THE GENETIC DISTANCE AMONG CATTLE BREEDS AS RELATED TO ANIMAL GENETIC RESOURCE CONSERVATION

BY

CHRISTIANE HANSEN

A Thesis
Submitted to the Faculty of Graduate Studies
in Partial Fulfillment of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

Department of Animal Science
University of Manitoba
Winnipeg, Manitoba

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION PAGE

THE GENETIC DISTANCE AMONG CATTLE BREEDS AS RELATED TO ANIMAL GENETIC RESOURCE CONSERVATION

BY

CHRISTIANE HANSEN

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirements of the degree

of

DOCTOR OF PHILOSOPHY

Suggested Short Title: GENETIC DISTANCE AMONG CATTLE BREEDS

THE GENETIC DISTANCE AMONG CATTLE BREEDS AS RELATED TO ANIMAL GENETIC RESOURCE CONSERVATION

ABSTRACT

Christiane Hansen Ph.D. Animal Science

Statistical and molecular genetic methodologies were used to determine the genetic variability within the Canadienne (CN), Holstein (HO), Jersey (JE) and Brown Swiss (BS) breeds of cattle and to determine the genetic distance among them. The results were intended to provide specific information to aid in the characterization and conservation of rare breeds of livestock in Canada.

An analysis of the pedigrees of each breed showed that the level of inbreeding in 1994 ranged from 0.68 to 4.8%. Sixty-one and 41% of animals in the CN and HO bull populations, respectively, were being sired by only ten bulls. The average effective population size in the CN breed, 86, and the number of annual registrations, 388, were sufficiently low to characterize the CN breed as endangered.

Genotyping 20 distantly related animals in each breed for 15 microsatellites and computing genetic distance estimates using Nei's standard genetic distance, the delta mu squared genetic distance and the Rst genetic distance formulae showed that the distance between individual breeds was significantly different from zero. Though not significant, CN cattle tended to be most closely related to HO and BS cattle, whereas the genetic distance

between BS and HO cattle tended to be smallest. In contrast, the JE breed tended to be the most genetically distant among the breeds examined.

An examination of the degree of sequence divergence in the D-loop region of mitochondrial DNA in 49 animals and a phylogenetic analysis, using a subset of the data, showed that the four breeds were not evolutionarily distinct. All four grouped together when a strict consensus tree was generated. Intra-breed variability in the D-loop sequence proved to be high for all breeds except the BS. The CN and BS as well as the BS and HO breeds showed the lowest degree of inter-breed variability. The variability seen between CN and JE cattle tended to be greatest among the four breeds.

ACKNOWLEDGEMENTS

I would like to acknowledge the following individuals for the substantial contribution they made toward the completion of these experiments:

my advisory committee that included Dr. R.J. Parker, Dr. J.N.B. Shrestha, Dr. G.H. Crow and Dr. P.J. McAlpine for their advice, enormous help and encouragement,

Dr. J. Derr for providing mitochondrial DNA D-loop primers, automatic sequencing services, several microsatellite primers as well as helpful advice,

Dr. C. Beattie for providing microsatellite primers for use in the experiments,

Dr. B. Benkel for his advice and for allowing me to use his laboratory,

Dr. A.J. Hackett and Mrs. Lorraine Robinson for their assistance during the collection of the blood samples and,

Dr. D. Petitclerc for providing DNA samples of Holstein cattle for use in the experiments,

Mr. J.-G. Bernier of the Société des Éleveurs de Bovins Canadiens for his assistance and for providing me with background information on the Canadienne breed,

all of the individual dairy farmers who generously allowed me to collect blood samples from animals in their herd.

Finally, I would like to gratefully acknowledge the support of my parents, Karl-Friedrich and Elke Hansen, who have always been behind me and who have encouraged me to pursue my goals.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF TABLES (CONT'D)

LIST OF ABBREVIATIONS

ATP	Adenosine triphosphate
ATPase	Adenosine triphosphatase
bp	Base pair
BS	Brown Swiss
BU	Butana
CLRC	Canadian Livestock Records Corporation
CN	Canadienne
dATP	2'-Deoxyadenosine 5'-triphosphate
dCTP	2'-Deoxycytidine 5'-triphosphate
dGTP	2'-Deoxyguanosine 5'-triphosphate
D-loop	Displacement loop
dNTPs	2'-Deoxynucleotide 5'-triphosphates
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
ddTTP	2',3'-Dideoxythymidine 5'-triphosphate
dTTP	2'-Deoxythymidine 5'-triphosphate
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
FAO	Food and Agriculture Organization
FR	Friesian
HA	Hariana
HE	Hereford
HO	Holstein
H strand	Heavy strand of mitochondrial DNA
HWE	Hardy-Weinberg equilibrium
JE	Jersey
kb	Kilobase
LINE	Long interspersed repeat element
L strand	Light strand of mitochondrial DNA
mtDNA	Mitochondrial deoxyribonucleic acid
mRNA	Messenger ribonucleic acid
NADH	Nicotinamide adenine dinucleotide, reduced form
ND	N'Dama
PAUP	Phylogenetic Analysis Using Parsimony
PCR	Polymerase chain reaction
QTL	Quantitative trait locus
RAM	Random access memory
RBC	Red blood cell
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid

