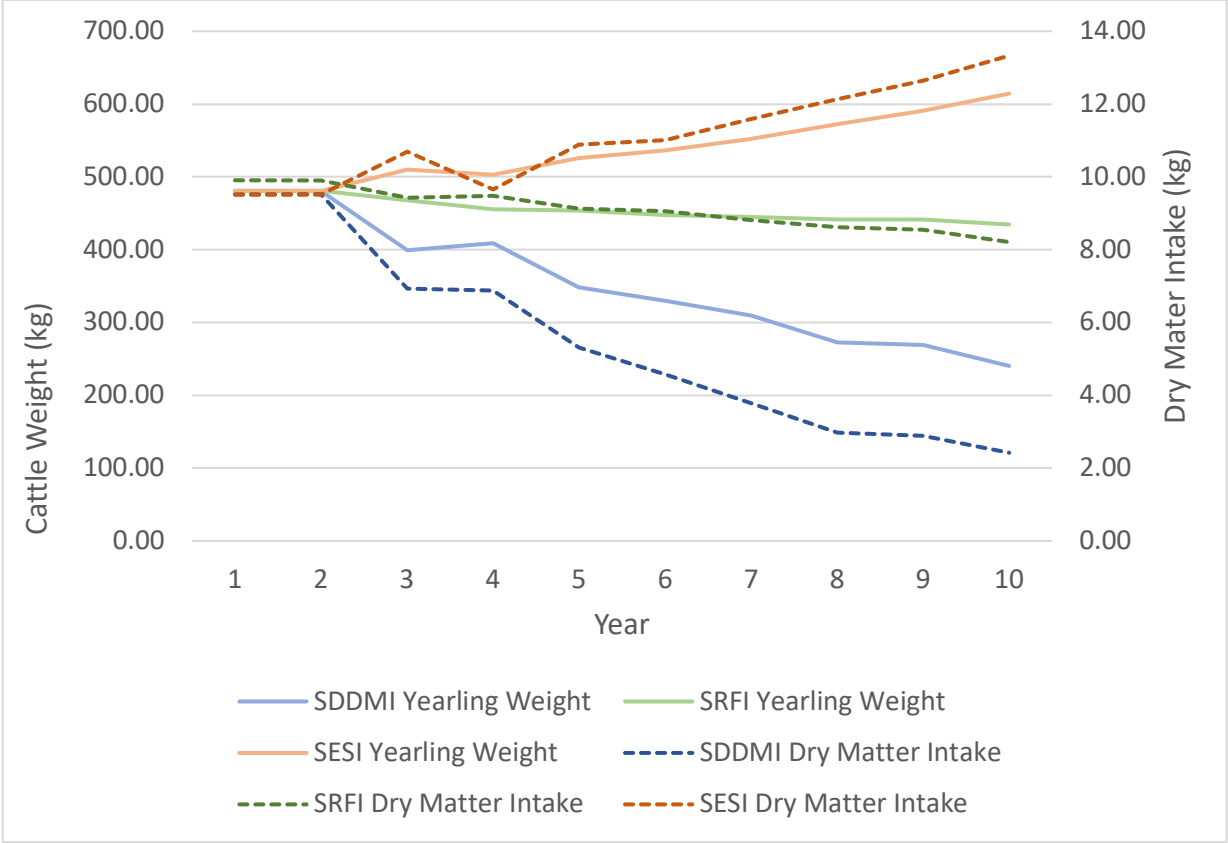

Figure 4.2. The simulated changes in phenotypic dry matter intake and yearling weight for three selection scenarios. S_{DDMI} was the single trait selection of dry matter intake and the response to the genetically correlated trait yearling weight. S_{RFI} was the selection of residual feed intake and the response of genetically correlated traits of dry matter intake and yearling weight. S_{ESI} was a simulation of selection using economic index values with genetically correlated traits of dry matter intake and yearling weight.

Table 4.6. The phenotypic average for each year's progeny from 3 simulation scenarios with selection on average daily dry matter intake (DDMI), residual feed intake (RFI) and an economic selection index. Included are the progeny averages of DDMI and genetically correlated traits included in each simulation (weaning weight and yearling weight).

Year	DDMI (kg)			Weaning Weight (kg)			Yearling Weight (kg)		
	S_{DDMI}[1]	**S_{RFI}[2]**	**S_{ESI}[3]**	**S_{DDMI}[1]**	**S_{RFI}[2]**	**S_{ESI}[3]**	**S_{DDMI}[1]**	**S_{RFI}[2]**	**S_{ESI}[3]**
1	9.51	9.91	9.52	248.10	248.50	248.10	480.50	481.00	480.50
2	9.52	9.91	9.52	248.20	248.30	248.20	480.50	480.60	480.50
3	6.94	9.42	10.69	203.30	239.90	264.30	398.90	467.50	509.80
4	6.88	9.47	9.66	211.67	237.20	250.40	409.20	455.30	503.20
5	5.31	9.14	10.89	179.55	236.10	269.00	348.20	453.90	525.70
6	4.57	9.05	11.00	170.46	235.00	271.40	329.40	447.70	536.10
7	3.78	8.81	11.59	158.60	234.00	280.80	309.50	444.70	552.60
8	2.97	8.62	12.14	141.44	234.20	290.90	272.76	441.50	572.50
9	2.88	8.54	12.64	139.00	234.60	300.00	268.90	441.60	591.00
10	2.42	8.21	13.33	128.96	233.40	313.00	240.32	434.50	614.30

S_{DDMI}[1] was the simulation of single trait selection for DDMI and correlated traits of weaning weight and yearling weight.

S_{RFI}[2] was the simulation of single trait selection for RFI and correlated traits of weaning weight and yearling weight.

S_{ESI}[3] was the simulated data from selection based on economic selection index values.

Table 4.7. Number of simulated progenies produced each year for three simulations scenarios.

Year	S_{DDMI}[1]	S_{RFI}[2]	S_{ESI}[3]
1	8964	8965	8964
2	8067	8069	8067
3	8960	8963	8935
4	8980	8960	8952
5	8903	8960	8917
6	8760	8958	8936
7	7190	8937	8824
8	6802	8957	8895
9	6527	8961	8862
10	4536	8959	8829

S_{DDMI}[1] was the simulation of single trait selection for DDMI and the response of correlated traits weaning weight and yearling weight.

S_{RFI}[2] was the simulation of single trait selection for residual feed intake and correlated traits of weaning weight and yearling weight.

S_{ESI}[3] was the simulated data from selection based on selection index values.

The number of progeny per year for S_{DDMI} began to decline after the fifth year. This was compared to the other simulation scenarios that had minimal loss in number of progeny. A decline in progeny number of this extent would significantly affect the profitability and sustainability of production operations.

The first simulation scenario performed as expected with regards to the decline in DDMI, WWT and YWT. Increasing output traits, such as yearling weight, will cause input traits, such as DDMI, to also increase (Meyer et al., 2008). It is reasonable to expect that the inverse is also true, decreased input would simultaneously decrease output as evident in S_{DDMI}. Published genetic correlations between feed intake and performance traits of weaning and yearling weight are limited; however, these few published correlations (Arthur et al., 2001a; Bouquet et al.,

2001) indicate a strong correlation between DDMI to weaning and yearling weights. This strong correlation coupled with a weighted heritability of 0.41 was evident in the rapid decline in DDMI, WWT and YWT for $S_{DDMI.}$

The simulation protocol for the selection for decreased RFI TBV demonstrated a less severe decrease in DDMI and YWT compared to S_{DDMI}. The decrease in DDMI for S_{RFI} was an average annual decrease of 0.19kg/year which was less of a decline than what was seen with S_{DDMI}. A decrease in weaning and yearling weights were also seen with an average annual decline of 1.64 kg/year and 5.13 kg/year, respectively. From year 1 to year 10, there was a decrease of 45.60 kg for average progeny yearling weight. This loss in progeny yearling weight was not as drastic as S_{DDMI} and as a result, there was minimal change in progeny numbers from year to year. When comparing S_{DDMI} and S_{RFI}, S_{RFI} suggests a better selection scenario for a production operation; however, for operations selling cattle based on yearling weight, there would still be a potential loss in profit given the decrease in YWT with only slight decreases in DDMI.

The popularity of RFI as a measure of feed efficiency has increased since it was proposed by Koch et al. (1963). The increased interest in RFI has been due to the zero phenotypic correlation of RFI to performance traits included in the calculation of RFI (Berry and Crowley, 2013). Typically, these performance traits include average daily gain, metabolic weight and back fat thickness but has no consideration of performance traits that most cattle are valued at, such as actual live or carcass weights. Herd et al. (2003) argued that an opportunity existed for the improvement of whole-herd production for efficiency through exploiting genetic variation in RFI in addition to genetic correlations with other performance traits. However, the computation of RFI give no direct consideration for underlying genetic regressions (Kennedy et al., 1993) and

there are genetic correlations to performance traits which would over time, affect the performance of those traits. Simulation 2 demonstrated the effect of genetic correlations between RFI and performance traits WWT and YWT leading to a decrease in performance for both weight traits. For beef cattle production systems whose end product is dependent on cattle weight, selection for RFI may not be optimal for these operations.

The final simulation (S_{ESI}) used economic index values for selection. Economic weights were placed on the TBV of DDMI, ADG, WWT and YWT with selection of replacement animals dependent on the animal's ranking of index values. An increase for DDMI, ADG, WWT and YWT was observed. An average annual increase of 0.42 kg/year occurred for DDMI with an increase of 3.81 kg from the progeny average from year 10 compared to year 1. Both weaning and yearling weight also increased by 7.11 kg/year and 14.84 kg/year, respectively. In contrast to S_{DDMI} and S_{RFI}, S_{ESI} was the only simulation that resulted in a gain in output traits such as WWT and YWT, but it consequently resulted in an elevation of DDMI.

When defining feed efficiency as a gain in output traits, such as WWT and YWT, while simultaneously reducing input traits (DDMI), S_{ESI} appears to not have met this requirement since there was an increase in DDMI. Simulations S_{DDMI} and S_{RFI} also failed to meet this requirement given that WWT and YWT both decreased. The implementation of a restricted selection index (Gibson and Kennedy, 1990) could be utilized to meet the requirements for the definition of feed efficiency by maintaining a genetic gain for DDMI to zero. A restricted selection index was constructed to maximize improvement in output traits (i.e. WWT and YWT) while limiting genetic change on an input trait (i.e. DMI; Eisen, 1997). Gibson and Kennedy (1990) concluded that constrained indexes would not be ideal economically as they would cause severe economic losses. The response to selection for an economic index was determined by the economic merit

of the traits included in the index. Within a restricted index, the economic merit was also restricted resulting in a loss of potential economic gain (Gibson and Kennedy, 1990).

To explore the economic merit for all 3 simulations, the average yearling price was estimated using the 10-year average cattle price and feed cost in presented in table 4.3. In addition, a 10-year average for days to finish was also estimated at 157 d (Focus on Feedlots Monthly Reports, https://www.asi.k-state.edu/about/newsletters/focus-on-feedlots/monthly-reports.html). Total feed intake cost for a finishing phase was estimated by multiplying the average DDMI to 157 days to finish to the total ration cost of $0.18/kg. An estimated selling price of cattle was calculated by using the yearling weight multiplied to the estimated 10-year average cattle price of $2.64/kg. The feed cost was subtracted from the estimated cattle price. The differences in profits were presented in figures 4.2 and 4.3. Figure 4.2 illustrates the profit or loss on a per head basis using the annual progeny average for each year. Figure 4.3 uses the same profit or loss as presented in figure 4.2 but also included the number of progenies for each year.

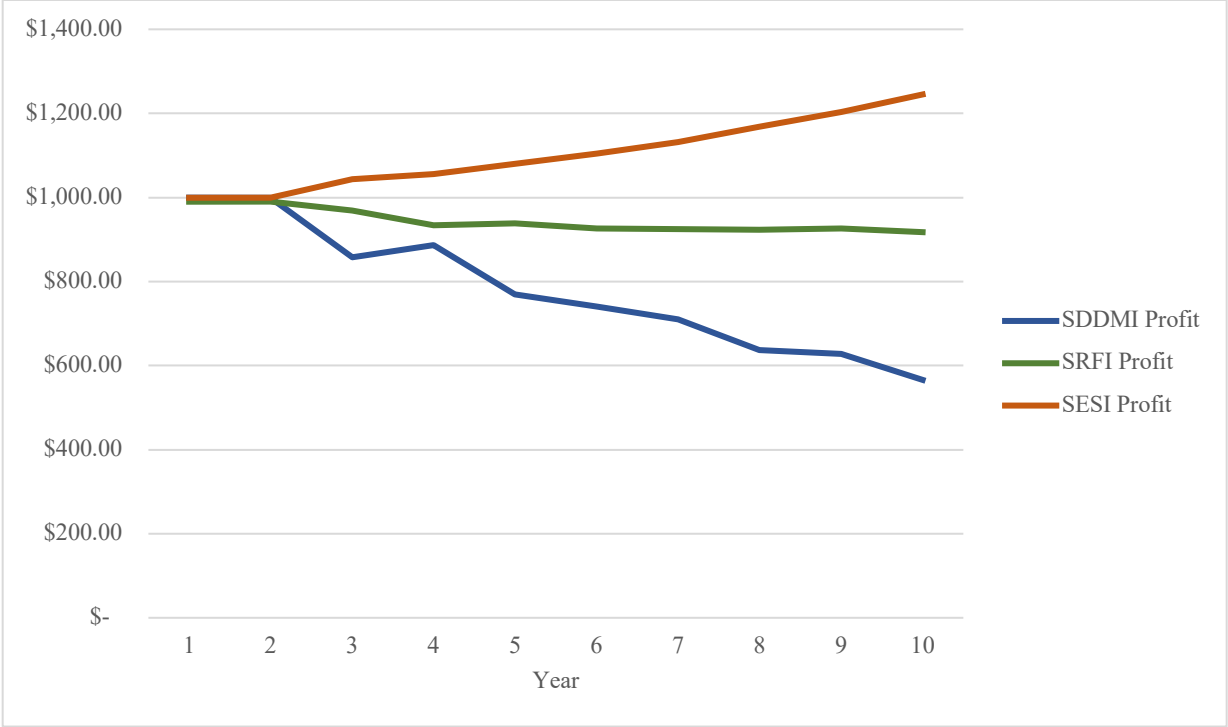

Figure 4.3. The average per head progeny profit or loss for three simulation scenarios. S_{DDMI} was a simulation for the single trait selection on average daily dry matter intake, S_{RFI} was the simulation for selection on residual feed intake, and S_{ESI} is the simulation for use of an economic index for yearling weight and feed intake.

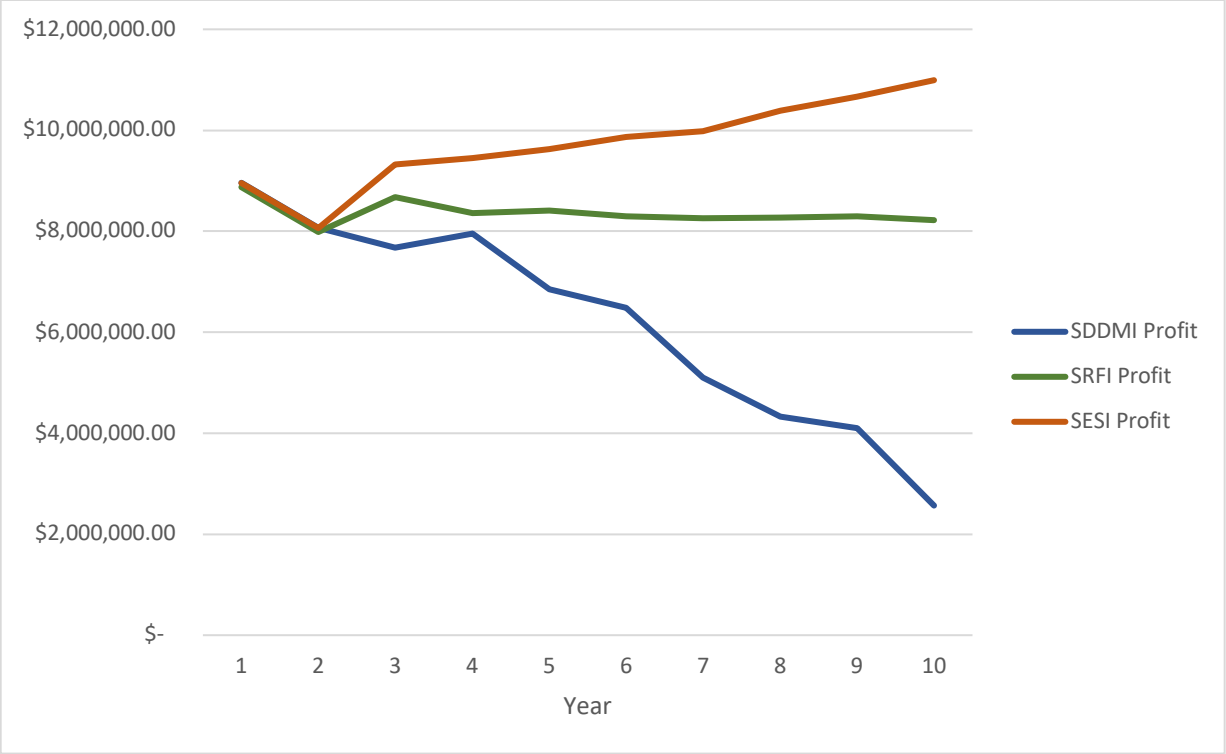

Figure 4.4 The profit or loss for the progeny from three simulation scenarios. S_{DDMI} was a simulation for the single trait selection on average daily dry matter intake, S_{RFI} was the simulation for selection on residual feed intake, and S_{ESI} is the simulation for use of an economic index for yearling weight and feed intake.

The comparison of the profitability on a per head basis of all three simulations was presented in table 4.7. The economic index simulation (S_{ESI}) was the only simulation that resulted in an increase of output (yearling weight) resulting in an increase in the average income from cattle prices. Simulation 3 also resulted in an increase in input costs (DDMI), the increase in yearling weight, and subsequent increase in average income from cattle prices, was significant enough to increase the profitability even with an increase in input costs. The results from S_{ESI} would support the recommendation of Nielsen et al. (2013) to incorporate feed intake into national cattle evaluations as an economically relevant trait within a selection index. Simulation 1 resulted in the largest decrease in DDMI, but it simultaneously resulted in a decrease in YWT. On a per head basis, the economic loss of S_{DDMI} was the largest compared to the other

simulations. When considering the entire population of simulated cattle, the economic loss of S_{DDMI} was even more drastic given the significant loss of progeny after year 5. The simulation for the single trait selection of RFI also showed a negative profitability after 10 years of selection but the degree of the loss was not as drastic as S_{DDMI}. Although S_{RFI} did decrease DDMI there was also a decrease in WWT and YWT through the genetic correlation of RFI to these traits, resulting in an overall economic loss for this simulation.

Table 4.8. Changes in estimated costs and incomes from the simulation of selection on average daily dry matter intake (S_{DDMI}), residual feed intake (S_{RFI}) or an economic index (S_{ESI}) on the average per head basis from year 1 to year 10.

	S_{DDMI}	S_{RFI}	S_{ESI}
DDMI [a]	$ (200.48)	$ (47.96)	$ 107.78
Cattle [b]	$ (634.08)	$ (120.38)	$ 353.23
Year 1 [c]	$ 999.60	$ 989.90	$ 999.60
Year 10 [d]	$ 566.00	$ 917.47	$ 1,245.05
Profit [e]	$ (433.60)	$ (72.43)	$ 245.45

DDMI [a] change in cost of feed from year 1 to year 10 for simulated data.
Cattle [b] was the change income from averaged per head cattle prices from year 1 to year 10 for simulated data.
Year 1 [c] was the difference from average cattle per head sale price and feed cost for year 1 of simulated data.
Year 10 [d] was the difference from average cattle per head sale price and feed cost for year 10 of simulated data.
Profit [e] was the change in profit from year 1 to year 10.
Values within parenthesis are negative dollar amounts.

4.4 Conclusion

Simulated data were used to illustrate the effects of selection for decreased feed intake and the response to this selection on correlated weight traits. Single trait selection for decreased DDMI or RFI resulted in decreased performance in output traits of WWT and YWT. As a consequence of decreased WWT and YWT, the overall production of these scenarios decreased. Although DDMI was reduced and therefore reduced production cost, this savings in cost was not important enough to account for the loss in revenue from decreased cattle weight. The third selection scenario used economic weights to account for the cost of production and revenue from cattle. Although simulation 3 resulted in an increase in DDMI, there were also increases in WWT and YWT resulting in higher revenues. Residual feed intake has increased in popularity as a feed efficiency trait that accounts for production traits, such as ADG. Although RFI has a zero phenotypic correlation to production traits, this study illustrated how selection for lowering RFI breeding values could affect production traits such as WWT and YWT and ultimately affect profitability of the herd. The results of this study also illustrated the application of an economic index as a method of multi-trait selection to increase revenue by increasing YWT but account for the cost of DDMI therefore increasing the profitability of an operation.

A weakness of this study was its dependence of published genetic and phenotypic estimates for RFI and DDMI and their correlations to production traits. The majority of published correlations with DDMI and RFI are to traits measured during feeding trails for measuring intake (i.e. metabolic mid-weight, average daily gain). Overall, there was a general lack of published genetic parameters of feed efficiency and its correlation to production traits, such as yearling weight. To improve production efficiency, a better understanding of the genetic

relationship between input traits (DDMI) and output traits (YWT) is required in order to select cattle that can improve production efficiency and therefore the overall profitability and sustainability of beef operations.

LITERATURE CITED

Anderson, R. V., R. J. Rasby, T. J. Klopfenstein, and R. T. Clark. 2005. An evaluation of production and economic efficiency of two beef systems from calving to slaughter. J. Anim. Sci. 83: 694-704.

Archer, J. A., P. F. Arthur, R. M. Herd, P. F. Parnell, and W. S. Pitchford. 1997. Optimum postweaning test for measurement of growing rate, feed intake, and feed efficiency in British breed cattle. J. Anim. Sci. 75: 2024-2032.

Arthur, P. F., J. A. Archer, D. J. Johnson, R. M. Herd, E. C. Richardson, and P. F. Parnell. 2001a. Genetic and phenotypic variance and covariance components for feed intake, feed efficiency and other postweaning traits in Angus cattle. J. Anim. Sci. 79:2805–2811

Arthur, P. R., J. A. Archer, R. M. Herd, and G. J. Melville. 2001b. Response to selection for net feed intake in beef cattle. Proc. Assoc. Advmt. Anim. Breed Genet. 14: 135-138.

Barwick, S. A., M. L. Wolcott, D. J. Johnston, H. M. Burrow and M. T. Sullivan. 2009. Genetics of steer daily and residual feed intake in two tropical beef genotypes, and relationships among intake, body composition, growth and other post-weaning mea- sures. Anim. Prod. Sci. 49:351–366.

Berry, D. P. and J. J. Crowley. 2013. Cell Biology Symposium: Genetics of feed efficiency in dairy and beef cattle. J. Anim. Sci. 91: 1594-1613.

Berry, D. P. and J. E. Pryce. 2013. Feed efficiency in growing and mature animals. Proc. 10th World Congr. Genet. Appl. to Livest. Prod., Vancouver, Canada.

Bouquet, A., M.-N. Fouilloux, G. Renand, and F. Phocas. 2010. Genetic parameters for growth, muscularity, feed efficiency and carcass traits of young beef bulls. Livest. Sci. 129:38–

48.

Bourdon, R.M. 1997. Understanding Animal Breeding. Prentice-Hall. p 280-281.

Ceacero, T. M., M. E. Z. Mercadante, J. N. S. G. Cyrillo, R. C. Canesin, S. F. M. Bonilha, L. G. de Albuquerque. 2016. Phenotypic and genetic correlations of feed efficiency traits with growth and carcass traits in Nellore cattle selected for postweaning weight. PLOS One 11:1-11.

Costa, R. B. et al. 2011. Estimation of genetic parameters for mature weight in Angus cattle. J. Anim. Sci. 89:2680-2686.

Crowley, J. J., M. McGee, D. A. Kenny, D. H. Crews Jr., R. D. Evans, and D. P. Berry. 2010. Phenotypic and genetic parameters for different measures of feed efficiency in different breeds of Irish performance tested beef bulls. J. Anim. Sci. 88:885–894.

Durunna, O. N., G. Plastow, F. D. N. Mujibi, J. Grant, J. Mah, J. A. Basarab, E. K. Okine, S. S. Moore, and Z. Wang. 2011. Genetic parameters and genotype × environment interaction for feed efficiency traits in steers fed grower and finisher diets. J. Anim. Sci. 89:158–167.

Eisen, E. J. 1997. Restricted selection index: An approach to selecting for feed efficiency. J. Anim. Sci. 44:958-972.

Fan, L. Q., D. R. C. Bailey, and N. H. Shannon. 1995. Genetic pa- rameter estimation of postweaning gain, feed intake and feed efficiency for Hereford and Angus bulls fed two different diets. J. Anim. Sci. 73:365–372.

Gibson, J. P. and B. W. Kennedy. 1990. The use of constrained selection indexes in breeding for economic merit. Theor. Appl. Genet. 80:801-805.

Herd, R. M., and S. C. Bishop. 2000. Genetic variation in residual feed intake and its association with other production traits in British Hereford cattle. Livest. Prod. Sci. 63:111–119.

Herd, R. M., J. A. Archer, and P. F. Arthur. 2003. Reducing the cost of beef production through genetic improvement in residual feed intake: Opportunity and challenges to application. J. Anim. Sci. 81:E9-E17.

Hoque, M. A., P. F. Arthur, K. Hiramoto, and T. Oikawa. 2006. Genetic relationship between different measures of feed ef- ficiency and its component traits in Japanese Black (Wagyu) bulls. Livest. Sci. 99:111–118.

Hoque, M. A., M. Hosono, T. Oikawa, and K. Suzuki. 2009. Genetic parameters for measures of energetic efficiency of bulls and their relationships with carcass traits of field progeny in Japanese Black cattle. J. Anim. Sci. 87:99–106.

Investing.com www.investing.com/commodities/live-cattle-historical-data (Accessed December 18, 2018)

Iowa State University Extension and Outreach. http://www.iowabeefcenter.org/calculators.html (Accessed November 25, 2018)

Iwaisaki, H., S. Tsuruta, I. Misztal, and J. K. Bertrand. 2005a. Estimation of correlation between maternal permanent environmental effects of related dams in beef cattle. J. Anim. Sci. 83:537-542.

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

Kennedy, B. W., J. H. J. van der Werf and T. H. E. Meuwissen. 1993. Genetic and statistical properties of residual feed intake. J. Anim. Sci. 71:3239-3250.

Koch, R. M., L. A. Swiger, D. Chambers and K. E. Gregory. 1963. Efficiency of feed use in beef cattle. J. Anim. Sci. 22:486-494.

Koots, K. R., J. P. Gibson, C. Smith, and J. W. Wilton. 1994a. Analyses of published genetic parameters estimates for beef production traits. 1. Heritability. Anim. Breed. Abstr. 65:309-338.

Koots, K. R. J. P. Gibson, and J. W. Wilton. 1994b. Analyses of published genetic parameter estimates for beef production traits. 2. Phenotypic and genetic correlations. Anim. Breed. Abstr. 62:825-853.

Korver, S., E. A. M. van Eekelen, H. Vos, G. J. Nieuwhof, and J. A. M. van Arendonk. 1991. Genetic parameters for feed intake and feed efficiency on growing dairy heifers. Livest. Prod. Sci. 29:49–59.

Lancaster, P. A., G. E. Carstens, D. H. Crews Jr., T. H. Welsh Jr., T. D. A. Forbes, D. W. Forrest, L. O. Tedeschi, R. D. Randel, and F. M. Rouquette. 2009. Phenotypic and genetic relationships of residual feed intake with performance and ultrasound carcass traits in Brangus heifers. J. Anim. Sci. 87:3887–3896.

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

Meyer, K. and S. P Smith. 1996. Resitricted maximum likelihood estimation for animal models using derivatives of the likelihood. Genet. Sel. Evol. 28:23-49.

Mujibi, F. D. N. , J. D. Nkrumah, O. N. Durunna, P. Stothard, J. Mah, Z. Wang, J. Basarab, G. Plastow, D. H. Crews, Jr. and S. S. Moore. 2011. Accuracy of genomic breeding values for residual feed intake in crossbred beef cattle. J. Anim. Sci. 89:3353-3361.

Mrode, R. A. 2005. Linear Models for the Prediction of Animal Breeding Values, 2nd Edition. CABI Publishing. p 98-99.

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

Nkrumah, J. D., J. A. Basarab, Z. Wang, C. Li, M. A. Price, E. K. Okine, D. H. Crews, Jr., and S. S. Moore. 2007. Genetic and phenotypic relationships of feed intake and measures of efficiency with growth and carcass merit of beef cattle. J. Anim. Sci. 85:2711-2720.

Phocas, F., and D. Laloe. 2004. Genetic parameters for birth and weaning traits in French specialized beef cattle breeds. Livest. Prod. Sci. 89:121-128.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Robinson, D. L., and V. H. Oddy. 2004. Genetic parameters for feed efficiency, fatness, muscle area and feeding behavior of feedlot finished beef cattle. Livest. Prod. Sci. 90:255-270.

Rolfe, K. M., W. M. Snelling, M. K. Nielsen, H. C. Freetly, C. L. Ferrell, and T. G. Jenkins. 2011. Genetic and phenotypic parame- ter estimates for feed intake and other traits in growing beef cattle, and opportunities for selection. J. Anim. Sci. 89:3452–3459.

Roughsedge, T., P. R. Amer, R. Thompson, and G. Simm. 2005. Genetic parameters for a maternal breeding goal in beef production. J. Anim. Sci. 83:2319-2329.

Schenkel, F. S., S. P. Miller, and J. W. Wilton. 2004. Genetic parameters and breed differences for feed efficiency, growth and body composition traits of young beef bulls. Can. J. Anim. Sci. 84:177–185.

Schaeffer, L. R Codes for Multiple Trait Models http://www.aps.uoguelph.ca/~lrs/Summer2012Full/MTiter.R, (Accessed September 15,

2018).

Steele, R. G. and J. H. Torrie. 1980. Principles and procedures of statistics. McGraw-Hill, New York, p 188-190.

USDA, Economic Research Service with data from USDA Agricultural Marketing Service and Grain and Feed Market News. https://marketnews.usda.gov/mnp/ls-report-config. (Accessed November 25, 2018).

USDA, National Agricultural Statistics Service. https://www.nass.usda.gov/Charts_and_Maps/graphics/data/pricecn.txt (Accessed November 25, 2018).

Van Vleck, L. D. and L. V. Cundiff. 1998. Sex effects on breed of sire differences for birth, weaning, and yearling weights. J. Anim. Sci. 76:1528-1534.

Williams, Y. J., J. E. Pryce, C. Grainger, W. J. Wales, N. Linden, M. Porker, and B. J. Hayes. 2011. Variation in residual feed intake in Holstein Friesian dairy heifers in southern Australia. J. Dairy Sci. 94:4715–4726.

Chapter V

ESTIMATION OF FORAGE INTAKE FOR GRAZING BEEF CATTLE USING EAR TAG ACCELEROMETER TECHNOLOGY

5.1 Introduction

Producing efficient beef cattle is a goal for animal breeders to meet the increasing demands of the world's growing population (9.7 billion by 2050; FAO, 2009) that relies on limited resources for food production. In addition, feed intake has been shown to be positively correlated to greenhouse gas emission from beef cattle (0.65 ± 0.02; Herd et al., 2014) and reducing the overall feed intake of the cow herd while maintaining production through genetic selection would reduce greenhouse gas emissions. Given that cattle finished in a feedlot spend 50 to 70% of their lifespan grazing forage prior to feedlot entry (Capper, 2011), a reduction of forage intake would likely contribute to the reduction of greenhouse gases. Furthermore, decreasing forage intake on rangeland has the potential of improving pasture quality. When cattle require less feed for production, stress placed on pastures due to drought, overgrazing or climate change could be mitigated.

Currently, individual feed intake measurements are collected in feedlot environments. The application of feedlot-measured intakes has a direct application to fed cattle, but the translation of these intakes to a cowherd grazing on rangeland is unknown. Feed costs are the largest variable expense to producers in the beef industry (Anderson et al., 2005) and

approximately 50% of which are attributed to the mature cow herd (Whisnet, USDA-NIFA-CRIS) with 70 to 75% of the total annual energy for maintenance (Ferrell and Jenkins, 1985). Selection for beef cattle with decreased feed costs without sacrificing production, would increase profitability for producers. However, to select for decreased feed intake, individual feed intake must be measured. To date, there is no technology to easily measure feed intake on a population of grazing cattle and the relationship between feed intake measured in a feedlot and intake from grazing cattle are unknown. Therefore, the objectives of this study were 1) to examine the use of an ear tag accelerometer (CowManger; Agis Automatisering BV, Harmelen, Netherlands) and corresponding data to develop a proxy for grazing intakes 2) to validate the data collected by the ear tag accelerometer, 3) to explore potential interference from other technologies, such as GrowSafe, with the ear tag accleerometer, 4) to examine the phenotypic correlation between grazing and feedlot intake and 5) to validate the use of a biomarker for the estimation of DMI.

5.2 Materials and Methods

Animals and Procedures. Institutional Animal Care and Use Committee at Colorado State University approved all animal procedures (approval number 17-7179A). Yearling Angus steers (n = 98) from the Colorado State University Beef Improvement Center (BIC) were placed in the Feed Intake Unit (FIU) at Colorado State University's Agriculture Research, Development and Education Center (ARDEC) located north of Fort Collins, CO mid-April 2017. The Feed Intake Unit was equipped with GrowSafe Feed Intake monitoring system (GrowSafe Systems, Ltd., Airdire, AB, Canada) that contains 6 pens capable of housing up to 35 animals per pen with 4 GrowSafe nodes per pen. Upon arrival, steers were equipped with an electronic ear tag (EID;

Allflex USA Inc., Dallas TX) for individual identification of animals by GrowSafe. In addition, cattle were also equipped with a CowManager ear tag accelerometer (CME) to monitor eating, rumination and activity behavior (Agis Automatisering BV, Harmelen, The Netherlands).

The CowManager ear tag is a 3-deminsional accelerometer that attaches to an EID and was placed in the center of the animal's ear. Based on the ear and jaw movement of cattle, the CME has developed a proprietary algorithm to calculate time spent eating, ruminating, resting and activity within each hour. These data were collected by CME and transmitted to a router with a 1,524 m radius. From the router, the information is sent to a coordinator connected to a laptop with an internet connection located at the research center. The raw data is transmitted to Agis Automatisering BV, Netherlands, where their algorithms transform the data for each behavior. Final behavioral data was downloaded to researchers each day in addition to the CowManager online application. Five behaviors were recorded by CME: eating (EAT), rumination (RUM), not active (NACT), active (ACT) and high active (HACT). The behavior of EAT measured the amount of time the animal spent consuming feed and RUM measured the amount of time when regurgitation occurred. Three levels of activity were observed by CME: NACT, ACT and HACT. The activity level of NACT was measured as the time the animal was inactive and not eating or ruminating. Active was measured as a low level of activity such as walking short distances, scratching or licking and HACT was measured as a higher level of activity such as mounting. Since each minute within an hour was allocated to each behavior, the summation of the reported behaviors within each hour was 60 minutes with no overlapping of behavioral times. The behavior times were summed to account for a 24-hour period. Times that did not sum to 1440 minutes (24 hours) were removed from the study.

Steers were weighed before entry into the FIU for a 21-d adaptation period, followed by a shortened 54-d performance test to measure feed intake in accordance with findings from Culbertson et al., (2015). Cattle were weighed on d 0, 14, 28, 42 and 54 with individual feed intakes collected by the GrowSafe Feed Intake monitoring system. Cattle were fed ration *ad libitum.* The ration (Table 5.1) consisted of NE_g of 51.50 mcal/cwt and a CP of 14.87%.

Based on DMI, the 15 lowest intake and 15 highest intake animals were identified. These low/high intake (LHI) animals were used to quantify grazing intake using the biomarker titanium dioxide (TiO_2). It should be noted that DMI was not adjusted in any way for weight or body size since the objective of this study was to identify an indicator for intake in a grazing setting.

Table 5.1 Composition of rations fed to cattle in Colorado State University Feed Intake Unit on an as fed basis.

Percentage	Ingredient
9.77%	Alfalfa Hay
38.00%	Corn Silage
8.00%	Wheat Straw
25.00%	Corn Dry Grain
14.48%	Distillers Grain
0.89%	Limestone
0.10%	Salt
3.75%	30% Hay Treat

Immediately following the performance test in the FIU, cattle were transported on July 1, 2017 to Colorado State University Eastern Colorado Research Center (ECRC) in Akron, CO where the cattle were placed on a 180 acre pasture comprised of western wheatgrass (*Pascopyrum smithii)*, sand bluestem (*Andropogon hallii*), blue gamma (*Bouteloua gracilis*), and prairie sand reed (*Calamovilfa longifolia*). Cattle had a 1-week adaptation period when placed on

pasture before data was collected using CME beginning on July 8, 2017. Cattle were maintained on pasture for a total of 43 days and were removed on August 13, 2017. The first router for CME was placed at the pasture's water source and the second router placed 957 m from the water source and 852 m from ECRC's main office where the coordinator and laptop were located in ECRC's main office (Figure 1; Google Earth, 2019).

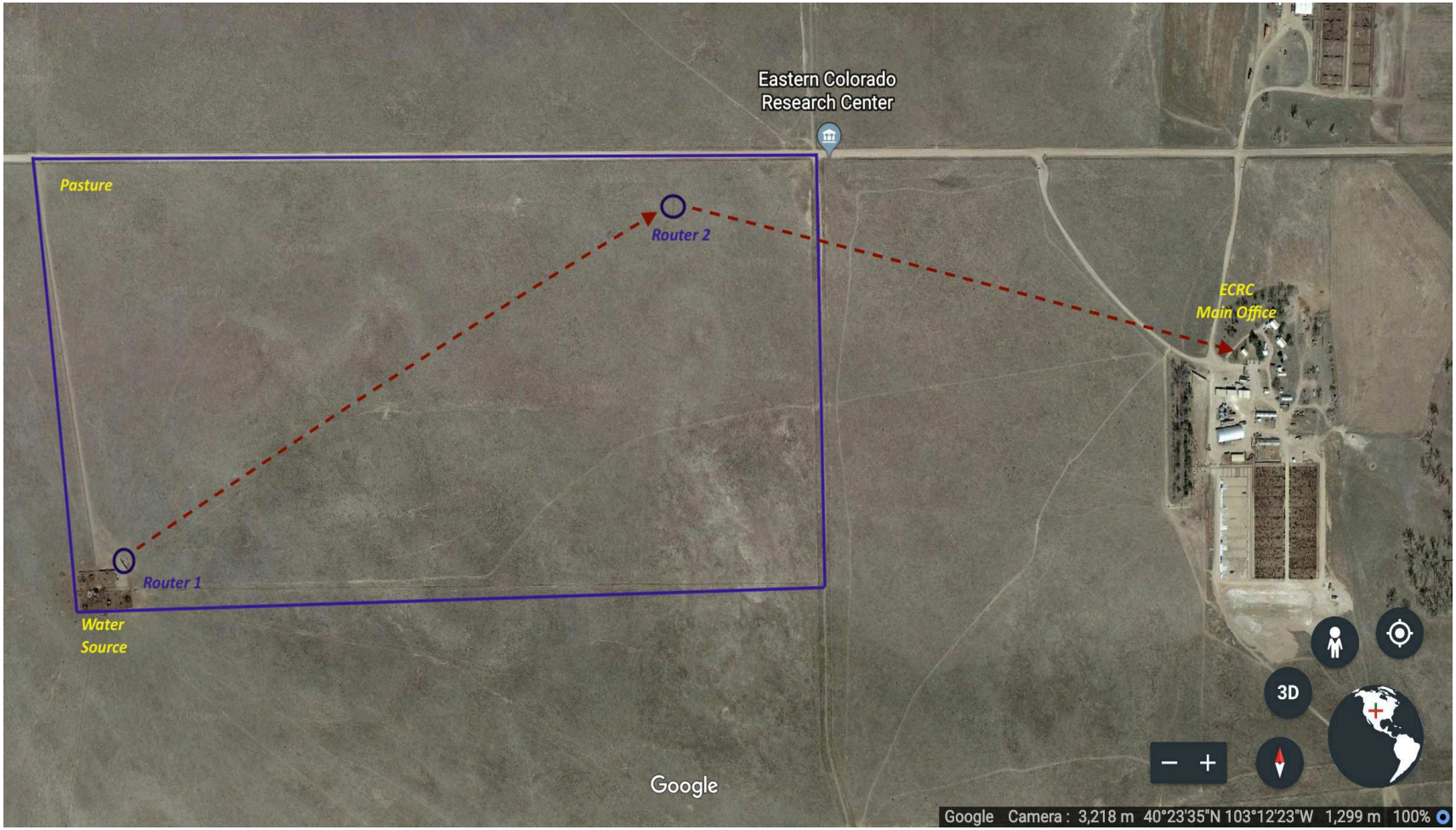

Figure 5.1. Picture of Colorado State University's Eastern Colorado Research Center diagraming the positioning of CowManager routers (Google Earth, 2019).

In order to quantify grazing intake, a TiO_2 biomarker was used (Meyer et al., 2004; Titgemeyer et al., 2001). For the final 20 d of maintaining cattle on pasture, the subset of LHI cattle (n = 30) were administer a 10 g of TiO_2 bolus each morning. It was required to bolus individual steers daily to ensure the adequate daily dose of TiO_2 was received. Each bolus contained 10 g of powdered TiO_2 within a gelatin capsule 7 cm in length and 1 cm radius (Torpac; Fairfeild, NJ). Steers were adapted to the TiO_2 for 14 d prior to fecal collections. Following the adaptation period, rectal fecal collections were conducted twice daily for 6 consecutive days with collections occurring 12 h apart. Every 24 h, time of collection was advanced 2 hours to minimize effects of diurnal variation (Meyer et al., 2004; Titgemeyer et al., 2001). As a result, each of the 30 steers sampled had 12 samples per steer.

Once fecal collections were completed, each fecal sample was placed in an aluminum pan and dried at 60°C in forced-air ovens for 48 h to 72 h. Once dry, samples were ground into fine particles using coffee bean grinders. For the 12 samples per animal, 10% of the dried weight of each sample per animal was combined into one final composite sample for each individual animal (12 samples were combined into 1 composite sample per animal). The final composite samples were analyzed for TiO_2 dioxide concentrations. Using the method proposed by Theurer (1996), grazing intake (GI) was estimated on a dry matter basis using equation 5.1:

$$GI = \left[{}^{FO}/_{DMD}\right] * 100 \qquad \text{(Eq. 5.1)}$$

In the above equation, GI was the dry matter grazing intake in kg/d, FO was fecal output in kg/d, and DMD was the *in vitro* dry matter disappearance of feed samples expressed as a percentage. The fecal output was determined using equation 5.2:

$$FO = {}^{DOSE}/_{FM}$$

(Eq. 5.2)

Where DOSE was the TiO_2 administered in g/d and FM was the fecal marker (mg/g).