LIST OF ABBREVIATIONS (CONT'D)

rRNA	Ribosomal ribonucleic acid
SAS	Statistical Analysis System
SDS	Sodium dodecylsulphate
SINE	Sinle interspersed repeat element
TBE	Tris-borate-ethylenediaminetetraacetic acid buffer
TE	Tris-ethylenediaminetetraacetic acid buffer
TP	Tharparkar
tRNA	Transfer ribonucleic acid
WBC	White blood cell

LIST OF APPENDIX FIGURES

LIST OF APPENDIX TABLES

LIST OF APPENDIX TABLES (CONT'D)

1. GENERAL INTRODUCTION

Cattle have been a species of significant economic and cultural importance to human societies for thousands of years. It has been estimated that approximately 780 different breeds of cattle exist world-wide (FAO, 1993). Many of these breeds are now threatened with extinction as a result of modern agricultural practices, human population pressures, the demands of the global market as well as by inadequate management and regulation of these important resources. This has led to growing concern about the narrowing of the genetic base and has resulted in the beginnings of a global initiative to study and combat this problem. The conservation of domestic animal genetic resources has thus gained increasing prominence as an issue on the national and international stage over the last several years.

The actual implementation of a sustainable program for the conservation of animal genetic resources requires a wide variety of approaches and technologies. One of the most basic and vital components of such a program on the national level involves the identification and characterization of all of the genetic resources found in a country. This process will allow national priorities to be set and informed decisions to be made concerning the allocation of financial resources. In Canada, the initial cataloguing of livestock and poultry resources is well underway. Steps must now be taken to characterize each breed on a more fundamental level so as to establish their uniqueness. Several approaches exist to tackle this problem

including pedigree analysis, the use of microsatellite markers and the analysis of mitochondrial DNA (mtDNA) sequence variation.

The establishment of herdbooks for the various breeds found in Canada and the existence of very specific registration criteria has ensured the maintenance of breed identity and, in many respects, breed purity. Through the use of pedigree records it is possible to trace the ancestry of any registered animal and know with certainty its family background and history. Analysis of such pedigree information, therefore, allows one to reliably choose unrelated or distantly related animals for studies in which a fairly complete spectrum of existing genotypes is desired. In addition, pedigree analysis allows one to gain an understanding of the rate and amount of inbreeding in a population as well as to gather important information related to population size and structure.

Inbreeding, which is the result of the mating of related individuals, has been shown repeatedly to have negative consequences on many performance and fitness traits (Hudson and Van Vleck, 1984; Hermas et al., 1987; Casanova et al., 1992; Miglior et al., 1995a,b). In general, increased homozygosity, redistribution of genetic variances, and a greater probability of the expression of lethal recessive genes have all been associated with inbreeding (Falconer, 1989). Increases in inbreeding as well as reductions in the genetic base of a breed or species can thus have devastating results in the long term and must be avoided if production levels are to be maintained and changing market

demands are to be addressed.

The Canadian dairy cattle industry, like those in many other parts of the world, is heavily dependent upon artificial insemination (A.I.) for breeding. In addition, the practice of embryo transfer is becoming more and more common and other newer reproductive technologies may soon see increased usage. This is especially true for the more dominant breeds of cattle, including the Holstein and Jersey breeds. Less popular breeds such as the Canadienne and Brown Swiss, however, also make significant and ever increasing use of these breeding and reproductive practices. The effects of such usage may be more pronounced in these breeds due to their limited genetic base. It has been shown that an important consequence of the widespread use of A.I. is that a single bull can service hundreds of thousands of females in its lifetime. In addition, only a few outstanding bulls will produce the majority of offspring in the breed population and often sons sired by these bulls become candidates for progeny testing and will end up entering the various A.I. studs themselves. The relationship among animals in the breed and among bulls in the A.I. studs will thus tend to increase over time, especially when this problem is not addressed effectively. Increases in the use of embryo transfer or embryo splitting will lead to an increase in the same phenomenon, although to a lesser degree, on the female side of the breeding equation. The size of the genetic base and inbreeding in cattle populations are thus of concern to many cattle geneticists.

In the past, blood protein polymorphisms were widely used to study the evolutionary relationships and genetic distance among breeds (eg. Namikawa et al., 1984; Graml et al., 1986; Gonzalez et al., 1988; Medjugorac et al., 1994). In recent years, however, the discovery and extensive mapping of microsatellite markers in the various animal genomes and the availability of the polymerase chain reaction (PCR) with its many applications has led to a change in the methods used for such characterizations.

Microsatellite markers have been found in the genomes of all known eukaryotes (Tautz, 1989; Weber and May, 1989; Quellar et al., 1993). They are simple sequence repeats, usually dimers or trimers, that are distributed randomly throughout the genome, have high mutation rates and are subject to replication slippage (Weber and May, 1989). It is still unresolved whether allelic distributions for microsatellites follow the infinite alleles model of mutation (Kimura and Crow, 1964) which speculates that any new mutations that occur are always different from the existing alleles in the population or the stepwise mutation model (Kimura and Ohta, 1978) which maintains that alleles of the same length but of differing descent may be produced by mutation. Their high degree of polymorphism and the ease with which they can be characterized using the PCR and polyacrylamide gel electrophoresis, however, lend them to use in genetic distance studies and thus to playing an important role in the conservation of genetic resources.