Validation Study. On January 16, 2018, 128 steers from BIC were placed in the FIU. The processing and testing procedures for the FIU, as stated previously, were repeated. During processing, animals were randomly allocated to 4 pens within the FIU resulting in 32 steers per pen. Individual DMI was measured using the GrowSafe system and data was collected via CowManager. On day 30, the variance of DMI within each pen was estimated. The pen with the most variance in DMI was identified and used for the validation of TiO_2.

The validation of TiO_2 was conducted by the addition of 320 g of TiO_2 (10g per head per day) into the steers' daily ration. The TiO_2 was mixed into the steers' ration daily when feedstuffs were mixed by the feed truck each day. The cattle were fed TiO_2 for 14 days as an adaptation period followed by 6 days of fecal sample collection. Fecal samples were collected following the same procedures performed in the previous year. The estimated intakes from the TiO_2 concentrations were compared to the intakes measured using the GrowSafe system.

The validation of the CME for the measurement of feeding behaviors was also conducted. Over the course of 3 days, 4 to 7 steers were chosen each day at random for a total of 18 steers. The number of steers per day varied due to the number of observers present. These steers were marked with chalk for ease of recognition. Observers were assigned 1 to 2 steers per observer. All observers were trained to record observations of rumination and eating by watching videos of steers expressing said behaviors. Some animals were assigned to more than 1 observer to

validate the observers' ability to record observations. The observers recorded the animal's eating, rumination and activity level for 4 hours. Observations made by 2 observers watching the same steer were compared to insure consistency between observers. Correlations over all behaviors observed for 1 animal between 2 observers less than 0.95 resulted in the removal of the observations from those observers from the analysis. As a result, 5 steers with observations remained for validation. These observations were then compared to the measurements recorded by CME.

To explore potential interference of GrowSafe with CME, the latter was placed in 5 finishing steers located at ARDEC not housed in the FIU. These steers were located in a 10 head pen with a concrete feed bunk. Unlike the GrowSafe system, the concrete bunk required the animals to completely lower their head down to the ground to access feed. When the GrowSafe feed bunks were filled with feed, cattle feeding from the GrowSafe bunks were not required to lower their heads in order to reach feed and could have caused a potential interference with how CME recorded eating or rumination (Figure 1). The observation of these 5 finishing steers allowed for a comparison of observations collected from steers within the GrowSafe system and outside of GrowSafe to examine any possible electronic interference of GrowSafe with CME. These 5 steers were equipped with CME 7 days prior to observations. Steers were observed for 2 hours by two observers. Each observer was assigned 3 steers each with 1 steer being observed by both observers. As with the previous observations of steers in the FIU, correlations of behaviors for the single steer observed by both observers was used to validate the consistency of the observations made between observers. A correlation of 0.98 was achieved and therefore the observations of both observers for all steers were included. Observations and CME measurements were compared for those 2 hours.

Figure 5.2 Picture of GrowSafe System feed bunks (right side of picture) and concrete feed bunk (left side of picture) at Colorado State University's Feed Intake Unit.

Given that the CME were initially developed for use in dairy cattle and differences with monitoring systems have been observed between dairy and beef cattle (Goldhawk et al., 2013), to further investigate potential differences between dairy and beef cattle, 4 dairy cows located at a dairy utilizing CME, were also observed for 2 hours. Two observers selected 4 dairy cows and observed eating, rumination and activity levels for 2 hours. The observational period began shortly after the cows returned from the milking parlor. Observation and CME measurements were compared.

Statistical Analysis. A regression analysis was used to examine the association between DMI and CME measurements. Each variable was the observation averaged over the time period of measurement. For example, DMI was the total dry matter intake consumed divided by the

number of days intake was measured. Variables measured by CME were the total minutes measured divided by the days of measurement for each behavior (EAT, RUM, NACT, ACT and HACT). Using the step function in R program (R Core Team, 2018), a stepwise model selection procedure was performed using Akaike information criteria (AIC) as the criteria for variable selection. The stepwise procedure uses an iterative process of adding and removing explanatory variables to find the subset of variables that optimized the predetermined model criteria, which in this case was AIC. Three stepwise model procedures were executed. For all three procedures, DMI was the dependent variable. For the first procedure, EAT, RUM, NACT, ACT and HACT were all included as the explanatory variables in the full model and a stepwise model selection reduced the model from the full model (equation 5.3).

$$Y_{DMI} = \beta_0 + \beta_{EAT}X_{EAT} + \beta_{RUM}X_{RUM} + \beta_{NACT}X_{NACT} + \beta_{ACT}X_{ACT} + \beta_{HACT}X_{HACT} + e \quad \text{(Eq. 5.3)}$$

where Y_{DMI} was the dependent variable of DMI, β_0 was the intercept, X_i was the explanatory variable for EAT, RUM, NACT, ACT or HACT, β_i was the regression coefficient for each explanatory variable and e is the model error term.

The second stepwise model selection procedure explored protentional higher order variables by including a quadratic term for each explanatory variable from the previous full model. For the final stepwise procedures, interactions for each explanatory variable were included in the full model.

The method used by CME for allocating time for each behavior creates a dependency for each behavior since the summation of combined behavioral time would have to equal 1440

minutes of a 24-hour period. As such, a principle component analysis (PCA) in addition to a principle component regression (PCR) was also performed. A principle component analysis was performed including EAT, RUM, NACT, ACT and HACT. The resulting principle components were used as regressors on DMI for the PCR.

5.3 Results and Discussion

Tables 5.2 and 5.3 present the summary statistics of the steers at the conclusion of the 2017 performance test. The summary statistics for subset of LHI animals were presented in table 5.4.

Table 5.2. Summary statistics for cattle (n=98) at the conclusion of the 2017 feed intake performance test.

	Age (d)	Weight (kg)	ADG (kg/d)	DMI (kg)
Mean	479	428.95	3.35	10.95
SD	21	58.27	0.52	1.72
Min	419	285.76	1.86	6.65
Max	523	555.65	4.27	14.72

Table 5.3. Summary statistics for CowManager ear tag accelerometer measurements of time (in minutes) for eating (EAT), rumination (RUM), not active (NACT), active (ACT) and high active (HACT) on Angus cattle (n=98) during the 2017 feed intake performance test.

	EAT	**RUM**	**NACT**	**ACT**	**HACT**
Mean	112.81	348.40	471.80	170.36	326.37
SD	39.78	54.02	62.98	38.77	49.45
Min	32.43	140.68	314.11	111.65	230.49
Max	246.67	514.80	641.52	322.86	437.67

Table 5.4. Summary statistics for the 15 lowest and highest intake animals, based on GrowSafe measures from the 2017 feed intake performance test.

	Low Intake Steers (n=15)			High Intake Steers (n=15)		
	Age (d)	**Weight (kg)**	**DMI (kg)**	**Age (d)**	**Weight (kg)**	**DMI (kg)**
Mean	465.67	377.09	8.54	481.07	483.23	12.83
SD	20.83	58.58	0.82	13.77	64.23	0.41
Min	419	285.76	5.95	452	349.27	12.18
Max	493	489.88	9.19	497	553.38	13.48

For the initial 2017 performance test, dry matter intake and measurements from CME were averaged over the 54-d feeding trial. Pearson's correlations for averaged DMI and CME measurements are presented in table 5.5. The correlations of CME measurements to DMI ranged from -0.11 to 0.12. The correlation between DMI to EAT was 0.09 (P = 0.36). Correlations for feeding time and DMI have been previously reported as moderate. Robinson and Oddy (2004) and Nkrumah et al. (2007) reported correlations of DMI to feeding time of 0.30 and 0.27, respectively. For both of these studies, eating time was measured with the same monitoring system that measured DMI. In this study, feeding time was measured by two different systems (CME and GrowSafe) which resulted in a low correlation. This low correlation is also in contradiction to Wolfger et al. (2015) which concluded that ear tag accelerometers were highly sensitive for measuring feeding time. In the study conducted by Wolfger et al. (2015),

investigators visually observed 18 yearling Hereford x Angus feedlot steers to validate the use of CME in beef cattle for the measurement of time spent eating and ruminating. Their conclusion of the sensitivity of CME for measuring feeding time resulted from 0.79 concordance correlation between observed time spent eating and CME measured time for eating. Cattle for the Wolfger et al. (2015) were slick bunk fed compared to the current study where steers were fed *ad libitum* from GrowSafe feed bunks. This difference in how cattle were fed was explored further the second year of the study.

To the author's knowledge, there are no reported studies examining the relationship between DMI and time spent ruminating in beef cattle; however, a few studies have been conducted in dairy cattle correlating DMI to rumination time. In studies by Krause et al. (2002), Schirmann et al. (2010) and Byskov et al. (2017), correlations between rumination time and DMI ranged from -0.28 to 0.61. Schirmann et al. (2010) and Byskov et al. (2017) both concluded that no relation between rumination time and DMI was evident with their study and the use of rumination time as an indicator of DMI was limited. Our study would support Schirmann et al. (2010) and Byskov et al. (2017) conclusions with a Pearson's correlation of -0.09 ($P = 0.39$) suggesting minimal relation between rumination time and DMI. Krause et al. (2002) estimated correlations between rumination time and DMI that ranged from -0.26 to 0.61. The variation of correlations was found to be dependent on the forage particle size. As particle size increased so did the correlation of rumination time to DMI indicating the use of rumination time as an indicator of DMI would be highly dependent on the ration and amount of forage included in the ration. This dependency of ration would be problematic for the use of rumination time for an indicator of DMI when comparing animals in a feedlot to those grazing rangeland.

The correlation between DMI with ACT and HACT were of the highest magnitude at -0.11 and 0.12, respectively; however, neither of these correlations were significant with *p*-values greater than 0.05. Although there were no significant ($P > 0.05$) correlations between DMI and CME measurements, correlations within the CME measurements were significant. The highest correlation estimate was that between RUM and NACT (-0.66, $P < 0.05$). The significant correlations between CME measurements were anticipated given how the accelerometer allocates each minute to a specific behavior.

Table 5.5. Pearson's correlations for CowManager ear tag accelerometer measurements of eating (EAT), rumination (RUM), not active (NACT), active (ACT), high active (HACT) and dry matter intake (DMI) measured by a GrowSafe system.

	DMI	**EAT**	**RUM**	**NACT**	**ACT**	**HACT**
DMI	1.00	0.09	-0.09	0.03	-0.11	0.12
EAT		1.00	0.14	-0.32	-0.38	-0.21
RUM			1.00	-0.66	-0.56	0.08
NACT				1.00	0.39	-0.55
ACT					1.00	-0.32
HACT						1.00

The results from the stepwise regression procedure for the full model found no significant CME variables as explanatory variables for feedlot DMI for any reduced models. This would be expected given the low correlations of CME measurements and DMI. The model with the lowest AIC was the “null” model which consisted of DMI and the intercept only (the average of DMI). The stepwise procedure performed for models including quadratic and interaction terms for explanatory variables found no significant variables and both resulted in the “null” model with the lowest AIC.

A principle component analysis (PCA) in addition to a principle component regression was also performed. Both analyses resulted in no significant variables for the estimation of DMI

using CME measurements (Figure 1). The principle component regression resulted in an adjusted R^2 of 0.01 in addition to none of the principle components being found to be significant explanatory variables for DMI with p-values above 0.05.

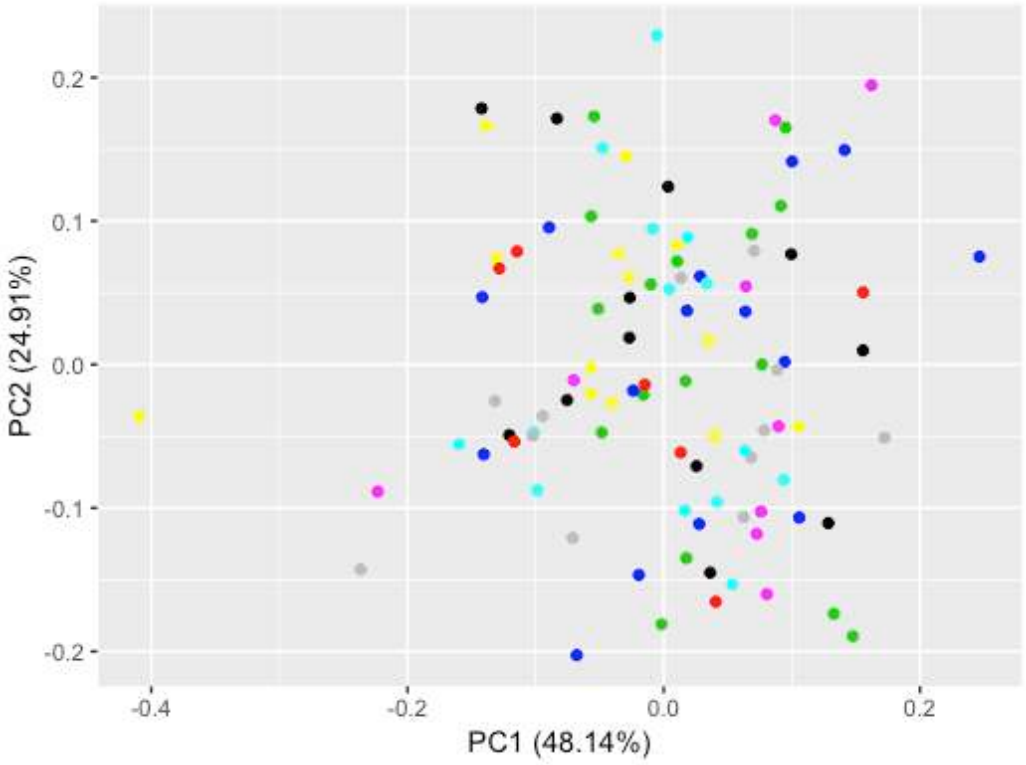

Figure 5.3. Plot from the principle component analysis for CowManager ear tag accelerometer measurements of eating (EAT), rumination (RUM), not active (NACT), active (ACT), high active (HACT).

To further investigate the potential relationship between DMI and CME measures, similar correlation analyses were performed on the LHI steers for the data collected while the steers were in the FIU. Table 5.6 presents the correlations for the LHI steers. The correlation between DMI and RUM was -0.39 (P = 0.04) for LHI steers which was larger than the same correlation from the entire cohort of steers measured in the FIU. However, the interpretation of these

correlations should be done with caution as they result from a subset of steers who represented the extremes of the distribution for DMI.

Table 5.6. Pearson's correlations for CowManager ear tag accelerometer measurements of eating (EAT), rumination (RUM), not active (NACT), active (ACT), high active (HACT) and dry matter intake (DMI) measured by a GrowSafe system for a subset of cattle with the highest and lowest intake values.

	DMI	EAT	RUM	NACT	ACT	HACT
DMI	1.00	0.00	-0.39	0.25	0.09	0.00
EAT		1.00	-0.25	0.13	-0.86	-0.12
RUM			1.00	-0.44	-0.03	-0.26
NACT				1.00	-0.34	0.02
ACT					1.00	-0.02
HACT						1.00

Once the feeding trial was completed in the FIU, animals were transported to ECRC and placed on pasture where rectal fecal samples were taken to estimate grazing intake. The average estimated grazing DMI from theTiO_2 analysis (GDMI) for the LHI steers was 12.03 $\pm$ 2.36 kg. This was compared to the average DMI measured in the feedlot on the same steers was 10.69 $\pm$ 2.27. The Pearson's correlation between GDMI and DMI measured in the feedlot was 0.84 $\pm$ 0.10 ($P < 0.05$) with a Spearman rank correlation of 0.99 $\pm$ 0.03 ($P < 0.05$). The Spearman's rank correlation would suggest minimal change in ranking of animals based on their DMI. The high Pearson's correlation suggest an important relationship between grazing and feedlot DMI; however, it was less than perfect (i.e. not equal to 1) suggesting some change between DMI and GDMI. A contributing factor to the less than perfect correlation may be attributed to the estimated GDMI was on average higher than DMI measured in a feedlot. This would be expected given that animals were moved from a diet of 25% concentrate to a straight forage diet. Dry matter intake was dependent on the energy content of the diet provided (NRC, 1996). In order to

meet their nutritional requirements while maintained on pasture, their intakes would have to increase.

There is little research exploring the differences in the effect of diet on DMI (Shike et al., 2016) and to our knowledge, there was no direct comparison of grazing intake to feedlot intake. Shike et al. (2016) estimated a 0.58 phenotypic correlation for DMI for steers fed grain diets to forage diets. This correlation was lower than the observations in our study. The diet fed in the FIU to the steers for this study was a 73% roughage diet which may account for the higher correlations estimated in this study since cattle were fed a higher roughage diet with a lower percentage of concentrate then was used for Shike et al. (2016).

The estimated GDMI were compared to measurements gathered from CME presented in table 5.7. Since the estimated GDMI was a measurement of intake for a 10 day period, the CME measurements used for comparison were averaged for the same 10 day period. The correlation between GDMI and EAT was the largest correlation of -0.22 ($P = 0.23$). This correlation suggested that intake would decrease as more time was spent eating. These results are in contrast to what was estimated by Robinson and Oddy (2004) who estimated a phenotypic estimate of 0.30 ± 0.03 in a feedlot environment.

Estimated correlation between GDMI and RUM was -0.11 (P=0.58), which would indicate that as intake increased time spent for rumination would decrease. As forage in a diet increased, the rate of passage decreased, thus increasing the time spent for rumination (Krause et al., 2002). None of these correlation estimates were statistically significant ($P > 0.05$).

Table 5.7 Phenotypic correlations for grazing intake estimated using TiO_2 (GDMI, kg) and measurements using CowManager ear accelerometers.

	GDMI	RUM[1]	EAT[2]	NACT[3]	ACT[4]	HACT[5]
GDMI	1.00	-0.11	-0.22	0.02	0.21	0.19
RUM		1.00	-0.25	-0.44	-0.03	-0.26
EAT			1.00	0.13	-0.86	-0.12
NACT				1.00	-0.34	0.02
ACT					1.00	-0.02
HACT						1.00

RUM[1] is rumination measured by CowManager in minutes.

EAT[2] is time spent eating measured by CowManager in minutes.

NACT[3] is the amount of time an animal was not active measured by CowManager in minutes

ACT[4] was the time the animal spent being active as measured by CowManager in minutes.

HACT[5] was high activity measured by CowManager in minutes.

Validation Study: For the animals selected for TiO_2 supplementation (n = 32) during the second year's performance test, summary statistics for DMI and estimated DMI using TiO_2 are presented in table 5.8. The resulting Pearson's correlation of DMI and TiO_2 estimated DMI was 0.98 ± 0.03 ($P < 0.05$), indicating a strong association. A Spearman's rank correlation was also estimated to be 0.97 ± 0.04 ($P < 0.05$), which suggested minimal change in ranking of animals. Due to the high correlations of DMI and TiO_2 estimated DMI, these results would indicate that TiO_2 was a reliable method for estimating intake. As such, these results also validate that estimated GDMI from the previous year as a reliable measure of grazing intake.

Table 5.8. Summary statistics for dry matter intake measured by a GrowSafe system and estimated dry matter intake with the use of a biomarker titanium dioxide on beef steers (n=32).

	[1]GrowSafe	[2]TiO_2
Average	9.29	9.61
SD	1.83	2.26
Min	5.63	4.96
Max	12.67	14.60

[1]Dry matter intake (kg) measured by GrowSafe

[2]Dry matter intake (kg) estimated with the use of titanium dioxide

The estimation of low correlations and lack of association between CME measurements and DMI were not expected. In particular, the correlation of EAT and DMI for cattle in the FIU was 0.09 and 0.00 for LHI steers. The correlation to EAT and GDMI was -0.22 which suggested an increase of intake while less time was spent eating which would be counterintuitive for cattle maintained on pasture. Correlations between EAT and DMI were anticipated to be positive and moderate. As previously mentioned, reported correlations to time spent eating and DMI was reported to be moderate and positive ranging from 0.27 to 0.30 (Nkrumah et al., 2007; Robinson and Oddy, 2004). To investigate these low correlations and lack of association between DMI and CME measurements, measurements recorded by CME were validated by observers recording the same measurements. A concordance correlation was used to evaluate the measurements between CME and the observers for EAT, RUM, NACT, ACT and HACT. The decision to use a concordance correlation was due to the parameter's ability to quantify the agreement between two variables and the repeatability of the measurement observations (Lin, 1989). In this case, the first of the two variables of interest were the measurements recorded by CME and the second

were the visual observations made by the observers. Our interest lies in how much the recorded observed behavior was in agreement with what CME measured. If these correlations were low, the reliability of CME as a measure of cattle behavior would be suspect.

The CowManager ear tag accelerometer measured HACT in all the steers but none of the observers noted any activity that would be categorized as HACT. One of the primary uses of CME is heat detection in dairy cattle. High activity measured by CME is designed to identify cows demonstrating standing estrus. This is achieved by the fact that cows in standing estrus are mounted or mounting in addition to increased walking. The steers in the FIU expressed little activity with minimal walking and no mounting. After communicating with CowManager, ACT and HACT were combined as one variable (AH).

The concordance correlations were 0.39 ± 0.53, 0.58 ± 0.47, 0.65 ± 0.44, and 0.19 ± 0.57 for EAT, RUM, NACT and AH, respectively. Large standard errors for these correlations were a result of the small sample size (n = 5). The correlations were lower than those reported by Bikker et al. (2014; dairy cattle) and Wolfger et al. (2015; beef cattle). In both studies, the CowManager Sensors were evaluated for the accuracy of measuring feeding behavior and activity. Bikker et al. (2014) used CME in dairy cattle while Wolfger et al. (2015) used CME in beef cattle in a feedlot. Both studies found a 0.75 concordance correlation for eating but Bikker et al., (2014) found a 0.93 concordance correlation for rumination whereas Wolfger et al. (2015) found a lower correlation of 0.41. In the current study, the concordance correlation for eating was 0.39 which was lower than those reported by either Bikker et al. (2014) and Wolfger et al. (2015). However, for rumination, the concordance correlation was 0.58 ± 0.47 which was lower than the 0.93 reported by Bikker et al. (2014) and higher than 0.41 which Wolfger et al. (2015) reported.

Visual assessment of steers eating brought to attention a potential interference of the GrowSafe system with CME. The feed within the feed bunks were higher than if placed in an average concrete feed bunk, resulting in a steer not having to lower their heads in order to eat. In addition, animals have to pass their heads through an opening to access the feed bunk. In some cases, the ear would rest on the opening and movement of the ear would be restricted. Since the allocation of time to a behavior is dependent on the ear and jaw movement of the steer, this potential restricted ear movement might have affected CME's ability to measure EAT.

For the observations of finished beef steers fed from concrete feed bunks and CME measurements, the estimated concordance correlations between observers and steers were 0.01 $\pm$ 0.58, -0.14 $\pm$ 0.57, 0.23 $\pm$ 0.56, and -0.37 $\pm$ 0.54 for EAT, RUM, NACT and HA, respectively. As with the previous observations, the large standard errors were a result of the small samples size (n = 5). These steers were fed in concrete bunks with none of the previously mentioned interference from GrowSafe. For the observation of rumination, only 1 of the 5 steers was observed to ruminate during the observational period. The CME measurement during the hour when the animal was observed to be ruminating recorded no time spent ruminating. In addition, the 4 remaining animals that had no rumination observed had CME measurements for RUM. These measurements and observations resulted in a -0.14 $\pm$ 0.57 concordance correlation. Wolfger et al. (2015) concluded that CME measurement of RUM was low for sensitivity but highly specific. The authors for the Wolfger et al. (2015) study argued that CME was not as accurate for measuring every minute of rumination but didn't inappropriately allocate minutes to rumination. However, results from the current study showed that this was not the case and minutes for RUM were allocated to steers when those steers did not in fact ruminate. These

results indicate a lack of accuracy for the measurement of eating and rumination time measured by CME.

Since CME was initially developed for use in dairy cattle, 4 dairy cows equipped with CME was observed for the comparison of how CME measures behaviors and how it compared to what was measured in beef cattle. Concordance correlations for the observations of dairy cattle and CME measurements of EAT, RUM, NACT and HA were 0.90 $\pm$ 0.31, 0.89 $\pm$ 0.32, 0.68 $\pm$ 0.52 and 0.24 $\pm$ 0.69, respectively. The results for EAT and RUM were much higher than those measured in the beef cattle. Goldhawk et al. (2013) noted that technology that has high accuracy in dairy cattle may not have the same accuracy in beef cattle. This was attributed to a difference in dietary and physical differences of dewlaps, musculature and skin thickness between dairy and beef cattle (Goldhawk et al., 2013). As previously noted in the beef cattle, HACT was measured by CME but observers did not record any activity that would be considered high activity while observing the dairy cows. Because CME compiled the raw data and their algorithms are proprietary, further investigation into the disagreement of CME and observed behaviors was not possible. These results would indicate a difference in accuracy between beef cattle compared to dairy cattle for the use of CME.

5.4 Conclusion

There have been numerous technological advancements in the application of remote sensors for monitoring cattle. To date, the largest application of this technology for cattle has been within the dairy industry with limited application of these technologies in the beef industry. This study was able to show the application of remote sensor technology for monitoring cattle

maintained on rangeland; however, the accuracy of measurements from CME failed to provide an indicator for DMI. A strong relationship between feedlot measured DMI and DMI for grazing cattle was established in this study. In order to truly understand the relationship between grazing and feedlot measured intake, grazing behavior must also be explored in addition to the relationship of these behaviors and their effects on grazing intake. There is great potential for the application of these technologies in the beef industry, however further exploration of the efficacy of these technologies in beef cattle needs to continue.

LITERATURE CITED

Anderson, R. V., R. J. Rasby, T. J. Klopfenstein, and R. T. Clark. 2005. An evaluation of production and economic efficiency

Bikker, J. P., H. van Laar, P. Rump, J. Doorenbos, K. van Meurs, G. M. Griffioen and J. Dijkstra. Technical note: Evaluation of an ear-attached movement sensor to record cow feeding behavior and activity. J. Dairy Sci. 97:2974-2979.

Byskov, M. V., A. Fogh and P. Lovendahl. 2017. Genetic parameters of rumination time and feed efficiency traits in primiparous Holstein cows under research and commercial conditions.

Capper, J. L. 2011. The environmental impact of beef production in the United States: 1977 compared with 2007. J. Anim. Sci. 89: 4249-4261.

Culbertson, M. M., S. E. Speidel, R. K. Peel, R. R. Cockrum, M. G. Thomas, and R. M. Enns. 2015. Optimum measurement period for evaluating feed intake traits in beef cattle. J. Anim. Sci. 93: 2482-2487.

FAO. 2009. How to feed the world in 2050. Food and Agriculture Organization of the United Nations, Rome, Italy.

Ferrell, C. L. and T. G. Jenkins. 1985. Cow type and the nutritional environment: Nutritional aspects. J. Anim. Sci. 3: 725-741.

Goldhawk, C., K. Schwartzkopf-Genswein and K. A. Beauchemin. 2013. Technical note: Validation of rumination collars for beef cattle. J. Anim. Sci. 91:2858-2862.

Google Earth. https://earth.google.com/web/@40.392738,-103.1980838,1297.2116285a,781.68900289d,35y,0h,45t,0r/data=ChUaEwoLL2cvMXRmeDVjeTkYAiABKAIoAg (Accessed February 15, 2019).

Herd, R. M., P. F. Arthur, K. A. Donoghue, S. H. Bird, T. Bird-Gardiner and R. S. Hegarty. 2014. Measures of methane production and their phenotypic relationships with dry matter intake, growth, and body composition traits in beef cattle. J. Anim. Sci. 92: 5267-5274.

Krause, K. M., D. K. Combs and K. A. Beauchemin. 2002. Effects of forage particle size on grain fermentation in midlactating cows. II. Ruminal pH and chewing activity. J. Dairy Sci. 85:1947-1957.

Lin, L. I., 1989. A concordance correlation to evaluate reproducibility. Biometrics. 45:255-268

Myers, W. D., P. A. Ludden, V. Nayigihugu and B. W. Hess. 2004. Technical Note: A procedure for the preparation and quantitative analysis of samples for titanium dioxide. J. Anim. Sci. 82:179-183.

Nkrumah, J. D., J. A. Basarab, Z. Wang, C. Li, M. A. Price, E. K. Okine, D. H. Crews, Jr., and S. S. Moore. 2007. Genetic and phenotypic relationships of feed intake and measures of efficiency with growth and carcass merit of beef cattle. J. Anim. Sci. 85:2711-2720.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Robinson, D. L. and V. H. Oddy. 2004. Genetic parameters for feed efficiency, fatness, muscle area and feeding behavior of feedlot beef cattle. Livest. Prod. Sci. 255-270.

Schirmann, K., N. Chapinal, D. M. Weary, W. Heuwieser and M. A. G. von Keyserlingk. 2010. Rumination and its relationship to feeding and lying behavior in Holstein dairy cows. J. Dairy Sci. 95:3212-3217.

Shike, D. W., C. J. Cassady, T. L. Felix and J. E. Beever. 2016. Effects of timing and duration of test period and diet type on intake and feed efficiency in Charolais-sired cattle. BIF Proc. 57-64.

Theurer 1996

Titgemeyer, E. C., C. K. Armendariz, D. J. Bindel, R. H. Greenwood and C. A. Loest. 2001. Evaluation of titanium dioxide as a digestibility marker for cattle. J. Anim. Sci. 79: 1059-1063.

Whisnant, C. USDA-NIFA-CRIS Research. http://cris.csrees.usda.gov/cgi-bin/starfinder/6244/crisassist.txt (Accessed April 30, 2017).

Wolfger, B., E. Timsti. E. A. Pajor, N. Cook, H. W. Barkema, and K. Orel. 2015. Technical note: Accuracy of an ear tag-attached accelerometer to monitor rumination and feeding behavior in feedlot cattle. J. Anim. Sci. 93: 3164-3168.

APPENDIX A: R CODE USED TO CREATE SIMULTED DATA FOR FEMALE BASE POPULATION

```
####Create a distribution of age for females
##R code to create base population of females with age distribution at equilibrium
##No selection for female base population
##20% Replacement

## ID for year 00 females (n=10000)
fid00 = c(1:10000)
yr00 <- data.frame(id = fid00, year = 00)
nam = 10000

## ID for year 01 replacement females
fid01 = c(10001:12000)
yr01 <- data.frame(id = fid01, year = 01)
nam1 = nam + length(fid01)

## Creating progeny from yr00 females in year 2
n_progeny = nam
set.seed(1234)
isex=(rbinom(nam, 1, 0.5))+1
prg_of_yr00 = data.frame(year = 02, sex = isex)
hfr = subset(prg_of_yr00, sex == 2)
hfr1<- hfr[sample(1:nrow(hfr), 2000, replace=FALSE),]
fid02 = c(12001:14000)
yr02 <- data.frame(id = fid02, year = hfr1$year)

##Random loss of 20% from yr00 females
n <- round(nrow(yr00)*0.8)
females00 <- yr00[sample(1:nrow(yr00), n, replace=FALSE),]
females = rbind(females00, yr01)

##Creating progeny for year 3
fid03 = c(14001:16000)
yr03 <- data.frame(id = fid03, year = 03)

##Random loss of 20% from females
n <- round(nrow(females)*0.8)
females00 <- females[sample(1:nrow(females), n, replace=FALSE),]
females = rbind(females00, yr01)

##Creating progeny for year 4
fid04 = c(16001:18000)
yr04 <- data.frame(id = fid04, year = 04)

##Random loss of 20% from females
n <- round(nrow(females)*0.8)
females00 <- females[sample(1:nrow(females), n, replace=FALSE),]
```

```
females = rbind(females00, yr02)

##Creating progeny for year 5
fid05 = c(18001:20000)
yr05 <- data.frame(id = fid05, year = 05)

##Random loss of 20% from females
n <- round(nrow(females)*0.8)
females00 <- females[sample(1:nrow(females), n, replace=FALSE),]
females = rbind(females00, yr03)

##Creating progeny for year 6
fid06 = c(20001:22000)
yr06 <- data.frame(id = fid06, year = 06)

##Random loss of 20% from females
n <- round(nrow(females)*0.8)
females00 <- females[sample(1:nrow(females), n, replace=FALSE),]
females = rbind(females00, yr04)

##Creating progeny for year 7
fid07 = c(22001:24000)
yr07 <- data.frame(id = fid07, year = 07)

##Random loss of 20% from females
n <- round(nrow(females)*0.8)
females00 <- females[sample(1:nrow(females), n, replace=FALSE),]
females = rbind(females00, yr05)

##Creating progeny for year 8
fid08 = c(24001:26000)
yr08 <- data.frame(id = fid08, year = 08)

##Random loss of 20% from females
n <- round(nrow(females)*0.8)
females00 <- females[sample(1:nrow(females), n, replace=FALSE),]
females = rbind(females00, yr06)

##Creating progeny for year 9
fid09 = c(26001:28000)
yr09 <- data.frame(id = fid09, year = 09)

##Random loss of 20% from females
n <- round(nrow(females)*0.8)
females00 <- females[sample(1:nrow(females), n, replace=FALSE),]
females = rbind(females00, yr07)
```

```
##Creating progeny for year 10
fid10 = c(28001:30000)
yr10 <- data.frame(id = fid10, year = 10)

##Random loss of 20% from females
n <- round(nrow(females)*0.8)
females00 <- females[sample(1:nrow(females), n, replace=FALSE),]
females = rbind(females00, yr08)

##Creating progeny for year 11
fid11 = c(30001:32000)
yr11 <- data.frame(id = fid11, year = 11)

##Random loss of 20% from females
n <- round(nrow(females)*0.8)
females00 <- females[sample(1:nrow(females), n, replace=FALSE),]
females = rbind(females00, yr09)

##Creating progeny for year 12
fid12 = c(32001:34000)
yr12 <- data.frame(id = fid12, year = 12)

##Random loss of 20% from females
n <- round(nrow(females)*0.8)
females00 <- females[sample(1:nrow(females), n, replace=FALSE),]
females = rbind(females00, yr10)

##Creating progeny for year 13
fid13 = c(34001:36000)
yr13 <- data.frame(id = fid13, year = 13)

##Random loss of 20% from females
n <- round(nrow(females)*0.8)
females00 <- females[sample(1:nrow(females), n, replace=FALSE),]
females = rbind(females00, yr11)

##Creating progeny for year 14
fid14 = c(36001:38000)
yr14 <- data.frame(id = fid14, year = 14)

##Random loss of 20% from females
n <- round(nrow(females)*0.8)
females00 <- females[sample(1:nrow(females), n, replace=FALSE),]
females = rbind(females00, yr12)
```

```
##Creating progeny for year 15
fid15 = c(38001:40000)
yr15 <- data.frame(id = fid15, year = 15)

##Random loss of 20% from females
n <- round(nrow(females)*0.8)
females00 <- females[sample(1:nrow(females), n, replace=FALSE),]
females = rbind(females00, yr13)

##Creating progeny for year 16
fid16 = c(40001:42000)
yr16 <- data.frame(id = fid16, year = 16)

##Random loss of 20% from females
n <- round(nrow(females)*0.8)
females00 <- females[sample(1:nrow(females), n, replace=FALSE),]
females = rbind(females00, yr14)

##Creating progeny for year 17
fid17 = c(42001:44000)
yr17 <- data.frame(id = fid17, year = 17)

##Remove animals from year 00 (these females are 16 years old)
young.fem = subset(females, year > 0)

##Random loss of 20% from females
females01 <- young.fem[sample(1:nrow(young.fem), 8000, replace=FALSE),]
fem = rbind(females01, yr15)

################################################################################
################################################################################
####Equilibrium reached once year 00 was removed.
####Using year 17 for base pop for females
####Remove females 16 years of age

table(unlist(fem$year))
#1   2   3   4   5   6   7   8   9  10  11  12  13  14  15
#135 114 152 167 206 268 376 426 529 692 837 1064 1381 1653 2000

##Recode file for simulation
fem1 <- fem[ order(fem$id), ]
fem.id <- c(1:10000)
base.females <- data.frame(id = fem.id, year = fem1$year, sex = 2)

################################################################################
####
```

APPENDIX B: R CODE USED TO CREATE SIMULTED DATA FOR SINGLE TRAIT SELECTION ON DRY MATTER INTAKE

```
###Final analysis with updated G and R matrices

#R code based on Dr. Larry Schaeffer's R code for multiple trait models && Hamad Saad's disertation
#Creating a simulated data (Selection on decreasing feed intake)
## Traits were: Feed intake (FI), 200 d weight (WWT), 400 d weight (YWT), and average daily gain (ADG)
# First creating the base population (No selection)
setwd("/Users/Randie/Documents/PhD Research/Simulation/SimRcode")

## To clear environment
#FISimData_v2
rm(list = ls())

##Run DistFem_USE.R first
source("DistFem_USE.R")

library(MASS)

#Herd effects (50 herds) on FI, ADG, WWT and YWT
set.seed(1234)
herdFI=(rnorm(50,0))
set.seed(1234)
herdWW=(rnorm(50,0))*2
set.seed(1234)
herdYW=(rnorm(50,0))*10
set.seed(1234)
herdADG=(rnorm(50,0))*0.25  ##Crews et al 2006 (SD for ADG)
herd=matrix(data = c(herdFI,herdADG,herdWW,herdYW),byrow = TRUE, nrow = 4)

# sex effects (50 herds) on FI, ADG, WWT, and YWT
##Sex effects for wwt and ywwt from Van Vleck & Cundiff 1998
##ADG sex effects used Hamad's
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)
# Residual matrix
R=matrix(data=c(1.05,0.049,0.528,17.749,0.049,0.014,0.287,1.982,0.528,0.287,552.75,465.63,17.749,1.982,465.63,1131.95),byrow=TRUE,nrow=4)
# G matrix
G=matrix(data=c(0.694,0.050,9.945,18.845,0.050,0.005,0.709,1.550,9.945,0.709,230.77,327.934,18.845,1.550,327.934,598.327),byrow=TRUE,nrow=4)
# Traits averages
mu=c(10.007,1.473,258.913,479.990) #weighted averages

##Matrices are positive definite
#G eigen values: 7.913849e+02 3.866650e+01 1.583454e-01 1.237949e-03
#R eigen values: 1390.9704349  294.3559224    0.9810670    0.0165757
```

```
#ID for animals (founders (10000 dams, 450 sires))
aid = c(1:10400)

##creating fields in the data file for sire (20 progeny each), dams (1 progeny each), herds (size of
200 each), sex (50% females and 50% males)
sid = c(numeric(10400),rep(1:400, by=1, each=20))
base.sire <- data.frame(id = c(10001:10400), year = 15, sex = 1)
did <- c(numeric(10400),1:10000)
bi=c(rep(1,10400),rep(0.5,10000))
set.seed(1234)
iherd=c(sample(rep(1:52, by=1, each=200),10400,replace=F))

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

nam = 10400
# Simulate true breeding values for all animals (founders and their F1 progeny)
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}

tbv = jd(nam,4)*0
for(i in 1:nam){
  x = LG %*% (sqrt(bi[i])*rnorm(4,0,1))
  if(sid[i]>0){
    ks=sid[i]
    kd=did[i]
    x = x + 0.5*(tbv[ks, ]+tbv[kd, ]) }
```

```
  tbv[i,]=x }
nrec=10400

#########################################################################
#########################################################################
##Creating base popuation with tbv but no observations
####Use base.females from DistFem_USE.R

base.pop <- rbind(base.females, base.sire)
simulateddata <- data.frame (id = base.pop$id, sire = c(rep(NA,10400)), dam =
c(rep(NA,10400)), sex = base.pop$sex, year = base.pop$year, herd = iherd, FI =
c(rep(NA,10400)), ADG = c(rep(NA,10400)), WWT = c(rep(NA,10400)), YWT =
c(rep(NA,10400)), FItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simulateddata)

animlist = c(simulateddata[ ,1])
detach(simulateddata)
uniqueanimlist = c(sort(unique(animlist)))
pedigree = data.frame(id = uniqueanimlist, sire = c(rep(NA,10400)), dam = c(rep(NA,10400)))
length(uniqueanimlist) #10400
length(pedigree[,1]) #10400

#Header: id sire dam sex year herd FI ADG WWT YWT FItbv ADGtbv WWTtbv YWTtbv
basepopdata <- simulateddata
attach(basepopdata)

#####################################################################
## creating P1 (progeny 1)
## Assuming 100% conception ????
## No selection of FI.  Must use whole population for first 2 generations
#####################################################################
## redefine the total number of all animals and number of base population in previous simulation
################################
#total number of all animals (base pop + P1)
nam = 20400
#number of sires and dams(founders) (400+10000) in base population
nbase = 10400

################################
## selecting top 5% sires and 80% dams (TBV for FI is the selection criteria)
## selection of sires
## average FI TBV for sires and dams
averages=by( basepopdata$FItbv, basepopdata$sex, mean)
ave1males=averages[1]
ave1females=averages[2]
## standard deviation of FI TBV for sires and dams
```

```
SDs=by( basepopdata$FItbv, basepopdata$sex, sd)
sd1males=SDs[1]
sd1females=SDs[2]

## selecting all sires from basepopdata
selectedmales <- subset(basepopdata, sex == 1)

#########################
## selection of all dams
## For first selection of cows from P1 only selected 100% of females to be mated to sires
females.base <- subset(basepopdata, sex == 2)
##90% Conception Rate
selectedallfemales <- females.base[sample(1:nrow(females.base), 9000, replace = FALSE),]

nrow(females.base)
nrow(selectedallfemales)
nrow(selectedallfemales)/nrow(females.base) #0.9

#########################
#### 400 selected males and 10000 selected females
n_sel_males = nrow(selectedmales)
n_sel_females = nrow(selectedallfemales)
n_progeny = n_sel_females
# number of females per sire = 22.5
n_females_per_sire = n_sel_females/n_sel_males
n_females_per_sire
# because number of dams per sire is 22.5, then sires will have different numbers of progeny
(some will have 22.5 and others will have 23 progeny)
rounded_n_females_per_sire = round(n_females_per_sire)
rounded_n_females_per_sire

if(rounded_n_females_per_sire < n_females_per_sire) {
  n1records_per_sire = (rounded_n_females_per_sire)
  n2records_per_sire = (n1records_per_sire)+1
} else {
  n1records_per_sire = (rounded_n_females_per_sire)-1
  n2records_per_sire = (rounded_n_females_per_sire)
}

#number of sires with 22 progeny (200 out of 400)
nsires_with_n1records = (n_sel_males)-((n_sel_females)-(n_sel_males*n1records_per_sire))
#number of sires with 23 progeny (200 out of 400)
nsires_with_n2records = (n_sel_males)-(nsires_with_n1records)
##pulling out the first 137 sires of sire list
own_record_sires_with_n1records = selectedmales[1:nsires_with_n1records, ]
## replicate each sire 22 times
```

```
selectedmales1=own_record_sires_with_n1records[rep(seq_len(nrow(own_record_sires_with_n
1records)), each=n1records_per_sire),]
##pulling out the remaining 200 sires of sire list
own_record_sires_with_n2records = selectedmales[((nsires_with_n1records)+1):n_sel_males, ]
## replicate each sire 24 times
selectedmales2=own_record_sires_with_n2records[rep(seq_len(nrow(own_record_sires_with_n
2records)), each=n2records_per_sire),]

## Create random list of sires with length of 9000 which is number of selected dams
allselectedmales=rbind(selectedmales1,selectedmales2)
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp1 = randomlymatedsires[ ,1]
damlistp1 = selectedallfemales[ ,1]

####################
selectedparents=rbind(selectedmales, selectedallfemales)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]
# pulling out selected parents and their pedigree
animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
subdata <- basepopdata[basepopdata[ ,1] %in% uniqueanimlist, ]
subdata <- data.matrix(subdata)
length(subdata[,1]) #9400
length(uniqueanimlist) #9400

## creating P1 pedigree which include(selected parents and their pedigree + new P1 ID with their
selected parents)
pedP1 =
rbind(subdata[,1:3],cbind(c((nam+1):(nam+n_progeny)),sirelistp1[1:n_progeny],damlistp1[1:n_
progeny]))
##################

###############
## P1 observations

library(MASS)

set.seed(1234)
herdFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
```

```
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(1.05,0.049,0.528,17.749,0.049,0.014,0.287,1.982,0.528,0.287,552.75,465.63,1
7.749,1.982,465.63,1131.95),byrow=TRUE,nrow=4)
# G matrix
G=matrix(data=c(0.694,0.050,9.945,18.845,0.050,0.005,0.709,1.550,9.945,0.709,230.77,327.934
,18.845,1.550,327.934,598.327),byrow=TRUE,nrow=4)
# Traits averages
mu=c(10.007,1.473,258.913,479.990) #weighted averages

anwr=c((nam+1):(nam+n_progeny))
aid = c(pedP1[ ,1])
sid = c(pedP1[ ,2])
did = c(pedP1[ ,3])

#number of herds
length(unique(selectedallfemales[ ,6]))
iherd=c(selectedallfemales[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
```

```
nam1=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam1){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
damtbv= selectedallfemales[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}
obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
pedigree1 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree1)
dataped <- pedigree1[pedigree1$id>nam,]
simdataP1 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2000, herd = iherd, FI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
FItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simdataP1)
nrow(simdataP1) #9000

## combine the data file of (base) with P1 data file
p1andbasepopdata <- rbind(basepopdata,simdataP1)
## Data file for basepop and P1 (has both observations and TBV)
p1andbasepopdata <- subset(p1andbasepopdata, p1andbasepopdata[ ,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1=mean(p1andbasepopdata$WWT)
WWTave_base_p1 #248.40
WWTsd_base_p1=sd(p1andbasepopdata$WWT)
WWTsd_base_p1 #37.81
```

```
WWTave_base_p1-(3*WWTsd_base_p1) #134.98
WWTave_base_p1+(3*WWTsd_base_p1) #361.83
p1andbasepopdata <- subset(p1andbasepopdata, WWT > (WWTave_base_p1-
(3*WWTsd_base_p1)))
p1andbasepopdata <- subset(p1andbasepopdata, WWT <
(WWTave_base_p1+(3*WWTsd_base_p1)))
summary(p1andbasepopdata$WWT)
nrow(p1andbasepopdata) #8965

# keeping YWT observations within 3 SD
YWTave_base_p1=mean(p1andbasepopdata$YWT)
YWTave_base_p1
YWTsd_base_p1=sd(p1andbasepopdata$YWT)
YWTsd_base_p1
YWTave_base_p1-(3*YWTsd_base_p1)
YWTave_base_p1+(3*YWTsd_base_p1)
p1andbasepop <- subset(p1andbasepopdata, YWT > (YWTave_base_p1-(3*YWTsd_base_p1)))
p1andbasepopdata <- subset(p1andbasepopdata, YWT <
(YWTave_base_p1+(3*YWTsd_base_p1)))
summary(p1andbasepopdata$YWT)
nrow(p1andbasepopdata) #8964

#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p1andbasepopdata <- subset(p1andbasepopdata, p1andbasepopdata[ ,10] - p1andbasepopdata[ ,9]
> 40)
nrow(p1andbasepopdata) #8964

simdataP1 <- subset(p1andbasepopdata, p1andbasepopdata[ ,1] > nam)
summary(simdataP1$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#136.0   222.5   247.5   248.1   272.9   361.4
summary(simdataP1$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#279.3   433.4   480.8   480.5   526.9   662.2

nrow(simdataP1) #8938
N_sires = length(c(sort(unique(simdataP1$sire))))
N_sires #400
basepop_p1_data <- rbind(basepopdata,simdataP1)
pedigreep1 <- data.frame (id = simdataP1$id, sire = simdataP1$sire, dam = simdataP1$dam)
pedigreep1andbase <- rbind(pedigree,pedigreep1)

##############################################################################
##############################################################################
## Creating P2 (Year 2)
```

```
#1 2 3 4 5 6 7 8 9 10 11 12 13 14
# id sire dam sex year herd CE BW ww pwg CEtbv BWtbv wwtbv pwgtbv
####Still no selection. No replacements.
#####################################################################
#####################################################################
## redefine (the total number of all animals) and (number of base population and P1) in previous
simulation
################################
#total number of all animals (Used unique IDs up to now)
nam2 = nam+nam1

################################
## selection of sires
## average FI TBV for sires and dams
averages=by( basepopdata$FItbv, basepopdata$sex, mean)
ave1males=averages[1]
ave1females=averages[2]
## standard deviation of FI TBV for sires and dams
SDs=by( basepopdata$FItbv, basepopdata$sex, sd)
sd1males=SDs[1]
sd1females=SDs[2]

## selecting all sires from basepopdata
selectedmales <- subset(basepopdata, sex == 1)

#########################
## selection of all dams
## Remove females from year 1 (older than 16 years)
females <- subset(selectedallfemales, year > 1)
selectedallfemales1 <- females[sample(1:nrow(females), 8100, replace = FALSE),]

nrow(selectedallfemales1)
nrow(selectedallfemales)
nrow(selectedallfemales1)/nrow(selectedallfemales) #0.90

#########################
#### 400 selected males and 8100 selected females
n_sel_males = nrow(selectedmales)
n_sel_females = nrow(selectedallfemales1)
n_progeny = n_sel_females
# number of females per sire = 20.25
n_females_per_sire = n_sel_females/n_sel_males
n_females_per_sire
# because number of dams per sire is 20.25, then sires will have different numbers of progeny
(some will have 20 and others will have 21 progeny)
rounded_n_females_per_sire = round(n_females_per_sire)
```

```
rounded_n_females_per_sire

if(rounded_n_females_per_sire < n_females_per_sire) {
  n1records_per_sire = (rounded_n_females_per_sire)
  n2records_per_sire = (n1records_per_sire)+1
} else {
  n1records_per_sire = (rounded_n_females_per_sire)-1
  n2records_per_sire = (rounded_n_females_per_sire)
}

#number of sires with 20 progeny (300 out of 400)
nsires_with_n1records = (n_sel_males)-((n_sel_females)-(n_sel_males*n1records_per_sire))
#number of sires with 21 progeny (100 out of 400)
nsires_with_n2records = (n_sel_males)-(nsires_with_n1records)
##pulling out the first 137 sires of sire list
own_record_sires_with_n1records = selectedmales[1:nsires_with_n1records, ]
## replicate each sire 25 times
selectedmales1=own_record_sires_with_n1records[rep(seq_len(nrow(own_record_sires_with_n
1records)), each=n1records_per_sire),]
##pulling out the remaining 263 sires of sire list
own_record_sires_with_n2records = selectedmales[((nsires_with_n1records)+1):n_sel_males, ]
## replicate each sire 24 times
selectedmales2=own_record_sires_with_n2records[rep(seq_len(nrow(own_record_sires_with_n
2records)), each=n2records_per_sire),]

## Create random list of sires with length of 8100 which is number of selected dams
allselectedmales=rbind(selectedmales1,selectedmales2)
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp2 = randomlymatedsires[ ,1]
damlistp2 = selectedallfemales1[ ,1]

####################
selectedparents=rbind(selectedmales, selectedallfemales1)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]
# pulling out selected parents and their pedigree
animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
#uniqueanimlist = uniqueanimlist[-1]
subdata <- basepopdata[basepopdata[ ,1] %in% uniqueanimlist, ]
subdata <- data.matrix(subdata)
length(subdata[,1]) #8500
length(uniqueanimlist) #8500

## creating P2 pedigree which include(selected parents and their pedigree + new P2 ID with their
selected parents)
```

```
pedP2 =
rbind(subdata[,1:3],cbind(c((nam2+1):(nam2+n_progeny)),sirelistp2[1:n_progeny],damlistp2[1:
n_progeny]))

###############
## For P2, repeating the same code used to create P1.

library(MASS)

set.seed(1234)
herdFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(1.05,0.049,0.528,17.749,0.049,0.014,0.287,1.982,0.528,0.287,552.75,465.63,1
7.749,1.982,465.63,1131.95),byrow=TRUE,nrow=4)
# G matrix
G=matrix(data=c(0.694,0.050,9.945,18.845,0.050,0.005,0.709,1.550,9.945,0.709,230.77,327.934
,18.845,1.550,327.934,598.327),byrow=TRUE,nrow=4)
# Traits averages
mu=c(10.007,1.473,258.913,479.990) #weighted averages

anwr=c((nam2+1):(nam2+n_progeny))
aid = c(pedP2[ ,1])
sid = c(pedP2[ ,2])
did <- c(pedP2[ ,3])

#number of herds
length(unique(selectedallfemales1[ ,6]))
iherd=c(selectedallfemales1[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2
```

```
# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam3=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam3){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
damtbv= selectedallfemales1[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}

obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
```

```
pedigree2 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree2)
dataped <- pedigree2[pedigree2$id>nam,]
simdataP2 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2001, herd = iherd, FI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
FItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simdataP2)
nrow(simdataP2) #8100

## combine the data file of (base and P1) with P2 data file
p2_p1_basepopdata <- rbind(p1andbasepopdata,simdataP2)
## Data file for basepop and P1 (has both observations and TBV)
p2_p1_basepopdata <- subset(p2_p1_basepopdata, p2_p1_basepopdata[ ,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1_p2=mean(p2_p1_basepopdata$WWT)
WWTave_base_p1_p2
WWTsd_base_p1_p2=sd(p2_p1_basepopdata$WWT)
WWTsd_base_p1_p2
WWTave_base_p1_p2-(3*WWTsd_base_p1_p2)
WWTave_base_p1_p2+(3*WWTsd_base_p1_p2)
p2_p1_basepopdata <- subset(p2_p1_basepopdata, WWT > (WWTave_base_p1_p2-
(3*WWTsd_base_p1_p2)))
p2_p1_basepopdata <- subset(p2_p1_basepopdata, WWT <
(WWTave_base_p1_p2+(3*WWTsd_base_p1_p2)))
summary(p2_p1_basepopdata$WWT)
nrow(p2_p1_basepopdata) #17023

#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p2_p1_basepopdata <- subset(p2_p1_basepopdata, p2_p1_basepopdata[ ,10] -
p2_p1_basepopdata[ ,9] > 40)
nrow(p2_p1_basepopdata) #17023

simdataP2 <- subset(p2_p1_basepopdata, p2_p1_basepopdata[ ,1] > nam2)
summary(simdataP2$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#138.5   223.2   247.2   248.2   272.6   358.0
summary(simdataP2$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
# 259.0   433.9   481.8   480.5   526.6   682.0

nrow(simdataP2) #8024
N_sires = length(c(sort(unique(simdataP2$sire))))
N_sires #400
basepop_p2_p1_data <- rbind(basepop_p1_data,simdataP2)
```

```
pedigreep2 <- data.frame (id = simdataP2$id, sire = simdataP2$sire, dam = simdataP2$dam)
pedigreep2p1andbase<- rbind(pedigreep1andbase,pedigreep2)

##############################################################################
##############################################################################
## Creating P3 (Year 3)
#1 2 3 4 5 6 7 8 9 10 11 12 13 14
# id sire dam sex year herd CE BW ww pwg CEtbv BWtbv wwtbv pwgtbv
# Begin selection for FI and use replacments from year 2000
#################################################################
#################################################################
## redefine (the total number of all animals) and (number of base population and P1) in previous
simulation
################################
#total number of all animals (Used unique IDs up to now): basepop, P1 and P2
nam4 = nam2+nam3

################################
## selecting top 10% sirs and 80% dams
## selection of sires
## sires selection changed to 10% otherwise there were not enough sires
averages=by( basepop_p1_data$FItbv, basepop_p1_data$sex, mean)
ave1males=averages[1]
ave1females=averages[2]
SDs=by( basepop_p1_data$FItbv, basepop_p1_data$sex, sd)
sd1males=SDs[1]
sd1females=SDs[2]

males3 <- subset(basepop_p1_data, sex == 1)
bulls3 <- males3[order(males3$FItbv), ]
selectedmales3 <- bulls3[1:450, ]
summary(selectedmales3[ ,7], na.rm=TRUE)

nrow(males3) #426
nrow(selectedmales3) #450
nrow(selectedmales3)/nrow(males3) #0.09

#########################
## selection of dams
## FIRST: Remove females from year 2 (older than 16 years)
## SECOND: Selection of cows and replace with 2000 heifers
## Assume 20% loss and 90% conception rate

females3 <- subset(basepop_p1_data, sex == 2 & year > 2 & year < 2000)
femalesp3 <- females3[sample(1:nrow(females3), 8000, replace=FALSE),]
hfs <- subset(basepop_p1_data, sex == 2 & year == 2000)
```

```
nrow(hfs) #4420

if(nrow(hfs) > 2000) {
  n = 2000
} else {
  n = nrow(hfs)
}

hfrs <- hfs[order(hfs$FItbv), ]
selectedhfrs <- hfrs[1:n, ]
fem3 <- rbind(femalesp3, selectedhfrs)
selectedallfemales3 <- fem3[sample(1:nrow(fem3), 9000, replace = FALSE),]
summary(selectedallfemales3[ ,7], na.rm=TRUE)

nrow(femalesp3) #8000
nrow(selectedallfemales3) #9000
damlistp3 = c(selectedallfemales3[ ,1])

##########################
#### 450 selected males and 9000 selected females

n_progeny = nrow(selectedallfemales3)

allselectedmales=selectedmales3[rep(seq_len(nrow(selectedmales3)), each=20),]
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp3 = randomlymatedsires[ ,1]

#####################
selectedparents=rbind(selectedmales3, selectedallfemales3)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]

animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
uniqueanimlist = uniqueanimlist[-1]
subdata <- basepop_p1_data[basepop_p1_data[ ,1] %in% uniqueanimlist, ]
subdata <- data.matrix(subdata)

length(subdata[,1]) #10446
length(uniqueanimlist) #10446
pedp3 = rbind(subdata[
,1:3],cbind(c((nam4+1):(nam4+n_progeny)),sirelistp3[1:n_progeny],damlistp3[1:n_progeny]))

##################
FIave=mean(selectedparents[ ,7], na.rm=TRUE)
FIave
ADGave=mean(selectedparents[ ,8], na.rm=TRUE)
```

```
ADGave
WWTave=mean(selectedparents[ ,9], na.rm=TRUE)
WWTave
YWTave=mean(selectedparents[ ,10], na.rm=TRUE)
YWTave

###############
###############
## P3 progeny, repeating the same code used to create P2.

library(MASS)

set.seed(1234)
herdFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(1.05,0.049,0.528,17.749,0.049,0.014,0.287,1.982,0.528,0.287,552.75,465.63,1
7.749,1.982,465.63,1131.95),byrow=TRUE,nrow=4)
# G matrix
G=matrix(data=c(0.694,0.050,9.945,18.845,0.050,0.005,0.709,1.550,9.945,0.709,230.77,327.934
,18.845,1.550,327.934,598.327),byrow=TRUE,nrow=4)

####averages:
mu=c(FIave,ADGave,WWTave,YWTave)
anwr=c((nam4+1):(nam4+n_progeny))
aid = c(pedp3[ ,1])
sid = c(pedp3[ ,2])
did <- c(pedp3[ ,3])

#number of herds
length(unique(selectedallfemales3[ ,6]))
iherd=c(selectedallfemales3[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
rd=diag(R)
```

```
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam5=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam5){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
damtbv= selectedallfemales3[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}

obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
```

```
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
pedigree3 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree3)
dataped <- pedigree3[pedigree3$id>nam4,]
simdataP3 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2002, herd = iherd, FI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
FItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simdataP3)
nrow(simdataP3) #9000

## combine the data file of (base, P1 and P2) with P3 data file
p3_p2_p1_basepopdata <- rbind(p2_p1_basepopdata,simdataP3)
## Data file for basepop and P1 (has both observations and TBV)
p3_p2_p1_basepopdata <- subset(p3_p2_p1_basepopdata, p3_p2_p1_basepopdata[ ,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1_p2_p3=mean(p3_p2_p1_basepopdata$WWT)
WWTave_base_p1_p2_p3
WWTsd_base_p1_p2_p3=sd(p3_p2_p1_basepopdata$WWT)
WWTsd_base_p1_p2_p3
WWTave_base_p1_p2_p3-(3*WWTsd_base_p1_p2_p3)
WWTave_base_p1_p2_p3+(3*WWTsd_base_p1_p2_p3)
p3_p2_p1_basepopdata <- subset(p3_p2_p1_basepopdata, WWT > (WWTave_base_p1_p2_p3-
(3*WWTsd_base_p1_p2_p3)))
p3_p2_p1_basepopdata <- subset(p3_p2_p1_basepopdata, WWT <
(WWTave_base_p1_p2_p3+(3*WWTsd_base_p1_p2_p3)))
summary(p3_p2_p1_basepopdata$WWT)
nrow(p3_p2_p1_basepopdata) #25994

# keeping YWT observations within 3 SD
YWTave_base_p1_p2_p3=mean(p3_p2_p1_basepopdata$YWT)
YWTave_base_p1_p2_p3
YWTsd_base_p1_p2_p3=sd(p3_p2_p1_basepopdata$YWT)
YWTsd_base_p1_p2_p3
YWTave_base_p1_p2_p3-(3*YWTsd_base_p1_p2_p3)
YWTave_base_p1_p2_p3+(3*YWTsd_base_p1_p2_p3)
p3_p2_p1_basepopdata <- subset(p3_p2_p1_basepopdata, YWT > (YWTave_base_p1_p2_p3-
(3*YWTsd_base_p1_p2_p3)))
p3_p2_p1_basepopdata <- subset(p3_p2_p1_basepopdata, YWT <
(YWTave_base_p1_p2_p3+(3*YWTsd_base_p1_p2_p3)))
summary(p3_p2_p1_basepopdata$YWT)
nrow(p3_p2_p1_basepopdata) #25984
```

```
#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p3_p2_p1_basepopdata <- subset(p3_p2_p1_basepopdata, p3_p2_p1_basepopdata[ ,10] -
p3_p2_p1_basepopdata[ ,9] > 40)
nrow(p3_p2_p1_basepopdata) #25984

#Remove animals with FI less than 0
p3_p2_p1_basepopdata <- subset(p3_p2_p1_basepopdata, p3_p2_p1_basepopdata[ ,7] > 1)
nrow(p3_p2_p1_basepopdata) #25968

simdataP3 <- subset(p3_p2_p1_basepopdata, p3_p2_p1_basepopdata[ ,1] > nam4)
summary(simdataP3$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#104.3   177.8   202.4   203.3   227.6   350.2
summary(simdataP3$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#232.1   350.7   398.9   399.0   445.7   618.6

nrow(simdataP3) #8960
N_sires = length(c(sort(unique(simdataP3$sire))))
N_sires #450
basepop_p1_p2_p3_data <- rbind(basepop_p2_p1_data,simdataP3)
pedigreep3 <- data.frame (id = simdataP3$id, sire = simdataP3$sire, dam = simdataP3$dam)
pedigreep3p2p1andbase <- rbind(pedigreep2p1andbase,pedigreep3)

##############################################################################
##############################################################################
## Creating P4 (Year 4)
#1 2 3 4 5 6 7 8 9 10 11 12 13
# id sire dam sex herd CE BW ww pwg CEtbv BWtbv wwtbv pwgtbv
#################################################################
#################################################################
## redefine (the total number of all animals) and (number of base population, P1 and P2) in
previous simulation
################################
#total number of all animals (Used unique IDs up to now): basepop, P1 and P2
nam6 = nam4+nam5

################################
## selecting top 5% sires (2000 & 2001) and top heifers for replacements
## selection of sires
averages=by( basepop_p2_p1_data$FItbv, basepop_p2_p1_data$sex, mean)
ave1males=averages[1]
ave1females=averages[2]
SDs=by( basepop_p2_p1_data$FItbv, basepop_p2_p1_data$sex, sd)
sd1males=SDs[1]
```

```
sd1females=SDs[2]

males4 <- subset(basepop_p2_p1_data, sex == 1 & year >= 2000)
bulls4 <- males4[order(males4$FItbv), ]
selectedmales4 <- bulls4[1:450, ]
summary(selectedmales4[ ,7], na.rm=TRUE)

nrow(males4) #8583
nrow(selectedmales4) #450
nrow(selectedmales4)/nrow(males4) #0.05

##########################
## selection of dams
## FIRST: Remove females from year 3 (older than 16 years)
## SECOND: Selection of cows and replace with 2000 heifers
## Assume 20% loss
## Assume 90% Conception Rate

females4 <- subset(selectedallfemales3, year > 3)
femalesp4 <- females4[sample(1:nrow(females4), 8000, replace=FALSE),]
heifers <- subset(basepop_p2_p1_data, sex == 2 & year == 2001)
nrow(heifers) #3959

if(nrow(hfs) > 2000) {
  n = 2000
} else {
  n = nrow(hfs)
}

hfrs <- heifers[order(heifers$FItbv), ]
selectedhfrs <- hfrs[1:n, ]
fem4 <- rbind(femalesp4, selectedhfrs)
selectedallfemales4 <- fem4[sample(1:nrow(fem4), 9000, replace = FALSE),]
summary(selectedallfemales4[ ,7], na.rm=TRUE)

nrow(femalesp4) #8000
nrow(selectedallfemales4) #9000
damlistp4 = c(selectedallfemales4[ ,1])

##########################
#### 450 selected males and 9000 selected females
n_progeny = nrow(selectedallfemales4)

allselectedmales=selectedmales4[rep(seq_len(nrow(selectedmales4)), each=20),]
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp4 = randomlymatedsires[ ,1]
```

```
####################
selectedparents=rbind(selectedmales4, selectedallfemales4)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]

animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
uniqueanimlist = uniqueanimlist[-1]
subdata <- basepop_p2_p1_data[basepop_p2_p1_data[ ,1] %in% uniqueanimlist, ]
subdata <- data.matrix(subdata)

length(subdata[,1]) #11179
length(uniqueanimlist) #11179
pedp4 = rbind(subdata[
,1:3],cbind(c((nam6+1):(nam6+n_progeny)),sirelistp4[1:n_progeny],damlistp4[1:n_progeny]))

###################
FIave=mean(selectedparents[ ,7], na.rm=TRUE)
FIave
ADGave=mean(selectedparents[ ,8], na.rm=TRUE)
ADGave
WWTave=mean(selectedparents[ ,9], na.rm=TRUE)
WWTave
YWTave=mean(selectedparents[ ,10], na.rm=TRUE)
YWTave

###############
###############
## P4 progeny, repeating the same code used to create P3.

library(MASS)

set.seed(1234)
herdFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(1.05,0.049,0.528,17.749,0.049,0.014,0.287,1.982,0.528,0.287,552.75,465.63,1
7.749,1.982,465.63,1131.95),byrow=TRUE,nrow=4)
```

```
# G matrix
G=matrix(data=c(0.694,0.050,9.945,18.845,0.050,0.005,0.709,1.550,9.945,0.709,230.77,327.934
,18.845,1.550,327.934,598.327),byrow=TRUE,nrow=4)
# Traits averages
mu=c(FIave,ADGave,WWTave,YWTave)

anwr=c((nam6+1):(nam6+n_progeny))
aid = c(pedp4[ ,1])
sid = c(pedp4[ ,2])
did <- c(pedp4[ ,3])

#number of herds
length(unique(selectedallfemales4[ ,6]))
iherd=c(selectedallfemales4[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam7=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam7){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
```

```
damtbv= selectedallfemales4[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}

obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
pedigree4 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree4)
dataped <- pedigree4[pedigree4$id>nam6,]
simdataP4 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2003, herd = iherd, FI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
FItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simdataP4)
nrow(simdataP4) #9000

## combine the data file of (base, P1, P2 and P3) with P4 data file
p4_p3_p2_p1_basepopdata <- rbind(p3_p2_p1_basepopdata,simdataP4)
## Data file for basepop, P1, P2, P3 and P4 (has both observations and TBV)
p4_p3_p2_p1_basepopdata <- subset(p4_p3_p2_p1_basepopdata, p4_p3_p2_p1_basepopdata[
,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1_p2_p3_p4=mean(p4_p3_p2_p1_basepopdata$WWT)
WWTave_base_p1_p2_p3_p4
WWTsd_base_p1_p2_p3_p4=sd(p4_p3_p2_p1_basepopdata$WWT)
WWTsd_base_p1_p2_p3_p4
WWTave_base_p1_p2_p3_p4-(3*WWTsd_base_p1_p2_p3_p4)
WWTave_base_p1_p2_p3_p4+(3*WWTsd_base_p1_p2_p3_p4)
```

```
p4_p3_p2_p1_basepopdata <- subset(p4_p3_p2_p1_basepopdata, WWT >
(WWTave_base_p1_p2_p3_p4-(3*WWTsd_base_p1_p2_p3_p4)))
p4_p3_p2_p1_basepopdata <- subset(p4_p3_p2_p1_basepopdata, WWT <
(WWTave_base_p1_p2_p3_p4+(3*WWTsd_base_p1_p2_p3_p4)))
summary(p4_p3_p2_p1_basepopdata$WWT)
nrow(p4_p3_p2_p1_basepopdata) #34928

# keeping YWT observations within 3 SD
YWTave_base_p1_p2_p3_p4=mean(p4_p3_p2_p1_basepopdata$YWT)
YWTave_base_p1_p2_p3_p4
YWTsd_base_p1_p2_p3_p4=sd(p4_p3_p2_p1_basepopdata$YWT)
YWTsd_base_p1_p2_p3_p4
YWTave_base_p1_p2_p3_p4-(3*YWTsd_base_p1_p2_p3_p4)
YWTave_base_p1_p2_p3_p4+(3*YWTsd_base_p1_p2_p3_p4)
p4_p3_p2_p1_basepopdata <- subset(p4_p3_p2_p1_basepopdata, YWT >
(YWTave_base_p1_p2_p3_p4-(3*YWTsd_base_p1_p2_p3_p4)))
p4_p3_p2_p1_basepopdata <- subset(p4_p3_p2_p1_basepopdata, YWT <
(YWTave_base_p1_p2_p3_p4+(3*YWTsd_base_p1_p2_p3_p4)))
summary(p4_p3_p2_p1_basepopdata$YWT)
nrow(p4_p3_p2_p1_basepopdata) #34923

#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p4_p3_p2_p1_basepopdata <- subset(p4_p3_p2_p1_basepopdata, p4_p3_p2_p1_basepopdata[
,10] - p4_p3_p2_p1_basepopdata[ ,9] > 40)
nrow(p4_p3_p2_p1_basepopdata) #34921

#Remove animals with FI less than 0
p4_p3_p2_p1_basepopdata <- subset(p4_p3_p2_p1_basepopdata, p4_p3_p2_p1_basepopdata[
,7] > 1)
nrow(p4_p3_p2_p1_basepopdata) #34901

simdataP4 <- subset(p4_p3_p2_p1_basepopdata, p4_p3_p2_p1_basepopdata[ ,1] > nam6)
summary(simdataP4$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#99.65  185.76  210.58  211.78  235.95  352.71
summary(simdataP4$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#221.4   360.9   410.0   409.5   456.5   626.7

nrow(simdataP4) #8961
N_sires = length(c(sort(unique(simdataP4$sire))))
N_sires #450
basepop_p1_p2_p3_p4_data <- rbind(basepop_p1_p2_p3_data,simdataP4)
pedigreep4 <- data.frame (id = simdataP4$id, sire = simdataP4$sire, dam = simdataP4$dam)
pedigreep4p3p2p1andbase <- rbind(pedigreep3p2p1andbase,pedigreep4)
```

```
#############################################################################
#############################################################################
## Creating P5 (Year 5)
#1 2 3 4 5 6 7 8 9 10 11 12 13
# id sire dam sex herd CE BW ww pwg CEtbv BWtbv wwtbv pwgtbv
#################################################################
#################################################################
## redefine (the total number of all animals) and (number of base population, P1, P2 and P3) in
previous simulation
###############################
#total number of all animals (Used unique IDs up to now): basepop, P1 and P2
nam8 = nam6+nam7

###############################
## selecting top 5% sires and 80% dams
## selection of sires
averages=by( basepop_p1_p2_p3_data$FItbv, basepop_p1_p2_p3_data$sex, mean)
ave1males=averages[1]
ave1females=averages[2]
SDs=by( basepop_p1_p2_p3_data$FItbv, basepop_p1_p2_p3_data$sex, sd)
sd1males=SDs[1]
sd1females=SDs[2]

males5 <- subset(basepop_p1_p2_p3_data, sex == 1 & year >= 2001)
bulls5 <- males5[order(males5$FItbv), ]
selectedmales5 <- bulls5[1:450, ]
summary(selectedmales5[ ,7], na.rm=TRUE)

nrow(males5) #8610
nrow(selectedmales5) #450
nrow(selectedmales5)/nrow(males5) #0.05

#########################
## selection of dams
## FIRST: Remove females from year 4 (older than 16 years)
## SECOND: Selection of cows and replace with 2000 heifers
## Assume 20% loss

females5 <- subset(selectedallfemales4, year > 4)
femalesp5 <- females5[sample(1:nrow(females5), 8000, replace=FALSE),]
heifers <- subset(basepop_p1_p2_p3_data, sex == 2 & year == 2002)
nrow(heifers) #4411

if(nrow(heifers) > 2000) {
  n = 2000
```

```
} else {
  n = nrow(heifers)
}

hfrs <- heifers[order(heifers$FItbv), ]
selectedhfrs <- hfrs[1:n, ]
fem5 <- rbind(femalesp5, selectedhfrs)
selectedallfemales5 <- fem5[sample(1:nrow(fem5), 9000, replace = FALSE),]
summary(selectedallfemales5[ ,7], na.rm=TRUE)

nrow(femalesp5) #8000
nrow(selectedallfemales5) #9000
damlistp5 = c(selectedallfemales5[ ,1])

#########################
#### 450 selected males and 9000 selected females
n_progeny = nrow(selectedallfemales5)

allselectedmales=selectedmales5[rep(seq_len(nrow(selectedmales5)), each=20),]
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp5 = randomlymatedsires[ ,1]

####################
selectedparents=rbind(selectedmales5, selectedallfemales5)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]

animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
uniqueanimlist = uniqueanimlist[-1]
subdata <- basepop_p1_p2_p3_data[basepop_p1_p2_p3_data[ ,1] %in% uniqueanimlist, ]
subdata <- data.matrix(subdata)

length(subdata[,1]) #12194
length(uniqueanimlist) #12194
pedp5 = rbind(subdata[
,1:3],cbind(c((nam8+1):(nam8+n_progeny)),sirelistp5[1:n_progeny],damlistp5[1:n_progeny]))

##################
FIave=mean(selectedparents[ ,7], na.rm=TRUE)
FIave
ADGave=mean(selectedparents[ ,8], na.rm=TRUE)
ADGave
WWTave=mean(selectedparents[ ,9], na.rm=TRUE)
WWTave
YWTave=mean(selectedparents[ ,10], na.rm=TRUE)
YWTave
```

```
##############
##############
## P5 progeny, repeating the same code used to create P4.

library(MASS)

set.seed(1234)
herdFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(1.05,0.049,0.528,17.749,0.049,0.014,0.287,1.982,0.528,0.287,552.75,465.63,1
7.749,1.982,465.63,1131.95),byrow=TRUE,nrow=4)
# G matrix
G=matrix(data=c(0.694,0.050,9.945,18.845,0.050,0.005,0.709,1.550,9.945,0.709,230.77,327.934
,18.845,1.550,327.934,598.327),byrow=TRUE,nrow=4)
# Traits averages
mu=c(FIave,ADGave,WWTave,YWTave)

anwr=c((nam8+1):(nam8+n_progeny))
aid = c(pedp5[ ,1])
sid = c(pedp5[ ,2])
did <- c(pedp5[ ,3])

#number of herds
length(unique(selectedallfemales5[ ,6]))
iherd=c(selectedallfemales5[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
```

```
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam9=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam9){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
damtbv= selectedallfemales5[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}

obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
pedigree5 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree5)
```

```
dataped <- pedigree5[pedigree5$id>nam8,]
simdataP5 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2004, herd = iherd, FI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
FItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simdataP5)
nrow(simdataP5) #9000

## combine the data file of (base, P1, P2, P3 and P4) with P5 data file
p5_p4_p3_p2_p1_basepopdata <- rbind(p4_p3_p2_p1_basepopdata,simdataP5)
## Data file for basepop and P's (has both observations and TBV)
p5_p4_p3_p2_p1_basepopdata <- subset(p5_p4_p3_p2_p1_basepopdata,
p5_p4_p3_p2_p1_basepopdata[ ,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1_p2_p3_p4_p5=mean(p5_p4_p3_p2_p1_basepopdata$WWT)
WWTave_base_p1_p2_p3_p4_p5
WWTsd_base_p1_p2_p3_p4_p5=sd(p5_p4_p3_p2_p1_basepopdata$WWT)
WWTsd_base_p1_p2_p3_p4_p5
WWTave_base_p1_p2_p3_p4_p5-(3*WWTsd_base_p1_p2_p3_p4_p5)
WWTave_base_p1_p2_p3_p4_p5+(3*WWTsd_base_p1_p2_p3_p4_p5)
p5_p4_p3_p2_p1_basepopdata <- subset(p5_p4_p3_p2_p1_basepopdata, WWT >
(WWTave_base_p1_p2_p3_p4_p5-(3*WWTsd_base_p1_p2_p3_p4_p5)))
p5_p4_p3_p2_p1_basepopdata <- subset(p5_p4_p3_p2_p1_basepopdata, WWT <
(WWTave_base_p1_p2_p3_p4_p5+(3*WWTsd_base_p1_p2_p3_p4_p5)))
summary(p5_p4_p3_p2_p1_basepopdata$WWT)
nrow(p5_p4_p3_p2_p1_basepopdata) #43854

# keeping YWT observations within 3 SD
YWTave_base_p1_p2_p3_p4_p5=mean(p5_p4_p3_p2_p1_basepopdata$YWT)
YWTave_base_p1_p2_p3_p4_p5
YWTsd_base_p1_p2_p3_p4_p5=sd(p5_p4_p3_p2_p1_basepopdata$YWT)
YWTsd_base_p1_p2_p3_p4_p5
YWTave_base_p1_p2_p3_p4_p5-(3*YWTsd_base_p1_p2_p3_p4_p5)
YWTave_base_p1_p2_p3_p4_p5+(3*YWTsd_base_p1_p2_p3_p4_p5)
p5_p4_p3_p2_p1_basepopdata <- subset(p5_p4_p3_p2_p1_basepopdata, YWT >
(YWTave_base_p1_p2_p3_p4_p5-(3*YWTsd_base_p1_p2_p3_p4_p5)))
p5_p4_p3_p2_p1_basepopdata <- subset(p5_p4_p3_p2_p1_basepopdata, YWT <
(YWTave_base_p1_p2_p3_p4_p5+(3*YWTsd_base_p1_p2_p3_p4_p5)))
summary(p5_p4_p3_p2_p1_basepopdata$YWT)
nrow(p5_p4_p3_p2_p1_basepopdata) #43837

#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p5_p4_p3_p2_p1_basepopdata <- subset(p5_p4_p3_p2_p1_basepopdata,
p5_p4_p3_p2_p1_basepopdata[ ,10] - p5_p4_p3_p2_p1_basepopdata[ ,9] > 40)
nrow(p5_p4_p3_p2_p1_basepopdata) #43817
```

```
#Remove animals with FI less than 0
p5_p4_p3_p2_p1_basepopdata <- subset(p5_p4_p3_p2_p1_basepopdata,
p5_p4_p3_p2_p1_basepopdata[ ,7] > 1)
nrow(p5_p4_p3_p2_p1_basepopdata) # 43499

simdataP5 <- subset(p5_p4_p3_p2_p1_basepopdata, p5_p4_p3_p2_p1_basepopdata[ ,1] >
nam8)
summary(simdataP5$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#79.7   154.9   179.8   180.8   205.4   322.3
summary(simdataP5$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#180.2   302.8   350.5   350.9   396.8   582.1

nrow(simdataP5) #8598
N_sires = length(c(sort(unique(simdataP5$sire))))
N_sires #450
basepop_p1_p2_p3_p4_p5_data <- rbind(basepop_p1_p2_p3_p4_data,simdataP5)
pedigreep5 <- data.frame (id = simdataP5$id, sire = simdataP5$sire, dam = simdataP5$dam)
pedigreep5p4p3p2p1andbase <- rbind(pedigreep4p3p2p1andbase,pedigreep5)

################################################################################
################################################################################
## Creating P6 (Year 6)
#1 2 3 4 5 6 7 8 9 10 11 12 13
# id sire dam sex herd CE BW ww pwg CEtbv BWtbv wwtbv pwgtbv
#################################################################
#################################################################
## redefine (the total number of all animals) and (number of base population, P1, P2, P3 and P4)
in previous simulation
################################