Only a limited number of studies have so far been carried out to

catalogue the microsatellite polymorphisms that exist between and within animal populations (eg. Buchanan et al., 1994; MacHugh et al., 1994). The Food and Agriculture Organization (FAO) has recently, however, begun trying to coordinate a global effort to do just that. Canadienne cattle, being a uniquely Canadian breed of cattle and given their current struggle for survival, are prime candidates for examination using this approach.

Finally, mtDNA contains a specific region called the displacement loop (D-loop), which is approximately 900 base pairs in length in cattle. This region has been shown to contain a significant amount of variation, including an area that appears to function as a molecular clock. It also contains an area known to act as the control region of mtDNA. By amplifying the D-loop region using specific flanking primers, sequencing the amplified portion and determining the sequence variations that exist, it is possible to estimate the genetic distance among breeds of animals and often to draw conclusions about their evolutionary relationship.

The studies described in the following pages utilize statistical and molecular genetic methodologies to measure the genetic distance among Canadienne, Holstein, Jersey and Brown Swiss cattle and examine the genetic variation within each of these breeds. Three different approaches were used to accomplish this: 1) pedigree analysis, 2) the use of hypervariable microsatellite markers from the nuclear genome and 3) the detection of sequence variations in the mitochondrial genome.

2. REVIEW OF THE LITERATURE

2.0 Genetic Resource Conservation

International and national concern for the world's biodiversity has been growing for several decades, including an increased concern for the diversity of our livestock populations. According to the FAO (1993), there are approximately 30 million species of plants and animals on this planet. Of these approximately 30 are used for agricultural purposes. Among these species a few, such as cattle and swine, dominate in their usefulness. Cattle, for example, produce 90% of the world's milk and 32% of the world's meat while pigs supply an additional 40% of global meat production (Cunningham, 1996). Other important species from a global perspective include buffalo, sheep, goats, horses, asses, camels and poultry. These species have been subjected to the influences of human selection and breeding practices for at least 10 000 years and as a result many distinct breeds have been formed within each animal species. We, therefore, have in excess of 4000 breeds at our disposal today for agricultural purposes, breeds that in many cases are uniquely adapted to their respective environment and to the community that they serve.

Biological diversity can be defined as the range of differences that exist between species, between breeds within a species and between individuals within a breed. For domestic animal conservation purposes, the type of

diversity we need to concern ourselves with is the range of genetic differences within each breed and across all breeds within that species.

The loss of animals to extinction, in other words the loss of animal biodiversity, has its roots in many of the man-made and naturally occurring phenomena we see in the world every day. Floods and droughts can cause animal breeds to be lost over a very short period of time. Wars can have the same impact on breeds, especially those that are already experiencing a decline in numbers. The most prevalent cause of loss of livestock and poultry breeds, however, remains the agricultural practices employed and, in many cases, the lack of understanding of the consequences of these practices. The breeds of livestock used commercially in most developed countries have been subjected to years of selective breeding and now generate truly remarkable production figures. In order to generate this production, however, they require a very high level of input and care. Such husbandry practices are often impossible to achieve in many developing countries. Producers in these countries, however, see the production records of these particular breeds and dream of raising such animals themselves. This leads to the importation of foreign breeds into these countries and often to indiscriminate crossbreeding with local breeds. Local breeds will frequently be lost completely during this process before some of their valuable characteristics can be recognized and exploited. Furthermore, these foreign breeds and their crosses often do not generate the level of production that was anticipated

because, as was mentioned earlier, the care that the average farmer can provide them with is not sufficient. Local breeds, despite their perceived inferiority, are, therefore, sometimes much better suited to local climatic and husbandry conditions. Likewise, the use of only a small number of high producing breeds, in any part of the world, leads to a loss of breeds that may carry traits for production characteristics or disease resistance not currently required but which may prove useful as market demands and requirements change. Breed loss before adequate assessments have been carried out can thus have disastrous consequences.

Many articles have been written on the subject of domestic animal conservation. These articles have listed and discussed the main arguments that justify why often economically inferior genetic material should be conserved. A review by Simon (1984) summarizes the five major arguments that have been put forth. The first of these arguments, based on Robertson's (1960) concept of selection limits, is that conservation is necessary for a maintenance of reserve populations that will make it possible to overcome possible selection limits within the present breeding populations and within the present environment. It is important to remember that domestic animal diversity, once lost, can't be replaced. The animals we have today are composed of approximately 50 000 to 100 000 genes which interact with each other and with the environment. Not only will a loss of this diversity hinder the ability of the animal breeder to change or improve the breeds and species

that we have at our disposal but furthermore, the technology we would need to artificially recreate the varied array of genotypes supplied by currently existing animal populations does not exist at the present time and may not exist for quite some time to come (FAO, 1993). Domestic animal diversity is thus also needed, as was alluded to earlier, in order to maintain the genetic variability required by animal breeders to meet future market or environmental requirements. For example, Simon (1984) lists possible changes in the nutritional requirements of man, new varieties of feed for animals that may need to be utilized, environmental changes and new types of diseases as possible realities that will have to be addressed. Plant breeders have already demonstrated how the use of genetic diversity from sources other than currently available commercial lines can benefit commercial agriculture and improve important crops. For example, genes that convey resistance to rust fungus have been introduced into commercial wheat cultivars from various *Triticum, Aegilops and Agropyron* species (Frankham, 1994). A third argument in favour of conservation is the need for a maintenance of genetic variability to enable animal production under less favourable conditions. This is especially important since much of the world, as was mentioned previously, is not capable of utilizing the intensive production systems prevalent in many developed countries and animals suited to these systems are often not superior, in terms of overall productivity, under extensive or low input conditions. Another important

reason for genetic resource conservation is that these resources are needed to improve our understanding of such subjects as evolution, domestication and the effects of natural and artificial selection. Without them scientists may never fully understand the processes that have shaped life as we know it. Finally, livestock biodiversity is an important component of our heritage. It has evolved over the past 10 000 years and as such should be nurtured and preserved for future generations.

Conservation of domestic animal genetic resources can be approached in a number of ways. The most obvious is through the maintenance of live populations. A second form of conservation involves the cryopreservation of semen, embryos and/or stem cells and DNA. Finally, another method that could be employed is the use of gene pools with a very broad genetic base and no artificial selection. All of these methods have advantages and disadvantages. For example, live populations are preferable from a purely emotional point of view. Seeing live animals in fields and on farms appeals to many individuals and is probably most effective in inspiring people to support their conservation. Such live populations are costly to maintain, however, and susceptible to hazards such as genetic bottlenecks, inbreeding depression, accumulating genetic drift as well as fires, natural disasters and disease (Smith, 1984). Frozen material, on the other hand, is much less costly to maintain and not as susceptible to some of the hazards that may afflict live populations. Several replicates of such material need to be stored, however,

to avoid accidental loss and regeneration of live animals from some types of material can be a problem. In addition, frozen genetic material may re-introduce diseases that were eradicated from a country or population after the material was frozen, but for which sufficiently sensitive tests did not exist at the time the sample was screened and stored.

Before one can begin to conserve genetic resources in any of the ways just mentioned, the status of all of the animal populations found in a region or country must be accurately evaluated. In order to accomplish this, a census of livestock populations must be carried out and trends within each population must be evaluated. It must then be decided whether a given population merits conservatory action. Numerous studies have been carried out and guidelines have been developed over the years to aid in this decision-making process. Maijala et al. (1984), for example, have established rough estimates for the size a livestock population should have in a developed country to be considered endangered. For cattle this population size is 1000 to 5000 animals and declining, with less than 1000 breeding females and less than 20 breeding males. In addition, the American Livestock Breeds Conservancy, a very active conservancy group in the United States, has established various categories to classify the status of a breed based on registration numbers. The three categories used by this group are critical, rare and watch. To be considered critical a breed must have less than 200 registrations annually in North America and an estimated global population

of less than 2000. Breeds categorized as rare must have less than 1000 registrations annually in North America and an estimated global population of less than 5000. Those breeds categorized as being under watch must have less than 2500 registrations annually in North America and an estimated global population of less than 10,000. Breeds that have been showing a substantial decline in numbers over the past 20 years are also included in the watch category (Bixby et al., 1994). The FAO has also published guidelines for breeds at risk. These are probably the ones most likely to be used in the future due to the FAO's efforts to harmonize the procedures being employed in different parts of the world, although each of the other methods or guidelines remain valuable in their own right as they provide a framework that will aid in decision-making. FAO (1993) categorizes breeds as critical, endangered or extinct. A breed is in the critical category if there are less than 100 breeding females or a total of 5 or less breeding males in the breed. A breed may also be considered critical if the total population size is just slightly greater than 100 and decreasing and less than 80% of females are being bred pure. A breed is classified as endangered if there are between 100 and 1000 breeding females in the breed or 6 to 20 breeding males. In addition, a breed is endangered if the total population of that breed is slightly less than 100 but increasing and more than 80% of females are being bred pure or if the total population is slightly more than 1000 and decreasing and less than 80% of females are being bred pure. Finally, a breed is extinct if one cannot recreate the original genetic base

of the breed and most of its genetic variation has been lost.

The conservation of animal genetic resources can be an expensive proposition if not approached with care. Decisions must be made throughout the process, each with important consequences. The type of preservation to use and the status of populations in the country, as were briefly discussed above, are only some of the choices that must be made and questions that must be addressed. Smith (1984) has done an excellent job of summarizing many of these choices. Briefly, an important decision one must make when conserving genetic resources is whether one wants to conserve individual genes or gene combinations. Gene combinations can be conserved most effectively in the form of pure lines, either live or as embryos. Individual genes, on the other hand, are best conserved in gene pools. Another important decision that must be made in a conservation program is which stocks are to be preserved. Since it is not economically feasible to store everything and since those stocks that might prove useful in the future are not self evident, selection must be based on certain criteria. One might, for example, select stocks on the basis of their genetic diversity since, quite logically, a higher level of diversity among the stocks would increase the probability of finding a gene or combination of genes that would allow one to meet some future requirement. This is where pedigree analysis and DNA analysis find their use and where such concepts as genetic distance, which will be discussed in more detail later, become important. Finally, sampling of

a given population is another aspect of a conservation program that must be given consideration. One must try to ensure that a representative sample of adequate size is preserved. The number of individuals sampled and the number of doses of semen or embryos stored will depend on what the stored material will eventually be used for and on the success that is routinely seen in utilizing this material. The use of pedigree analysis and molecular markers to identify unrelated animals can again be useful aids in the sampling process.