#total number of all animals (Used unique IDs up to now): basepop, P1, P2, P3, P4, P5 and P5
nam10 = nam8+nam9

################################
## selecting top 5% sirs and 80% dams
## selection of sires

males6 <- subset(basepop_p1_p2_p3_p4_data, sex == 1 & year >= 2002)
bulls6 <- males6[order(males6$FItbv), ]
selectedmales6 <- bulls6[1:450, ]
summary(selectedmales6[ ,7], na.rm=TRUE)

nrow(males6) #9071
nrow(selectedmales6) #450
```

```
nrow(selectedmales6)/nrow(males6) #0.05

#########################
## selection of dams
## FIRST: Remove females from year 5 (older than 16 years)
## SECOND: Selection of cows and replace with 2000 heifers
## Assume 20% loss

females6 <- subset(selectedallfemales5, sex == 2 & year > 5)
femalesp6 <- females6[sample(1:nrow(females6), 8000, replace=FALSE),]
heifers <- subset(basepop_p1_p2_p3_p4_data, sex == 2 & year == 2003)
nrow(heifers) #4425

if(nrow(heifers) > 2000) {
  n = 2000
} else {
  n = nrow(heifers)
}

hfrs <- heifers[order(heifers$FItbv), ]
selectedhfrs <- hfrs[1:n, ]
nrow(selectedhfrs)
fem6 <- rbind(femalesp6, selectedhfrs)
selectedallfemales6 <- fem6[sample(1:nrow(fem6), 9000, replace = FALSE),]
summary(selectedallfemales6[ ,7], na.rm=TRUE)

nrow(femalesp6) #8000
nrow(selectedallfemales6) #9000
damlistp6 = c(selectedallfemales6[ ,1])

#########################
#### 450 selected males and 9000 selected females
n_progeny = nrow(selectedallfemales6)

allselectedmales=selectedmales6[rep(seq_len(nrow(selectedmales6)), each=20),]
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp6 = randomlymatedsires[ ,1]

#####################
selectedparents=rbind(selectedmales6, selectedallfemales6)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]

animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
uniqueanimlist = uniqueanimlist[-1]
```

```
subdata <- basepop_p1_p2_p3_p4_data[basepop_p1_p2_p3_p4_data[ ,1] %in% uniqueanimlist,
]
subdata <- data.matrix(subdata)

length(subdata[,1]) #12964
length(uniqueanimlist) #12964
pedp6 = rbind(subdata[
,1:3],cbind(c((nam10+1):(nam10+n_progeny)),sirelistp6[1:n_progeny],damlistp6[1:n_progeny]))

###################
FIave=mean(selectedparents[ ,7], na.rm=TRUE)
FIave
ADGave=mean(selectedparents[ ,8], na.rm=TRUE)
ADGave
WWTave=mean(selectedparents[ ,9], na.rm=TRUE)
WWTave
YWTave=mean(selectedparents[ ,10], na.rm=TRUE)
YWTave

###############
###############
## P6 progeny, repeating the same code used to create P5.

library(MASS)

set.seed(1234)
herdFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(1.05,0.049,0.528,17.749,0.049,0.014,0.287,1.982,0.528,0.287,552.75,465.63,1
7.749,1.982,465.63,1131.95),byrow=TRUE,nrow=4)
# G matrix
G=matrix(data=c(0.694,0.050,9.945,18.845,0.050,0.005,0.709,1.550,9.945,0.709,230.77,327.934
,18.845,1.550,327.934,598.327),byrow=TRUE,nrow=4)
# Traits averages
mu=c(FIave,ADGave,WWTave,YWTave)

anwr=c((nam10+1):(nam10+n_progeny))
```

```
aid = c(pedp6[ ,1])
sid = c(pedp6[ ,2])
did <- c(pedp6[ ,3])

#number of herds
length(unique(selectedallfemales6[ ,6]))
iherd=c(selectedallfemales6[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam11=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam11){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
damtbv= selectedallfemales6[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
```

```
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}

obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
pedigree6 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree6)
dataped <- pedigree6[pedigree6$id>nam10,]
simdataP6 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2005, herd = iherd, FI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
FItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simdataP6)
nrow(simdataP6) #9000

## combine the data file of (base, P1, P2, P3, P4 and P5) with P6 data file
p6_p5_p4_p3_p2_p1_basepopdata <- rbind(p5_p4_p3_p2_p1_basepopdata,simdataP6)
## Data file for basepop and P's (has both observations and TBV)
p6_p5_p4_p3_p2_p1_basepopdata <- subset(p6_p5_p4_p3_p2_p1_basepopdata,
p6_p5_p4_p3_p2_p1_basepopdata[ ,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1_p2_p3_p4_p5_p6=mean(p6_p5_p4_p3_p2_p1_basepopdata$WWT)
WWTave_base_p1_p2_p3_p4_p5_p6
WWTsd_base_p1_p2_p3_p4_p5_p6=sd(p6_p5_p4_p3_p2_p1_basepopdata$WWT)
WWTsd_base_p1_p2_p3_p4_p5_p6
WWTave_base_p1_p2_p3_p4_p5_p6-(3*WWTsd_base_p1_p2_p3_p4_p5_p6)
WWTave_base_p1_p2_p3_p4_p5_p6+(3*WWTsd_base_p1_p2_p3_p4_p5_p6)
p6_p5_p4_p3_p2_p1_basepopdata <- subset(p6_p5_p4_p3_p2_p1_basepopdata, WWT >
(WWTave_base_p1_p2_p3_p4_p5_p6-(3*WWTsd_base_p1_p2_p3_p4_p5_p6)))
p6_p5_p4_p3_p2_p1_basepopdata <- subset(p6_p5_p4_p3_p2_p1_basepopdata, WWT <
(WWTave_base_p1_p2_p3_p4_p5_p6+(3*WWTsd_base_p1_p2_p3_p4_p5_p6)))
summary(p6_p5_p4_p3_p2_p1_basepopdata$WWT)
nrow(p6_p5_p4_p3_p2_p1_basepopdata) #52468

# keeping YWT observations within 3 SD
```

```
YWTave_base_p1_p2_p3_p4_p5_p6=mean(p6_p5_p4_p3_p2_p1_basepopdata$YWT)
YWTave_base_p1_p2_p3_p4_p5_p6
YWTsd_base_p1_p2_p3_p4_p5_p6=sd(p6_p5_p4_p3_p2_p1_basepopdata$YWT)
YWTsd_base_p1_p2_p3_p4_p5_p6
YWTave_base_p1_p2_p3_p4_p5_p6-(3*YWTsd_base_p1_p2_p3_p4_p5_p6)
YWTave_base_p1_p2_p3_p4_p5_p6+(3*YWTsd_base_p1_p2_p3_p4_p5_p6)
p6_p5_p4_p3_p2_p1_basepopdata <- subset(p6_p5_p4_p3_p2_p1_basepopdata, YWT >
(YWTave_base_p1_p2_p3_p4_p5_p6-(3*YWTsd_base_p1_p2_p3_p4_p5_p6)))
p6_p5_p4_p3_p2_p1_basepopdata <- subset(p6_p5_p4_p3_p2_p1_basepopdata, YWT <
(YWTave_base_p1_p2_p3_p4_p5_p6+(3*YWTsd_base_p1_p2_p3_p4_p5_p6)))
summary(p6_p5_p4_p3_p2_p1_basepopdata$YWT)
nrow(p6_p5_p4_p3_p2_p1_basepopdata) #52460

#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p6_p5_p4_p3_p2_p1_basepopdata <- subset(p6_p5_p4_p3_p2_p1_basepopdata,
p6_p5_p4_p3_p2_p1_basepopdata[ ,10] - p6_p5_p4_p3_p2_p1_basepopdata[ ,9] > 40)
nrow(p6_p5_p4_p3_p2_p1_basepopdata) #52414

#Remove animals with FI less than 0
p6_p5_p4_p3_p2_p1_basepopdata <- subset(p6_p5_p4_p3_p2_p1_basepopdata,
p6_p5_p4_p3_p2_p1_basepopdata[ ,7] > 1)
nrow(p6_p5_p4_p3_p2_p1_basepopdata) #51445

simdataP6 <- subset(p6_p5_p4_p3_p2_p1_basepopdata, p6_p5_p4_p3_p2_p1_basepopdata[ ,1]
> nam10)
summary(simdataP6$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#67.07  148.03  171.83  173.50  197.87  311.98
summary(simdataP6$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#155.1   287.2   334.7   334.9   381.0   552.4

nrow(simdataP6) #7955
N_sires = length(c(sort(unique(simdataP6$sire))))
N_sires #450
basepop_p1_p2_p3_p4_p5_p6_data <- rbind(basepop_p1_p2_p3_p4_p5_data,simdataP6)
pedigreep6 <- data.frame (id = simdataP6$id, sire = simdataP6$sire, dam = simdataP6$dam)
pedigreep6p5p4p3p2p1andbase <- rbind(pedigreep5p4p3p2p1andbase,pedigreep6)

##############################################################################
##############################################################################
## Creating P7 (Year 7)
#1 2 3 4 5 6 7 8 9 10 11 12 13
# id sire dam sex herd CE BW ww pwg CEtbv BWtbv wwtbv pwgtbv
##############################################################
```

```
#####################################################################
## redefine (the total number of all animals) and (number of base population, P1, P2, P3, P4 and P5) in previous simulation
################################
#total number of all animals (Used unique IDs up to now): basepop, P1, P2, P3, P4, P5, P6 and P7
nam12 = nam10+nam11

################################
## selecting top 5% sirs and 80% dams
## selection of sires

averages=by( basepop_p1_p2_p3_p4_p5_data$FItbv, basepop_p1_p2_p3_p4_p5_data$sex, mean)
ave1males=averages[1]
ave1females=averages[2]
SDs=by( basepop_p1_p2_p3_p4_p5_data$FItbv, basepop_p1_p2_p3_p4_p5_data$sex, sd)
sd1males=SDs[1]
sd1females=SDs[2]

males7 <- subset(basepop_p1_p2_p3_p4_p5_data, sex == 1 & year >= 2003)
bulls7 <- males7[order(males7$FItbv), ]
selectedmales7 <- bulls7[1:450, ]
summary(selectedmales7[ ,7], na.rm=TRUE)

nrow(males7) #8911
nrow(selectedmales7) #450
nrow(selectedmales7)/nrow(males7) #0.05

#########################
## selection of dams
## FIRST: Remove females from year 6 (older than 16 years)
## SECOND: Selection of cows and replace with 2000 heifers
## Assume 20% loss

females7 <- subset(selectedallfemales6, sex == 2 & year > 6)
femalesp7 <- females7[sample(1:nrow(females7), 8000, replace=FALSE),]
heifers <- subset(basepop_p1_p2_p3_p4_p5_data, sex == 2 & year == 2004)
nrow(heifers)

if(nrow(heifers) > 2000) {
  n = 2000
} else {
  n = nrow(heifers)
}
```

```
hfrs <- heifers[order(heifers$FItbv), ]
selectedhfrs <- hfrs[1:n, ]
nrow(selectedhfrs)
fem7 <- rbind(femalesp7, selectedhfrs)
selectedallfemales7 <- fem7[sample(1:nrow(fem7), 9000, replace = FALSE),]
summary(selectedallfemales7[ ,7], na.rm=TRUE)

nrow(femalesp7) #8000
nrow(selectedallfemales7) #9000
damlistp7 = c(selectedallfemales7[ ,1])

##########################
#### 450 selected males and 9000 selected females
n_progeny = nrow(selectedallfemales7)

allselectedmales=selectedmales7[rep(seq_len(nrow(selectedmales7)), each=20),]
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp7= randomlymatedsires[ ,1]

#####################
selectedparents=rbind(selectedmales7, selectedallfemales7)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]

animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
uniqueanimlist = uniqueanimlist[-1]
subdata <- basepop_p1_p2_p3_p4_p5_data[basepop_p1_p2_p3_p4_p5_data[ ,1] %in%
uniqueanimlist, ]
subdata <- data.matrix(subdata)

length(subdata[,1]) #13831
length(uniqueanimlist) #13831
pedp7 = rbind(subdata[
,1:3],cbind(c((nam12+1):(nam12+n_progeny)),sirelistp7[1:n_progeny],damlistp7[1:n_progeny]))

##################
FIave=mean(selectedparents[ ,7], na.rm=TRUE)
FIave
ADGave=mean(selectedparents[ ,8], na.rm=TRUE)
ADGave
WWTave=mean(selectedparents[ ,9], na.rm=TRUE)
WWTave
YWTave=mean(selectedparents[ ,10], na.rm=TRUE)
YWTave

###############
```

```
###############
## P7 progeny, repeating the same code used to create P6.

library(MASS)

set.seed(1234)
herdFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(1.05,0.049,0.528,17.749,0.049,0.014,0.287,1.982,0.528,0.287,552.75,465.63,1
7.749,1.982,465.63,1131.95),byrow=TRUE,nrow=4)
# G matrix
G=matrix(data=c(0.694,0.050,9.945,18.845,0.050,0.005,0.709,1.550,9.945,0.709,230.77,327.934
,18.845,1.550,327.934,598.327),byrow=TRUE,nrow=4)
# Traits averages
mu=c(FIave,ADGave,WWTave,YWTave)

anwr=c((nam12+1):(nam12+n_progeny))
aid = c(pedp7[ ,1])
sid = c(pedp7[ ,2])
did <- c(pedp7[ ,3])

#number of herds
length(unique(selectedallfemales7[ ,6]))
iherd=c(selectedallfemales7[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
```

```
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam13=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam13){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
damtbv= selectedallfemales7[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}

obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
pedigree7 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree7)
dataped <- pedigree7[pedigree7$id>nam12,]
```

```
simdataP7 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2006, herd = iherd, FI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
FItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simdataP7)
nrow(simdataP7) #9000

## combine the data file of (base, P1, P2, P3, P4, P5 and P6) with P7 data file
p7_p6_p5_p4_p3_p2_p1_basepopdata <- rbind(p6_p5_p4_p3_p2_p1_basepopdata,simdataP7)
## Data file for basepop, P1, P2, P3, P4, P5, P6 and P7 (has both observations and TBV)
p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p7_p6_p5_p4_p3_p2_p1_basepopdata,
p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1_p2_p3_p4_p5_p6_p7=mean(p7_p6_p5_p4_p3_p2_p1_basepopdata$WWT)
WWTave_base_p1_p2_p3_p4_p5_p6_p7
WWTsd_base_p1_p2_p3_p4_p5_p6_p7=sd(p7_p6_p5_p4_p3_p2_p1_basepopdata$WWT)
WWTsd_base_p1_p2_p3_p4_p5_p6_p7
WWTave_base_p1_p2_p3_p4_p5_p6_p7-(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7)
WWTave_base_p1_p2_p3_p4_p5_p6_p7+(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7)
p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p7_p6_p5_p4_p3_p2_p1_basepopdata, WWT >
(WWTave_base_p1_p2_p3_p4_p5_p6_p7-(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7)))
p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p7_p6_p5_p4_p3_p2_p1_basepopdata, WWT <
(WWTave_base_p1_p2_p3_p4_p5_p6_p7+(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7)))
summary(p7_p6_p5_p4_p3_p2_p1_basepopdata$WWT)
nrow(p7_p6_p5_p4_p3_p2_p1_basepopdata) #60414

# keeping YWT observations within 3 SD
YWTave_base_p1_p2_p3_p4_p5_p6_p7=mean(p7_p6_p5_p4_p3_p2_p1_basepopdata$YWT)
YWTave_base_p1_p2_p3_p4
YWTsd_base_p1_p2_p3_p4_p5_p6_p7=sd(p7_p6_p5_p4_p3_p2_p1_basepopdata$YWT)
YWTsd_base_p1_p2_p3_p4
YWTave_base_p1_p2_p3_p4-(3*YWTsd_base_p1_p2_p3_p4)
YWTave_base_p1_p2_p3_p4+(3*YWTsd_base_p1_p2_p3_p4)
p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p7_p6_p5_p4_p3_p2_p1_basepopdata, YWT >
(YWTave_base_p1_p2_p3_p4-(3*YWTsd_base_p1_p2_p3_p4)))
p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p7_p6_p5_p4_p3_p2_p1_basepopdata, YWT <
(YWTave_base_p1_p2_p3_p4+(3*YWTsd_base_p1_p2_p3_p4)))
summary(p7_p6_p5_p4_p3_p2_p1_basepopdata$YWT)
nrow(p7_p6_p5_p4_p3_p2_p1_basepopdata) #58898

#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p7_p6_p5_p4_p3_p2_p1_basepopdata,
p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,10] - p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,9] > 40)
nrow(p7_p6_p5_p4_p3_p2_p1_basepopdata) #57183
```

```
#Remove animals with FI less than 0
p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p7_p6_p5_p4_p3_p2_p1_basepopdata,
p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,7] > 1)
nrow(p7_p6_p5_p4_p3_p2_p1_basepopdata) #57147

simdataP7 <- subset(p7_p6_p5_p4_p3_p2_p1_basepopdata,
p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,1] > nam12)
summary(simdataP7$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#54.03  139.43  162.99  164.54  187.24  311.78
summary(simdataP7$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#219.8   273.2   315.3   317.6   357.5   518.8

nrow(simdataP7) #5958
N_sires = length(c(sort(unique(simdataP7$sire))))
N_sires #450
basepop_p1_p2_p3_p4_p5_p6_p7_data <- rbind(basepop_p1_p2_p3_p4_p5_p6_data,simdataP7)
pedigreep7 <- data.frame (id = simdataP7$id, sire = simdataP7$sire, dam = simdataP7$dam)
pedigreep7p6p5p4p3p2p1andbase <- rbind(pedigreep6p5p4p3p2p1andbase,pedigreep7)

##############################################################################
##############################################################################
## Creating P8 (Year 8)
#1 2 3 4 5 6 7 8 9 10 11 12 13
# id sire dam sex herd CE BW ww pwg CEtbv BWtbv wwtbv pwgtbv
#####################################################################
#####################################################################
## redefine (the total number of all animals) and (number of base population, P1, P2, P3, P4, P5
and P6) in previous simulation
################################
#total number of all animals (Used unique IDs up to now): basepop, P1, P2, P3, P4, P5, P6, P7
and P8
nam14 = nam12+nam13

################################
## selecting top 5% sires
## selection of sires

males8 <- subset(basepop_p1_p2_p3_p4_p5_p6_data, sex == 1 & year >= 2004)
bulls8 <- males8[order(males8$FItbv), ]
selectedmales8 <- bulls8[1:450, ]
summary(selectedmales8[ ,7], na.rm=TRUE)

nrow(males8) #8413
nrow(selectedmales8) #450
```

```
nrow(selectedmales8)/nrow(males8) #0.05

#########################
## selection of dams
## FIRST: Remove females from year 7 (older than 16 years)
## SECOND: Selection of cows and replace with 2000 heifers
## Assume 20% loss

females8 <- subset(selectedallfemales7, sex == 2 & year > 7)
femalesp8 <- females8[sample(1:nrow(females8), 8000, replace=FALSE),]
heifers <- subset(basepop_p1_p2_p3_p4_p5_p6_data, sex == 2 & year == 2005)
nrow(heifers) #3917

if(nrow(heifers) > 2000) {
  n = 2000
} else {
  n = nrow(heifers)
}

hfrs <- heifers[order(heifers$FItbv), ]
selectedhfrs <- hfrs[1:n, ]
nrow(selectedhfrs)
fem8 <- rbind(femalesp8, selectedhfrs)
selectedallfemales8 <- fem8[sample(1:nrow(fem8), 9000, replace = FALSE),]
summary(selectedallfemales8[ ,7], na.rm=TRUE)

nrow(femalesp8) #8000
nrow(selectedallfemales8) #9000
damlistp8 = c(selectedallfemales8[ ,1])

#########################
#### 450 selected males and 9000 selected females
n_progeny = nrow(selectedallfemales8)

allselectedmales=selectedmales8[rep(seq_len(nrow(selectedmales8)), each=20),]
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp8 = randomlymatedsires[ ,1]

#####################
selectedparents=rbind(selectedmales8, selectedallfemales8)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]

animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
uniqueanimlist = uniqueanimlist[-1]
```

```
subdata <- basepop_p1_p2_p3_p4_p5_p6_data[basepop_p1_p2_p3_p4_p5_p6_data[ ,1] %in%
uniqueanimlist, ]
subdata <- data.matrix(subdata)

length(subdata[,1]) #14482
length(uniqueanimlist) #14482
pedp8 = rbind(subdata[
,1:3],cbind(c((nam14+1):(nam14+n_progeny)),sirelistp8[1:n_progeny],damlistp8[1:n_progeny]))

###################
FIave=mean(selectedparents[ ,7], na.rm=TRUE)
FIave
ADGave=mean(selectedparents[ ,8], na.rm=TRUE)
ADGave
WWTave=mean(selectedparents[ ,9], na.rm=TRUE)
WWTave
YWTave=mean(selectedparents[ ,10], na.rm=TRUE)
YWTave

###############
###############
## P8 progeny, repeating the same code used to create P7.

library(MASS)

set.seed(1234)
herdFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(1.05,0.049,0.528,17.749,0.049,0.014,0.287,1.982,0.528,0.287,552.75,465.63,1
7.749,1.982,465.63,1131.95),byrow=TRUE,nrow=4)
# G matrix
G=matrix(data=c(0.694,0.050,9.945,18.845,0.050,0.005,0.709,1.550,9.945,0.709,230.77,327.934
,18.845,1.550,327.934,598.327),byrow=TRUE,nrow=4)
# Traits averages
mu=c(FIave,ADGave,WWTave,YWTave)

anwr=c((nam14+1):(nam14+n_progeny))
```

```
aid = c(pedp8[ ,1])
sid = c(pedp8[ ,2])
did <- c(pedp8[ ,3])

#number of herds
length(unique(selectedallfemales8[ ,6]))
iherd=c(selectedallfemales8[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam15=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam15){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
damtbv= selectedallfemales8[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
```

```
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}

obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
pedigree8 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree8)
dataped <- pedigree8[pedigree8$id>nam14,]
simdataP8 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2007, herd = iherd, FI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
FItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simdataP8)
nrow(simdataP8) #9000

## combine the data file of (base, P1, P2, P3, P4, P5, P6 and P7) with P8 data file
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
rbind(p7_p6_p5_p4_p3_p2_p1_basepopdata,simdataP8)
## Data file for basepop and P's (has both observations and TBV)
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8=mean(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$
WWT)
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8
WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8=sd(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$WW
T)
WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8-(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8)
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8+(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8)
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
WWT > (WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8-
(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8)))
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
WWT <
```

```
(WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8+(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8))
)
summary(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$WWT)
nrow(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #66112

# keeping YWT observations within 3 SD
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8=mean(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$
YWT)
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8
YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8=sd(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$YWT
)
YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8-(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8)
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8+(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8)
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
YWT > (YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8-
(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8)))
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
YWT <
(YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8+(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8)))
summary(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$YWT)
nrow(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #66102

#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,10] - p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,9]
> 40)
nrow(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #65787

#Remove animals with FI less than 0
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,7] > 1)
nrow(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #61787

simdataP8 <- subset(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,1] > nam14)
summary(simdataP8$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#42.85  128.13  151.16  152.80  176.31  312.48
summary(simdataP8$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#127.7   246.3   293.4   292.4   336.2   485.6

nrow(simdataP8) #4640 Loss of animal due to FI < 1
N_sires = length(c(sort(unique(simdataP8$sire))))
```

```
N_sires #450
basepop_p1_p2_p3_p4_p5_p6_p7_p8_data <-
rbind(basepop_p1_p2_p3_p4_p5_p6_p7_data,simdataP8)
pedigreep8 <- data.frame (id = simdataP8$id, sire = simdataP8$sire, dam = simdataP8$dam)
pedigreep8p7p6p5p4p3p2p1andbase <- rbind(pedigreep7p6p5p4p3p2p1andbase,pedigreep8)

##############################################################################
##############################################################################
## Creating P9 (Year 9)
#1 2 3 4 5 6 7 8 9 10 11 12 13
# id sire dam sex herd CE BW ww pwg CEtbv BWtbv wwtbv pwgtbv
#####################################################################
#####################################################################
## redefine (the total number of all animals) and (number of base population, P1, P2, P3, P4, P5,
P6, P7 and P8) in previous simulation
################################
#total number of all animals (Used unique IDs up to now): basepop, P1, P2, P3, P4, P5, P6, P7,
P8 and P9
nam16 = nam14+nam15

################################
## selecting top 5% sires
## selection of sires

males9 <- subset(basepop_p1_p2_p3_p4_p5_p6_p7_data, sex == 1 & year >= 2005)
bulls9 <- males9[order(males9$FItbv), ]
selectedmales9 <- bulls9[1:450, ]
summary(selectedmales9[ ,7], na.rm=TRUE)

nrow(males9) #7324
nrow(selectedmales9) #450
nrow(selectedmales9)/nrow(males9) #0.06

#########################
## selection of dams
## FIRST: Remove females from year 8 (older than 16 years)
## SECOND: Selection of cows and replace with 2000 heifers
## Assume 20% loss

females9 <- subset(selectedallfemales8, sex == 2 & year > 8)
femalesp9 <- females9[sample(1:nrow(females9), 8000, replace=FALSE),]
heifers <- subset(basepop_p1_p2_p3_p4_p5_p6_p7_data, sex == 2 & year == 2006)
nrow(heifers)

if(nrow(heifers) > 2000) {
  n = 2000
```

```
} else {
  n = nrow(heifers)
}

hfrs <- heifers[order(heifers$FItbv), ]
selectedhfrs <- hfrs[1:n, ]
nrow(selectedhfrs)
fem9 <- rbind(femalesp8, selectedhfrs)
selectedallfemales9 <- fem8[sample(1:nrow(fem9), 9000, replace = FALSE),]
summary(selectedallfemales9[ ,7], na.rm=TRUE)

nrow(femalesp9) #8000
nrow(selectedallfemales9) #10000
damlistp9 = c(selectedallfemales9[ ,1])

##########################
#### 450 selected males and 9000 selected females
n_progeny = nrow(selectedallfemales9)

allselectedmales=selectedmales9[rep(seq_len(nrow(selectedmales9)), each=20),]
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp9 = randomlymatedsires[ ,1]

#####################
selectedparents=rbind(selectedmales9, selectedallfemales9)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]

animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
uniqueanimlist = uniqueanimlist[-1]
subdata <- basepop_p1_p2_p3_p4_p5_p6_p7_data[basepop_p1_p2_p3_p4_p5_p6_p7_data[ ,1]
%in% uniqueanimlist, ]
subdata <- data.matrix(subdata)

length(subdata[,1]) #14565
length(uniqueanimlist) #14565
pedp9 = rbind(subdata[
,1:3],cbind(c((nam16+1):(nam16+n_progeny)),sirelistp9[1:n_progeny],damlistp9[1:n_progeny]))

##################
FIave=mean(selectedparents[ ,7], na.rm=TRUE)
FIave
ADGave=mean(selectedparents[ ,8], na.rm=TRUE)
ADGave
WWTave=mean(selectedparents[ ,9], na.rm=TRUE)
WWTave
```

```
YWTave=mean(selectedparents[ ,10], na.rm=TRUE)
YWTave

###############
###############
## P9 progeny, repeating the same code used to create P8.

library(MASS)

set.seed(1234)
herdFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(1.05,0.049,0.528,17.749,0.049,0.014,0.287,1.982,0.528,0.287,552.75,465.63,1
7.749,1.982,465.63,1131.95),byrow=TRUE,nrow=4)
# G matrix
G=matrix(data=c(0.694,0.050,9.945,18.845,0.050,0.005,0.709,1.550,9.945,0.709,230.77,327.934
,18.845,1.550,327.934,598.327),byrow=TRUE,nrow=4)

####averages:
mu=c(FIave,ADGave,WWTave,YWTave)
anwr=c((nam16+1):(nam16+n_progeny))
aid = c(pedp9[ ,1])
sid = c(pedp9[ ,2])
did <- c(pedp9[ ,3])

#number of herds
length(unique(selectedallfemales9[ ,6]))
iherd=c(selectedallfemales9[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2
```

```
# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam17=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam1){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
damtbv= selectedallfemales9[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}

obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
```

```
pedigree9 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree9)
dataped <- pedigree9[pedigree9$id>nam16,]
simdataP9 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2008, herd = iherd, FI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
FItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simdataP9)
nrow(simdataP9) #9000

## combine the data file of (base, P1, P2, P3, P4, P5, P6, P7 and P8) with P9 data file
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
rbind(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,simdataP9)
## Data file for basepop and P's (has both observations and TBV)
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9=mean(p9_p8_p7_p6_p5_p4_p3_p2_p1_basep
opdata$WWT)
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9
WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9=sd(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdat
a$WWT)
WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9-
(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9)
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9+(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p
8_p9)
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata, WWT >
(WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9-
(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9)))
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata, WWT <
(WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9+(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_
p8_p9)))
summary(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$WWT)
nrow(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #70753

# keeping YWT observations within 3 SD
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9=mean(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepo
pdata$YWT)
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9
YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9=sd(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdat
a$YWT)
YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9
```

```
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9-
(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9)
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9+(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8
_p9)
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata, YWT >
(YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9-
(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9)))
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata, YWT <
(YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9+(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p
8_p9)))
summary(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$YWT)
nrow(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #70746

#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,10] -
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,9] > 40)
nrow(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #70408

#Remove animals with FI less than 0
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,7] > 1)
nrow(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #66280

simdataP9 <- subset(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,1] > nam16)
summary(simdataP9$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#37.79  127.37  150.44  152.12  175.37  313.02
summary(simdataP9$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#133.1   244.6   291.9   290.8   334.7   500.6

nrow(simdataP9) #4504
N_sires = length(c(sort(unique(simdataP9$sire))))
N_sires # 450
basepop_p1_p2_p3_p4_p5_p6_p7_p8_p9_data <-
rbind(basepop_p1_p2_p3_p4_p5_p6_p7_p8_data,simdataP9)
pedigreep9 <- data.frame (id = simdataP9$id, sire = simdataP9$sire, dam = simdataP9$dam)
pedigreep9p8p7p6p5p4p3p2p1andbase <-
rbind(pedigreep8p7p6p5p4p3p2p1andbase,pedigreep9)
```

```
#############################################################################
#############################################################################
## Creating P10 (Year 10)
#1 2 3 4 5 6 7 8 9 10 11 12 13
# id sire dam sex herd CE BW ww pwg CEtbv BWtbv wwtbv pwgtbv
#####################################################################
#####################################################################
## redefine (the total number of all animals) and (number of base population, P1, P2, P3, P4, P5,
P6, P7 and P8) in previous simulation
################################
#total number of all animals (Used unique IDs up to now): basepop, P1, P2, P3, P4, P5, P6, P7,
P8 and P9
nam18 = nam16+nam17

################################
## selecting top 5% sires
## selection of sires
averages=by( basepop_p1_p2_p3_p4_p5_p6_p7_p8_data$FItbv,
basepop_p1_p2_p3_p4_p5_p6_p7_p8_data$sex, mean)
ave1males=averages[1]
ave1females=averages[2]
SDs=by( basepop_p1_p2_p3_p4_p5_p6_p7_p8_data$FItbv,
basepop_p1_p2_p3_p4_p5_p6_p7_p8_data$sex, sd)
sd1males=SDs[1]
sd1females=SDs[2]

males10 <- subset(basepop_p1_p2_p3_p4_p5_p6_p7_p8_data, sex == 1 & year >= 2006)
bulls10 <- males10[order(males10$FItbv), ]
selectedmales10 <- bulls10[1:450, ]
summary(selectedmales10[ ,7], na.rm=TRUE)

nrow(males10) #5687
nrow(selectedmales10) #450
nrow(selectedmales10)/nrow(males10) #0.08

#########################
## selection of dams
## FIRST: Remove females from year 9 (older than 16 years)
## SECOND: Selection of cows and replace with 2000 heifers
## Assume 20% loss

females10 <- subset(selectedallfemales9, sex == 2 & year > 9)
femalesp10 <- females10[sample(1:nrow(females10), 8000, replace=FALSE),]
heifers <- subset(basepop_p1_p2_p3_p4_p5_p6_p7_p8_data, sex == 2 & year == 2007)
nrow(heifers)
```

```
if(nrow(heifers) > 2000) {
  n = 2000
} else {
  n = nrow(heifers)
}

hfrs <- heifers[order(heifers$FItbv), ]
selectedhfrs <- hfrs[1:n, ]
fem10 <- rbind(femalesp10, heifers)
selectedallfemales10 <- fem10[sample(1:nrow(fem10), 9000, replace = FALSE),]

nrow(femalesp10) #8000
nrow(selectedallfemales10) #10000
damlistp10 = c(selectedallfemales10[ ,1])

##########################
#### 450 selected males and 9000 selected females
n_progeny = nrow(selectedallfemales10)

allselectedmales=selectedmales10[rep(seq_len(nrow(selectedmales10)), each=20),]
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp10 = randomlymatedsires[ ,1]

#####################
selectedparents=rbind(selectedmales10, selectedallfemales10)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]

animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
uniqueanimlist = uniqueanimlist[-1]
subdata <-
basepop_p1_p2_p3_p4_p5_p6_p7_p8_data[basepop_p1_p2_p3_p4_p5_p6_p7_p8_data[ ,1]
%in% uniqueanimlist, ]
subdata <- data.matrix(subdata)

length(subdata[,1]) #15073
length(uniqueanimlist) #15073
pedp10 = rbind(subdata[
,1:3],cbind(c((nam18+1):(nam18+n_progeny)),sirelistp10[1:n_progeny],damlistp10[1:n_progeny
]))

##################
FIave=mean(selectedparents[ ,7], na.rm=TRUE)
FIave
ADGave=mean(selectedparents[ ,8], na.rm=TRUE)
```

```
ADGave
WWTave=mean(selectedparents[ ,9], na.rm=TRUE)
WWTave
YWTave=mean(selectedparents[ ,10], na.rm=TRUE)
YWTave

###############
###############
## P10 progeny, repeating the same code used to create P9.

set.seed(1234)
herdFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(1.05,0.049,0.528,17.749,0.049,0.014,0.287,1.982,0.528,0.287,552.75,465.63,1
7.749,1.982,465.63,1131.95),byrow=TRUE,nrow=4)
# G matrix
G=matrix(data=c(0.694,0.050,9.945,18.845,0.050,0.005,0.709,1.550,9.945,0.709,230.77,327.934
,18.845,1.550,327.934,598.327),byrow=TRUE,nrow=4)

####averages:
mu=c(FIave,ADGave,WWTave,YWTave)
anwr=c((nam18+1):(nam18+n_progeny))
aid = c(pedp10[ ,1])
sid = c(pedp10[ ,2])
did <- c(pedp10[ ,3])

#number of herds
length(unique(selectedallfemales10[ ,6]))
iherd=c(selectedallfemales10[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2
```

```
# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam19=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam1){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
damtbv= selectedallfemales10[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}

obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)
```

```
#########################################################################
pedigree10 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree10)
dataped <- pedigree10[pedigree10$id>nam18,]
simdataP10 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2009, herd = iherd, FI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
FItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simdataP10)
nrow(simdataP10) #9000

## combine the data file of (base, P1, P2, P3, P4, P5, P6, P7, P8 and P9) with P10 data file
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
rbind(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,simdataP10)
## Data file for basepop and P's (has both observations and TBV)
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10=mean(p10_p9_p8_p7_p6_p5_p4_p3_p2_
p1_basepopdata$WWT)
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10
WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10=sd(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_b
asepopdata$WWT)
WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10-
(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10)
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10+(3*WWTsd_base_p1_p2_p3_p4_p5_p6_
p7_p8_p9_p10)
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata, WWT >
(WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10-
(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10)))
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata, WWT <
(WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10+(3*WWTsd_base_p1_p2_p3_p4_p5_p6
_p7_p8_p9_p10)))
summary(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$WWT)
nrow(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #75255

# keeping YWT observations within 3 SD
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10=mean(p10_p9_p8_p7_p6_p5_p4_p3_p2_
p1_basepopdata$YWT)
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10
YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10=sd(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_b
asepopdata$YWT)
```

```
YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10-
(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10)
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10+(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p
7_p8_p9_p10)
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata, YWT >
(YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10-
(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10)))
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata, YWT <
(YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10+(3*YWTsd_base_p1_p2_p3_p4_p5_p6_
p7_p8_p9_p10)))
summary(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$YWT)
nrow(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #75245

#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,9] -
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,8] > 40)
nrow(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #75203

#Remove animals with FI less than 0
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,7] > 1)
nrow(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #69756

simdataP10 <- subset(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,1] > nam18)
summary(simdataP10$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
# 40.7   120.8   144.6   146.3   171.0   284.6
summary(simdataP10$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#101.2   226.3   272.5   272.3   316.8   472.7

nrow(simdataP10) #3477
N_sires = length(c(sort(unique(simdataP10$sire))))
N_sires #450
basepop_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10_data <-
rbind(basepop_p1_p2_p3_p4_p5_p6_p7_p8_p9_data,simdataP10)
pedigreep10 <- data.frame (id = simdataP10$id, sire = simdataP10$sire, dam =
simdataP10$dam)
```

```
pedigreep10p9p8p7p6p5p4p3p2p1andbase <-
rbind(pedigreep9p8p7p6p5p4p3p2p1andbase,pedigreep10)

##############################################################################
##############################################################################
## see how many records per sire
try <- subset(basepop_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10_data,
basepop_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10_data[ ,1] > nbase)
nrow(try)
ones = c(rep(1,(nrow(try))))
try = data.matrix(try)
try = cbind(try,ones)
ham=sort(by( try[ ,14], try[ ,2], length))
head(ham)

##############################################
## Final data files
## pedigree file
## ****data file for animals with records (will be used for ****ASREML3.0****)
data_anim_with_record <- subset(basepop_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10_data,
basepop_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10_data[ ,1] > nbase)
summary(data_anim_with_record$FI)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#-0.930   4.400   6.160   6.218   7.990  14.260
summary(data_anim_with_record$YWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#76.69  257.10  302.50  304.20  350.70  527.00
summary(data_anim_with_record$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#84.4   171.1   196.3   197.6   222.8   325.9
summary(data_anim_with_record$ADG)
# Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#-0.7200  0.4400  0.7200  0.7306  1.0100  2.3100
```