2.1 Animal Breeding and Inbreeding

The field of animal breeding has, over the years, made use of numerous genetic techniques in its quest for animal improvement. Inbred lines, which lend themselves to a wide variety of genetic studies and which are used extensively in many areas of genetics, have, however, had only limited use in practical breeding programs for large animals. Although inbreeding combined with selection were used historically to establish the foundation populations used by future generations of animal breeders to create the stocks we have today (Young, 1984), inbreeding is generally something to be avoided in large animal breeding.

Inbreeding can be defined quite simply as matings between related individuals. According to Wright (1931), the general formula for estimating the increase in inbreeding in a population of finite size due to the chance

mating of related individuals is $\frac{1}{8Nm} + \frac{1}{8Nf}$ where Nm and Nf are the

effective numbers of males and females. Since Nf is usually much larger than Nm, it is often ignored and the increase in inbreeding is thus approximated by $\frac{1}{8Nm}$. It is important to note, therefore, that it is the effective number of animals in a population, especially males, that is of concern when discussing inbreeding not the total number of males or females and this must be kept in mind when analyzing a population for risk of loss.

The effects of inbreeding are summarized by Falconer (1989) as being increased homozygosity, redistribution of genetic variances, a reduction in the performance of inbred individuals, especially in the areas of fertility, reproduction and health, and a greater probability of the expression of lethal recessive genes. These effects are well documented in the literature. Recently, Miglior et al. (1995b) showed that the inbreeding depression for milk, fat, and protein production and fat and protein percentages was -25, -0.9 and -0.8 kg and 0.05% and 0.05% per 1% increase in inbreeding, respectively, in Holstein cattle. Likewise, Casanova et al. (1992) showed that there was a decrease of 26 kg for every percentage increase in inbreeding in the phenotypic performance for milk yield in Swiss Braunvieh cattle. In addition, Hudson and Van Vleck (1984) showed that the regressions of milk, fat, stayability and

calving interval on inbreeding coefficients in Ayrshire cattle were -23 kg, -1 kg, -0.008 pts., and -0.095 days, respectively, per 1% increase in inbreeding. The effect of mild inbreeding on reproductive performance in first lactation Guernsey cattle was demonstrated by Hermas et al. (1987) who showed that an increase of 1% in inbreeding increased the number of services by 0.17, decreased the conception rate by 3.3%, increased the service period by 1 day and increased days open by 2 days. Miglior et al. (1995a) meanwhile provided evidence that inbreeding is related to disease prevalence in dairy populations by demonstrating that, on average, inbred animals tended to have higher lactation somatic cell scores.

Over the last 25 or 30 years, artificial insemination (A.I.) has had an ever increasing impact on dairy populations. The widespread use of genetically superior A.I. sires has enabled breeders to achieve very rapid genetic progress over the short term. More recently, the use of embryo transfer has also increased substantially and in the future such technologies as embryo splitting and in vitro embryo maturation and production may become widespread. All of these technologies have led, and will continue to lead, to an increase in the number of genetically identical animals, full sibs and three-quarter sibs in the dairy cattle population (Miglior et al., 1995b). There is today, therefore, a concern that the use of only a few very outstanding sires will eventually narrow the genetic base of the different cattle populations and one will start to observe more and more of the detrimental effects of

inbreeding over the long term. Such concerns are, in theory, not unfounded. After all, if all A.I. stations use only specific bulls as sires of sons, then all genetically superior bulls will eventually be related to each other (Young, 1984). From studies that have been carried out on some of the dairy cattle populations, one will tend to conclude that the problem may be increasing in significance. For example, studies done by Lush et al. (1936), Young (1973) and Bonczek and Young (1980) indicate that inbreeding in the U.S. Holstein population increased gradually up to 1931. After 1931 and until 1976, which was the limit of the data in these studies, inbreeding did not appear to change dramatically, however. While the average relationship of these animals was greater in the 1970s than in earlier years, it was concluded by Young (1984) that this relationship was still too small and increasing at too slow a rate to present serious problems before the end of the century. A similar conclusion was drawn by Hudson and Van Vleck (1984) from a study of artificially bred dairy cattle in the Northeastern United States. By the year 1988, Young et al. were, however, starting to sound a little less optimistic. In their study involving registered Holsteins in the United States, they found that while inbreeding per se seemed to be remaining fairly steady at 4.7%, 3.8% and 4.3% for 1970, 1976 and 1982, respectively, inter se relationship increased from 3.4% in 1931 to 9.8% in 1982. These authors concluded that this increase in relationship without an accompanying increase in inbreeding indicated a change in mating strategies over time. By 1982, the majority of breeders

seemed to be mating animals that were less related than if mating was random. Earlier, there had appeared to be more of a tendency for breeders to mate animals within the same line. Young and colleagues thus suggested that the Holstein breed needed to be monitored regularly because they felt it was only a matter of time before the increase in inter se relationship led to an increase in inbreeding. Finally, a study done on Canadian Holstein cattle by Miglior and Burnside (1995) showed that inbreeding in this population has steadily increased at a rate of 0.07% and 0.21% annually for cows and bulls, respectively, for the period from 1987 to 1990. They thus suggest that new breeding strategies to control inbreeding are needed for the long-term.

2.2 Canadienne Cattle

Canadienne cattle are one of the only uniquely Canadian breeds of cattle that currently exist in this country. Their history is fairly extensive for a North American breed. In fact, the Canadienne has been suggested to be one of the oldest breed of cattle in the western hemisphere (St-Pierre, 1936). Canadienne cattle can trace their origins back to the very first settlements that sprang up in New France in the 17th century. Their ancestors came mainly from the Normandy and Brittany regions of France and it is widely believed that they, therefore, stem from the same general lineage as the Jersey, Guernsey and Kerry cattle (Fortin, 1940). Over the years, their relative isolation in what is today the province of Québec and the effects of natural

selection in a harsh environment resulted in the development of an extremely hardy, self-sufficient breed of cattle. They were extremely well suited to pioneer life, requiring very little care and being adept at foraging for feed (Grisdale, 1909). They provided early settlers with the power they required for draught and produced both milk and meat for their consumption. In fact, by the year 1901 the Canadienne had evolved to such an extent that it was recognized as the most economical milk producing breed at the Pan-American show in Buffalo, N.Y. and had been widely shown to be the most profitable dairy breed in Canada (Fortin, 1940).

Canadienne cattle are small, compact dairy animals. Females generally weigh about 27 kg at birth and mature to weigh 450 to 550 kg. Males, meanwhile, are born weighing approximately 32 kg and reach 700 to 900 kg at maturity. These animals may be black, brown, tawny or russet in colour and generally display lighter tones along their dorsal line, around the muzzle and in the vicinity of the udder. Their skin is black and their udder is of good texture (Bernier, 1995).

In the year 1850, a census of the Québec cattle population showed that, with a few exceptions, all of the cattle found in the province were of the Canadienne breed (Couture, 1909). Unfortunately, however, it became fashionable in the subsequent years to crossbreed or replace the Canadienne with newly imported British breeds of cattle in the often mistaken belief that this would produce animals that would be more suitable and productive.

This was done with little consideration for local conditions and requirements. The average farmer at that time in Québec was not a wealthy or even educated person. Animals were thus not exposed to the improved nutritional and husbandry conditions that were being developed and increasingly applied in other parts of the world. Canadienne cattle were very small animals at that time, smaller than they are today, and their appearance was not as pleasing as that of many of the imported European cattle. As a result, "upgrading" to other breeds or crossbreeding was so widely encouraged by the influential people of the time, individuals who for various reasons were determined to denigrate the Canadienne breed, that in 1882 the Conseil d'Agriculture du Québec erroneously declared that the Canadienne breed had ceased to exist. In fact, a census performed in 1883 revealed that 75% of cattle in the province were still purebred Canadienne (Couture, 1909).

The popularity of the Canadienne among the general population was not to last, however, despite the efforts of some of its very dedicated supporters. The economic weight and influence of countries such as Great Britain and the U.S.A., the economic power of English Canada and the exclusion of the Canadienne from most English Canadian agricultural publications around the turn of the century led to the widespread use of breeds such as the Ayrshire and later the Holstein in Canadian agriculture. These breeds were further helped by the trend toward industrial type milk production that created an economic advantage for large, high producing

breeds and by the establishment of the milk quota and pricing system in Canada that, until 1992, tended to discriminate against breeds producing milk with a high solids content (Bernier, 1995). Furthermore, the Canadienne, being a local breed, did not have a herdbook until 1885. By that time many of the British breeds already had herdbooks and thus established pedigrees. This was used for many years to exclude the Canadienne from agricultural shows or to force them to compete in the "crossbred" category at these shows. Many Canadienne breeders were thus discouraged from competing and thereby exposing their cattle to wider audiences. In fact, even after the herdbook was established, advocates of the Canadienne breed had to fight for the chance to compete at agricultural exhibitions. The lack of use of A.I. in the Canadienne breed for many years also contributed to its decline by delaying genetic improvement through artificial selection in comparison to other breeds.

The Canadian government became involved in maintaining the Canadienne breed in the early 1900s. After the Société des Éleveurs de Bovins Canadiens was incorporated under the Animal Pedigree Act in 1905, authority for the herdbook was transferred to what is today Agriculture and Agri-Food Canada (Bernier, 1995). A herd was purchased by the Department of Agriculture and maintained at Ottawa's Central Experimental Farm until 1925. At that time, the herd was transferred to the Experimental Station at Cap Rouge, Québec, where it remained until 1940. During the following years, the Québec government supported the Canadienne breed by

establishing herds at Ste-Anne de la Pocatière, St-Hyacinthe and Deschambault. Through a rigorous selection and breeding program, first at St-Hyacinthe and from 1940 to 1970 at Deschambault, significant improvements were made in the breed. In 1969 however, due to increasing concerns about inbreeding in the Canadienne population, the Québec Ministry of Agriculture decided to crossbreed the Canadienne with Brown Swiss cattle in an effort to accelerate improvement of the breed (Bernier, 1995). Breeders of Canadienne cattle were encouraged to cooperate in this effort. While these crossbreeding efforts did much to increase milk production and improve various other production and conformation traits as well as to decrease inbreeding in the population, the crossbreeding led to a severe decline in the number of purebred Canadienne.