APPENDIX C: R CODE USED TO CREATE SIMULATED DATA FOR SINGLE TRAIT SELECTION ON RESIDUAL FEED INTAKE

```
##Finalized simulation for RFI selection

#R code based on Dr. Larry Schaeffer's R code for multiple trait models && Hamad Saad's
disertation
#Creating a simulated data (Selection on decreasing residual feed intake)
## Traits were: Feed intake (RFI), 200 d weight (WWT), 400 d weight (YWT), and average
daily gain (ADG)
# First creating the base population (No selection)
setwd("/Users/Randie/Documents/PhD Research/Simulation")

## To clear environment
#RFISimData_v3
rm(list = ls())

##Run DistFem_USE.R first
source("DistFem_USE.R")

library(MASS)

#Herd effects (50 herds) on RFI, ADG, WWT and YWT
set.seed(1234)
herdRFI=(rnorm(50,0))
set.seed(1234)
herdWW=(rnorm(50,0))*2
set.seed(1234)
herdYW=(rnorm(50,0))*10
set.seed(1234)
herdADG=(rnorm(50,0))*0.25  ##Crews et al 2006 (SD for ADG)
herd=matrix(data = c(herdRFI,herdADG,herdWW,herdYW),byrow = TRUE, nrow = 4)

# sex effects (52) herds) on RFI, ADG, WWT, and YWT
##Sex effects for wwt and ywwt from Van Vleck & Cundiff 1998
##ADG sex effects used Hamad's
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)
# Residual matrix
R=matrix(data=c(0.311,-0.009,0.219,0.579,-
0.009,0.014,0.548,1.231,0.219,0.548,552.753,465.634,0.579,1.231,465.634,1131.95),byrow=TR
UE,nrow=4)
# G matrix
G=matrix(data=c(0.18,0.0005,-1.372,-2.249,0.0005,0.01,1.014,2.217,-
1.372,1.014,230.77,327.931,-2.249,2.217,327.931,598.33),byrow=TRUE,nrow=4)
# Traits averages
mu=c(-0.011,1.473,258.913,479.990) #weighted averages from literature estimates

#Matrices are positive definite
#eigen values for R: 1.390698e+03 2.940065e+02 3.110134e-01 1.234581e-02
```

```
#eigen values for G: 7.904835e+02 3.863431e+01 1.715661e-01 5.907012e-04

#ID for animals (founders (10000 dams, 450 sires))
aid = c(1:10400)

##creating fields in the data file for sire (20 progeny each), dams (1 progeny each), herds (size of
200 each), sex (50% females and 50% males)
sid = c(numeric(10400),rep(1:400, by=1, each=20))
base.sire <- data.frame(id = c(10001:10400), year = 15, sex = 1)
did <- c(numeric(10400),1:10000)
bi=c(rep(1,10400),rep(0.5,10000))
set.seed(1234)
iherd=c(sample(rep(1:52, by=1, each=200),10400,replace=F))

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

nam = 10400
# Simulate true breeding values for all animals (founders and their F1 progeny)
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}

tbv = jd(nam,4)*0
for(i in 1:nam){
  x = LG %*% (sqrt(bi[i])*rnorm(4,0,1))
  if(sid[i]>0){
    ks=sid[i]
    kd=did[i]
```

```
   x = x + 0.5*(tbv[ks, ]+tbv[kd, ]) }
  tbv[i,]=x }
nrec=10400

#########################################################################
#########################################################################
##Creating base popuation with tbv but no observations
####Use base.females from DistFem_USE.R

base.pop <- rbind(base.females, base.sire)
simulateddata <- data.frame (id = base.pop$id, sire = c(rep(NA,10400)), dam =
c(rep(NA,10400)), sex = base.pop$sex, year = base.pop$year, herd = iherd, RFI =
c(rep(NA,10400)), ADG = c(rep(NA,10400)), WWT = c(rep(NA,10400)), YWT =
c(rep(NA,10400)), RFItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[
,4])
attach(simulateddata)

animlist = c(simulateddata[ ,1])
detach(simulateddata)
uniqueanimlist = c(sort(unique(animlist)))
pedigree = data.frame(id = uniqueanimlist, sire = c(rep(NA,10400)), dam = c(rep(NA,10400)))
length(uniqueanimlist) #10400
length(pedigree[,1]) #10400

#Header: id sire dam sex year herd RFI ADG WWT YWT RFItbv ADGtbv WWTtbv YWTtbv
basepopdata <- simulateddata
attach(basepopdata)

#####################################################################
## creating P1 (progeny 1)
## Assuming 100% conception
## No selection of RFI.  Must use whole population for first 2 generations
#####################################################################
## redefine the total number of all animals and number of base population in previous simulation
################################
#total number of all animals (base pop + P1)
nam = 20400
#number of sires and dams(founders) (400+10000) in base population
nbase = 10400

################################
## selecting top 5% sirs and 80% dams (TBV for RFI is the selection criteria)
## selection of sires
## average RFI TBV for sires and dams
averages=by( basepopdata$RFItbv, basepopdata$sex, mean)
ave1males=averages[1]
```

```
ave1females=averages[2]
## standard deviation of RFI TBV for sires and dams
SDs=by( basepopdata$RFItbv, basepopdata$sex, sd)
sd1males=SDs[1]
sd1females=SDs[2]

## selecting all sires from basepopdata
selectedmales <- subset(basepopdata, sex == 1)

#########################
## selection of all dams
## For first selection of cows from P1 only selected 100% of females to be mated to sires
females.base <- subset(basepopdata, sex == 2)
##90% Conception Rate
selectedallfemales <- females.base[sample(1:nrow(females.base), 9000, replace = FALSE),]

nrow(females.base)
nrow(selectedallfemales)
nrow(selectedallfemales)/nrow(females.base) #0.9

#########################
#### 400 selected males and 10000 selected females
n_sel_males = nrow(selectedmales)
n_sel_females = nrow(selectedallfemales)
n_progeny = n_sel_females
# number of females per sire = 22.5
n_females_per_sire = n_sel_females/n_sel_males
n_females_per_sire
# because number of dams per sire is 22.5, then sires will have different numbers of progeny
(some will have 22.5 and others will have 23 progeny)
rounded_n_females_per_sire = round(n_females_per_sire)
rounded_n_females_per_sire

if(rounded_n_females_per_sire < n_females_per_sire) {
  n1records_per_sire = (rounded_n_females_per_sire)
  n2records_per_sire = (n1records_per_sire)+1
} else {
  n1records_per_sire = (rounded_n_females_per_sire)-1
  n2records_per_sire = (rounded_n_females_per_sire)
}

#number of sires with 22 progeny (200 out of 400)
nsires_with_n1records = (n_sel_males)-((n_sel_females)-(n_sel_males*n1records_per_sire))
#number of sires with 23 progeny (200 out of 400)
nsires_with_n2records = (n_sel_males)-(nsires_with_n1records)
##pulling out the first 137 sires of sire list
```

```
own_record_sires_with_n1records = selectedmales[1:nsires_with_n1records, ]
## replicate each sire 22 times
selectedmales1=own_record_sires_with_n1records[rep(seq_len(nrow(own_record_sires_with_n
1records)), each=n1records_per_sire),]
##pulling out the remaining 200 sires of sire list
own_record_sires_with_n2records = selectedmales[((nsires_with_n1records)+1):n_sel_males, ]
## replicate each sire 24 times
selectedmales2=own_record_sires_with_n2records[rep(seq_len(nrow(own_record_sires_with_n
2records)), each=n2records_per_sire),]

## Create random list of sires with length of 10000 which is number of selected dams
allselectedmales=rbind(selectedmales1,selectedmales2)
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp1 = randomlymatedsires[ ,1]
damlistp1 = selectedallfemales[ ,1]

####################
selectedparents=rbind(selectedmales, selectedallfemales)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]
# pulling out selected parents and their pedigree
animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
subdata <- basepopdata[basepopdata[ ,1] %in% uniqueanimlist, ]
subdata <- data.matrix(subdata)
length(subdata[,1]) #9400
length(uniqueanimlist) #9400

## creating P1 pedigree which include(selected parents and their pedigree + new P1 ID with their
selected parents)
pedP1 =
rbind(subdata[,1:3],cbind(c((nam+1):(nam+n_progeny)),sirelistp1[1:n_progeny],damlistp1[1:n_
progeny]))
###################

###############
## P1 observations

library(MASS)

set.seed(1234)
herdRFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
```

```
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdRFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(0.311,-0.040,0.681,1.806,-
0.04,0.014,0.548,1.231,0.681,0.548,552.753,465.634,1.806,1.231,465.634,1131.95),byrow=TRU
E,nrow=4)
# G matrix
G=matrix(data=c(0.18,0.002,-4.265,-6.994,0.002,0.01,1.014,2.217,-
4.265,1.014,230.77,327.931,-6.994,2.217,327.931,598.33),byrow=TRUE,nrow=4)
# Traits averages
mu=c(-0.011,1.473,258.913,479.990)

anwr=c((nam+1):(nam+n_progeny))
aid = c(pedP1[ ,1])
sid = c(pedP1[ ,2])
did <- c(pedP1[ ,3])

#number of herds
length(unique(selectedallfemales[ ,6]))
iherd=c(selectedallfemales[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
```

```
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam1=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam1){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
damtbv= selectedallfemales[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}
obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
pedigree1 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree1)
dataped <- pedigree1[pedigree1$id>nam,]
simdataP1 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2000, herd = iherd, RFI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
RFItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simdataP1)
nrow(simdataP1) #9000

## combine the data file of (base) with P1 data file
p1andbasepopdata <- rbind(basepopdata,simdataP1)
## Data file for basepop and P1 (has both observations and TBV)
p1andbasepopdata <- subset(p1andbasepopdata, p1andbasepopdata[ ,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1=mean(p1andbasepopdata$WWT)
```

```
WWTave_base_p1
WWTsd_base_p1=sd(p1andbasepopdata$WWT)
WWTsd_base_p1
WWTave_base_p1-(3*WWTsd_base_p1)
WWTave_base_p1+(3*WWTsd_base_p1)
p1andbasepopdata <- subset(p1andbasepopdata, WWT > (WWTave_base_p1-
(3*WWTsd_base_p1)))
p1andbasepopdata <- subset(p1andbasepopdata, WWT <
(WWTave_base_p1+(3*WWTsd_base_p1)))
summary(p1andbasepopdata$WWT)
nrow(p1andbasepopdata) #8960

# keeping YWT observations within 3 SD
YWTave_base_p1=mean(p1andbasepopdata$YWT)
YWTave_base_p1
YWTsd_base_p1=sd(p1andbasepopdata$YWT)
YWTsd_base_p1
YWTave_base_p1-(3*YWTsd_base_p1)
YWTave_base_p1+(3*YWTsd_base_p1)
p1andbasepop <- subset(p1andbasepopdata, YWT > (YWTave_base_p1-(3*YWTsd_base_p1)))
p1andbasepopdata <- subset(p1andbasepopdata, YWT <
(YWTave_base_p1+(3*YWTsd_base_p1)))
summary(p1andbasepopdata$YWT)
nrow(p1andbasepopdata) #8960

#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p1andbasepopdata <- subset(p1andbasepopdata, p1andbasepopdata[ ,10] - p1andbasepopdata[ ,9]
> 40)
nrow(p1andbasepopdata) #8960

simdataP1 <- subset(p1andbasepopdata, p1andbasepopdata[ ,1] > nam)
summary(simdataP1$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#139.3   222.9   247.8   248.4   273.1   359.4
summary(simdataP1$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#253.5   434.8   479.7   480.9   527.6   666.5

nrow(simdataP1) #8960
N_sires = length(c(sort(unique(simdataP1$sire))))
N_sires #400
basepop_p1_data <- rbind(basepopdata,simdataP1)
pedigreep1 <- data.frame (id = simdataP1$id, sire = simdataP1$sire, dam = simdataP1$dam)
pedigreep1andbase <- rbind(pedigree,pedigreep1)
```

```
##############################################################################
##############################################################################
## Creating P2 (Year 2)
#1 2 3 4 5 6 7 8 9 10 11 12 13 14
# id sire dam sex year herd CE BW ww pwg CEtbv BWtbv wwtbv pwgtbv
####Still no selection. No replacements.
#################################################################
#################################################################
## redefine (the total number of all animals) and (number of base population and P1) in previous
simulation
################################
#total number of all animals (Used unique IDs up to now)
nam2 = nam+nam1

################################
## selection of sires
## average RFI TBV for sires and dams
averages=by( basepopdata$RFItbv, basepopdata$sex, mean)
ave1males=averages[1]
ave1females=averages[2]
## standard deviation of RFI TBV for sires and dams
SDs=by( basepopdata$RFItbv, basepopdata$sex, sd)
sd1males=SDs[1]
sd1females=SDs[2]

## selecting all sires from basepopdata
selectedmales <- subset(basepopdata, sex == 1)

#########################
## selection of all dams
## Remove females from year 1 (older than 16 years)
females <- subset(selectedallfemales, year > 1)
selectedallfemales1 <- females[sample(1:nrow(females), 8100, replace = FALSE),]

nrow(selectedallfemales1)
nrow(selectedallfemales)
nrow(selectedallfemales1)/nrow(selectedallfemales) #0.90

#########################
#### 400 selected males and 8100 selected females
n_sel_males = nrow(selectedmales)
n_sel_females = nrow(selectedallfemales1)
n_progeny = n_sel_females
# number of females per sire = 20.25
n_females_per_sire = n_sel_females/n_sel_males
n_females_per_sire
```

```
# because number of dams per sire is 20.25, then sires will have different numbers of progeny
(some will have 20 and others will have 21 progeny)
rounded_n_females_per_sire = round(n_females_per_sire)
rounded_n_females_per_sire

if(rounded_n_females_per_sire < n_females_per_sire) {
  n1records_per_sire = (rounded_n_females_per_sire)
  n2records_per_sire = (n1records_per_sire)+1
} else {
  n1records_per_sire = (rounded_n_females_per_sire)-1
  n2records_per_sire = (rounded_n_females_per_sire)
}

#number of sires with 20 progeny (300 out of 400)
nsires_with_n1records = (n_sel_males)-((n_sel_females)-(n_sel_males*n1records_per_sire))
#number of sires with 21 progeny (100 out of 400)
nsires_with_n2records = (n_sel_males)-(nsires_with_n1records)
##pulling out the first 137 sires of sire list
own_record_sires_with_n1records = selectedmales[1:nsires_with_n1records, ]
## replicate each sire 25 times
selectedmales1=own_record_sires_with_n1records[rep(seq_len(nrow(own_record_sires_with_n
1records)), each=n1records_per_sire),]
##pulling out the remaining 263 sires of sire list
own_record_sires_with_n2records = selectedmales[((nsires_with_n1records)+1):n_sel_males, ]
## replicate each sire 24 times
selectedmales2=own_record_sires_with_n2records[rep(seq_len(nrow(own_record_sires_with_n
2records)), each=n2records_per_sire),]

## Create random list of sires with length of 8100 which is number of selected dams
allselectedmales=rbind(selectedmales1,selectedmales2)
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp2 = randomlymatedsires[ ,1]
damlistp2 = selectedallfemales1[ ,1]

######################
selectedparents=rbind(selectedmales, selectedallfemales1)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]
# pulling out selected parents and their pedigree
animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
#uniqueanimlist = uniqueanimlist[-1]
subdata <- basepopdata[basepopdata[ ,1] %in% uniqueanimlist, ]
subdata <- data.matrix(subdata)
length(subdata[,1]) #8500
length(uniqueanimlist) #8500
```

```
## creating P2 pedigree which include(selected parents and their pedigree + new P2 ID with their selected parents)
pedP2 =
rbind(subdata[,1:3],cbind(c((nam2+1):(nam2+n_progeny)),sirelistp2[1:n_progeny],damlistp2[1:n_progeny]))

###############
## For P2, repeating the same code used to create P1.

library(MASS)

set.seed(1234)
herdRFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdRFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(0.311,-0.009,0.219,0.579,-
0.009,0.014,0.548,1.231,0.219,0.548,552.753,465.634,0.579,1.231,465.634,1131.95),byrow=TR
UE,nrow=4)
# G matrix
G=matrix(data=c(0.18,0.0005,-1.372,-2.249,0.0005,0.01,1.014,2.217,-
1.372,1.014,230.77,327.931,-2.249,2.217,327.931,598.33),byrow=TRUE,nrow=4)
# Traits averages
mu=c(-0.011,1.473,258.913,479.990)

anwr=c((nam2+1):(nam2+n_progeny))
aid = c(pedP2[ ,1])
sid = c(pedP2[ ,2])
did <- c(pedP2[ ,3])

#number of herds
length(unique(selectedallfemales1[ ,6]))
iherd=c(selectedallfemales1[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
rd=diag(R)
```

```
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam3=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam3){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
damtbv= selectedallfemales1[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}

obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
```

```
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
pedigree2 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree2)
dataped <- pedigree2[pedigree2$id>nam,]
simdataP2 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2001, herd = iherd, RFI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
RFItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simdataP2)
nrow(simdataP2) #8100

## combine the data file of (base and P1) with P2 data file
p2_p1_basepopdata <- rbind(p1andbasepopdata,simdataP2)
## Data file for basepop and P1 (has both observations and TBV)
p2_p1_basepopdata <- subset(p2_p1_basepopdata, p2_p1_basepopdata[ ,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1_p2=mean(p2_p1_basepopdata$WWT)
WWTave_base_p1_p2
WWTsd_base_p1_p2=sd(p2_p1_basepopdata$WWT)
WWTsd_base_p1_p2
WWTave_base_p1_p2-(3*WWTsd_base_p1_p2)
WWTave_base_p1_p2+(3*WWTsd_base_p1_p2)
p2_p1_basepopdata <- subset(p2_p1_basepopdata, WWT > (WWTave_base_p1_p2-
(3*WWTsd_base_p1_p2)))
p2_p1_basepopdata <- subset(p2_p1_basepopdata, WWT <
(WWTave_base_p1_p2+(3*WWTsd_base_p1_p2)))
summary(p2_p1_basepopdata$WWT)
nrow(p2_p1_basepopdata) #17021

#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p2_p1_basepopdata <- subset(p2_p1_basepopdata, p2_p1_basepopdata[ ,10] -
p2_p1_basepopdata[ ,9] > 40)
nrow(p2_p1_basepopdata) #17021

simdataP2 <- subset(p2_p1_basepopdata, p2_p1_basepopdata[ ,1] > nam2)
summary(simdataP2$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#137.5   223.0   247.9   248.3   272.5   360.2
summary(simdataP2$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#286.5   434.2   481.4   480.4   526.4   680.6

nrow(simdataP2) #8069
```

```
N_sires = length(c(sort(unique(simdataP2$sire))))
N_sires #400
basepop_p2_p1_data <- rbind(basepop_p1_data,simdataP2)
pedigreep2 <- data.frame (id = simdataP2$id, sire = simdataP2$sire, dam = simdataP2$dam)
pedigreep2p1andbase<- rbind(pedigreep1andbase,pedigreep2)

##############################################################################
##############################################################################
## Creating P3 (Year 3)
#1 2 3 4 5 6 7 8 9 10 11 12 13 14
# id sire dam sex year herd CE BW ww pwg CEtbv BWtbv wwtbv pwgtbv
# Begin selection for RFI and use replacments from year 2000
#################################################################
#################################################################
## redefine (the total number of all animals) and (number of base population and P1) in previous
simulation
################################
#total number of all animals (Used unique IDs up to now): basepop, P1 and P2
nam4 = nam2+nam3

################################
## selecting top 10% sirs and 80% dams
## selection of sires
## sires selection changed to 10% otherwise there were not enough sires
averages=by( basepop_p1_data$RFItbv, basepop_p1_data$sex, mean)
ave1males=averages[1]
ave1females=averages[2]
SDs=by( basepop_p1_data$RFItbv, basepop_p1_data$sex, sd)
sd1males=SDs[1]
sd1females=SDs[2]

males3 <- subset(basepop_p1_data, sex == 1)
bulls3 <- males3[order(males3$RFItbv), ]
selectedmales3 <- bulls3[1:450, ]
summary(selectedmales3[ ,7], na.rm=TRUE)

nrow(males3) #4924
nrow(selectedmales3) #450
nrow(selectedmales3)/nrow(males3) #0.09

#########################
## selection of dams
## FIRST: Remove females from year 2 (older than 16 years)
## SECOND: Selection of cows and replace with 2000 heifers
## Assume 20% loss and 90% conception rate
```

```
females3 <- subset(basepop_p1_data, sex == 2 & year > 2 & year < 2000)
femalesp3 <- females3[sample(1:nrow(females3), 8000, replace=FALSE),]
hfs <- subset(basepop_p1_data, sex == 2 & year == 2000)
nrow(hfs) #4437

if(nrow(hfs) > 2000) {
  n = 2000
} else {
  n = nrow(hfs)
}

hfrs <- hfs[order(hfs$RFItbv), ]
selectedhfrs <- hfrs[1:n, ]
fem3 <- rbind(femalesp3, selectedhfrs)
selectedallfemales3 <- fem3[sample(1:nrow(fem3), 9000, replace = FALSE),]
summary(selectedallfemales3[ ,7], na.rm=TRUE)

nrow(femalesp3) #8000
nrow(selectedallfemales3) #9000
damlistp3 = c(selectedallfemales3[ ,1])

##########################
#### 450 selected males and 9000 selected females

n_progeny = nrow(selectedallfemales3)

allselectedmales=selectedmales3[rep(seq_len(nrow(selectedmales3)), each=20),]
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp3 = randomlymatedsires[ ,1]

####################
selectedparents=rbind(selectedmales3, selectedallfemales3)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]

animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
uniqueanimlist = uniqueanimlist[-1]
subdata <- basepop_p1_data[basepop_p1_data[ ,1] %in% uniqueanimlist, ]
subdata <- data.matrix(subdata)

length(subdata[,1]) #10445
length(uniqueanimlist) #10445
pedp3 = rbind(subdata[
,1:3],cbind(c((nam4+1):(nam4+n_progeny)),sirelistp3[1:n_progeny],damlistp3[1:n_progeny]))

##################
```

```
RFIave=mean(selectedparents[ ,7], na.rm=TRUE)
RFIave
ADGave=mean(selectedparents[ ,8], na.rm=TRUE)
ADGave
WWTave=mean(selectedparents[ ,9], na.rm=TRUE)
WWTave
YWTave=mean(selectedparents[ ,10], na.rm=TRUE)
YWTave

###############
###############
## P3 progeny, repeating the same code used to create P2.

library(MASS)

set.seed(1234)
herdRFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdRFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(0.311,-0.009,0.219,0.579,-
0.009,0.014,0.548,1.231,0.219,0.548,552.753,465.634,0.579,1.231,465.634,1131.95),byrow=TR
UE,nrow=4)
# G matrix
G=matrix(data=c(0.18,0.0005,-1.372,-2.249,0.0005,0.01,1.014,2.217,-
1.372,1.014,230.77,327.931,-2.249,2.217,327.931,598.33),byrow=TRUE,nrow=4)

####averages:
mu=c(RFIave,ADGave,WWTave,YWTave)
anwr=c((nam4+1):(nam4+n_progeny))
aid = c(pedp3[ ,1])
sid = c(pedp3[ ,2])
did <- c(pedp3[ ,3])

#number of herds
length(unique(selectedallfemales3[ ,6]))
iherd=c(selectedallfemales3[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1
```

```
# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam5=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam5){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
damtbv= selectedallfemales3[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}
```

```
obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
pedigree3 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree3)
dataped <- pedigree3[pedigree3$id>nam4,]
simdataP3 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2002, herd = iherd, RFI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
RFItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simdataP3)
nrow(simdataP3) #9000

## combine the data file of (base, P1 and P2) with P3 data file
p3_p2_p1_basepopdata <- rbind(p2_p1_basepopdata,simdataP3)
## Data file for basepop and P1 (has both observations and TBV)
p3_p2_p1_basepopdata <- subset(p3_p2_p1_basepopdata, p3_p2_p1_basepopdata[ ,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1_p2_p3=mean(p3_p2_p1_basepopdata$WWT)
WWTave_base_p1_p2_p3
WWTsd_base_p1_p2_p3=sd(p3_p2_p1_basepopdata$WWT)
WWTsd_base_p1_p2_p3
WWTave_base_p1_p2_p3-(3*WWTsd_base_p1_p2_p3)
WWTave_base_p1_p2_p3+(3*WWTsd_base_p1_p2_p3)
p3_p2_p1_basepopdata <- subset(p3_p2_p1_basepopdata, WWT > (WWTave_base_p1_p2_p3-
(3*WWTsd_base_p1_p2_p3)))
p3_p2_p1_basepopdata <- subset(p3_p2_p1_basepopdata, WWT <
(WWTave_base_p1_p2_p3+(3*WWTsd_base_p1_p2_p3)))
summary(p3_p2_p1_basepopdata$WWT)
nrow(p3_p2_p1_basepopdata) #25974

# keeping YWT observations within 3 SD
YWTave_base_p1_p2_p3=mean(p3_p2_p1_basepopdata$YWT)
YWTave_base_p1_p2_p3
YWTsd_base_p1_p2_p3=sd(p3_p2_p1_basepopdata$YWT)
YWTsd_base_p1_p2_p3
YWTave_base_p1_p2_p3-(3*YWTsd_base_p1_p2_p3)
YWTave_base_p1_p2_p3+(3*YWTsd_base_p1_p2_p3)
p3_p2_p1_basepopdata <- subset(p3_p2_p1_basepopdata, YWT > (YWTave_base_p1_p2_p3-
(3*YWTsd_base_p1_p2_p3)))
p3_p2_p1_basepopdata <- subset(p3_p2_p1_basepopdata, YWT <
(YWTave_base_p1_p2_p3+(3*YWTsd_base_p1_p2_p3)))
```

```
summary(p3_p2_p1_basepopdata$YWT)
nrow(p3_p2_p1_basepopdata) #25961

#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p3_p2_p1_basepopdata <- subset(p3_p2_p1_basepopdata, p3_p2_p1_basepopdata[ ,10] -
p3_p2_p1_basepopdata[ ,9] > 40)
nrow(p3_p2_p1_basepopdata) #25967

simdataP3 <- subset(p3_p2_p1_basepopdata, p3_p2_p1_basepopdata[ ,1] > nam4)
summary(simdataP3$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#134.1   214.3   238.9   239.8   264.4   358.0
summary(simdataP3$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
# 286.3   422.1   468.1   468.9   516.7   660.3

nrow(simdataP3) #8951
N_sires = length(c(sort(unique(simdataP3$sire))))
N_sires #450
basepop_p1_p2_p3_data <- rbind(basepop_p2_p1_data,simdataP3)
pedigreep3 <- data.frame (id = simdataP3$id, sire = simdataP3$sire, dam = simdataP3$dam)
pedigreep3p2p1andbase <- rbind(pedigreep2p1andbase,pedigreep3)

################################################################################
################################################################################
## Creating P4 (Year 4)
#1 2 3 4 5 6 7 8 9 10 11 12 13
# id sire dam sex herd CE BW ww pwg CEtbv BWtbv wwtbv pwgtbv
#####################################################################
#####################################################################
## redefine (the total number of all animals) and (number of base population, P1 and P2) in
previous simulation
################################
#total number of all animals (Used unique IDs up to now): basepop, P1 and P2
nam6 = nam4+nam5

################################
## selecting top 5% sires (2000 & 2001) and top heifers for replacements
## selection of sires
averages=by( basepop_p2_p1_data$RFItbv, basepop_p2_p1_data$sex, mean)
ave1males=averages[1]
ave1females=averages[2]
SDs=by( basepop_p2_p1_data$RFItbv, basepop_p2_p1_data$sex, sd)
sd1males=SDs[1]
sd1females=SDs[2]
```

```
males4 <- subset(basepop_p2_p1_data, sex == 1 & year >= 2000)
bulls4 <- males4[order(males4$RFItbv), ]
selectedmales4 <- bulls4[1:450, ]
summary(selectedmales4[ ,7], na.rm=TRUE)

nrow(males4) #8593
nrow(selectedmales4) #450
nrow(selectedmales4)/nrow(males4) #0.05

##########################
## selection of dams
## FIRST: Remove females from year 3 (older than 16 years)
## SECOND: Selection of cows and replace with 2000 heifers
## Assume 20% loss
## Assume 90% Conception Rate

females4 <- subset(selectedallfemales3, year > 3)
femalesp4 <- females4[sample(1:nrow(females4), 8000, replace=FALSE),]
heifers <- subset(basepop_p2_p1_data, sex == 2 & year == 2001)
nrow(heifers) #3991

if(nrow(hfs) > 2000) {
  n = 2000
} else {
  n = nrow(hfs)
}

hfrs <- heifers[order(heifers$RFItbv), ]
selectedhfrs <- hfrs[1:n, ]
fem4 <- rbind(femalesp4, selectedhfrs)
selectedallfemales4 <- fem4[sample(1:nrow(fem4), 9000, replace = FALSE),]
summary(selectedallfemales4[ ,7], na.rm=TRUE)

nrow(femalesp4) #8000
nrow(selectedallfemales4) #9000
damlistp4 = c(selectedallfemales4[ ,1])

##########################
#### 450 selected males and 9000 selected females
n_progeny = nrow(selectedallfemales4)

allselectedmales=selectedmales4[rep(seq_len(nrow(selectedmales4)), each=20),]
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp4 = randomlymatedsires[ ,1]
```

```
####################
selectedparents=rbind(selectedmales4, selectedallfemales4)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]

animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
uniqueanimlist = uniqueanimlist[-1]
subdata <- basepop_p2_p1_data[basepop_p2_p1_data[ ,1] %in% uniqueanimlist, ]
subdata <- data.matrix(subdata)

length(subdata[,1]) #11231
length(uniqueanimlist) #11231
pedp4 = rbind(subdata[
,1:3],cbind(c((nam6+1):(nam6+n_progeny)),sirelistp4[1:n_progeny],damlistp4[1:n_progeny]))

###################
RFIave=mean(selectedparents[ ,7], na.rm=TRUE)
RFIave
ADGave=mean(selectedparents[ ,8], na.rm=TRUE)
ADGave
WWTave=mean(selectedparents[ ,9], na.rm=TRUE)
WWTave
YWTave=mean(selectedparents[ ,10], na.rm=TRUE)
YWTave

###############
###############
## P4 progeny, repeating the same code used to create P3.

library(MASS)

set.seed(1234)
herdRFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdRFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(0.311,-0.009,0.219,0.579,-
0.009,0.014,0.548,1.231,0.219,0.548,552.753,465.634,0.579,1.231,465.634,1131.95),byrow=TR
UE,nrow=4)
```

```
# G matrix
G=matrix(data=c(0.18,0.0005,-1.372,-2.249,0.0005,0.01,1.014,2.217,-
1.372,1.014,230.77,327.931,-2.249,2.217,327.931,598.33),byrow=TRUE,nrow=4)
# Traits averages
mu=c(RFIave,ADGave,WWTave,YWTave)

anwr=c((nam6+1):(nam6+n_progeny))
aid = c(pedp4[ ,1])
sid = c(pedp4[ ,2])
did <- c(pedp4[ ,3])

#number of herds
length(unique(selectedallfemales4[ ,6]))
iherd=c(selectedallfemales4[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam7=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam7){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
```

```
damtbv= selectedallfemales4[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}

obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
pedigree4 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree4)
dataped <- pedigree4[pedigree4$id>nam6,]
simdataP4 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2003, herd = iherd, RFI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
RFItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simdataP4)
nrow(simdataP4) #9000

## combine the data file of (base, P1, P2 and P3) with P4 data file
p4_p3_p2_p1_basepopdata <- rbind(p3_p2_p1_basepopdata,simdataP4)
## Data file for basepop, P1, P2, P3 and P4 (has both observations and TBV)
p4_p3_p2_p1_basepopdata <- subset(p4_p3_p2_p1_basepopdata, p4_p3_p2_p1_basepopdata[
,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1_p2_p3_p4=mean(p4_p3_p2_p1_basepopdata$WWT)
WWTave_base_p1_p2_p3_p4
WWTsd_base_p1_p2_p3_p4=sd(p4_p3_p2_p1_basepopdata$WWT)
WWTsd_base_p1_p2_p3_p4
WWTave_base_p1_p2_p3_p4-(3*WWTsd_base_p1_p2_p3_p4)
WWTave_base_p1_p2_p3_p4+(3*WWTsd_base_p1_p2_p3_p4)
```

```
p4_p3_p2_p1_basepopdata <- subset(p4_p3_p2_p1_basepopdata, WWT >
(WWTave_base_p1_p2_p3_p4-(3*WWTsd_base_p1_p2_p3_p4)))
p4_p3_p2_p1_basepopdata <- subset(p4_p3_p2_p1_basepopdata, WWT <
(WWTave_base_p1_p2_p3_p4+(3*WWTsd_base_p1_p2_p3_p4)))
summary(p4_p3_p2_p1_basepopdata$WWT)
nrow(p4_p3_p2_p1_basepopdata) #34911

# keeping YWT observations within 3 SD
YWTave_base_p1_p2_p3_p4=mean(p4_p3_p2_p1_basepopdata$YWT)
YWTave_base_p1_p2_p3_p4
YWTsd_base_p1_p2_p3_p4=sd(p4_p3_p2_p1_basepopdata$YWT)
YWTsd_base_p1_p2_p3_p4
YWTave_base_p1_p2_p3_p4-(3*YWTsd_base_p1_p2_p3_p4)
YWTave_base_p1_p2_p3_p4+(3*YWTsd_base_p1_p2_p3_p4)
p4_p3_p2_p1_basepopdata <- subset(p4_p3_p2_p1_basepopdata, YWT >
(YWTave_base_p1_p2_p3_p4-(3*YWTsd_base_p1_p2_p3_p4)))
p4_p3_p2_p1_basepopdata <- subset(p4_p3_p2_p1_basepopdata, YWT <
(YWTave_base_p1_p2_p3_p4+(3*YWTsd_base_p1_p2_p3_p4)))
summary(p4_p3_p2_p1_basepopdata$YWT)
nrow(p4_p3_p2_p1_basepopdata) #34897

#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p4_p3_p2_p1_basepopdata <- subset(p4_p3_p2_p1_basepopdata, p4_p3_p2_p1_basepopdata[
,10] - p4_p3_p2_p1_basepopdata[ ,9] > 40)
nrow(p4_p3_p2_p1_basepopdata) #34897

simdataP4 <- subset(p4_p3_p2_p1_basepopdata, p4_p3_p2_p1_basepopdata[ ,1] > nam6)
summary(simdataP4$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#129.8   209.6   235.0   235.6   260.3   356.1
summary(simdataP4$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#280.2   408.0   454.8   455.1   501.7   661.5

nrow(simdataP4) #8951
N_sires = length(c(sort(unique(simdataP4$sire))))
N_sires #450
basepop_p1_p2_p3_p4_data <- rbind(basepop_p1_p2_p3_data,simdataP4)
pedigreep4 <- data.frame (id = simdataP4$id, sire = simdataP4$sire, dam = simdataP4$dam)
pedigreep4p3p2p1andbase <- rbind(pedigreep3p2p1andbase,pedigreep4)

################################################################################
################################################################################
## Creating P5 (Year 5)
#1 2 3 4 5 6 7 8 9 10 11 12 13
```

```
# id sire dam sex herd CE BW ww pwg CEtbv BWtbv wwtbv pwgtbv
#####################################################################
#####################################################################
## redefine (the total number of all animals) and (number of base population, P1, P2 and P3) in
previous simulation
################################
#total number of all animals (Used unique IDs up to now): basepop, P1 and P2
nam8 = nam6+nam7

################################
## selecting top 5% sires and 80% dams
## selection of sires
averages=by( basepop_p1_p2_p3_data$RFItbv, basepop_p1_p2_p3_data$sex, mean)
ave1males=averages[1]
ave1females=averages[2]
SDs=by( basepop_p1_p2_p3_data$RFItbv, basepop_p1_p2_p3_data$sex, sd)
sd1males=SDs[1]
sd1females=SDs[2]

males5 <- subset(basepop_p1_p2_p3_data, sex == 1 & year >= 2001)
bulls5 <- males5[order(males5$RFItbv), ]
selectedmales5 <- bulls5[1:450, ]
summary(selectedmales5[ ,7], na.rm=TRUE)

nrow(males5) #8593
nrow(selectedmales5) #450
nrow(selectedmales5)/nrow(males5) #0.05

##########################
## selection of dams
## FIRST: Remove females from year 4 (older than 16 years)
## SECOND: Selection of cows and replace with 2000 heifers
## Assume 20% loss

females5 <- subset(selectedallfemales4, year > 4)
femalesp5 <- females5[sample(1:nrow(females5), 8000, replace=FALSE),]
heifers <- subset(basepop_p1_p2_p3_data, sex == 2 & year == 2002)
nrow(heifers) #4427

if(nrow(heifers) > 2000) {
  n = 2000
} else {
  n = nrow(heifers)
}

hfrs <- heifers[order(heifers$RFItbv), ]
```

```
selectedhfrs <- hfrs[1:n, ]
fem5 <- rbind(femalesp5, selectedhfrs)
selectedallfemales5 <- fem5[sample(1:nrow(fem5), 9000, replace = FALSE),]
summary(selectedallfemales5[ ,7], na.rm=TRUE)

nrow(femalesp5) #8000
nrow(selectedallfemales5) #9000
damlistp5 = c(selectedallfemales5[ ,1])

##########################
#### 450 selected males and 9000 selected females
n_progeny = nrow(selectedallfemales5)

allselectedmales=selectedmales5[rep(seq_len(nrow(selectedmales5)), each=20),]
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp5 = randomlymatedsires[ ,1]

####################
selectedparents=rbind(selectedmales5, selectedallfemales5)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]

animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
uniqueanimlist = uniqueanimlist[-1]
subdata <- basepop_p1_p2_p3_data[basepop_p1_p2_p3_data[ ,1] %in% uniqueanimlist, ]
subdata <- data.matrix(subdata)

length(subdata[,1]) #12222
length(uniqueanimlist) #12222
pedp5 = rbind(subdata[
,1:3],cbind(c((nam8+1):(nam8+n_progeny)),sirelistp5[1:n_progeny],damlistp5[1:n_progeny]))

##################
RFIave=mean(selectedparents[ ,7], na.rm=TRUE)
RFIave
ADGave=mean(selectedparents[ ,8], na.rm=TRUE)
ADGave
WWTave=mean(selectedparents[ ,9], na.rm=TRUE)
WWTave
YWTave=mean(selectedparents[ ,10], na.rm=TRUE)
YWTave

###############
###############
## P5 progeny, repeating the same code used to create P4.
```

```
library(MASS)

set.seed(1234)
herdRFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdRFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(0.311,-0.009,0.219,0.579,-
0.009,0.014,0.548,1.231,0.219,0.548,552.753,465.634,0.579,1.231,465.634,1131.95),byrow=TR
UE,nrow=4)
# G matrix
G=matrix(data=c(0.18,0.0005,-1.372,-2.249,0.0005,0.01,1.014,2.217,-
1.372,1.014,230.77,327.931,-2.249,2.217,327.931,598.33),byrow=TRUE,nrow=4)
# Traits averages
mu=c(RFIave,ADGave,WWTave,YWTave)

anwr=c((nam8+1):(nam8+n_progeny))
aid = c(pedp5[ ,1])
sid = c(pedp5[ ,2])
did <- c(pedp5[ ,3])

#number of herds
length(unique(selectedallfemales5[ ,6]))
iherd=c(selectedallfemales5[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
```

```
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam9=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam9){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
damtbv= selectedallfemales5[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}

obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
pedigree5 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree5)
dataped <- pedigree5[pedigree5$id>nam8,]
simdataP5 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2004, herd = iherd, RFI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
RFItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
```

```
attach(simdataP5)
nrow(simdataP5) #9000

## combine the data file of (base, P1, P2, P3 and P4) with P5 data file
p5_p4_p3_p2_p1_basepopdata <- rbind(p4_p3_p2_p1_basepopdata,simdataP5)
## Data file for basepop and P's (has both observations and TBV)
p5_p4_p3_p2_p1_basepopdata <- subset(p5_p4_p3_p2_p1_basepopdata,
p5_p4_p3_p2_p1_basepopdata[ ,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1_p2_p3_p4_p5=mean(p5_p4_p3_p2_p1_basepopdata$WWT)
WWTave_base_p1_p2_p3_p4_p5
WWTsd_base_p1_p2_p3_p4_p5=sd(p5_p4_p3_p2_p1_basepopdata$WWT)
WWTsd_base_p1_p2_p3_p4_p5
WWTave_base_p1_p2_p3_p4_p5-(3*WWTsd_base_p1_p2_p3_p4_p5)
WWTave_base_p1_p2_p3_p4_p5+(3*WWTsd_base_p1_p2_p3_p4_p5)
p5_p4_p3_p2_p1_basepopdata <- subset(p5_p4_p3_p2_p1_basepopdata, WWT >
(WWTave_base_p1_p2_p3_p4_p5-(3*WWTsd_base_p1_p2_p3_p4_p5)))
p5_p4_p3_p2_p1_basepopdata <- subset(p5_p4_p3_p2_p1_basepopdata, WWT <
(WWTave_base_p1_p2_p3_p4_p5+(3*WWTsd_base_p1_p2_p3_p4_p5)))
summary(p5_p4_p3_p2_p1_basepopdata$WWT)
nrow(p5_p4_p3_p2_p1_basepopdata) #43847

# keeping YWT observations within 3 SD
YWTave_base_p1_p2_p3_p4_p5=mean(p5_p4_p3_p2_p1_basepopdata$YWT)
YWTave_base_p1_p2_p3_p4_p5
YWTsd_base_p1_p2_p3_p4_p5=sd(p5_p4_p3_p2_p1_basepopdata$YWT)
YWTsd_base_p1_p2_p3_p4_p5
YWTave_base_p1_p2_p3_p4_p5-(3*YWTsd_base_p1_p2_p3_p4_p5)
YWTave_base_p1_p2_p3_p4_p5+(3*YWTsd_base_p1_p2_p3_p4_p5)
p5_p4_p3_p2_p1_basepopdata <- subset(p5_p4_p3_p2_p1_basepopdata, YWT >
(YWTave_base_p1_p2_p3_p4_p5-(3*YWTsd_base_p1_p2_p3_p4_p5)))
p5_p4_p3_p2_p1_basepopdata <- subset(p5_p4_p3_p2_p1_basepopdata, YWT <
(YWTave_base_p1_p2_p3_p4_p5+(3*YWTsd_base_p1_p2_p3_p4_p5)))
summary(p5_p4_p3_p2_p1_basepopdata$YWT)
nrow(p5_p4_p3_p2_p1_basepopdata) #43836

#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p5_p4_p3_p2_p1_basepopdata <- subset(p5_p4_p3_p2_p1_basepopdata,
p5_p4_p3_p2_p1_basepopdata[ ,10] - p5_p4_p3_p2_p1_basepopdata[ ,9] > 40)
nrow(p5_p4_p3_p2_p1_basepopdata) #42627

simdataP5 <- subset(p5_p4_p3_p2_p1_basepopdata, p5_p4_p3_p2_p1_basepopdata[ ,1] >
nam8)
summary(simdataP5$WWT)
```

```
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#129.4   210.1   235.2   236.0   261.2   353.9
summary(simdataP5$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#274.8   409.8   455.5   456.1   502.3   649.9

nrow(simdataP5) #8955
N_sires = length(c(sort(unique(simdataP5$sire))))
N_sires #450
basepop_p1_p2_p3_p4_p5_data <- rbind(basepop_p1_p2_p3_p4_data,simdataP5)
pedigreep5 <- data.frame (id = simdataP5$id, sire = simdataP5$sire, dam = simdataP5$dam)
pedigreep5p4p3p2p1andbase <- rbind(pedigreep4p3p2p1andbase,pedigreep5)

##############################################################################
##############################################################################
## Creating P6 (Year 6)
#1 2 3 4 5 6 7 8 9 10 11 12 13
# id sire dam sex herd CE BW ww pwg CEtbv BWtbv wwtbv pwgtbv
#####################################################################
#####################################################################
## redefine (the total number of all animals) and (number of base population, P1, P2, P3 and P4)
in previous simulation
################################
#total number of all animals (Used unique IDs up to now): basepop, P1, P2, P3, P4, P5 and P5
nam10 = nam8+nam9

################################
## selecting top 5% sirs and 80% dams
## selection of sires

males6 <- subset(basepop_p1_p2_p3_p4_data, sex == 1 & year >= 2002)
bulls6 <- males6[order(males6$RFItbv), ]
selectedmales6 <- bulls6[1:450, ]
summary(selectedmales6[ ,7], na.rm=TRUE)

nrow(males6) #9080
nrow(selectedmales6) #450
nrow(selectedmales6)/nrow(males6) #0.05

#########################
## selection of dams
## FIRST: Remove females from year 5 (older than 16 years)
## SECOND: Selection of cows and replace with 2000 heifers
## Assume 20% loss

females6 <- subset(selectedallfemales5, sex == 2 & year > 5)
```

```
femalesp6 <- females6[sample(1:nrow(females6), 8000, replace=FALSE),]
heifers <- subset(basepop_p1_p2_p3_p4_data, sex == 2 & year == 2003)
nrow(heifers)

if(nrow(heifers) > 2000) {
  n = 2000
} else {
  n = nrow(heifers)
}

hfrs <- heifers[order(heifers$RFItbv), ]
selectedhfrs <- hfrs[1:n, ]
nrow(selectedhfrs)
fem6 <- rbind(femalesp6, selectedhfrs)
selectedallfemales6 <- fem6[sample(1:nrow(fem6), 9000, replace = FALSE),]
summary(selectedallfemales6[ ,7], na.rm=TRUE)

nrow(femalesp6) #8000
nrow(selectedallfemales6) #9000
damlistp6 = c(selectedallfemales6[ ,1])

##########################
#### 450 selected males and 9000 selected females
n_progeny = nrow(selectedallfemales3)

allselectedmales=selectedmales6[rep(seq_len(nrow(selectedmales6)), each=20),]
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp6 = randomlymatedsires[ ,1]

####################
selectedparents=rbind(selectedmales6, selectedallfemales6)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]

animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
uniqueanimlist = uniqueanimlist[-1]
subdata <- basepop_p1_p2_p3_p4_data[basepop_p1_p2_p3_p4_data[ ,1] %in% uniqueanimlist,
]
subdata <- data.matrix(subdata)

length(subdata[,1]) #12970
length(uniqueanimlist) #12970
pedp6 = rbind(subdata[
,1:3],cbind(c((nam10+1):(nam10+n_progeny)),sirelistp6[1:n_progeny],damlistp6[1:n_progeny]))

##################
```

```
RFIave=mean(selectedparents[ ,7], na.rm=TRUE)
RFIave
ADGave=mean(selectedparents[ ,8], na.rm=TRUE)
ADGave
WWTave=mean(selectedparents[ ,9], na.rm=TRUE)
WWTave
YWTave=mean(selectedparents[ ,10], na.rm=TRUE)
YWTave

###############
###############
## P6 progeny, repeating the same code used to create P5.

library(MASS)

set.seed(1234)
herdRFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdRFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(0.311,-0.009,0.219,0.579,-
0.009,0.014,0.548,1.231,0.219,0.548,552.753,465.634,0.579,1.231,465.634,1131.95),byrow=TR
UE,nrow=4)
# G matrix
G=matrix(data=c(0.18,0.0005,-1.372,-2.249,0.0005,0.01,1.014,2.217,-
1.372,1.014,230.77,327.931,-2.249,2.217,327.931,598.33),byrow=TRUE,nrow=4)
# Traits averages
mu=c(RFIave,ADGave,WWTave,YWTave)

anwr=c((nam10+1):(nam10+n_progeny))
aid = c(pedp6[ ,1])
sid = c(pedp6[ ,2])
did <- c(pedp6[ ,3])

#number of herds
length(unique(selectedallfemales6[ ,6]))
iherd=c(selectedallfemales6[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1
```

```
# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam11=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam11){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
damtbv= selectedallfemales6[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}
```

```
obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
pedigree6 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree6)
dataped <- pedigree6[pedigree6$id>nam10,]
simdataP6 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2005, herd = iherd, RFI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
RFItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simdataP6)
nrow(simdataP6) #9000

## combine the data file of (base, P1, P2, P3, P4 and P5) with P6 data file
p6_p5_p4_p3_p2_p1_basepopdata <- rbind(p5_p4_p3_p2_p1_basepopdata,simdataP6)
## Data file for basepop and P's (has both observations and TBV)
p6_p5_p4_p3_p2_p1_basepopdata <- subset(p6_p5_p4_p3_p2_p1_basepopdata,
p6_p5_p4_p3_p2_p1_basepopdata[ ,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1_p2_p3_p4_p5_p6=mean(p6_p5_p4_p3_p2_p1_basepopdata$WWT)
WWTave_base_p1_p2_p3_p4_p5_p6
WWTsd_base_p1_p2_p3_p4_p5_p6=sd(p6_p5_p4_p3_p2_p1_basepopdata$WWT)
WWTsd_base_p1_p2_p3_p4_p5_p6
WWTave_base_p1_p2_p3_p4_p5_p6-(3*WWTsd_base_p1_p2_p3_p4_p5_p6)
WWTave_base_p1_p2_p3_p4_p5_p6+(3*WWTsd_base_p1_p2_p3_p4_p5_p6)
p6_p5_p4_p3_p2_p1_basepopdata <- subset(p6_p5_p4_p3_p2_p1_basepopdata, WWT >
(WWTave_base_p1_p2_p3_p4_p5_p6-(3*WWTsd_base_p1_p2_p3_p4_p5_p6)))
p6_p5_p4_p3_p2_p1_basepopdata <- subset(p6_p5_p4_p3_p2_p1_basepopdata, WWT <
(WWTave_base_p1_p2_p3_p4_p5_p6+(3*WWTsd_base_p1_p2_p3_p4_p5_p6)))
summary(p6_p5_p4_p3_p2_p1_basepopdata$WWT)
nrow(p6_p5_p4_p3_p2_p1_basepopdata) #52789

# keeping YWT observations within 3 SD
YWTave_base_p1_p2_p3_p4_p5_p6=mean(p6_p5_p4_p3_p2_p1_basepopdata$YWT)
YWTave_base_p1_p2_p3_p4_p5_p6
YWTsd_base_p1_p2_p3_p4_p5_p6=sd(p6_p5_p4_p3_p2_p1_basepopdata$YWT)
YWTsd_base_p1_p2_p3_p4_p5_p6
YWTave_base_p1_p2_p3_p4_p5_p6-(3*YWTsd_base_p1_p2_p3_p4_p5_p6)
YWTave_base_p1_p2_p3_p4_p5_p6+(3*YWTsd_base_p1_p2_p3_p4_p5_p6)
p6_p5_p4_p3_p2_p1_basepopdata <- subset(p6_p5_p4_p3_p2_p1_basepopdata, YWT >
(YWTave_base_p1_p2_p3_p4_p5_p6-(3*YWTsd_base_p1_p2_p3_p4_p5_p6)))
```

```
p6_p5_p4_p3_p2_p1_basepopdata <- subset(p6_p5_p4_p3_p2_p1_basepopdata, YWT <
(YWTave_base_p1_p2_p3_p4_p5_p6+(3*YWTsd_base_p1_p2_p3_p4_p5_p6)))
summary(p6_p5_p4_p3_p2_p1_basepopdata$YWT)
nrow(p6_p5_p4_p3_p2_p1_basepopdata) #52779

#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p6_p5_p4_p3_p2_p1_basepopdata <- subset(p6_p5_p4_p3_p2_p1_basepopdata,
p6_p5_p4_p3_p2_p1_basepopdata[ ,10] - p6_p5_p4_p3_p2_p1_basepopdata[ ,9] > 40)
nrow(p6_p5_p4_p3_p2_p1_basepopdata) #57483

simdataP6 <- subset(p6_p5_p4_p3_p2_p1_basepopdata, p6_p5_p4_p3_p2_p1_basepopdata[ ,1]
> nam10)
summary(simdataP6$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#126.8   208.1   233.8   234.5   259.9   353.4
summary(simdataP6$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#270.8   403.3   450.1   450.3   497.4   651.9

nrow(simdataP6) #8953
N_sires = length(c(sort(unique(simdataP6$sire))))
N_sires #450
basepop_p1_p2_p3_p4_p5_p6_data <- rbind(basepop_p1_p2_p3_p4_p5_data,simdataP6)
pedigreep6 <- data.frame (id = simdataP6$id, sire = simdataP6$sire, dam = simdataP6$dam)
pedigreep6p5p4p3p2p1andbase <- rbind(pedigreep5p4p3p2p1andbase,pedigreep6)

################################################################################
################################################################################
## Creating P7 (Year 7)
#1 2 3 4 5 6 7 8 9 10 11 12 13
# id sire dam sex herd CE BW ww pwg CEtbv BWtbv wwtbv pwgtbv
################################################################
################################################################
## redefine (the total number of all animals) and (number of base population, P1, P2, P3, P4 and
P5) in previous simulation
################################
#total number of all animals (Used unique IDs up to now): basepop, P1, P2, P3, P4, P5, P6 and
P7
nam12 = nam10+nam11

################################
## selecting top 5% sirs and 80% dams
## selection of sires
```

```
averages=by( basepop_p1_p2_p3_p4_p5_data$RFItbv, basepop_p1_p2_p3_p4_p5_data$sex,
mean)
ave1males=averages[1]
ave1females=averages[2]
SDs=by( basepop_p1_p2_p3_p4_p5_data$RFItbv, basepop_p1_p2_p3_p4_p5_data$sex, sd)
sd1males=SDs[1]
sd1females=SDs[2]

males7 <- subset(basepop_p1_p2_p3_p4_p5_data, sex == 1 & year >= 2003)
bulls7 <- males7[order(males7$RFItbv), ]
selectedmales7 <- bulls7[1:450, ]
summary(selectedmales7[ ,7], na.rm=TRUE)

nrow(males7) #9067
nrow(selectedmales7) #450
nrow(selectedmales7)/nrow(males7) #0.05

##########################
## selection of dams
## FIRST: Remove females from year 6 (older than 16 years)
## SECOND: Selection of cows and replace with 2000 heifers
## Assume 20% loss

females7 <- subset(selectedallfemales6, sex == 2 & year > 6)
femalesp7 <- females7[sample(1:nrow(females7), 8000, replace=FALSE),]
heifers <- subset(basepop_p1_p2_p3_p4_p5_data, sex == 2 & year == 2004)
nrow(heifers)

if(nrow(heifers) > 2000) {
  n = 2000
} else {
  n = nrow(heifers)
}

hfrs <- heifers[order(heifers$RFItbv), ]
selectedhfrs <- hfrs[1:n, ]
nrow(selectedhfrs)
fem7 <- rbind(femalesp7, selectedhfrs)
selectedallfemales7 <- fem7[sample(1:nrow(fem7), 9000, replace = FALSE),]
summary(selectedallfemales7[ ,7], na.rm=TRUE)

nrow(femalesp7) #8000
nrow(selectedallfemales7) #9000
damlistp7 = c(selectedallfemales7[ ,1])

##########################
```

```
#### 450 selected males and 9000 selected females
n_progeny = nrow(selectedallfemales7)

allselectedmales=selectedmales7[rep(seq_len(nrow(selectedmales7)), each=20),]
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp7= randomlymatedsires[ ,1]

######################
selectedparents=rbind(selectedmales7, selectedallfemales7)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]

animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
uniqueanimlist = uniqueanimlist[-1]
subdata <- basepop_p1_p2_p3_p4_p5_data[basepop_p1_p2_p3_p4_p5_data[ ,1] %in%
uniqueanimlist, ]
subdata <- data.matrix(subdata)

length(subdata[,1]) #13850
length(uniqueanimlist) #13850
pedp7 = rbind(subdata[
,1:3],cbind(c((nam12+1):(nam12+n_progeny)),sirelistp7[1:n_progeny],damlistp7[1:n_progeny]))

###################
RFIave=mean(selectedparents[ ,7], na.rm=TRUE)
RFIave
ADGave=mean(selectedparents[ ,8], na.rm=TRUE)
ADGave
WWTave=mean(selectedparents[ ,9], na.rm=TRUE)
WWTave
YWTave=mean(selectedparents[ ,10], na.rm=TRUE)
YWTave

###############
###############
## P7 progeny, repeating the same code used to create P6.

library(MASS)

set.seed(1234)
herdRFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
```

```
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdRFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(0.311,-0.009,0.219,0.579,-
0.009,0.014,0.548,1.231,0.219,0.548,552.753,465.634,0.579,1.231,465.634,1131.95),byrow=TR
UE,nrow=4)
# G matrix
G=matrix(data=c(0.18,0.0005,-1.372,-2.249,0.0005,0.01,1.014,2.217,-
1.372,1.014,230.77,327.931,-2.249,2.217,327.931,598.33),byrow=TRUE,nrow=4)
# Traits averages
mu=c(RFIave,ADGave,WWTave,YWTave)

anwr=c((nam12+1):(nam12+n_progeny))
aid = c(pedp7[ ,1])
sid = c(pedp7[ ,2])
did <- c(pedp7[ ,3])

#number of herds
length(unique(selectedallfemales7[ ,6]))
iherd=c(selectedallfemales7[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
```

```
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam13=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam13){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
damtbv= selectedallfemales7[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}

obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
pedigree7 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree7)
dataped <- pedigree7[pedigree7$id>nam12,]
simdataP7 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2006, herd = iherd, RFI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
RFItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simdataP7)
nrow(simdataP7) #9000

## combine the data file of (base, P1, P2, P3, P4, P5 and P6) with P7 data file
p7_p6_p5_p4_p3_p2_p1_basepopdata <- rbind(p6_p5_p4_p3_p2_p1_basepopdata,simdataP7)
## Data file for basepop, P1, P2, P3, P4, P5, P6 and P7 (has both observations and TBV)
p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p7_p6_p5_p4_p3_p2_p1_basepopdata,
p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,1] > nbase)
```

```
# keeping WWT observations within 3 SD
WWTave_base_p1_p2_p3_p4_p5_p6_p7=mean(p7_p6_p5_p4_p3_p2_p1_basepopdata$WWT)
WWTave_base_p1_p2_p3_p4_p5_p6_p7
WWTsd_base_p1_p2_p3_p4_p5_p6_p7=sd(p7_p6_p5_p4_p3_p2_p1_basepopdata$WWT)
WWTsd_base_p1_p2_p3_p4_p5_p6_p7
WWTave_base_p1_p2_p3_p4_p5_p6_p7-(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7)
WWTave_base_p1_p2_p3_p4_p5_p6_p7+(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7)
p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p7_p6_p5_p4_p3_p2_p1_basepopdata, WWT >
(WWTave_base_p1_p2_p3_p4_p5_p6_p7-(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7)))
p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p7_p6_p5_p4_p3_p2_p1_basepopdata, WWT <
(WWTave_base_p1_p2_p3_p4_p5_p6_p7+(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7)))
summary(p7_p6_p5_p4_p3_p2_p1_basepopdata$WWT)
nrow(p7_p6_p5_p4_p3_p2_p1_basepopdata) #61743

# keeping YWT observations within 3 SD
YWTave_base_p1_p2_p3_p4_p5_p6_p7=mean(p7_p6_p5_p4_p3_p2_p1_basepopdata$YWT)
YWTave_base_p1_p2_p3_p4
YWTsd_base_p1_p2_p3_p4_p5_p6_p7=sd(p7_p6_p5_p4_p3_p2_p1_basepopdata$YWT)
YWTsd_base_p1_p2_p3_p4
YWTave_base_p1_p2_p3_p4-(3*YWTsd_base_p1_p2_p3_p4)
YWTave_base_p1_p2_p3_p4+(3*YWTsd_base_p1_p2_p3_p4)
p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p7_p6_p5_p4_p3_p2_p1_basepopdata, YWT >
(YWTave_base_p1_p2_p3_p4-(3*YWTsd_base_p1_p2_p3_p4)))
p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p7_p6_p5_p4_p3_p2_p1_basepopdata, YWT <
(YWTave_base_p1_p2_p3_p4+(3*YWTsd_base_p1_p2_p3_p4)))
summary(p7_p6_p5_p4_p3_p2_p1_basepopdata$YWT)
nrow(p7_p6_p5_p4_p3_p2_p1_basepopdata) #61726

#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p7_p6_p5_p4_p3_p2_p1_basepopdata,
p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,10] - p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,9] > 40)
nrow(p7_p6_p5_p4_p3_p2_p1_basepopdata) #61723

simdataP7 <- subset(p7_p6_p5_p4_p3_p2_p1_basepopdata,
p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,1] > nam12)
summary(simdataP7$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#126.5   208.0   233.8   234.5   259.3   352.8
summary(simdataP7$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#278.5   401.4   449.6   449.5   496.8   648.1

nrow(simdataP7) #8954
N_sires = length(c(sort(unique(simdataP7$sire))))
N_sires #450
```

```
basepop_p1_p2_p3_p4_p5_p6_p7_data <- rbind(basepop_p1_p2_p3_p4_p5_p6_data,simdataP7)
pedigreep7 <- data.frame (id = simdataP7$id, sire = simdataP7$sire, dam = simdataP7$dam)
pedigreep7p6p5p4p3p2p1andbase <- rbind(pedigreep6p5p4p3p2p1andbase,pedigreep7)

##############################################################################
##############################################################################
## Creating P8 (Year 8)
#1 2 3 4 5 6 7 8 9 10 11 12 13
# id sire dam sex herd CE BW ww pwg CEtbv BWtbv wwtbv pwgtbv
#################################################################
#################################################################
## redefine (the total number of all animals) and (number of base population, P1, P2, P3, P4, P5
and P6) in previous simulation
################################
#total number of all animals (Used unique IDs up to now): basepop, P1, P2, P3, P4, P5, P6, P7
and P8
nam14 = nam12+nam13

################################
## selecting top 5% sires
## selection of sires

males8 <- subset(basepop_p1_p2_p3_p4_p5_p6_data, sex == 1 & year >= 2004)
bulls8 <- males8[order(males8$RFItbv), ]
selectedmales8 <- bulls8[1:450, ]
summary(selectedmales8[ ,7], na.rm=TRUE)

nrow(males8) #9070
nrow(selectedmales8) #450
nrow(selectedmales8)/nrow(males8) #0.05

#########################
## selection of dams
## FIRST: Remove females from year 7 (older than 16 years)
## SECOND: Selection of cows and replace with 2000 heifers
## Assume 20% loss

females8 <- subset(selectedallfemales7, sex == 2 & year > 7)
femalesp8 <- females8[sample(1:nrow(females8), 8000, replace=FALSE),]
heifers <- subset(basepop_p1_p2_p3_p4_p5_p6_data, sex == 2 & year == 2005)
nrow(heifers)

if(nrow(heifers) > 2000) {
  n = 2000
} else {
  n = nrow(heifers)
```

```
}

hfrs <- heifers[order(heifers$RFItbv), ]
selectedhfrs <- hfrs[1:n, ]
nrow(selectedhfrs)
fem8 <- rbind(femalesp8, selectedhfrs)
selectedallfemales8 <- fem8[sample(1:nrow(fem8), 9000, replace = FALSE),]
summary(selectedallfemales8[ ,7], na.rm=TRUE)

nrow(femalesp8) #8000
nrow(selectedallfemales8) #9000
damlistp8 = c(selectedallfemales8[ ,1])

##########################
#### 450 selected males and 9000 selected females
n_progeny = nrow(selectedallfemales8)

allselectedmales=selectedmales8[rep(seq_len(nrow(selectedmales8)), each=20),]
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp8 = randomlymatedsires[ ,1]

#####################
selectedparents=rbind(selectedmales8, selectedallfemales8)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]

animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
uniqueanimlist = uniqueanimlist[-1]
subdata <- basepop_p1_p2_p3_p4_p5_p6_data[basepop_p1_p2_p3_p4_p5_p6_data[ ,1] %in%
uniqueanimlist, ]
subdata <- data.matrix(subdata)

length(subdata[,1]) #14525
length(uniqueanimlist) #14525
pedp8 = rbind(subdata[
,1:3],cbind(c((nam14+1):(nam14+n_progeny)),sirelistp8[1:n_progeny],damlistp8[1:n_progeny]))

##################
RFIave=mean(selectedparents[ ,7], na.rm=TRUE)
RFIave
ADGave=mean(selectedparents[ ,8], na.rm=TRUE)
ADGave
WWTave=mean(selectedparents[ ,9], na.rm=TRUE)
WWTave
YWTave=mean(selectedparents[ ,10], na.rm=TRUE)
YWTave
```

```
###############
###############
## P8 progeny, repeating the same code used to create P7.

library(MASS)

set.seed(1234)
herdRFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdRFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(0.311,-0.009,0.219,0.579,-
0.009,0.014,0.548,1.231,0.219,0.548,552.753,465.634,0.579,1.231,465.634,1131.95),byrow=TR
UE,nrow=4)
# G matrix
G=matrix(data=c(0.18,0.0005,-1.372,-2.249,0.0005,0.01,1.014,2.217,-
1.372,1.014,230.77,327.931,-2.249,2.217,327.931,598.33),byrow=TRUE,nrow=4)
# Traits averages
mu=c(RFIave,ADGave,WWTave,YWTave)

anwr=c((nam14+1):(nam14+n_progeny))
aid = c(pedp8[ ,1])
sid = c(pedp8[ ,2])
did <- c(pedp8[ ,3])

#number of herds
length(unique(selectedallfemales8[ ,6]))
iherd=c(selectedallfemales8[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
```

```
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam15=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam15){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
damtbv= selectedallfemales8[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}

obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
pedigree8 <- data.frame (id = aid, sire = sid, dam = did)
```

```
attach(pedigree8)
dataped <- pedigree8[pedigree8$id>nam14,]
simdataP8 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2007, herd = iherd, RFI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
RFItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simdataP8)
nrow(simdataP8) #9000

## combine the data file of (base, P1, P2, P3, P4, P5, P6 and P7) with P8 data file
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
rbind(p7_p6_p5_p4_p3_p2_p1_basepopdata,simdataP8)
## Data file for basepop and P's (has both observations and TBV)
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8=mean(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$
WWT)
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8
WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8=sd(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$WW
T)
WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8-(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8)
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8+(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8)
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
WWT > (WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8-
(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8)))
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
WWT <
(WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8+(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8))
)
summary(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$WWT)
nrow(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #70676

# keeping YWT observations within 3 SD
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8=mean(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$
YWT)
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8
YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8=sd(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$YWT
)
YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8-(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8)
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8+(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8)
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
YWT > (YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8-
(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8)))
```

```
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
YWT <
(YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8+(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8)))
summary(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$YWT)
nrow(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #66130

#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,10] - p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,9]
> 40)
nrow(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #70664

simdataP8 <- subset(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,1] > nam14)
summary(simdataP8$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#125.5   208.1   232.8   233.7   258.7   351.8
summary(simdataP8$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#266.4   400.1   446.7   447.4   495.0   641.0

nrow(simdataP8) #8951
N_sires = length(c(sort(unique(simdataP8$sire))))
N_sires #450
basepop_p1_p2_p3_p4_p5_p6_p7_p8_data <-
rbind(basepop_p1_p2_p3_p4_p5_p6_p7_data,simdataP8)
pedigreep8 <- data.frame (id = simdataP8$id, sire = simdataP8$sire, dam = simdataP8$dam)
pedigreep8p7p6p5p4p3p2p1andbase <- rbind(pedigreep7p6p5p4p3p2p1andbase,pedigreep8)

################################################################################
################################################################################
## Creating P9 (Year 9)
#1 2 3 4 5 6 7 8 9 10 11 12 13
# id sire dam sex herd CE BW ww pwg CEtbv BWtbv wwtbv pwgtbv
################################################################
################################################################
## redefine (the total number of all animals) and (number of base population, P1, P2, P3, P4, P5,
P6, P7 and P8) in previous simulation
################################
#total number of all animals (Used unique IDs up to now): basepop, P1, P2, P3, P4, P5, P6, P7,
P8 and P9
nam16 = nam14+nam15

################################
## selecting top 5% sires
```

```
## selection of sires

males9 <- subset(basepop_p1_p2_p3_p4_p5_p6_p7_data, sex == 1 & year >= 2005)
bulls9 <- males9[order(males9$RFItbv), ]
selectedmales9 <- bulls9[1:450, ]
summary(selectedmales9[ ,7], na.rm=TRUE)

nrow(males9) #9065
nrow(selectedmales9) #450
nrow(selectedmales9)/nrow(males9) #0.05

##########################
## selection of dams
## FIRST: Remove females from year 8 (older than 16 years)
## SECOND: Selection of cows and replace with 2000 heifers
## Assume 20% loss

females9 <- subset(selectedallfemales8, sex == 2 & year > 8)
femalesp9 <- females9[sample(1:nrow(females9), 8000, replace=FALSE),]
heifers <- subset(basepop_p1_p2_p3_p4_p5_p6_p7_data, sex == 2 & year == 2006)
nrow(heifers)

if(nrow(heifers) > 2000) {
  n = 2000
} else {
  n = nrow(heifers)
}

hfrs <- heifers[order(heifers$RFItbv), ]
selectedhfrs <- hfrs[1:n, ]
nrow(selectedhfrs)
fem9 <- rbind(femalesp8, selectedhfrs)
selectedallfemales9 <- fem8[sample(1:nrow(fem9), 9000, replace = FALSE),]
summary(selectedallfemales9[ ,7], na.rm=TRUE)

nrow(femalesp9) #8000
nrow(selectedallfemales9) #10000
damlistp9 = c(selectedallfemales9[ ,1])

##########################
#### 450 selected males and 9000 selected females
n_progeny = nrow(selectedallfemales9)

allselectedmales=selectedmales9[rep(seq_len(nrow(selectedmales9)), each=20),]
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp9 = randomlymatedsires[ ,1]
```

```
####################
selectedparents=rbind(selectedmales9, selectedallfemales9)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]

animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
uniqueanimlist = uniqueanimlist[-1]
subdata <- basepop_p1_p2_p3_p4_p5_p6_p7_data[basepop_p1_p2_p3_p4_p5_p6_p7_data[ ,1]
%in% uniqueanimlist, ]
subdata <- data.matrix(subdata)

length(subdata[,1]) #14636
length(uniqueanimlist) #14636
pedp9 = rbind(subdata[
,1:3],cbind(c((nam16+1):(nam16+n_progeny)),sirelistp9[1:n_progeny],damlistp9[1:n_progeny]))

###################
RFIave=mean(selectedparents[ ,7], na.rm=TRUE)
RFIave
ADGave=mean(selectedparents[ ,8], na.rm=TRUE)
ADGave
WWTave=mean(selectedparents[ ,9], na.rm=TRUE)
WWTave
YWTave=mean(selectedparents[ ,10], na.rm=TRUE)
YWTave

###############
###############
## P9 progeny, repeating the same code used to create P8.

library(MASS)

set.seed(1234)
herdRFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdRFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
```

```
R=matrix(data=c(0.311,-0.009,0.219,0.579,-
0.009,0.014,0.548,1.231,0.219,0.548,552.753,465.634,0.579,1.231,465.634,1131.95),byrow=TR
UE,nrow=4)
# G matrix
G=matrix(data=c(0.18,0.0005,-1.372,-2.249,0.0005,0.01,1.014,2.217,-
1.372,1.014,230.77,327.931,-2.249,2.217,327.931,598.33),byrow=TRUE,nrow=4)

####averages:
mu=c(RFIave,ADGave,WWTave,YWTave)
anwr=c((nam16+1):(nam16+n_progeny))
aid = c(pedp9[ ,1])
sid = c(pedp9[ ,2])
did <- c(pedp9[ ,3])

#number of herds
length(unique(selectedallfemales9[ ,6]))
iherd=c(selectedallfemales9[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam17=n_progeny
mendelian = jd(nam1,4)*0
```

```
for(i in 1:nam1){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
damtbv= selectedallfemales9[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}

obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
pedigree9 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree9)
dataped <- pedigree9[pedigree9$id>nam16,]
simdataP9 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2008, herd = iherd, RFI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
RFItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simdataP9)
nrow(simdataP9) #9000

## combine the data file of (base, P1, P2, P3, P4, P5, P6, P7 and P8) with P9 data file
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
rbind(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,simdataP9)
## Data file for basepop and P's (has both observations and TBV)
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9=mean(p9_p8_p7_p6_p5_p4_p3_p2_p1_basep
opdata$WWT)
```

```
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9
WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9=sd(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdat
a$WWT)
WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9-
(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9)
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9+(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p
8_p9)
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata, WWT >
(WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9-
(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9)))
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata, WWT <
(WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9+(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_
p8_p9)))
summary(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$WWT)
nrow(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #79617

# keeping YWT observations within 3 SD
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9=mean(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepo
pdata$YWT)
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9
YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9=sd(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdat
a$YWT)
YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9-
(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9)
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9+(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8
_p9)
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata, YWT >
(YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9-
(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9)))
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata, YWT <
(YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9+(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p
8_p9)))
summary(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$YWT)
nrow(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #79605

#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
```

```
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,10] -
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,9] > 40)
nrow(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #79604

simdataP9 <- subset(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,1] > nam16)
summary(simdataP9$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#125.5   209.1   233.8   234.4   258.9   350.7
summary(simdataP9$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#268.1   401.7   447.8   449.2   497.0   639.2

nrow(simdataP9) #8950
N_sires = length(c(sort(unique(simdataP9$sire))))
N_sires # 450
basepop_p1_p2_p3_p4_p5_p6_p7_p8_p9_data <-
rbind(basepop_p1_p2_p3_p4_p5_p6_p7_p8_data,simdataP9)
pedigreep9 <- data.frame (id = simdataP9$id, sire = simdataP9$sire, dam = simdataP9$dam)
pedigreep9p8p7p6p5p4p3p2p1andbase <-
rbind(pedigreep8p7p6p5p4p3p2p1andbase,pedigreep9)

##############################################################################
##############################################################################
## Creating P10 (Year 10)
#1 2 3 4 5 6 7 8 9 10 11 12 13
# id sire dam sex herd CE BW ww pwg CEtbv BWtbv wwtbv pwgtbv
#################################################################
#################################################################
## redefine (the total number of all animals) and (number of base population, P1, P2, P3, P4, P5,
P6, P7 and P8) in previous simulation
################################
#total number of all animals (Used unique IDs up to now): basepop, P1, P2, P3, P4, P5, P6, P7,
P8 and P9
nam18 = nam16+nam17

################################
## selecting top 5% sires
## selection of sires
averages=by( basepop_p1_p2_p3_p4_p5_p6_p7_p8_data$RFItbv,
basepop_p1_p2_p3_p4_p5_p6_p7_p8_data$sex, mean)
ave1males=averages[1]
ave1females=averages[2]
SDs=by( basepop_p1_p2_p3_p4_p5_p6_p7_p8_data$RFItbv,
basepop_p1_p2_p3_p4_p5_p6_p7_p8_data$sex, sd)
sd1males=SDs[1]
```

```
sd1females=SDs[2]

males10 <- subset(basepop_p1_p2_p3_p4_p5_p6_p7_p8_data, sex == 1 & year >= 2006)
bulls10 <- males10[order(males10$RFItbv), ]
selectedmales10 <- bulls10[1:450, ]
summary(selectedmales10[ ,7], na.rm=TRUE)

nrow(males10) #9059
nrow(selectedmales10) #450
nrow(selectedmales10)/nrow(males10) #0.05

##########################
## selection of dams
## FIRST: Remove females from year 9 (older than 16 years)
## SECOND: Selection of cows and replace with 2000 heifers
## Assume 20% loss

females10 <- subset(selectedallfemales9, sex == 2 & year > 9)
femalesp10 <- females10[sample(1:nrow(females10), 8000, replace=FALSE),]
heifers <- subset(basepop_p1_p2_p3_p4_p5_p6_p7_p8_data, sex == 2 & year == 2007)
nrow(heifers)

if(nrow(heifers) > 2000) {
  n = 2000
} else {
  n = nrow(heifers)
}

hfrs <- heifers[order(heifers$RFItbv), ]
selectedhfrs <- hfrs[1:n, ]
fem10 <- rbind(femalesp10, heifers)
selectedallfemales10 <- fem10[sample(1:nrow(fem10), 9000, replace = FALSE),]

nrow(femalesp10) #8000
nrow(selectedallfemales10) #10000
damlistp10 = c(selectedallfemales10[ ,1])

##########################
#### 450 selected males and 9000 selected females
n_progeny = nrow(selectedallfemales10)

allselectedmales=selectedmales10[rep(seq_len(nrow(selectedmales10)), each=20),]
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp10 = randomlymatedsires[ ,1]

#####################
```

```
selectedparents=rbind(selectedmales10, selectedallfemales10)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]

animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
uniqueanimlist = uniqueanimlist[-1]
subdata <-
basepop_p1_p2_p3_p4_p5_p6_p7_p8_data[basepop_p1_p2_p3_p4_p5_p6_p7_p8_data[ ,1]
%in% uniqueanimlist, ]
subdata <- data.matrix(subdata)

length(subdata[,1]) #15097
length(uniqueanimlist) #15097
pedp10 = rbind(subdata[
,1:3],cbind(c((nam18+1):(nam18+n_progeny)),sirelistp10[1:n_progeny],damlistp10[1:n_progeny
]))

###################
RFIave=mean(selectedparents[ ,7], na.rm=TRUE)
RFIave
ADGave=mean(selectedparents[ ,8], na.rm=TRUE)
ADGave
WWTave=mean(selectedparents[ ,9], na.rm=TRUE)
WWTave
YWTave=mean(selectedparents[ ,10], na.rm=TRUE)
YWTave

###############
###############
## P10 progeny, repeating the same code used to create P9.

set.seed(1234)
herdRFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdRFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(0.311,-0.009,0.219,0.579,-
0.009,0.014,0.548,1.231,0.219,0.548,552.753,465.634,0.579,1.231,465.634,1131.95),byrow=TR
UE,nrow=4)
```

```
# G matrix
G=matrix(data=c(0.18,0.0005,-1.372,-2.249,0.0005,0.01,1.014,2.217,-
1.372,1.014,230.77,327.931,-2.249,2.217,327.931,598.33),byrow=TRUE,nrow=4)

####averages:
mu=c(RFIave,ADGave,WWTave,YWTave)
anwr=c((nam18+1):(nam18+n_progeny))
aid = c(pedp10[ ,1])
sid = c(pedp10[ ,2])
did <- c(pedp10[ ,3])

#number of herds
length(unique(selectedallfemales10[ ,6]))
iherd=c(selectedallfemales10[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam19=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam1){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
```

```
damtbv= selectedallfemales10[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}

obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
pedigree10 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree10)
dataped <- pedigree10[pedigree10$id>nam18,]
simdataP10 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2009, herd = iherd, RFI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
RFItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simdataP10)
nrow(simdataP10) #9000

## combine the data file of (base, P1, P2, P3, P4, P5, P6, P7, P8 and P9) with P10 data file
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
rbind(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,simdataP10)
## Data file for basepop and P's (has both observations and TBV)
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10=mean(p10_p9_p8_p7_p6_p5_p4_p3_p2_
p1_basepopdata$WWT)
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10
WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10=sd(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_b
asepopdata$WWT)
```

```
WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10-
(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10)
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10+(3*WWTsd_base_p1_p2_p3_p4_p5_p6_
p7_p8_p9_p10)
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata, WWT >
(WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10-
(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10)))
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata, WWT <
(WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10+(3*WWTsd_base_p1_p2_p3_p4_p5_p6
_p7_p8_p9_p10)))
summary(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$WWT)
nrow(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #88567

# keeping YWT observations within 3 SD
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10=mean(p10_p9_p8_p7_p6_p5_p4_p3_p2_
p1_basepopdata$YWT)
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10
YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10=sd(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_b
asepopdata$YWT)
YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10-
(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10)
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10+(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p
7_p8_p9_p10)
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata, YWT >
(YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10-
(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10)))
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata, YWT <
(YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10+(3*YWTsd_base_p1_p2_p3_p4_p5_p6_
p7_p8_p9_p10)))
summary(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$YWT)
nrow(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #88558

#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,9] -
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,8] > 40)
nrow(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #88558
```

```
simdataP10 <- subset(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,1] > nam18)
summary(simdataP10$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#124.2   207.2   232.6   233.4   258.8   351.3
summary(simdataP10$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#264.1   396.6   443.7   443.9   493.0   639.9

nrow(simdataP10) #8962
N_sires = length(c(sort(unique(simdataP10$sire))))
N_sires #450
basepop_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10_data <-
rbind(basepop_p1_p2_p3_p4_p5_p6_p7_p8_p9_data,simdataP10)
pedigreep10 <- data.frame (id = simdataP10$id, sire = simdataP10$sire, dam =
simdataP10$dam)
pedigreep10p9p8p7p6p5p4p3p2p1andbase <-
rbind(pedigreep9p8p7p6p5p4p3p2p1andbase,pedigreep10)

##############################################################################
##############################################################################
## see how many records per sire
try <- subset(basepop_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10_data,
basepop_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10_data[ ,1] > nbase)
nrow(try)
ones = c(rep(1,(nrow(try))))
try = data.matrix(try)
try = cbind(try,ones)
ham=sort(by( try[ ,14], try[ ,2], length))
head(ham)

##############################################
## Final data files
## pedigree file
## ****data file for animals with records (will be used for ****ASREML3.0****)
data_anim_with_record <- subset(basepop_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10_data,
basepop_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10_data[ ,1] > nbase)
summary(data_anim_with_record$RFI)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#-25.310 -11.440  -7.510  -7.425  -3.410   9.020
summary(data_anim_with_record$YWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#253.5   410.4   457.4   457.9   505.7   680.6
summary(data_anim_with_record$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#124.2   211.7   237.0   237.8   262.9   360.2
```

```
summary(data_anim_with_record$ADG)
# Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#-0.120   0.920   1.130   1.143   1.350   2.730
```