Today, the Canadienne remains a breed in considerable danger of becoming extinct. The Québec government gave up its herd at La Pocatière in 1979 and the herd at Deschambault was lost to a fire in 1983. The Société des Éleveurs de Bovins Canadiens has not received any government funding since 1993 and financial support for its goal of preserving and improving the Canadienne breed must, therefore, come from its members and from special fund raising events. The records indicate that the total number of Canadienne cattle that had been registered by December 31, 1995 was 65,547. There were only 209 registrations of Canadienne cattle in 1995 of which only 105 were purebred animals. All of the registrations except six (three pure and

three non-pure) were from the province of Québec. Of the registrations outside of Québec, the three pure and one of the non-pure animals were from Ontario and the remaining two non-pure registrations were from the United States (CLRC, 1995).

2.3 The Nuclear Genome and Microsatellites

Molecular technologies provide the tools for the extensive characterization of animal populations. Such characterization is not only important for the facilitation of animal breeding and improvement, but also plays an important role in the process of conservation of genetic diversity and has applications in terms of parentage verification of valuable stocks. The nuclear genomes of the various livestock species, including cattle, have thus come under ever increasing scrutiny over the years.

The organization of nuclear DNA has been studied extensively. It has been demonstrated that eukaryotic genes are what is referred to as interrupted genes. Briefly, they are composed of alternating exons and introns. Only the exons are commonly represented in RNA. The number and size of a given intron or exon varies between genes, however, and genes tend to be split, not dispersed. Furthermore, transcription of a given region of DNA can be quite complex and thus requires rather complex control systems. One gene's intron may be another gene's exon, for example (Lewin, 1990). In addition, alternate splicing of gene transcripts can allow different products to be produced by

single genes and sometimes genes are even found encoded within other genes (Watson et al., 1992). It follows, therefore, that the chromosomes of eukaryotes are fairly complex in terms of arrangement, more so than those of prokaryotes. They contain large areas of repetitive DNA. This repetitive DNA may make up as much as 20 to 50 percent of the eukaryotic genome. The first evidence of the existence of repetitive DNA came from density gradient studies of different types of DNA. Centrifugation of prokaryotic DNA, such as that of *E. coli,* to equilibrium in a CsCl density gradient will yield a single band of DNA in the gradient. Kit (1961) first showed that when mouse DNA is centrifuged to equilibrium in such a density gradient there is a small band of DNA in addition to the main band. This additional band of DNA is referred to as satellite DNA and one or more such bands are always found when eukaryotic DNA is analyzed in this manner. When the different satellite DNAs were isolated and analyzed, they were shown to contain repeating DNA sequences of varying length and composition. Waring and Britten (1966) showed, using $C_o t$ curve analysis, that mouse satellite DNA represents a 300-400 base pair (bp) sequence repeated approximately 10^6 times. $C_o t$ curve analysis is a type of analysis based on the fact that the fraction of the original single stranded DNA remaining in a renaturation reaction mixture at any given time is a function of the original concentration. The more repetitive a sequence is, therefore, the lower will be its $C_o t$ value. Several different types of DNA thus occur in all eukaryotic species, with the possible

exception of yeast. These include, in order of decreasing $C_o t$ value, low or non-repetitive sequences such as those that would encode for polypeptides, middle or moderately repetitive sequence such as those encoding for histones, transfer RNAs (tRNAs) and ribosomal RNAs (rRNAs) and highly repetitive sequences such minisatellite DNA, microsatellite DNA, short interspersed repeat elements (SINEs) and long interspersed repeat elements (LINEs). While the function of these repetitive sequences has yet to be determined, they have proved extremely useful for the study of animal genomes.

Microsatellite DNAs, the type of repetitive DNAs used in one of the following studies, are regions of nuclear DNA that contain very simple sequence repeats. These repeats may be tandem arrays of di-, tri- or tetranucleotides, although dinucleotide repeats are the most common. They have been shown to be extremely abundant in the genomes of all the complex organisms that have been studied, including livestock species (Rafalski and Tingey, 1993). Estimates have shown that microsatellites containing the $(GT)_n$ repetitive sequence, for example, occur approximately every 30 kb in euchromatic regions of the chromosomes. These same repeats appear to be under represented in the centric heterochromatin (Stallings et al., 1991). Portions of different microsatellite repeats have been shown to be well conserved between closely related species and to a lesser extent between more distantly related species. Microsatellite polymorphisms have proven to be

very useful in genetic characterization and microsatellite analysis using the PCR has become a popular technique used for linkage studies and physical mapping. Microsatellites have also been useful for the study of quantitative trait loci (QTL) in animals and this use promises to increase as genetic maps are improved.

The role of microsatellite DNA has not been satisfactorily established to date. Hamada et al. (1984) have suggested that some repetitive sequences may enhance the transcriptional activity of genes and Pardue et al. (1987) have demonstrated that there may be a functional role for these sequences in genetic recombination. Stallings et al. (1991), however, argue that the lack of extensive conservation of the $(GT)_n$ repetitive sequences that they studied in evolutionarily distant species suggests that they probably don't play a very precise role in gene regulation. In addition, they have shown that there does not appear to be a correlation between the number of $(GT)_n$ repeats and recombination activity. They, therefore, suggest that these microsatellites may be necessary but not sufficient for the occurrence of recombinational hotspots. Finally, Stallings et al. suggest that since bacteria have been shown to lack $(GT)_n$ repeats (Gross and Garrard, 1986), these sequences may play a role in DNA condensation and packaging in eukaryotic chromosomes. Much of what is known about the functional importance of microsatellites, therefore, remains speculative.