APPENDIX D: R CODE USED TO CREATE ECONOMIC SELECTION INDEX PARAMETERS

```
#R code based on Dr. Larry Schaeffer's R code for multiple trait models && Hamad Saad's
disertation
#Creating a simulated data (Selection on decreasing feed intake)
## Traits were: Feed intake (FI), 200 d weight (WWT), 400 d weight (YWT), and average daily
gain (ADG)
# First creating the base population (No selection)
setwd("/Users/Randie/Documents/PhD Research/Simulation/SimRCode")

## To clear environment
#FISimData_v2
rm(list = ls())

##Run DistFem_USE.R first
source("DistFem_USE.R")

#Herd effects (52 herds) on FI, ADG, WWT and YWT
##Simulate observations using G and R matrices for FI.  NO SELECTION FOR FI!!!
library(MASS)

set.seed(1234)
herdFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(1.05,0.049,0.528,17.749,0.049,0.014,0.287,1.982,0.528,0.287,552.75,465.63,1
7.749,1.982,465.63,1131.95),byrow=TRUE,nrow=4)
# G matrix
G=matrix(data=c(0.694,0.050,9.945,18.845,0.050,0.005,0.709,1.550,9.945,0.709,230.77,327.934
,18.845,1.550,327.934,598.327),byrow=TRUE,nrow=4)
# Traits averages
mu=c(10.007,1.473,258.913,479.990) #weighted averages

#ID for animals (founders (10000 dams, 400 sires))
aid = c(1:10400)

##creating fields in the data file for sire (20 progeny each), dams (1 progeny each), herds (size of
200 each), sex (50% females and 50% males)
sid = c(numeric(10400),rep(1:400, by=1, each=20))
base.sire <- data.frame(id = c(10001:10400), year = 00, sex = 1)
```

```
did <- c(numeric(10400),1:10000)
bi=c(rep(1,10400),rep(0.5,10000))
set.seed(1234)
isex=(rbinom(10000, 1, 0.5))+1
set.seed(1234)
iherd=c(sample(rep(1:52, by=1, each=200),10400,replace=F))

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

nam = 10400

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}

tbv = jd(nam,4)*0
for(i in 1:nam){
  x = LG %*% (sqrt(bi[i])*rnorm(4,0,1))
  if(sid[i]>0){
    ks=sid[i]
    kd=did[i]
    x = x + 0.5*(tbv[ks, ]+tbv[kd, ]) }
  tbv[i,]=x }
nrec=10400

#########################################################################
#########################################################################
```

```
##Creating base popuation with tbv but no observations
####Use base.females from DistFem_USE.R

base.pop <- rbind(base.females, base.sire)
simulateddata <- data.frame (id = base.pop$id, sire = c(rep(NA,10400)), dam =
c(rep(NA,10400)), sex = base.pop$sex, year = base.pop$year, herd = iherd, FI =
c(rep(NA,10400)), ADG = c(rep(NA,10400)), WWT = c(rep(NA,10400)), YWT =
c(rep(NA,10400)), FItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simulateddata)

animlist = c(simulateddata[ ,1])
detach(simulateddata)
uniqueanimlist = c(sort(unique(animlist)))
pedigree = data.frame(id = uniqueanimlist, sire = c(rep(NA,10400)), dam = c(rep(NA,10400)))
length(uniqueanimlist) #10400
length(pedigree[,1]) #10400

#Header: id sire dam sex year herd FI ADG WWT YWT FItbv ADGtbv WWTtbv YWTtbv
basepopdata <- simulateddata
attach(basepopdata)

#####################################################################
## creating P1 (progeny 1)
## Assuming 100% conception
## No selection of FI.  Must use whole population for first 2 generations
#####################################################################
## redefine the total number of all animals and number of base population in previous simulation
################################
#total number of all animals (base pop + P1)
nam = 20400
#number of sires and dams(founders) (400+10000) in base population
nbase = 10400

################################
## No Selection for sires to create base pop for SI values.
## Select 5% of sires randomly
## selection of sires

## selecting all sires from basepopdata
selectedmales <- subset(basepopdata, sex == 1)

#########################
## selection of all dams
## For first selection of cows from P1 only selected 100% of females to be mated to sires
selectedfemales <- subset(basepopdata, sex == 2)
```

```
#########################
#### 400 selected males and 10000 selected females
## Create random list of sires with length of 10000 which is number of selected dams
allselectedmales=selectedmales[rep(seq_len(nrow(selectedmales)), each=25),]
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp1 = randomlymatedsires[ ,1]
damlistp1 = selectedfemales[ ,1]

####################
selectedparents=rbind(selectedmales, selectedfemales)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]
# pulling out selected parents and their pedigree
animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
subdata <- basepopdata[basepopdata[ ,1] %in% uniqueanimlist, ]
subdata <- data.matrix(subdata)
length(subdata[,1]) #10400
length(uniqueanimlist) #10400

## creating BP (base population for SI values) pedigree which include(selected parents and their
pedigree + new BP ID with their selected parents)
pedBP =
rbind(subdata[,1:3],cbind(c((nam+1):(nam+n_progeny)),sirelistp1[1:n_progeny],damlistp1[1:n_
progeny]))
##################

###############
## P1 observations
library(MASS)

set.seed(1234)
herdFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(1.05,0.049,0.528,17.749,0.049,0.014,0.287,1.982,0.528,0.287,552.75,465.63,1
7.749,1.982,465.63,1131.95),byrow=TRUE,nrow=4)
# G matrix
```

```
G=matrix(data=c(0.694,0.050,9.945,18.845,0.050,0.005,0.709,1.550,9.945,0.709,230.77,327.934
,18.845,1.550,327.934,598.327),byrow=TRUE,nrow=4)
# Traits averages
mu=c(10.007,1.473,258.913,479.990) #weighted averages

anwr=c((nam+1):(nam+n_progeny))
aid = c(pedBP[ ,1])
sid = c(pedBP[ ,2])
did <- c(pedBP[ ,3])

#number of herds
length(unique(selectedfemales[ ,6]))
iherd=c(selectedfemales[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam1=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam1){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
damtbv= selectedfemales[ ,11:14]
```

```
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}
obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
pedigreeBP <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigreeBP)
dataped <- pedigreeBP[pedigreeBP$id>nam,]
base.pop <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
herd = iherd, FI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4], FItbv = tbv[ ,1],
ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])

## Use base.pop to develop SI values.
## Use only males assuming all were finished in a feedlot
## 10 year averages were used to estimate cattle prices ($/kg), feed cost ($/kg) and average days
on feed
## Cattle price: $2.64/kg
## Feed Cost: $0.18/kg
## Day On Feed: 157 d

males <- subset(base.pop, sex == 1)
y <- males$YWT*2.64 - males$FI*157*0.18
x.FI <- males$FItbv
x.ADG <- males$ADGtbv
x.WWT <- males$WWTtbv
x.YWT <- males$YWTtbv

SI.reg <- lm(y ~ x.FI + x.ADG + x.WWT + x.YWT)
summary(SI.reg)
#Call:
```

```
#lm(formula = y ~ x.FI + x.ADG + x.WWT + x.YWT)
#
#Residuals:
#  Min       1Q   Median       3Q      Max
#-281.585  -52.141    1.002   52.992  298.335
#
#Coefficients:
#  Estimate Std. Error  t value Pr(>|t|)
#(Intercept) 1119.42324    1.10451 1013.503   <2e-16 ***
#  x.FI        -29.66932    2.90774  -10.204   <2e-16 ***
#  x.ADG        21.99373   37.28831    0.590    0.555
#  x.WWT         0.09652    0.16289    0.593    0.554
#  x.YWT         2.55104    0.20261   12.591   <2e-16 ***
#  ---
#  Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
#
#Residual standard error: 78.5 on 5052 degrees of freedom
#Multiple R-squared:  0.3065,        Adjusted R-squared:  0.306
#F-statistic: 558.3 on 4 and 5052 DF,  p-value: < 2.2e-16

###  SI = -29.66932FItbv + 21.99373ADGtbv + 0.09652WWTtbv + 2.55104YWTtbv


##Sensitivity Testing
install.packages("caret")
library(caret)

konfound(SI.reg, x.FI, to_return = "print", test_all = TRUE)
# A tibble: 4 x 8
#var_name      t    df action       inference          pct_bias_to_change_inference   itcv r_con
#<chr>     <dbl> <dbl> <chr>        <chr>                                      <dbl>  <dbl> <dbl>
#  1 x.FI     -10.2    5053 to_invalidate reject_null                          80.6  0.118 0.343
#2 x.ADG      0.590  5053 to_sustain    fail_to_reject_null                     69.9 -0.019 0.137
#3 x.WWT      0.595  5053 to_sustain    fail_to_reject_null                     69.6 -0.019 0.137
#4 x.YWT     12.6    5053 to_invalidate reject_null                          84.2  0.151 0.388
>
```

APPENDIX E: R CODE USED TO CREATE SIMULATE DATA FOR SELECTION USING AN ECONOMIC SELECTION INDEX

```
#R code based on Dr. Larry Schaeffer's R code for multiple trait models && Hamad Saad's
disertation
#Creating a simulated data (Selection on decreasing feed intake)
## Traits were: Feed intake (FI), 200 d weight (WWT), 400 d weight (YWT), and average daily
gain (ADG)
# First creating the base population (No selection)
setwd("/Users/Randie/Documents/PhD Research/Simulation")

## To clear environment
#FISimData_v2
rm(list = ls())

##Run DistFem_USE.R first
source("DistFem_USE.R")

library(MASS)

#Herd effects (50 herds) on FI, ADG, WWT and YWT
set.seed(1234)
herdFI=(rnorm(50,0))
set.seed(1234)
herdWW=(rnorm(50,0))*2
set.seed(1234)
herdYW=(rnorm(50,0))*10
set.seed(1234)
herdADG=(rnorm(50,0))*0.25  ##Crews et al 2006 (SD for ADG)
herd=matrix(data = c(herdFI,herdADG,herdWW,herdYW),byrow = TRUE, nrow = 4)

# sex effects (52) herds) on FI, ADG, WWT, and YWT
##Sex effects for wwt and ywwt from Van Vleck & Cundiff 1998
##ADG sex effects used Hamad's
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)
# Residual matrix
R=matrix(data=c(1.05,0.049,0.528,17.749,0.049,0.014,0.287,1.982,0.528,0.287,552.75,465.63,1
7.749,1.982,465.63,1131.95),byrow=TRUE,nrow=4)
# G matrix
G=matrix(data=c(0.694,0.050,9.945,18.845,0.050,0.005,0.709,1.550,9.945,0.709,230.77,327.934
,18.845,1.550,327.934,598.327),byrow=TRUE,nrow=4)
# Traits averages
mu=c(10.007,1.473,258.913,479.990) #weighted averages

#ID for animals (founders (10000 dams, 450 sires))
aid = c(1:10400)

##creating fields in the data file for sire (20 progeny each), dams (1 progeny each), herds (size of
200 each), sex (50% females and 50% males)
```

```
sid = c(numeric(10400),rep(1:400, by=1, each=20))
base.sire <- data.frame(id = c(10001:10400), year = 15, sex = 1)
did <- c(numeric(10400),1:10000)
bi=c(rep(1,10400),rep(0.5,10000))
set.seed(1234)
iherd=c(sample(rep(1:52, by=1, each=200),10400,replace=F))

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

nam = 10400
# Simulate true breeding values for all animals (founders and their F1 progeny)
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}

tbv = jd(nam,4)*0
for(i in 1:nam){
  x = LG %*% (sqrt(bi[i])*rnorm(4,0,1))
  if(sid[i]>0){
    ks=sid[i]
    kd=did[i]
    x = x + 0.5*(tbv[ks, ]+tbv[kd, ]) }
  tbv[i,]=x }
nrec=10400

#########################################################################
#########################################################################
##Creating base popuation with tbv but no observations
```

```
####Use base.females from DistFem_USE.R

base.pop <- rbind(base.females, base.sire)
simulateddata <- data.frame (id = base.pop$id, sire = c(rep(NA,10400)), dam =
c(rep(NA,10400)), sex = base.pop$sex, year = base.pop$year, herd = iherd, FI =
c(rep(NA,10400)), ADG = c(rep(NA,10400)), WWT = c(rep(NA,10400)), YWT =
c(rep(NA,10400)), FItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simulateddata)

animlist = c(simulateddata[ ,1])
detach(simulateddata)
uniqueanimlist = c(sort(unique(animlist)))
pedigree = data.frame(id = uniqueanimlist, sire = c(rep(NA,10400)), dam = c(rep(NA,10400)))
length(uniqueanimlist) #10400
length(pedigree[,1]) #10400

#Header: id sire dam sex year herd FI ADG WWT YWT FItbv ADGtbv WWTtbv YWTtbv
basepopdata <- simulateddata
attach(basepopdata)

#####################################################################
## creating P1 (progeny 1)
## Assuming 100% conception
## No selection.  Must use whole population for first 2 generations
#####################################################################
## redefine the total number of all animals and number of base population in previous simulation
################################
#total number of all animals (base pop + P1)
nam = 20400
#number of sires and dams(founders) (400+10000) in base population
nbase = 10400

################################
## selecting top 5% sirs and 80% dams (TBV for FI is the selection criteria)
## selection of sires
## average FI TBV for sires and dams
averages=by( basepopdata$FItbv, basepopdata$sex, mean)
ave1males=averages[1]
ave1females=averages[2]
## standard deviation of FI TBV for sires and dams
SDs=by( basepopdata$FItbv, basepopdata$sex, sd)
sd1males=SDs[1]
sd1females=SDs[2]

## selecting all sires from basepopdata
selectedmales <- subset(basepopdata, sex == 1)
```

```
#########################
## selection of all dams
## For first selection of cows from P1 only selected 100% of females to be mated to sires
females.base <- subset(basepopdata, sex == 2)
##90% Conception Rate
selectedallfemales <- females.base[sample(1:nrow(females.base), 9000, replace = FALSE),]

nrow(females.base)
nrow(selectedallfemales)
nrow(selectedallfemales)/nrow(females.base) #0.9

#########################
#### 400 selected males and 10000 selected females
n_sel_males = nrow(selectedmales)
n_sel_females = nrow(selectedallfemales)
n_progeny = n_sel_females
# number of females per sire = 22.5
n_females_per_sire = n_sel_females/n_sel_males
n_females_per_sire

# because number of dams per sire is 22.5, then sires will have different numbers of progeny
(some will have 22.5 and others will have 23 progeny)
rounded_n_females_per_sire = round(n_females_per_sire)
rounded_n_females_per_sire

if(rounded_n_females_per_sire < n_females_per_sire) {
  n1records_per_sire = (rounded_n_females_per_sire)
  n2records_per_sire = (n1records_per_sire)+1
} else {
  n1records_per_sire = (rounded_n_females_per_sire)-1
  n2records_per_sire = (rounded_n_females_per_sire)
}

#number of sires with 22 progeny (200 out of 400)
nsires_with_n1records = (n_sel_males)-((n_sel_females)-(n_sel_males*n1records_per_sire))
#number of sires with 23 progeny (200 out of 400)
nsires_with_n2records = (n_sel_males)-(nsires_with_n1records)
##pulling out the first 137 sires of sire list
own_record_sires_with_n1records = selectedmales[1:nsires_with_n1records, ]
## replicate each sire 22 times
selectedmales1=own_record_sires_with_n1records[rep(seq_len(nrow(own_record_sires_with_n
1records)), each=n1records_per_sire),]
##pulling out the remaining 200 sires of sire list
own_record_sires_with_n2records = selectedmales[((nsires_with_n1records)+1):n_sel_males, ]
## replicate each sire 24 times
```

```
selectedmales2=own_record_sires_with_n2records[rep(seq_len(nrow(own_record_sires_with_n
2records)), each=n2records_per_sire),]

## Create random list of sires with length of 10000 which is number of selected dams
allselectedmales=rbind(selectedmales1,selectedmales2)
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp1 = randomlymatedsires[ ,1]
damlistp1 = selectedallfemales[ ,1]

####################
selectedparents=rbind(selectedmales, selectedallfemales)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]
# pulling out selected parents and their pedigree
animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
subdata <- basepopdata[basepopdata[ ,1] %in% uniqueanimlist, ]
subdata <- data.matrix(subdata)
length(subdata[,1]) #9400
length(uniqueanimlist) #9400

## creating P1 pedigree which include(selected parents and their pedigree + new P1 ID with their
selected parents)
pedP1 =
rbind(subdata[,1:3],cbind(c((nam+1):(nam+n_progeny)),sirelistp1[1:n_progeny],damlistp1[1:n_
progeny]))
###################

##############
## P1 observations

library(MASS)

set.seed(1234)
herdFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(1.05,0.049,0.528,17.749,0.049,0.014,0.287,1.982,0.528,0.287,552.75,465.63,1
7.749,1.982,465.63,1131.95),byrow=TRUE,nrow=4)
```

```
# G matrix
G=matrix(data=c(0.694,0.050,9.945,18.845,0.050,0.005,0.709,1.550,9.945,0.709,230.77,327.934
,18.845,1.550,327.934,598.327),byrow=TRUE,nrow=4)
# Traits averages
mu=c(10.007,1.473,258.913,479.990) #weighted averages

anwr=c((nam+1):(nam+n_progeny))
aid = c(pedP1[ ,1])
sid = c(pedP1[ ,2])
did <- c(pedP1[ ,3])

#number of herds
length(unique(selectedallfemales[ ,6]))
iherd=c(selectedallfemales[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam1=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam1){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
```

```
damtbv= selectedallfemales[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}
obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
pedigree1 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree1)
dataped <- pedigree1[pedigree1$id>nam,]
simdataP1 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2000, herd = iherd, FI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
FItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simdataP1)
nrow(simdataP1) #9000

## combine the data file of (base) with P1 data file
p1andbasepopdata <- rbind(basepopdata,simdataP1)
## Data file for basepop and P1 (has both observations and TBV)
p1andbasepopdata <- subset(p1andbasepopdata, p1andbasepopdata[ ,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1=mean(p1andbasepopdata$WWT)
WWTave_base_p1 #231.39
WWTsd_base_p1=sd(p1andbasepopdata$WWT)
WWTsd_base_p1 #31.63
WWTave_base_p1-(3*WWTsd_base_p1) #136.49
WWTave_base_p1+(3*WWTsd_base_p1) #326.29
p1andbasepopdata <- subset(p1andbasepopdata, WWT > (WWTave_base_p1-
(3*WWTsd_base_p1)))
```

```
p1andbasepopdata <- subset(p1andbasepopdata, WWT <
(WWTave_base_p1+(3*WWTsd_base_p1)))
summary(p1andbasepopdata$WWT)
nrow(p1andbasepopdata) #8954

# keeping YWT observations within 3 SD
YWTave_base_p1=mean(p1andbasepopdata$YWT)
YWTave_base_p1 #370.82
YWTsd_base_p1=sd(p1andbasepopdata$YWT)
YWTsd_base_p1 #54.09
YWTave_base_p1-(3*YWTsd_base_p1) #208.55
YWTave_base_p1+(3*YWTsd_base_p1) #533.08
p1andbasepop <- subset(p1andbasepopdata, YWT > (YWTave_base_p1-(3*YWTsd_base_p1)))
p1andbasepopdata <- subset(p1andbasepopdata, YWT <
(YWTave_base_p1+(3*YWTsd_base_p1)))
summary(p1andbasepopdata$YWT)
nrow(p1andbasepopdata) #8954

#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p1andbasepopdata <- subset(p1andbasepopdata, p1andbasepopdata[ ,10] - p1andbasepopdata[ ,9]
> 40)
nrow(p1andbasepopdata) #8938

simdataP1 <- subset(p1andbasepopdata, p1andbasepopdata[ ,1] > nam)
summary(simdataP1$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#136.6   222.5   247.5   248.1   272.9   361.4
summary(simdataP1$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#280.1   433.2   480.8   480.5   527.0   660.9

nrow(simdataP1) #8938
N_sires = length(c(sort(unique(simdataP1$sire))))
N_sires #400
basepop_p1_data <- rbind(basepopdata,simdataP1)
pedigreep1 <- data.frame (id = simdataP1$id, sire = simdataP1$sire, dam = simdataP1$dam)
pedigreep1andbase <- rbind(pedigree,pedigreep1)

##############################################################################
##############################################################################
## Creating P2 (Year 2)
#1 2 3 4 5 6 7 8 9 10 11 12 13 14
# id sire dam sex year herd CE BW ww pwg CEtbv BWtbv wwtbv pwgtbv
####Still no selection. No replacements.
#################################################################
```

```
#####################################################################
## redefine (the total number of all animals) and (number of base population and P1) in previous simulation
#################################
#total number of all animals (Used unique IDs up to now)
nam2 = nam+nam1

#################################
## selection of sires
## average FI TBV for sires and dams
averages=by( basepopdata$FItbv, basepopdata$sex, mean)
ave1males=averages[1]
ave1females=averages[2]
## standard deviation of FI TBV for sires and dams
SDs=by( basepopdata$FItbv, basepopdata$sex, sd)
sd1males=SDs[1]
sd1females=SDs[2]

## selecting all sires from basepopdata
selectedmales <- subset(basepopdata, sex == 1)

#########################
## selection of all dams
## Remove females from year 1 (older than 16 years)
females <- subset(selectedallfemales, year > 1)
selectedallfemales1 <- females[sample(1:nrow(females), 8100, replace = FALSE),]

nrow(selectedallfemales1)
nrow(selectedallfemales)
nrow(selectedallfemales1)/nrow(selectedallfemales) #0.90

#########################
#### 400 selected males and 8100 selected females
n_sel_males = nrow(selectedmales)
n_sel_females = nrow(selectedallfemales1)
n_progeny = n_sel_females
# number of females per sire = 20.25
n_females_per_sire = n_sel_females/n_sel_males
n_females_per_sire
# because number of dams per sire is 20.25, then sires will have different numbers of progeny (some will have 20 and others will have 21 progeny)
rounded_n_females_per_sire = round(n_females_per_sire)
rounded_n_females_per_sire

if(rounded_n_females_per_sire < n_females_per_sire) {
  n1records_per_sire = (rounded_n_females_per_sire)
```

```
  n2records_per_sire = (n1records_per_sire)+1
} else {
  n1records_per_sire = (rounded_n_females_per_sire)-1
  n2records_per_sire = (rounded_n_females_per_sire)
}

#number of sires with 20 progeny (300 out of 400)
nsires_with_n1records = (n_sel_males)-((n_sel_females)-(n_sel_males*n1records_per_sire))
#number of sires with 21 progeny (100 out of 400)
nsires_with_n2records = (n_sel_males)-(nsires_with_n1records)
##pulling out the first 137 sires of sire list
own_record_sires_with_n1records = selectedmales[1:nsires_with_n1records, ]
## replicate each sire 25 times
selectedmales1=own_record_sires_with_n1records[rep(seq_len(nrow(own_record_sires_with_n
1records)), each=n1records_per_sire),]
##pulling out the remaining 263 sires of sire list
own_record_sires_with_n2records = selectedmales[((nsires_with_n1records)+1):n_sel_males, ]
## replicate each sire 24 times
selectedmales2=own_record_sires_with_n2records[rep(seq_len(nrow(own_record_sires_with_n
2records)), each=n2records_per_sire),]

## Create random list of sires with length of 8100 which is number of selected dams
allselectedmales=rbind(selectedmales1,selectedmales2)
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp2 = randomlymatedsires[ ,1]
damlistp2 = selectedallfemales1[ ,1]

####################
selectedparents=rbind(selectedmales, selectedallfemales1)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]
# pulling out selected parents and their pedigree
animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
#uniqueanimlist = uniqueanimlist[-1]
subdata <- basepopdata[basepopdata[ ,1] %in% uniqueanimlist, ]
subdata <- data.matrix(subdata)
length(subdata[,1]) #8500
length(uniqueanimlist) #8500

## creating P2 pedigree which include(selected parents and their pedigree + new P2 ID with their
selected parents)
pedP2 =
rbind(subdata[,1:3],cbind(c((nam2+1):(nam2+n_progeny)),sirelistp2[1:n_progeny],damlistp2[1:
n_progeny]))

###############
```

```
## For P2, repeating the same code used to create P1.

library(MASS)

set.seed(1234)
herdFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(1.05,0.049,0.528,17.749,0.049,0.014,0.287,1.982,0.528,0.287,552.75,465.63,1
7.749,1.982,465.63,1131.95),byrow=TRUE,nrow=4)
# G matrix
G=matrix(data=c(0.694,0.050,9.945,18.845,0.050,0.005,0.709,1.550,9.945,0.709,230.77,327.934
,18.845,1.550,327.934,598.327),byrow=TRUE,nrow=4)
# Traits averages
mu=c(10.007,1.473,258.913,479.990) #weighted averages

anwr=c((nam2+1):(nam2+n_progeny))
aid = c(pedP2[ ,1])
sid = c(pedP2[ ,2])
did <- c(pedP2[ ,3])

#number of herds
length(unique(selectedallfemales1[ ,6]))
iherd=c(selectedallfemales1[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
```

```
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam3=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam3){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
damtbv= selectedallfemales1[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}

obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
pedigree2 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree2)
dataped <- pedigree2[pedigree2$id>nam,]
```

```
simdataP2 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2001, herd = iherd, FI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
FItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simdataP2)
nrow(simdataP2) #8100

## combine the data file of (base and P1) with P2 data file
p2_p1_basepopdata <- rbind(p1andbasepopdata,simdataP2)
## Data file for basepop and P1 (has both observations and TBV)
p2_p1_basepopdata <- subset(p2_p1_basepopdata, p2_p1_basepopdata[ ,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1_p2=mean(p2_p1_basepopdata$WWT)
WWTave_base_p1_p2
WWTsd_base_p1_p2=sd(p2_p1_basepopdata$WWT)
WWTsd_base_p1_p2
WWTave_base_p1_p2-(3*WWTsd_base_p1_p2)
WWTave_base_p1_p2+(3*WWTsd_base_p1_p2)
p2_p1_basepopdata <- subset(p2_p1_basepopdata, WWT > (WWTave_base_p1_p2-
(3*WWTsd_base_p1_p2)))
p2_p1_basepopdata <- subset(p2_p1_basepopdata, WWT <
(WWTave_base_p1_p2+(3*WWTsd_base_p1_p2)))
summary(p2_p1_basepopdata$WWT)
nrow(p2_p1_basepopdata) #17022

#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p2_p1_basepopdata <- subset(p2_p1_basepopdata, p2_p1_basepopdata[ ,10] -
p2_p1_basepopdata[ ,9] > 40)
nrow(p2_p1_basepopdata) #17022

simdataP2 <- subset(p2_p1_basepopdata, p2_p1_basepopdata[ ,1] > nam2)
summary(simdataP2$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
# 138.8   223.2   247.2   248.2   272.5   357.9
summary(simdataP2$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#258.1   433.8   482.0   480.5   526.7   682.6

nrow(simdataP2) #8067
N_sires = length(c(sort(unique(simdataP2$sire))))
N_sires #400
basepop_p2_p1_data <- rbind(basepop_p1_data,simdataP2)
pedigreep2 <- data.frame (id = simdataP2$id, sire = simdataP2$sire, dam = simdataP2$dam)
pedigreep2p1andbase<- rbind(pedigreep1andbase,pedigreep2)
```

```
##############################################################################
##############################################################################
## Creating P3 (Year 3)
#1 2 3 4 5 6 7 8 9 10 11 12 13 14
# id sire dam sex year herd CE BW ww pwg CEtbv BWtbv wwtbv pwgtbv
# Begin selection for based on Selection Index and use replacments from year 2000
##################################################################
##################################################################
## redefine (the total number of all animals) and (number of base population and P1) in previous
simulation
################################
#total number of all animals (Used unique IDs up to now): basepop, P1 and P2
nam4 = nam2+nam3

################################
## selecting top 10% sires and 80% dams
## selection of sires
## sires selection changed to 10% otherwise there were not enough sires
males3 <- subset(basepop_p1_data, sex == 1)
si.males3 <- data.frame(males3$id, si = -29.66932*males3$FItbv + 21.99373*males3$ADGtbv
+ 0.09652*males3$WWTtbv + 2.55104*males3$YWTtbv)
males3.si <- cbind(males3, si.males3)
bulls3 <- males3.si[order(-males3.si$si), ]
selectedmales3 <- bulls3[1:450, ]
selectedmales3$males3.id <- NULL
mean(selectedmales3[ ,15], na.rm=TRUE)
selectedmales3$si <- NULL


nrow(males3) #4918
nrow(selectedmales3) #450
nrow(selectedmales3)/nrow(males3) #0.09

#########################
## selection of dams
## FIRST: Remove females from year 2 (older than 16 years)
## SECOND: Selection of cows and replace with 2000 heifers
## Assume 20% loss and 90% conception rate

females3 <- subset(basepop_p1_data, sex == 2 & year > 2 & year < 2000)
femalesp3 <- females3[sample(1:nrow(females3), 8000, replace=FALSE),]
hfr <- subset(basepop_p1_data, sex == 2 & year == 2000)
nrow(hfr) #4420

if(nrow(hfr) > 2000) {
  n = 2000
```

```
} else {
  n = nrow(hfr)
}

si.hfr <- data.frame(hfr$id, si = -29.66932*hfr$FItbv + 21.99373*hfr$ADGtbv +
0.09652*hfr$WWTtbv + 2.55104*hfr$YWTtbv)
hfs.si <- cbind(hfr, si.hfr)
hfrs <- hfs.si[order(-hfs.si$si), ]
selectedhfrs <- hfrs[1:n, ]
mean(selectedhfrs$si)
selectedhfrs$hfr.id <- NULL
selectedhfrs$si <- NULL
fem3 <- rbind(femalesp3, selectedhfrs)
selectedallfemales3 <- fem3[sample(1:nrow(fem3), 9000, replace = FALSE),]

nrow(femalesp3) #8000
nrow(selectedallfemales3) #9000
damlistp3 = c(selectedallfemales3[ ,1])

##########################
#### 450 selected males and 9000 selected females

n_progeny = nrow(selectedallfemales3)

allselectedmales=selectedmales3[rep(seq_len(nrow(selectedmales3)), each=20),]
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp3 = randomlymatedsires[ ,1]

#####################
selectedparents=rbind(selectedmales3, selectedallfemales3)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]

animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
uniqueanimlist = uniqueanimlist[-1]
subdata <- basepop_p1_data[basepop_p1_data[ ,1] %in% uniqueanimlist, ]
subdata <- data.matrix(subdata)

length(subdata[,1]) #10436
length(uniqueanimlist) #10436
pedp3 = rbind(subdata[
,1:3],cbind(c((nam4+1):(nam4+n_progeny)),sirelistp3[1:n_progeny],damlistp3[1:n_progeny]))

##################
FIave=mean(selectedparents[ ,7], na.rm=TRUE)
FIave
```

```
ADGave=mean(selectedparents[ ,8], na.rm=TRUE)
ADGave
WWTave=mean(selectedparents[ ,9], na.rm=TRUE)
WWTave
YWTave=mean(selectedparents[ ,10], na.rm=TRUE)
YWTave

###############
###############
## P3 progeny, repeating the same code used to create P2.

library(MASS)

set.seed(1234)
herdFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(1.05,0.049,0.528,17.749,0.049,0.014,0.287,1.982,0.528,0.287,552.75,465.63,1
7.749,1.982,465.63,1131.95),byrow=TRUE,nrow=4)
# G matrix
G=matrix(data=c(0.694,0.050,9.945,18.845,0.050,0.005,0.709,1.550,9.945,0.709,230.77,327.934
,18.845,1.550,327.934,598.327),byrow=TRUE,nrow=4)

####averages:
mu=c(FIave,ADGave,WWTave,YWTave)
anwr=c((nam4+1):(nam4+n_progeny))
aid = c(pedp3[ ,1])
sid = c(pedp3[ ,2])
did <- c(pedp3[ ,3])

#number of herds
length(unique(selectedallfemales3[ ,6]))
iherd=c(selectedallfemales3[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
```

```
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam5=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam5){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
damtbv= selectedallfemales3[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}

obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
```

```
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
pedigree3 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree3)
dataped <- pedigree3[pedigree3$id>nam4,]
simdataP3 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2002, herd = iherd, FI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
FItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simdataP3)
nrow(simdataP3) #9000

## combine the data file of (base, P1 and P2) with P3 data file
p3_p2_p1_basepopdata <- rbind(p2_p1_basepopdata,simdataP3)
## Data file for basepop and P1 (has both observations and TBV)
p3_p2_p1_basepopdata <- subset(p3_p2_p1_basepopdata, p3_p2_p1_basepopdata[ ,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1_p2_p3=mean(p3_p2_p1_basepopdata$WWT)
WWTave_base_p1_p2_p3
WWTsd_base_p1_p2_p3=sd(p3_p2_p1_basepopdata$WWT)
WWTsd_base_p1_p2_p3
WWTave_base_p1_p2_p3-(3*WWTsd_base_p1_p2_p3)
WWTave_base_p1_p2_p3+(3*WWTsd_base_p1_p2_p3)
p3_p2_p1_basepopdata <- subset(p3_p2_p1_basepopdata, WWT > (WWTave_base_p1_p2_p3-
(3*WWTsd_base_p1_p2_p3)))
p3_p2_p1_basepopdata <- subset(p3_p2_p1_basepopdata, WWT <
(WWTave_base_p1_p2_p3+(3*WWTsd_base_p1_p2_p3)))
summary(p3_p2_p1_basepopdata$WWT)
nrow(p3_p2_p1_basepopdata) #25953

# keeping YWT observations within 3 SD
YWTave_base_p1_p2_p3=mean(p3_p2_p1_basepopdata$YWT)
YWTave_base_p1_p2_p3
YWTsd_base_p1_p2_p3=sd(p3_p2_p1_basepopdata$YWT)
YWTsd_base_p1_p2_p3
YWTave_base_p1_p2_p3-(3*YWTsd_base_p1_p2_p3)
YWTave_base_p1_p2_p3+(3*YWTsd_base_p1_p2_p3)
p3_p2_p1_basepopdata <- subset(p3_p2_p1_basepopdata, YWT > (YWTave_base_p1_p2_p3-
(3*YWTsd_base_p1_p2_p3)))
p3_p2_p1_basepopdata <- subset(p3_p2_p1_basepopdata, YWT <
(YWTave_base_p1_p2_p3+(3*YWTsd_base_p1_p2_p3)))
summary(p3_p2_p1_basepopdata$YWT)
nrow(p3_p2_p1_basepopdata) #25937
```

```
#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p3_p2_p1_basepopdata <- subset(p3_p2_p1_basepopdata, p3_p2_p1_basepopdata[ ,10] -
p3_p2_p1_basepopdata[ ,9] > 40)
nrow(p3_p2_p1_basepopdata) #25937

simdataP3 <- subset(p3_p2_p1_basepopdata, p3_p2_p1_basepopdata[ ,1] > nam4)
summary(simdataP3$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#145.9   239.4   263.6   264.3   289.1   366.8
summary(simdataP3$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#310.8   463.3   510.0   509.8   556.3   680.4

nrow(simdataP3) #8935
N_sires = length(c(sort(unique(simdataP3$sire))))
N_sires #450
basepop_p1_p2_p3_data <- rbind(basepop_p2_p1_data,simdataP3)
pedigreep3 <- data.frame (id = simdataP3$id, sire = simdataP3$sire, dam = simdataP3$dam)
pedigreep3p2p1andbase <- rbind(pedigreep2p1andbase,pedigreep3)

################################################################################
################################################################################
## Creating P4 (Year 4)
#1 2 3 4 5 6 7 8 9 10 11 12 13
# id sire dam sex herd CE BW ww pwg CEtbv BWtbv wwtbv pwgtbv
#################################################################
#################################################################
## redefine (the total number of all animals) and (number of base population, P1 and P2) in
previous simulation
################################
#total number of all animals (Used unique IDs up to now): basepop, P1 and P2
nam6 = nam4+nam5

################################
## selecting top 5% sires (2000 & 2001) and top heifers for replacements
## selection of sires
males4 <- subset(basepop_p2_p1_data, sex == 1)
si.males4 <- data.frame(males4$id, si = -29.66932*males4$FItbv + 21.99373*males4$ADGtbv
+ 0.09652*males4$WWTtbv + 2.55104*males4$YWTtbv)
males4.si <- cbind(males4, si.males4)
bulls4 <- males4.si[order(-males4.si$si), ]
selectedmales4 <- bulls4[1:450, ]
selectedmales4$males4.id <- NULL
mean(selectedmales4[ ,15], na.rm=TRUE)
selectedmales4$si <- NULL
```

```
nrow(males4) #9001
nrow(selectedmales4) #450
nrow(selectedmales4)/nrow(males4) #0.05

#########################
## selection of dams
## FIRST: Remove females from year 3 (older than 16 years)
## SECOND: Selection of cows and replace with 2000 heifers
## Assume 20% loss
## Assume 90% Conception Rate

females4 <- subset(selectedallfemales3, year > 3)
femalesp4 <- females4[sample(1:nrow(females4), 8000, replace=FALSE),]
hfr <- subset(basepop_p2_p1_data, sex == 2 & year == 2001)
nrow(hfr) #3992

if(nrow(hfr) > 2000) {
  n = 2000
} else {
  n = nrow(hfr)
}

si.hfr <- data.frame(hfr$id, si = -29.66932*hfr$FItbv + 21.99373*hfr$ADGtbv +
0.09652*hfr$WWTtbv + 2.55104*hfr$YWTtbv)
hfs.si <- cbind(hfr, si.hfr)
hfrs <- hfs.si[order(-hfs.si$si), ]
selectedhfrs <- hfrs[1:n, ]
mean(selectedhfrs$si)
selectedhfrs$hfr.id <- NULL
selectedhfrs$si <- NULL
fem4 <- rbind(femalesp4, selectedhfrs)
selectedallfemales4 <- fem4[sample(1:nrow(fem4), 9000, replace = FALSE),]

nrow(femalesp4) #8000
nrow(selectedallfemales4) #9000
damlistp4 = c(selectedallfemales4[ ,1])

#########################
#### 450 selected males and 9000 selected females
n_progeny = nrow(selectedallfemales4)

allselectedmales=selectedmales4[rep(seq_len(nrow(selectedmales4)), each=20),]
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp4 = randomlymatedsires[ ,1]
```

```
####################
selectedparents=rbind(selectedmales4, selectedallfemales4)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]

animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
uniqueanimlist = uniqueanimlist[-1]
subdata <- basepop_p2_p1_data[basepop_p2_p1_data[ ,1] %in% uniqueanimlist, ]
subdata <- data.matrix(subdata)

length(subdata[,1]) #11229
length(uniqueanimlist) #11229
pedp4 = rbind(subdata[
,1:3],cbind(c((nam6+1):(nam6+n_progeny)),sirelistp4[1:n_progeny],damlistp4[1:n_progeny]))

###################
FIave=mean(selectedparents[ ,7], na.rm=TRUE)
FIave
ADGave=mean(selectedparents[ ,8], na.rm=TRUE)
ADGave
WWTave=mean(selectedparents[ ,9], na.rm=TRUE)
WWTave
YWTave=mean(selectedparents[ ,10], na.rm=TRUE)
YWTave

###############
###############
## P4 progeny, repeating the same code used to create P3.

library(MASS)

set.seed(1234)
herdFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(1.05,0.049,0.528,17.749,0.049,0.014,0.287,1.982,0.528,0.287,552.75,465.63,1
7.749,1.982,465.63,1131.95),byrow=TRUE,nrow=4)
# G matrix
```

```
G=matrix(data=c(0.694,0.050,9.945,18.845,0.050,0.005,0.709,1.550,9.945,0.709,230.77,327.934
,18.845,1.550,327.934,598.327),byrow=TRUE,nrow=4)
# Traits averages
mu=c(FIave,ADGave,WWTave,YWTave)

anwr=c((nam6+1):(nam6+n_progeny))
aid = c(pedp4[ ,1])
sid = c(pedp4[ ,2])
did <- c(pedp4[ ,3])

#number of herds
length(unique(selectedallfemales4[ ,6]))
iherd=c(selectedallfemales4[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam7=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam7){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
damtbv= selectedallfemales4[ ,11:14]
```

```
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}

obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
pedigree4 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree4)
dataped <- pedigree4[pedigree4$id>nam6,]
simdataP4 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2003, herd = iherd, FI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
FItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simdataP4)
nrow(simdataP4) #9000

## combine the data file of (base, P1, P2 and P3) with P4 data file
p4_p3_p2_p1_basepopdata <- rbind(p3_p2_p1_basepopdata,simdataP4)
## Data file for basepop, P1, P2, P3 and P4 (has both observations and TBV)
p4_p3_p2_p1_basepopdata <- subset(p4_p3_p2_p1_basepopdata, p4_p3_p2_p1_basepopdata[
,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1_p2_p3_p4=mean(p4_p3_p2_p1_basepopdata$WWT)
WWTave_base_p1_p2_p3_p4
WWTsd_base_p1_p2_p3_p4=sd(p4_p3_p2_p1_basepopdata$WWT)
WWTsd_base_p1_p2_p3_p4
WWTave_base_p1_p2_p3_p4-(3*WWTsd_base_p1_p2_p3_p4)
WWTave_base_p1_p2_p3_p4+(3*WWTsd_base_p1_p2_p3_p4)
p4_p3_p2_p1_basepopdata <- subset(p4_p3_p2_p1_basepopdata, WWT >
(WWTave_base_p1_p2_p3_p4-(3*WWTsd_base_p1_p2_p3_p4)))
```

```
p4_p3_p2_p1_basepopdata <- subset(p4_p3_p2_p1_basepopdata, WWT <
(WWTave_base_p1_p2_p3_p4+(3*WWTsd_base_p1_p2_p3_p4)))
summary(p4_p3_p2_p1_basepopdata$WWT)
nrow(p4_p3_p2_p1_basepopdata) #34893

# keeping YWT observations within 3 SD
YWTave_base_p1_p2_p3_p4=mean(p4_p3_p2_p1_basepopdata$YWT)
YWTave_base_p1_p2_p3_p4
YWTsd_base_p1_p2_p3_p4=sd(p4_p3_p2_p1_basepopdata$YWT)
YWTsd_base_p1_p2_p3_p4
YWTave_base_p1_p2_p3_p4-(3*YWTsd_base_p1_p2_p3_p4)
YWTave_base_p1_p2_p3_p4+(3*YWTsd_base_p1_p2_p3_p4)
p4_p3_p2_p1_basepopdata <- subset(p4_p3_p2_p1_basepopdata, YWT >
(YWTave_base_p1_p2_p3_p4-(3*YWTsd_base_p1_p2_p3_p4)))
p4_p3_p2_p1_basepopdata <- subset(p4_p3_p2_p1_basepopdata, YWT <
(YWTave_base_p1_p2_p3_p4+(3*YWTsd_base_p1_p2_p3_p4)))
summary(p4_p3_p2_p1_basepopdata$YWT)
nrow(p4_p3_p2_p1_basepopdata) #34882

#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p4_p3_p2_p1_basepopdata <- subset(p4_p3_p2_p1_basepopdata, p4_p3_p2_p1_basepopdata[
,10] - p4_p3_p2_p1_basepopdata[ ,9] > 40)
nrow(p4_p3_p2_p1_basepopdata) #34835

simdataP4 <- subset(p4_p3_p2_p1_basepopdata, p4_p3_p2_p1_basepopdata[ ,1] > nam6)
summary(simdataP4$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#139.8   223.7   249.7   250.4   276.2   365.9
summary(simdataP4$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#303.1   456.2   503.1   503.2   549.7   678.0

nrow(simdataP4) #8952
N_sires = length(c(sort(unique(simdataP4$sire))))
N_sires #450
basepop_p1_p2_p3_p4_data <- rbind(basepop_p1_p2_p3_data,simdataP4)
pedigreep4 <- data.frame (id = simdataP4$id, sire = simdataP4$sire, dam = simdataP4$dam)
pedigreep4p3p2p1andbase <- rbind(pedigreep3p2p1andbase,pedigreep4)

##############################################################################
##############################################################################
## Creating P5 (Year 5)
#1 2 3 4 5 6 7 8 9 10 11 12 13
# id sire dam sex herd CE BW ww pwg CEtbv BWtbv wwtbv pwgtbv
##############################################################
```

```
#####################################################################
## redefine (the total number of all animals) and (number of base population, P1, P2 and P3) in previous simulation
#################################
#total number of all animals (Used unique IDs up to now): basepop, P1 and P2
nam8 = nam6+nam7

#################################
## selecting top 5% sires and 80% dams
## selection of sires
males5 <- subset(basepop_p1_p2_p3_data, sex == 1 & year >= 2001)
si.males5 <- data.frame(males5$id, si = -29.66932*males5$FItbv + 21.99373*males5$ADGtbv
+ 0.09652*males5$WWTtbv + 2.55104*males5$YWTtbv)
males5.si <- cbind(males5, si.males5)
bulls5 <- males5.si[order(-males5.si$si), ]
selectedmales5 <- bulls5[1:450, ]
selectedmales5$males5.id <- NULL
mean(selectedmales5[ ,15], na.rm=TRUE)
selectedmales5$si <- NULL

nrow(males5) #8571
nrow(selectedmales5) #450
nrow(selectedmales5)/nrow(males5) #0.05

#########################
## selection of dams
## FIRST: Remove females from year 4 (older than 16 years)
## SECOND: Selection of cows and replace with 2000 heifers
## Assume 20% loss

females5 <- subset(selectedallfemales4, year > 4)
femalesp5 <- females5[sample(1:nrow(females5), 8000, replace=FALSE),]
hfr <- subset(basepop_p1_p2_p3_data, sex == 2 & year == 2002)
nrow(hfr) #4431

if(nrow(hfr) > 2000) {
  n = 2000
} else {
  n = nrow(hfr)
}

si.hfs <- data.frame(hfr$id, si = -29.66932*hfr$FItbv + 21.99373*hfr$ADGtbv +
0.09652*hfr$WWTtbv + 2.55104*hfr$YWTtbv)
hfs.si <- cbind(hfr, si.hfs)
hfrs <- hfs.si[order(-hfs.si$si), ]
selectedhfrs <- hfrs[1:n, ]
```

```
mean(selectedhfrs$si) #72.53
selectedhfrs$hfr.id <- NULL
selectedhfrs$si <- NULL
fem5 <- rbind(femalesp5, selectedhfrs)
selectedallfemales5 <- fem5[sample(1:nrow(fem5), 9000, replace = FALSE),]

nrow(femalesp5) #8000
nrow(selectedallfemales5) #9000
damlistp5 = c(selectedallfemales5[ ,1])

##########################
#### 450 selected males and 9000 selected females
n_progeny = nrow(selectedallfemales5)

allselectedmales=selectedmales5[rep(seq_len(nrow(selectedmales5)), each=20),]
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp5 = randomlymatedsires[ ,1]

#####################
selectedparents=rbind(selectedmales5, selectedallfemales5)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]

animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
uniqueanimlist = uniqueanimlist[-1]
subdata <- basepop_p1_p2_p3_data[basepop_p1_p2_p3_data[ ,1] %in% uniqueanimlist, ]
subdata <- data.matrix(subdata)

length(subdata[,1]) #12293
length(uniqueanimlist) #12293
pedp5 = rbind(subdata[
,1:3],cbind(c((nam8+1):(nam8+n_progeny)),sirelistp5[1:n_progeny],damlistp5[1:n_progeny]))

##################
FIave=mean(selectedparents[ ,7], na.rm=TRUE)
FIave
ADGave=mean(selectedparents[ ,8], na.rm=TRUE)
ADGave
WWTave=mean(selectedparents[ ,9], na.rm=TRUE)
WWTave
YWTave=mean(selectedparents[ ,10], na.rm=TRUE)
YWTave

##############
##############
## P5 progeny, repeating the same code used to create P4.
```

```
library(MASS)

set.seed(1234)
herdFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(1.05,0.049,0.528,17.749,0.049,0.014,0.287,1.982,0.528,0.287,552.75,465.63,1
7.749,1.982,465.63,1131.95),byrow=TRUE,nrow=4)
# G matrix
G=matrix(data=c(0.694,0.050,9.945,18.845,0.050,0.005,0.709,1.550,9.945,0.709,230.77,327.934
,18.845,1.550,327.934,598.327),byrow=TRUE,nrow=4)
# Traits averages
mu=c(FIave,ADGave,WWTave,YWTave)

anwr=c((nam8+1):(nam8+n_progeny))
aid = c(pedp5[ ,1])
sid = c(pedp5[ ,2])
did <- c(pedp5[ ,3])

#number of herds
length(unique(selectedallfemales5[ ,6]))
iherd=c(selectedallfemales5[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
```

```
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam9=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam9){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
damtbv= selectedallfemales5[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}

obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
pedigree5 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree5)
dataped <- pedigree5[pedigree5$id>nam8,]
simdataP5 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2004, herd = iherd, FI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
FItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
```

```
attach(simdataP5)
nrow(simdataP5) #9000

## combine the data file of (base, P1, P2, P3 and P4) with P5 data file
p5_p4_p3_p2_p1_basepopdata <- rbind(p4_p3_p2_p1_basepopdata,simdataP5)
## Data file for basepop and P's (has both observations and TBV)
p5_p4_p3_p2_p1_basepopdata <- subset(p5_p4_p3_p2_p1_basepopdata,
p5_p4_p3_p2_p1_basepopdata[ ,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1_p2_p3_p4_p5=mean(p5_p4_p3_p2_p1_basepopdata$WWT)
WWTave_base_p1_p2_p3_p4_p5
WWTsd_base_p1_p2_p3_p4_p5=sd(p5_p4_p3_p2_p1_basepopdata$WWT)
WWTsd_base_p1_p2_p3_p4_p5
WWTave_base_p1_p2_p3_p4_p5-(3*WWTsd_base_p1_p2_p3_p4_p5)
WWTave_base_p1_p2_p3_p4_p5+(3*WWTsd_base_p1_p2_p3_p4_p5)
p5_p4_p3_p2_p1_basepopdata <- subset(p5_p4_p3_p2_p1_basepopdata, WWT >
(WWTave_base_p1_p2_p3_p4_p5-(3*WWTsd_base_p1_p2_p3_p4_p5)))
p5_p4_p3_p2_p1_basepopdata <- subset(p5_p4_p3_p2_p1_basepopdata, WWT <
(WWTave_base_p1_p2_p3_p4_p5+(3*WWTsd_base_p1_p2_p3_p4_p5)))
summary(p5_p4_p3_p2_p1_basepopdata$WWT)
nrow(p5_p4_p3_p2_p1_basepopdata) #43812

# keeping YWT observations within 3 SD
YWTave_base_p1_p2_p3_p4_p5=mean(p5_p4_p3_p2_p1_basepopdata$YWT)
YWTave_base_p1_p2_p3_p4_p5
YWTsd_base_p1_p2_p3_p4_p5=sd(p5_p4_p3_p2_p1_basepopdata$YWT)
YWTsd_base_p1_p2_p3_p4_p5
YWTave_base_p1_p2_p3_p4_p5-(3*YWTsd_base_p1_p2_p3_p4_p5)
YWTave_base_p1_p2_p3_p4_p5+(3*YWTsd_base_p1_p2_p3_p4_p5)
p5_p4_p3_p2_p1_basepopdata <- subset(p5_p4_p3_p2_p1_basepopdata, YWT >
(YWTave_base_p1_p2_p3_p4_p5-(3*YWTsd_base_p1_p2_p3_p4_p5)))
p5_p4_p3_p2_p1_basepopdata <- subset(p5_p4_p3_p2_p1_basepopdata, YWT <
(YWTave_base_p1_p2_p3_p4_p5+(3*YWTsd_base_p1_p2_p3_p4_p5)))
summary(p5_p4_p3_p2_p1_basepopdata$YWT)
nrow(p5_p4_p3_p2_p1_basepopdata) #43787

#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p5_p4_p3_p2_p1_basepopdata <- subset(p5_p4_p3_p2_p1_basepopdata,
p5_p4_p3_p2_p1_basepopdata[ ,10] - p5_p4_p3_p2_p1_basepopdata[ ,9] > 40)
nrow(p5_p4_p3_p2_p1_basepopdata) #43747

simdataP5 <- subset(p5_p4_p3_p2_p1_basepopdata, p5_p4_p3_p2_p1_basepopdata[ ,1] >
nam8)
summary(simdataP5$WWT)
```

```
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#140.9   243.5   268.6   269.0   294.8   372.1
summary(simdataP5$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#328.0   478.1   526.3   525.7   572.1   694.6

nrow(simdataP5) #8917
N_sires = length(c(sort(unique(simdataP5$sire))))
N_sires #450
basepop_p1_p2_p3_p4_p5_data <- rbind(basepop_p1_p2_p3_p4_data,simdataP5)
pedigreep5 <- data.frame (id = simdataP5$id, sire = simdataP5$sire, dam = simdataP5$dam)
pedigreep5p4p3p2p1andbase <- rbind(pedigreep4p3p2p1andbase,pedigreep5)

##############################################################################
##############################################################################
## Creating P6 (Year 6)
#1 2 3 4 5 6 7 8 9 10 11 12 13
# id sire dam sex herd CE BW ww pwg CEtbv BWtbv wwtbv pwgtbv
#####################################################################
#####################################################################
## redefine (the total number of all animals) and (number of base population, P1, P2, P3 and P4)
in previous simulation
################################
#total number of all animals (Used unique IDs up to now): basepop, P1, P2, P3, P4, P5 and P5
nam10 = nam8+nam9

################################
## selecting top 5% sirs and 80% dams
## selection of sires

males6 <- subset(basepop_p1_p2_p3_p4_data, sex == 1 & year >= 2002)
si.males6 <- data.frame(males6$id, si = -29.66932*males6$FItbv + 21.99373*males6$ADGtbv
+ 0.09652*males6$WWTtbv + 2.55104*males6$YWTtbv)
males6.si <- cbind(males6, si.males6)
bulls6 <- males6.si[order(-males6.si$si), ]
selectedmales6 <- bulls6[1:450, ]
selectedmales6$males6.id <- NULL
mean(selectedmales6[ ,15], na.rm=TRUE)
selectedmales6$si <- NULL

nrow(males6) #9013
nrow(selectedmales6) #450
nrow(selectedmales6)/nrow(males6) #0.05

#########################
## selection of dams
```

```
## FIRST: Remove females from year 5 (older than 16 years)
## SECOND: Selection of cows and replace with 2000 heifers
## Assume 20% loss

females6 <- subset(selectedallfemales5, sex == 2 & year > 5)
femalesp6 <- females6[sample(1:nrow(females6), 8000, replace=FALSE),]
hfr <- subset(basepop_p1_p2_p3_p4_data, sex == 2 & year == 2003)
nrow(hfr) #4424

if(nrow(hfr) > 2000) {
  n = 2000
} else {
  n = nrow(hfr)
}

si.hfr <- data.frame(hfr$id, si = -29.66932*hfr$FItbv + 21.99373*hfr$ADGtbv +
0.09652*hfr$WWTtbv + 2.55104*hfr$YWTtbv)
hfs.si <- cbind(hfr, si.hfr)
hfrs <- hfs.si[order(-hfs.si$si), ]
selectedhfrs <- hfrs[1:n, ]
mean(selectedhfrs$si)
selectedhfrs$hfr.id <- NULL
selectedhfrs$si <- NULL
fem6 <- rbind(femalesp6, selectedhfrs)
selectedallfemales6 <- fem6[sample(1:nrow(fem6), 9000, replace = FALSE),]

nrow(femalesp6) #8000
nrow(selectedallfemales6) #9000
damlistp6 = c(selectedallfemales6[ ,1])

##########################
#### 450 selected males and 9000 selected females
n_progeny = nrow(selectedallfemales6)

allselectedmales=selectedmales6[rep(seq_len(nrow(selectedmales6)), each=20),]
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp6 = randomlymatedsires[ ,1]

####################
selectedparents=rbind(selectedmales6, selectedallfemales6)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]

animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
uniqueanimlist = uniqueanimlist[-1]
```

```
subdata <- basepop_p1_p2_p3_p4_data[basepop_p1_p2_p3_p4_data[ ,1] %in% uniqueanimlist,
]
subdata <- data.matrix(subdata)

length(subdata[,1]) #13098
length(uniqueanimlist) #13098
pedp6 = rbind(subdata[
,1:3],cbind(c((nam10+1):(nam10+n_progeny)),sirelistp6[1:n_progeny],damlistp6[1:n_progeny]))

###################
FIave=mean(selectedparents[ ,7], na.rm=TRUE)
FIave
ADGave=mean(selectedparents[ ,8], na.rm=TRUE)
ADGave
WWTave=mean(selectedparents[ ,9], na.rm=TRUE)
WWTave
YWTave=mean(selectedparents[ ,10], na.rm=TRUE)
YWTave

###############
###############
## P6 progeny, repeating the same code used to create P5.

library(MASS)

set.seed(1234)
herdFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(1.05,0.049,0.528,17.749,0.049,0.014,0.287,1.982,0.528,0.287,552.75,465.63,1
7.749,1.982,465.63,1131.95),byrow=TRUE,nrow=4)
# G matrix
G=matrix(data=c(0.694,0.050,9.945,18.845,0.050,0.005,0.709,1.550,9.945,0.709,230.77,327.934
,18.845,1.550,327.934,598.327),byrow=TRUE,nrow=4)
# Traits averages
mu=c(FIave,ADGave,WWTave,YWTave)

anwr=c((nam10+1):(nam10+n_progeny))
```

```
aid = c(pedp6[ ,1])
sid = c(pedp6[ ,2])
did <- c(pedp6[ ,3])

#number of herds
length(unique(selectedallfemales6[ ,6]))
iherd=c(selectedallfemales6[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam11=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam11){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
damtbv= selectedallfemales6[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
```

```
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}

obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
pedigree6 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree6)
dataped <- pedigree6[pedigree6$id>nam10,]
simdataP6 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2005, herd = iherd, FI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
FItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simdataP6)
nrow(simdataP6) #9000

## combine the data file of (base, P1, P2, P3, P4 and P5) with P6 data file
p6_p5_p4_p3_p2_p1_basepopdata <- rbind(p5_p4_p3_p2_p1_basepopdata,simdataP6)
## Data file for basepop and P's (has both observations and TBV)
p6_p5_p4_p3_p2_p1_basepopdata <- subset(p6_p5_p4_p3_p2_p1_basepopdata,
p6_p5_p4_p3_p2_p1_basepopdata[ ,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1_p2_p3_p4_p5_p6=mean(p6_p5_p4_p3_p2_p1_basepopdata$WWT)
WWTave_base_p1_p2_p3_p4_p5_p6
WWTsd_base_p1_p2_p3_p4_p5_p6=sd(p6_p5_p4_p3_p2_p1_basepopdata$WWT)
WWTsd_base_p1_p2_p3_p4_p5_p6
WWTave_base_p1_p2_p3_p4_p5_p6-(3*WWTsd_base_p1_p2_p3_p4_p5_p6)
WWTave_base_p1_p2_p3_p4_p5_p6+(3*WWTsd_base_p1_p2_p3_p4_p5_p6)
p6_p5_p4_p3_p2_p1_basepopdata <- subset(p6_p5_p4_p3_p2_p1_basepopdata, WWT >
(WWTave_base_p1_p2_p3_p4_p5_p6-(3*WWTsd_base_p1_p2_p3_p4_p5_p6)))
p6_p5_p4_p3_p2_p1_basepopdata <- subset(p6_p5_p4_p3_p2_p1_basepopdata, WWT <
(WWTave_base_p1_p2_p3_p4_p5_p6+(3*WWTsd_base_p1_p2_p3_p4_p5_p6)))
summary(p6_p5_p4_p3_p2_p1_basepopdata$WWT)
nrow(p6_p5_p4_p3_p2_p1_basepopdata) #52728

# keeping YWT observations within 3 SD
```

```
YWTave_base_p1_p2_p3_p4_p5_p6=mean(p6_p5_p4_p3_p2_p1_basepopdata$YWT)
YWTave_base_p1_p2_p3_p4_p5_p6
YWTsd_base_p1_p2_p3_p4_p5_p6=sd(p6_p5_p4_p3_p2_p1_basepopdata$YWT)
YWTsd_base_p1_p2_p3_p4_p5_p6
YWTave_base_p1_p2_p3_p4_p5_p6-(3*YWTsd_base_p1_p2_p3_p4_p5_p6)
YWTave_base_p1_p2_p3_p4_p5_p6+(3*YWTsd_base_p1_p2_p3_p4_p5_p6)
p6_p5_p4_p3_p2_p1_basepopdata <- subset(p6_p5_p4_p3_p2_p1_basepopdata, YWT >
(YWTave_base_p1_p2_p3_p4_p5_p6-(3*YWTsd_base_p1_p2_p3_p4_p5_p6)))
p6_p5_p4_p3_p2_p1_basepopdata <- subset(p6_p5_p4_p3_p2_p1_basepopdata, YWT <
(YWTave_base_p1_p2_p3_p4_p5_p6+(3*YWTsd_base_p1_p2_p3_p4_p5_p6)))
summary(p6_p5_p4_p3_p2_p1_basepopdata$YWT)
nrow(p6_p5_p4_p3_p2_p1_basepopdata) #52714

#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p6_p5_p4_p3_p2_p1_basepopdata <- subset(p6_p5_p4_p3_p2_p1_basepopdata,
p6_p5_p4_p3_p2_p1_basepopdata[ ,10] - p6_p5_p4_p3_p2_p1_basepopdata[ ,9] > 40)
nrow(p6_p5_p4_p3_p2_p1_basepopdata) #52714

simdataP6 <- subset(p6_p5_p4_p3_p2_p1_basepopdata, p6_p5_p4_p3_p2_p1_basepopdata[ ,1]
> nam10)
summary(simdataP6$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#144.6   245.7   270.4   271.4   297.3   375.0
summary(simdataP6$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#322.2   489.0   535.7   536.1   583.5   704.4

nrow(simdataP6) #8936
N_sires = length(c(sort(unique(simdataP6$sire))))
N_sires #450
basepop_p1_p2_p3_p4_p5_p6_data <- rbind(basepop_p1_p2_p3_p4_p5_data,simdataP6)
pedigreep6 <- data.frame (id = simdataP6$id, sire = simdataP6$sire, dam = simdataP6$dam)
pedigreep6p5p4p3p2p1andbase <- rbind(pedigreep5p4p3p2p1andbase,pedigreep6)

##############################################################################
##############################################################################
## Creating P7 (Year 7)
#1 2 3 4 5 6 7 8 9 10 11 12 13
# id sire dam sex herd CE BW ww pwg CEtbv BWtbv wwtbv pwgtbv
#################################################################
#################################################################
## redefine (the total number of all animals) and (number of base population, P1, P2, P3, P4 and
P5) in previous simulation
################################
```

```
#total number of all animals (Used unique IDs up to now): basepop, P1, P2, P3, P4, P5, P6 and
P7
nam12 = nam10+nam11

################################
## selecting top 5% sires and 80% dams
## selection of sires

males7 <- subset(basepop_p1_p2_p3_p4_p5_data, sex == 1 & year >= 2003)
si.males7 <- data.frame(males7$id, si = -29.66932*males7$FItbv + 21.99373*males7$ADGtbv
+ 0.09652*males7$WWTtbv + 2.55104*males7$YWTtbv)
males7.si <- cbind(males7, si.males7)
bulls7 <- males7.si[order(-males7.si$si), ]
selectedmales7 <- bulls7[1:450, ]
selectedmales7$males7.id <- NULL
mean(selectedmales7[ ,15], na.rm=TRUE) #142.43
selectedmales7$si <- NULL

nrow(males7) #9007
nrow(selectedmales7) #450
nrow(selectedmales7)/nrow(males7) #0.05

#########################
## selection of dams
## FIRST: Remove females from year 5 (older than 16 years)
## SECOND: Selection of cows and replace with 2000 heifers
## Assume 20% loss

females7 <- subset(selectedallfemales6, sex == 2 & year > 6)
femalesp7 <- females7[sample(1:nrow(females7), 8000, replace=FALSE),]
hfr <- subset(basepop_p1_p2_p3_p4_p5_data, sex == 2 & year == 2004)
nrow(hfr) #4413

if(nrow(hfr) > 2000) {
  n = 2000
} else {
  n = nrow(hfr)
}

si.hfr <- data.frame(hfr$id, si = -29.66932*hfr$FItbv + 21.99373*hfr$ADGtbv +
0.09652*hfr$WWTtbv + 2.55104*hfr$YWTtbv)
hfs.si <- cbind(hfr, si.hfr)
hfrs <- hfs.si[order(-hfs.si$si), ]
selectedhfrs <- hfrs[1:n, ]
mean(selectedhfrs$si)
selectedhfrs$hfr.id <- NULL
```

```
selectedhfrs$si <- NULL
fem7 <- rbind(femalesp7, selectedhfrs)
selectedallfemales7 <- fem7[sample(1:nrow(fem7), 9000, replace = FALSE),]

nrow(femalesp7) #8000
nrow(selectedallfemales7) #9000
damlistp7 = c(selectedallfemales7[ ,1])

#########################
#### 450 selected males and 9000 selected females
n_progeny = nrow(selectedallfemales7)

allselectedmales=selectedmales7[rep(seq_len(nrow(selectedmales7)), each=20),]
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp7= randomlymatedsires[ ,1]

####################
selectedparents=rbind(selectedmales7, selectedallfemales7)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]

animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
uniqueanimlist = uniqueanimlist[-1]
subdata <- basepop_p1_p2_p3_p4_p5_data[basepop_p1_p2_p3_p4_p5_data[ ,1] %in%
uniqueanimlist, ]
subdata <- data.matrix(subdata)

length(subdata[,1]) #13856
length(uniqueanimlist) #13856
pedp7 = rbind(subdata[
,1:3],cbind(c((nam12+1):(nam12+n_progeny)),sirelistp7[1:n_progeny],damlistp7[1:n_progeny]))

##################
FIave=mean(selectedparents[ ,7], na.rm=TRUE)
FIave
ADGave=mean(selectedparents[ ,8], na.rm=TRUE)
ADGave
WWTave=mean(selectedparents[ ,9], na.rm=TRUE)
WWTave
YWTave=mean(selectedparents[ ,10], na.rm=TRUE)
YWTave

###############
###############
## P7 progeny, repeating the same code used to create P6.
```

```
library(MASS)

set.seed(1234)
herdFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(1.05,0.049,0.528,17.749,0.049,0.014,0.287,1.982,0.528,0.287,552.75,465.63,1
7.749,1.982,465.63,1131.95),byrow=TRUE,nrow=4)
# G matrix
G=matrix(data=c(0.694,0.050,9.945,18.845,0.050,0.005,0.709,1.550,9.945,0.709,230.77,327.934
,18.845,1.550,327.934,598.327),byrow=TRUE,nrow=4)
# Traits averages
mu=c(FIave,ADGave,WWTave,YWTave)

anwr=c((nam12+1):(nam12+n_progeny))
aid = c(pedp7[ ,1])
sid = c(pedp7[ ,2])
did <- c(pedp7[ ,3])

#number of herds
length(unique(selectedallfemales7[ ,6]))
iherd=c(selectedallfemales7[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
```

```
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam13=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam13){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
damtbv= selectedallfemales7[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}

obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
pedigree7 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree7)
dataped <- pedigree7[pedigree7$id>nam12,]
simdataP7 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2006, herd = iherd, FI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
FItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simdataP7)
```

```
nrow(simdataP7) #9000

## combine the data file of (base, P1, P2, P3, P4, P5 and P6) with P7 data file
p7_p6_p5_p4_p3_p2_p1_basepopdata <- rbind(p6_p5_p4_p3_p2_p1_basepopdata,simdataP7)
## Data file for basepop, P1, P2, P3, P4, P5, P6 and P7 (has both observations and TBV)
p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p7_p6_p5_p4_p3_p2_p1_basepopdata,
p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1_p2_p3_p4_p5_p6_p7=mean(p7_p6_p5_p4_p3_p2_p1_basepopdata$WWT)
WWTave_base_p1_p2_p3_p4_p5_p6_p7
WWTsd_base_p1_p2_p3_p4_p5_p6_p7=sd(p7_p6_p5_p4_p3_p2_p1_basepopdata$WWT)
WWTsd_base_p1_p2_p3_p4_p5_p6_p7
WWTave_base_p1_p2_p3_p4_p5_p6_p7-(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7)
WWTave_base_p1_p2_p3_p4_p5_p6_p7+(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7)
p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p7_p6_p5_p4_p3_p2_p1_basepopdata, WWT >
(WWTave_base_p1_p2_p3_p4_p5_p6_p7-(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7)))
p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p7_p6_p5_p4_p3_p2_p1_basepopdata, WWT <
(WWTave_base_p1_p2_p3_p4_p5_p6_p7+(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7)))
summary(p7_p6_p5_p4_p3_p2_p1_basepopdata$WWT)
nrow(p7_p6_p5_p4_p3_p2_p1_basepopdata) #61639

# keeping YWT observations within 3 SD
YWTave_base_p1_p2_p3_p4_p5_p6_p7=mean(p7_p6_p5_p4_p3_p2_p1_basepopdata$YWT)
YWTave_base_p1_p2_p3_p4
YWTsd_base_p1_p2_p3_p4_p5_p6_p7=sd(p7_p6_p5_p4_p3_p2_p1_basepopdata$YWT)
YWTsd_base_p1_p2_p3_p4
YWTave_base_p1_p2_p3_p4-(3*YWTsd_base_p1_p2_p3_p4)
YWTave_base_p1_p2_p3_p4+(3*YWTsd_base_p1_p2_p3_p4)
p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p7_p6_p5_p4_p3_p2_p1_basepopdata, YWT >
(YWTave_base_p1_p2_p3_p4-(3*YWTsd_base_p1_p2_p3_p4)))
p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p7_p6_p5_p4_p3_p2_p1_basepopdata, YWT <
(YWTave_base_p1_p2_p3_p4+(3*YWTsd_base_p1_p2_p3_p4)))
summary(p7_p6_p5_p4_p3_p2_p1_basepopdata$YWT)
nrow(p7_p6_p5_p4_p3_p2_p1_basepopdata) #61494

#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p7_p6_p5_p4_p3_p2_p1_basepopdata,
p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,10] - p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,9] > 40)
nrow(p7_p6_p5_p4_p3_p2_p1_basepopdata) #61494

simdataP7 <- subset(p7_p6_p5_p4_p3_p2_p1_basepopdata,
p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,1] > nam12)
summary(simdataP7$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
```

```
#149.8   255.5   280.6   280.8   306.2   380.6
summary(simdataP7$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#308.7   506.1   552.6   552.6   600.9   685.9

nrow(simdataP7) #8824
N_sires = length(c(sort(unique(simdataP7$sire))))
N_sires #450
basepop_p1_p2_p3_p4_p5_p6_p7_data <- rbind(basepop_p1_p2_p3_p4_p5_p6_data,simdataP7)
pedigreep7 <- data.frame (id = simdataP7$id, sire = simdataP7$sire, dam = simdataP7$dam)
pedigreep7p6p5p4p3p2p1andbase <- rbind(pedigreep6p5p4p3p2p1andbase,pedigreep7)

#############################################################################
#############################################################################
## Creating P8 (Year 8)
#1 2 3 4 5 6 7 8 9 10 11 12 13
# id sire dam sex herd CE BW ww pwg CEtbv BWtbv wwtbv pwgtbv
#################################################################
#################################################################
## redefine (the total number of all animals) and (number of base population, P1, P2, P3, P4, P5
and P6) in previous simulation
################################
#total number of all animals (Used unique IDs up to now): basepop, P1, P2, P3, P4, P5, P6, P7
and P8
nam14 = nam12+nam13

################################
## selecting top 5% sires
## selection of sires
males8 <- subset(basepop_p1_p2_p3_p4_p5_p6_data, sex == 1 & year >= 2004)
si.males8 <- data.frame(males8$id, si = -29.66932*males8$FItbv + 21.99373*males8$ADGtbv
+ 0.09652*males8$WWTtbv + 2.55104*males8$YWTtbv)
males8.si <- cbind(males8, si.males8)
bulls8 <- males8.si[order(-males8.si$si), ]
selectedmales8 <- bulls8[1:450, ]
selectedmales8$males8.id <- NULL
mean(selectedmales8[ ,15], na.rm=TRUE)
selectedmales8$si <- NULL

nrow(males8) #8997
nrow(selectedmales8) #450
nrow(selectedmales8)/nrow(males8) #0.05

#########################
## selection of dams
## FIRST: Remove females from year 7 (older than 16 years)
```

```
## SECOND: Selection of cows and replace with 2000 heifers
## Assume 20% loss

females8 <- subset(selectedallfemales7, sex == 2 & year > 7)
femalesp8 <- females8[sample(1:nrow(females8), 8000, replace=FALSE),]
hfr <- subset(basepop_p1_p2_p3_p4_p5_p6_data, sex == 2 & year == 2005)
nrow(hfr) #4419

if(nrow(hfr) > 2000) {
  n = 2000
} else {
  n = nrow(hfr)
}

si.hfr <- data.frame(hfr$id, si = -29.66932*hfr$FItbv + 21.99373*hfr$ADGtbv +
0.09652*hfr$WWTtbv + 2.55104*hfr$YWTtbv)
hfs.si <- cbind(hfr, si.hfr)
hfrs <- hfs.si[order(-hfs.si$si), ]
selectedhfrs <- hfrs[1:n, ]
mean(selectedhfrs$si)
selectedhfrs$hfr.id <- NULL
selectedhfrs$si <- NULL
fem8 <- rbind(femalesp8, selectedhfrs)
selectedallfemales8 <- fem8[sample(1:nrow(fem8), 9000, replace = FALSE),]

nrow(femalesp8) #8000
nrow(selectedallfemales8) #9000
damlistp8 = c(selectedallfemales8[ ,1])

##########################
#### 450 selected males and 9000 selected females
n_progeny = nrow(selectedallfemales8)

allselectedmales=selectedmales8[rep(seq_len(nrow(selectedmales8)), each=20),]
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp8 = randomlymatedsires[ ,1]

#####################
selectedparents=rbind(selectedmales8, selectedallfemales8)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]

animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
uniqueanimlist = uniqueanimlist[-1]
subdata <- basepop_p1_p2_p3_p4_p5_p6_data[basepop_p1_p2_p3_p4_p5_p6_data[ ,1] %in%
uniqueanimlist, ]
```

```
subdata <- data.matrix(subdata)

length(subdata[,1]) #14556
length(uniqueanimlist) #14556
pedp8 = rbind(subdata[
,1:3],cbind(c((nam14+1):(nam14+n_progeny)),sirelistp8[1:n_progeny],damlistp8[1:n_progeny]))

###################
FIave=mean(selectedparents[ ,7], na.rm=TRUE)
FIave
ADGave=mean(selectedparents[ ,8], na.rm=TRUE)
ADGave
WWTave=mean(selectedparents[ ,9], na.rm=TRUE)
WWTave
YWTave=mean(selectedparents[ ,10], na.rm=TRUE)
YWTave

###############
###############
## P8 progeny, repeating the same code used to create P7.

library(MASS)

set.seed(1234)
herdFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(1.05,0.049,0.528,17.749,0.049,0.014,0.287,1.982,0.528,0.287,552.75,465.63,1
7.749,1.982,465.63,1131.95),byrow=TRUE,nrow=4)
# G matrix
G=matrix(data=c(0.694,0.050,9.945,18.845,0.050,0.005,0.709,1.550,9.945,0.709,230.77,327.934
,18.845,1.550,327.934,598.327),byrow=TRUE,nrow=4)
# Traits averages
mu=c(FIave,ADGave,WWTave,YWTave)

anwr=c((nam14+1):(nam14+n_progeny))
aid = c(pedp8[ ,1])
sid = c(pedp8[ ,2])
```

```
did <- c(pedp8[ ,3])

#number of herds
length(unique(selectedallfemales8[ ,6]))
iherd=c(selectedallfemales8[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam15=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam15){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
damtbv= selectedallfemales8[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
obser = jd(nrec,4)*0
for(k in 1:nrec){
```

```
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}

obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
pedigree8 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree8)
dataped <- pedigree8[pedigree8$id>nam14,]
simdataP8 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2007, herd = iherd, FI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
FItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simdataP8)
nrow(simdataP8) #9000

## combine the data file of (base, P1, P2, P3, P4, P5, P6 and P7) with P8 data file
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
rbind(p7_p6_p5_p4_p3_p2_p1_basepopdata,simdataP8)
## Data file for basepop and P's (has both observations and TBV)
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8=mean(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$
WWT)
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8
WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8=sd(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$WW
T)
WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8-(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8)
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8+(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8)
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
WWT > (WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8-
(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8)))
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
WWT <
(WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8+(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8))
)
```

```
summary(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$WWT)
nrow(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #70411

# keeping YWT observations within 3 SD
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8=mean(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$
YWT)
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8
YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8=sd(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$YWT
)
YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8-(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8)
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8+(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8)
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
YWT > (YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8-
(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8)))
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
YWT <
(YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8+(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8)))
summary(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$YWT)
nrow(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #70378

#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <- subset(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,10] - p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,9]
> 40)
nrow(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #70378

simdataP8 <- subset(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,1] > nam14)
summary(simdataP8$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#150.6   265.3   290.1   290.9   316.4   386.5
summary(simdataP8$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#346.1   526.2   573.0   572.5   620.2   729.9

nrow(simdataP8) #8895
N_sires = length(c(sort(unique(simdataP8$sire))))
N_sires #450
basepop_p1_p2_p3_p4_p5_p6_p7_p8_data <-
rbind(basepop_p1_p2_p3_p4_p5_p6_p7_data,simdataP8)
pedigreep8 <- data.frame (id = simdataP8$id, sire = simdataP8$sire, dam = simdataP8$dam)
pedigreep8p7p6p5p4p3p2p1andbase <- rbind(pedigreep7p6p5p4p3p2p1andbase,pedigreep8)

################################################################################
```

```
#############################################################################
## Creating P9 (Year 9)
#1 2 3 4 5 6 7 8 9 10 11 12 13
# id sire dam sex herd CE BW ww pwg CEtbv BWtbv wwtbv pwgtbv
#################################################################
#################################################################
## redefine (the total number of all animals) and (number of base population, P1, P2, P3, P4, P5,
P6, P7 and P8) in previous simulation
################################
#total number of all animals (Used unique IDs up to now): basepop, P1, P2, P3, P4, P5, P6, P7,
P8 and P9
nam16 = nam14+nam15

################################
## selecting top 5% sires
## selection of sires
males9 <- subset(basepop_p1_p2_p3_p4_p5_p6_p7_data, sex == 1 & year >= 2005)
si.males9 <- data.frame(males9$id, si = -29.66932*males9$FItbv + 21.99373*males9$ADGtbv
+ 0.09652*males9$WWTtbv + 2.55104*males9$YWTtbv)
males9.si <- cbind(males9, si.males9)
bulls9 <- males9.si[order(-males9.si$si), ]
selectedmales9 <- bulls9[1:450, ]
selectedmales9$males9.id <- NULL
mean(selectedmales9[ ,15], na.rm=TRUE)
selectedmales9$si <- NULL

nrow(males9) #8899
nrow(selectedmales9) #450
nrow(selectedmales9)/nrow(males9) #0.05

#########################
## selection of dams
## FIRST: Remove females from year 8 (older than 16 years)
## SECOND: Selection of cows and replace with 2000 heifers
## Assume 20% loss

females9 <- subset(selectedallfemales8, sex == 2 & year > 8)
femalesp9 <- females9[sample(1:nrow(females9), 8000, replace=FALSE),]
hfr <- subset(basepop_p1_p2_p3_p4_p5_p6_p7_data, sex == 2 & year == 2006)
nrow(hfr) #4434

if(nrow(hfr) > 2000) {
  n = 2000
} else {
  n = nrow(hfr)
}
```

```
si.hfr <- data.frame(hfr$id, si = -29.66932*hfr$FItbv + 21.99373*hfr$ADGtbv +
0.09652*hfr$WWTtbv + 2.55104*hfr$YWTtbv)
hfs.si <- cbind(hfr, si.hfr)
hfrs <- hfs.si[order(-hfs.si$si), ]
selectedhfrs <- hfrs[1:n, ]
mean(selectedhfrs$si)
selectedhfrs$hfr.id <- NULL
selectedhfrs$si <- NULL
fem9 <- rbind(femalesp9, selectedhfrs)
selectedallfemales9 <- fem9[sample(1:nrow(fem9), 9000, replace = FALSE),]

nrow(femalesp9) #8000
nrow(selectedallfemales9) #9000
damlistp9 = c(selectedallfemales9[ ,1])

#########################
#### 450 selected males and 9000 selected females
n_progeny = nrow(selectedallfemales9)

allselectedmales=selectedmales9[rep(seq_len(nrow(selectedmales9)), each=20),]
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp9 = randomlymatedsires[ ,1]

####################
selectedparents=rbind(selectedmales9, selectedallfemales9)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]

animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
uniqueanimlist = uniqueanimlist[-1]
subdata <- basepop_p1_p2_p3_p4_p5_p6_p7_data[basepop_p1_p2_p3_p4_p5_p6_p7_data[ ,1]
%in% uniqueanimlist, ]
subdata <- data.matrix(subdata)

length(subdata[,1]) #15070
length(uniqueanimlist) #15070
pedp9 = rbind(subdata[
,1:3],cbind(c((nam16+1):(nam16+n_progeny)),sirelistp9[1:n_progeny],damlistp9[1:n_progeny]))

##################
FIave=mean(selectedparents[ ,7], na.rm=TRUE)
FIave
ADGave=mean(selectedparents[ ,8], na.rm=TRUE)
ADGave
WWTave=mean(selectedparents[ ,9], na.rm=TRUE)
```

```
WWTave
YWTave=mean(selectedparents[ ,10], na.rm=TRUE)
YWTave

###############
###############
## P9 progeny, repeating the same code used to create P8.

library(MASS)

set.seed(1234)
herdFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(1.05,0.049,0.528,17.749,0.049,0.014,0.287,1.982,0.528,0.287,552.75,465.63,1
7.749,1.982,465.63,1131.95),byrow=TRUE,nrow=4)
# G matrix
G=matrix(data=c(0.694,0.050,9.945,18.845,0.050,0.005,0.709,1.550,9.945,0.709,230.77,327.934
,18.845,1.550,327.934,598.327),byrow=TRUE,nrow=4)

####averages:
mu=c(FIave,ADGave,WWTave,YWTave)
anwr=c((nam16+1):(nam16+n_progeny))
aid = c(pedp9[ ,1])
sid = c(pedp9[ ,2])
did <- c(pedp9[ ,3])

#number of herds
length(unique(selectedallfemales9[ ,6]))
iherd=c(selectedallfemales9[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2
```

```
# correlations among traits
# Function to calculate correlations from a covariance matrix
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam17=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam1){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
damtbv= selectedallfemales9[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}

obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)
```

```
#########################################################################
pedigree9 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree9)
dataped <- pedigree9[pedigree9$id>nam16,]
simdataP9 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2008, herd = iherd, FI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
FItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simdataP9)
nrow(simdataP9) #9000

## combine the data file of (base, P1, P2, P3, P4, P5, P6, P7 and P8) with P9 data file
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
rbind(p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,simdataP9)
## Data file for basepop and P's (has both observations and TBV)
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9=mean(p9_p8_p7_p6_p5_p4_p3_p2_p1_basep
opdata$WWT)
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9
WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9=sd(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdat
a$WWT)
WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9-
(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9)
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9+(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p
8_p9)
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata, WWT >
(WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9-
(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9)))
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata, WWT <
(WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9+(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_
p8_p9)))
summary(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$WWT)
nrow(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #79271

# keeping YWT observations within 3 SD
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9=mean(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepo
pdata$YWT)
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9
YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9=sd(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdat
a$YWT)
```

```
YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9-
(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9)
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9+(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8
_p9)
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata, YWT >
(YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9-
(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9)))
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata, YWT <
(YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9+(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p
8_p9)))
summary(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$YWT)
nrow(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #79239

#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,10] -
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,9] > 40)
nrow(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #79409

simdataP9 <- subset(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,1] > nam16)
summary(simdataP9$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#161.5   275.4   300.2   300.0   325.4   393.6
summary(simdataP9$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#342.0   544.8   591.9   591.0   638.8   746.2

nrow(simdataP9) #8862
N_sires = length(c(sort(unique(simdataP9$sire))))
N_sires # 450
basepop_p1_p2_p3_p4_p5_p6_p7_p8_p9_data <-
rbind(basepop_p1_p2_p3_p4_p5_p6_p7_p8_data,simdataP9)
pedigreep9 <- data.frame (id = simdataP9$id, sire = simdataP9$sire, dam = simdataP9$dam)
pedigreep9p8p7p6p5p4p3p2p1andbase <-
rbind(pedigreep8p7p6p5p4p3p2p1andbase,pedigreep9)

##############################################################################
##############################################################################
## Creating P10 (Year 10)
#1 2 3 4 5 6 7 8 9 10 11 12 13
```

```
# id sire dam sex herd CE BW ww pwg CEtbv BWtbv wwtbv pwgtbv
#####################################################################
#####################################################################
## redefine (the total number of all animals) and (number of base population, P1, P2, P3, P4, P5, P6, P7 and P8) in previous simulation
################################
#total number of all animals (Used unique IDs up to now): basepop, P1, P2, P3, P4, P5, P6, P7, P8 and P9
nam18 = nam16+nam17

################################
## selecting top 5% sires and 80% dams
## selection of sires
males10 <- subset(basepop_p1_p2_p3_p4_p5_p6_p7_p8_data, sex == 1 & year >= 2006)
si.males10 <- data.frame(males10$id, si = -29.66932*males10$FItbv +
21.99373*males10$ADGtbv + 0.09652*males10$WWTtbv + 2.55104*males10$YWTtbv)
males10.si <- cbind(males10, si.males10)
bulls10 <- males10.si[order(-males10.si$si), ]
selectedmales10 <- bulls10[1:450, ]
selectedmales10$males10.id <- NULL
mean(selectedmales10[ ,15], na.rm=TRUE)
selectedmales10$si <- NULL

nrow(males10) #8866
nrow(selectedmales10) #450
nrow(selectedmales10)/nrow(males10) #0.05

#########################
## selection of dams
## FIRST: Remove females from year 9 (older than 16 years)
## SECOND: Selection of cows and replace with 2000 heifers
## Assume 20% loss

females10 <- subset(selectedallfemales9, sex == 2 & year > 9)
femalesp10 <- females10[sample(1:nrow(females10), 8000, replace=FALSE),]
hfr <- subset(basepop_p1_p2_p3_p4_p5_p6_p7_p8_data, sex == 2 & year == 2007)
nrow(hfr)

if(nrow(hfr) > 2000) {
  n = 2000
} else {
  n = nrow(hfr)
}

si.hfr <- data.frame(hfr$id, si = -29.66932*hfr$FItbv + 21.99373*hfr$ADGtbv +
0.09652*hfr$WWTtbv + 2.55104*hfr$YWTtbv)
```

```
hfs.si <- cbind(hfr, si.hfr)
hfrs <- hfs.si[order(-hfs.si$si), ]
selectedhfrs <- hfrs[1:n, ]
mean(selectedhfrs$si)
selectedhfrs$hfr.id <- NULL
selectedhfrs$si <- NULL
fem10 <- rbind(femalesp10, selectedhfrs)
selectedallfemales10 <- fem10[sample(1:nrow(fem10), 9000, replace = FALSE),]

nrow(femalesp10) #8000
nrow(selectedallfemales10) #9000
damlistp10 = c(selectedallfemales10[ ,1])

##########################
#### 450 selected males and 9000 selected females
n_progeny = nrow(selectedallfemales10)

allselectedmales=selectedmales10[rep(seq_len(nrow(selectedmales10)), each=20),]
randomlymatedsires=allselectedmales[sample(nrow(allselectedmales)),]
sirelistp10 = randomlymatedsires[ ,1]

#####################
selectedparents=rbind(selectedmales10, selectedallfemales10)
sortedselectedparents <- selectedparents[order(selectedparents[ ,1]),]

animlist = c(selectedparents[ ,1],selectedparents[ ,2], selectedparents[ ,3])
uniqueanimlist = c(sort(unique(animlist)))
uniqueanimlist = uniqueanimlist[-1]
subdata <-
basepop_p1_p2_p3_p4_p5_p6_p7_p8_data[basepop_p1_p2_p3_p4_p5_p6_p7_p8_data[ ,1]
%in% uniqueanimlist, ]
subdata <- data.matrix(subdata)

length(subdata[,1]) #15624
length(uniqueanimlist) #15624
pedp10 = rbind(subdata[
,1:3],cbind(c((nam18+1):(nam18+n_progeny)),sirelistp10[1:n_progeny],damlistp10[1:n_progeny
]))

##################
FIave=mean(selectedparents[ ,7], na.rm=TRUE)
FIave
ADGave=mean(selectedparents[ ,8], na.rm=TRUE)
ADGave
WWTave=mean(selectedparents[ ,9], na.rm=TRUE)
WWTave
```

```
YWTave=mean(selectedparents[ ,10], na.rm=TRUE)
YWTave

###############
###############
## P10 progeny, repeating the same code used to create P9.

set.seed(1234)
herdFI=(rnorm(52,0))
set.seed(1234)
herdWW=(rnorm(52,0))*25
set.seed(1234)
herdYW=(rnorm(52,0))*10
set.seed(1234)
herdADG=(rnorm(52,0))*0.25
herd=matrix(data=c(herdFI,herdADG,herdWW,herdYW),byrow=TRUE,nrow=4)
sex=matrix(data=c(0,0,0.060,-0.040,10,-7,50,-40),byrow=TRUE,nrow=4)

# Residual matrix
R=matrix(data=c(1.05,0.049,0.528,17.749,0.049,0.014,0.287,1.982,0.528,0.287,552.75,465.63,1
7.749,1.982,465.63,1131.95),byrow=TRUE,nrow=4)
# G matrix
G=matrix(data=c(0.694,0.050,9.945,18.845,0.050,0.005,0.709,1.550,9.945,0.709,230.77,327.934
,18.845,1.550,327.934,598.327),byrow=TRUE,nrow=4)

####averages:
mu=c(FIave,ADGave,WWTave,YWTave)
anwr=c((nam18+1):(nam18+n_progeny))
aid = c(pedp10[ ,1])
sid = c(pedp10[ ,2])
did <- c(pedp10[ ,3])

#number of herds
length(unique(selectedallfemales10[ ,6]))
iherd=c(selectedallfemales10[ ,6])
set.seed(123)
isex=(rbinom(n_progeny, 1, 0.5))+1

# heritability of traits
gd=diag(G)
rd=diag(R)
h2=gd/(gd+rd)
h2

# correlations among traits
# Function to calculate correlations from a covariance matrix
```

```
CORMAT=function(Q) {
  D = diag(Q)
  D = sqrt(D)
  B = diag(1/D)
  HC = B %*% Q %*% B
  HC }
CORMAT(R)
CORMAT(G)
# Get cholesky decompositions of G and R
LG = t(chol(G))
LR = t(chol(R))

# Simulate true breeding values for all animals
# J MATRIX FUNCTION
jd = function(n,m){
  matrix(c(1),nrow=n,ncol=m)}
nam19=n_progeny
mendelian = jd(nam1,4)*0
for(i in 1:nam1){
  mendelian[i, ] = LG %*% (rnorm(4,0,1))
}
damtbv= selectedallfemales10[ ,11:14]
siretbv= randomlymatedsires[ ,11:14]
parentsaveragetbv=0.5*(damtbv + siretbv)
tbv=mendelian+parentsaveragetbv

nrec=n_progeny
# Make an observation for all traits for all animals
obser = jd(nrec,4)*0
for(k in 1:nrec){
  kherd=iherd[k]
  ksex=isex[k]
  obser[k, ]=mu + herd[ ,kherd] + sex[ ,ksex]
  res = LR %*% rnorm(4,0,1)
  obser[k, ]=obser[k, ] + res
}

obs = obser + tbv
obs[ ,1] = round(obs[ ,1], digits=2)
obs[ ,2] = round(obs[ ,2], digits=2)
obs[ ,3] = round(obs[ ,3], digits=2)
obs[ ,4] = round(obs[ ,4], digits = 2)

#########################################################################
pedigree10 <- data.frame (id = aid, sire = sid, dam = did)
attach(pedigree10)
```

```
dataped <- pedigree10[pedigree10$id>nam18,]
simdataP10 <- data.frame (id = dataped$id, sire = dataped$sire, dam = dataped$dam, sex = isex,
year = 2009, herd = iherd, FI = obs[ ,1], ADG = obs[ ,2], WWT = obs[ ,3], YWT = obs[ ,4],
FItbv = tbv[ ,1], ADGtbv = tbv[ ,2], WWTtbv = tbv[ ,3], YWTtbv = tbv[ ,4])
attach(simdataP10)
nrow(simdataP10) #9000

## combine the data file of (base, P1, P2, P3, P4, P5, P6, P7, P8 and P9) with P10 data file
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
rbind(p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,simdataP10)
## Data file for basepop and P's (has both observations and TBV)
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,1] > nbase)

# keeping WWT observations within 3 SD
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10=mean(p10_p9_p8_p7_p6_p5_p4_p3_p2_
p1_basepopdata$WWT)
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10
WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10=sd(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_b
asepopdata$WWT)
WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10-
(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10)
WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10+(3*WWTsd_base_p1_p2_p3_p4_p5_p6_
p7_p8_p9_p10)
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata, WWT >
(WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10-
(3*WWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10)))
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata, WWT <
(WWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10+(3*WWTsd_base_p1_p2_p3_p4_p5_p6
_p7_p8_p9_p10)))
summary(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$WWT)
nrow(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #88106

# keeping YWT observations within 3 SD
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10=mean(p10_p9_p8_p7_p6_p5_p4_p3_p2_
p1_basepopdata$YWT)
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10
YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10=sd(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_b
asepopdata$YWT)
YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10-
(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10)
```

```
YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10+(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p
7_p8_p9_p10)
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata, YWT >
(YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10-
(3*YWTsd_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10)))
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata, YWT <
(YWTave_base_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10+(3*YWTsd_base_p1_p2_p3_p4_p5_p6_
p7_p8_p9_p10)))
summary(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata$YWT)
nrow(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #88068

#Remove animals with less than 40kg (88lbs) gain from wwt to ywt (might be too small to
reproduce anyway)
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata <-
subset(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,9] -
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,8] > 40)
nrow(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata) #88312

simdataP10 <- subset(p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata,
p10_p9_p8_p7_p6_p5_p4_p3_p2_p1_basepopdata[ ,1] > nam18)
summary(simdataP10$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#167.9   288.2   313.0   313.0   338.8   403.4
summary(simdataP10$YWT)
#  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#381.2   568.3   615.1   614.3   661.3   765.6

nrow(simdataP10) #8829
N_sires = length(c(sort(unique(simdataP10$sire))))
N_sires #450
basepop_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10_data <-
rbind(basepop_p1_p2_p3_p4_p5_p6_p7_p8_p9_data,simdataP10)
pedigreep10 <- data.frame (id = simdataP10$id, sire = simdataP10$sire, dam =
simdataP10$dam)
pedigreep10p9p8p7p6p5p4p3p2p1andbase <-
rbind(pedigreep9p8p7p6p5p4p3p2p1andbase,pedigreep10)

################################################################################
################################################################################
## see how many records per sire
try <- subset(basepop_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10_data,
basepop_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10_data[ ,1] > nbase)
nrow(try)
```

```
ones = c(rep(1,(nrow(try))))
try = data.matrix(try)
try = cbind(try,ones)
ham=sort(by( try[ ,14], try[ ,2], length))
head(ham)

##############################################
## Final data files
## pedigree file
data_anim_with_record <- subset(basepop_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10_data,
basepop_p1_p2_p3_p4_p5_p6_p7_p8_p9_p10_data[ ,1] > nbase)
summary(data_anim_with_record$FI)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#2.96    9.64   11.07   11.11   12.55   20.10
summary(data_anim_with_record$YWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#258.1   482.3   536.5   537.0   590.6   765.6
summary(data_anim_with_record$WWT)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#136.6   244.1   273.0   273.8   303.1   403.4
summary(data_anim_with_record$ADG)
# Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#0.340   1.200   1.360   1.381   1.540   2.610


##SI for each year's progeny
P1.si <- data.frame(id = simdataP1$id, si = -29.66932*simdataP1$FItbv +
21.99373*simdataP1$ADGtbv + 0.09652*simdataP1$WWTtbv + 2.55104*simdataP1$YWTtbv)
P2.si <- data.frame(id = simdataP2$id, si = -29.66932*simdataP2$FItbv +
21.99373*simdataP2$ADGtbv + 0.09652*simdataP2$WWTtbv + 2.55104*simdataP2$YWTtbv)
P3.si <- data.frame(id = simdataP3$id, si = -29.66932*simdataP3$FItbv +
21.99373*simdataP3$ADGtbv + 0.09652*simdataP3$WWTtbv + 2.55104*simdataP$YWTtbv)
P4.si <- data.frame(id = simdataP4$id, si = -29.66932*simdataP4$FItbv +
21.99373*simdataP4$ADGtbv + 0.09652*simdataP4$WWTtbv + 2.55104*simdataP4$YWTtbv)
P5.si <- data.frame(id = simdataP5$id, si = -29.66932*simdataP5$FItbv +
21.99373*simdataP5$ADGtbv + 0.09652*simdataP5$WWTtbv + 2.55104*simdataP5$YWTtbv)
P6.si <- data.frame(id = simdataP6$id, si = -29.66932*simdataP6$FItbv +
21.99373*simdataP6$ADGtbv + 0.09652*simdataP6$WWTtbv + 2.55104*simdataP6$YWTtbv)
P7.si <- data.frame(id = simdataP7$id, si = -29.66932*simdataP7$FItbv +
21.99373*simdataP7$ADGtbv + 0.09652*simdataP7$WWTtbv + 2.55104*simdataP7$YWTtbv)
P8.si <- data.frame(id = simdataP8$id, si = -29.66932*simdataP8$FItbv +
21.99373*simdataP8$ADGtbv + 0.09652*simdataP8$WWTtbv + 2.55104*simdataP8$YWTtbv)
P9.si <- data.frame(id = simdataP9$id, si = -29.66932*simdataP9$FItbv +
21.99373*simdataP9$ADGtbv + 0.09652*simdataP9$WWTtbv + 2.55104*simdataP9$YWTtbv)
```

```
P10.si <- data.frame(id = simdataP10$id, si = -29.66932*simdataP10$FItbv +
21.99373*simdataP10$ADGtbv + 0.09652*simdataP10$WWTtbv +
2.55104*simdataP10$YWTtbv)
```