Microsatellite analysis, such as that used for the characterization of

animals and populations, involves the PCR amplification of the short genomic region of DNA containing the repeated sequence. Fluorescent or radioactively labelled primers are used in the PCR reaction to repeatedly amplify the region of DNA located between the two primers. The size of the repeated sequences is then estimated using high resolution gel electrophoresis of the labelled strands. This method of analysis is rapid, taking only one to two days to complete and does not depend on restriction enzymes for the detection of polymorphisms. Microsatellite analysis has the advantage of requiring only very small quantities of DNA (less than 100 ng). These types of markers show an extremely high degree of polymorphism in that the number of alleles one can observe in a population may be as much as 10 or more (Tuggle, 1994). They, therefore, lend themselves extremely well to genetic distance studies (eg. Buchanan et al., 1994; MacHugh et al., 1994), as will be described later, and have been targeted for use by the FAO as part of the global effort to characterize and conserve genetic diversity (Barker, 1994).

2.4 Mitochondrial DNA

Today, mtDNA has become one of the favourite molecules for the study of population genetics and the genetic divergence of organisms. The reasons for this are multiple. The mitochondrial genome of animals is fairly small and relatively uniform in size among vertebrates and invertebrates (Brown et al., 1979). Unlike nuclear DNA, whose sequence analysis is greatly

complicated by the fact that recombination rearranges the genes and that very few sequence differences accumulate even over many generations, mtDNA is primarily maternally inherited and does not generally recombine. It has also been shown to accumulate nucleotide substitutions extremely quickly, having an evolution rate that is approximately ten times that of nuclear DNA (Brown et al., 1979). This fast evolution and unique manner of inheritance thus provide a very interesting perspective on the evolutionary or phylogenic history of an organism (Cann et al., 1987).

Mitochondria in animal and plant cells are thought to have their origins in the microbial world. It is believed that the so called protomitochondria were oxygen-respiring, rod shaped bacteria probably related to the *Paracoccus denitrificans*, a type of purple bacteria we see today (Margulis, 1993). It is not known for sure how these organisms came to be inside other organisms that, in turn, went on to develop into eukaryotic cells. There are no known prokaryotes with phagocytotic capabilities of this sort. However, rRNA gene sequences have definitely shown that the mitochondrial genome differs from the nuclear genome of a cell and that it closely resembles that of eubacteria. The symbiotic relationship that developed between these organisms has thus proved to be lasting, and today mitochondria and their cellular "host" are intricately dependent on one another. Mitochondria do, however, retain some of the independent attributes they must have possessed as bacteria as evidenced by their ability to autonomously replicate their DNA.

Due to the small size of the mitochondrial genome it has been possible to sequence it in several different species. Much has thus come to be known about the organization of mtDNA. The human mitochondrial genome was one of the first mammalian mitochondrial genomes to be sequenced (Anderson et al., 1981). It was followed closely by that of the mouse (Bibb et al., 1981) and the bovine (Anderson et al., 1982). Since then the complete mtDNA sequences of many other animals including the rat (Gadaleta et al., 1989), fin whale (Arnason et al., 1991), harbor seal (Arnason and Johnsson, 1992) and the grey seal (Arnason et al., 1993) have been elucidated. In addition, partial sequences are available for animals such as the dolphin (Southern et al., 1988), pig (Mackay et al., 1993) and duck (Ramirez et al., 1993).

The mtDNA of all mammals consists of a closed circle of double stranded DNA that comprises approximately 16.5 kbp. The two strands of DNA are known as the heavy (H) and light (L) strands based on their density. The organization of the mammalian mitochondrial genome is extremely compact compared to that of the nuclear genome. The genes and their organization appear to have been widely conserved between species. There are few non-coding bases in mtDNA, except in the D-loop which will be discussed in more detail a little later, no introns and some genes have been shown to overlap (Anderson et al., 1982).

Like other mammalian mtDNAs that have been studied, bovine mtDNA contains genes coding for cytochrome b, three subunits of

cytochrome oxidase, two subunits of ATPase, and seven subunits of NADH dehydrogenase. In addition there are genes coding for 2 rRNA and 22 tRNAs. The majority of the genes are oriented in the same direction and in most cases tRNA genes are interspersed between genes coding for other proteins or rRNAs. These tRNA genes appear to act as cleavage sites during transcription. The D-loop contains the promoters for DNA transcription of both strands. Transcription of the H strand has been shown to proceed from a site in the D-loop bordering the tRNA located just before the 12S rRNA gene all the way to the D-loop again. The primary transcript is then cleaved on either side of the tRNA sites to give 12S and 16S rRNA, tRNAs and several mRNAs for the other proteins. These mRNAs are not capped but they are polyadenylated (Anderson et al., 1981). L strand transcription appears to proceed in a similar manner with the large primary transcript being cleaved and degraded some time after transcription to yield appropriate tRNA and some mRNA sequences (Lewin, 1990). The origin of replication is also contained within the D-loop and replication proceeds by H strand synthesis. Synthesis of the L strand does not occur until the L strand has been exposed by H stand synthesis (Anderson et al., 1981).

The D-loop of mtDNA is the main regulatory region of the genome since, as mentioned earlier, it contains the origin of H strand replication as well as promoters for the transcription of both the H and the L strands. It is also the most variable and rapidly evolving part of the mitochondrial

genome and contains a region that has been shown to behave as a molecular clock (Saccone et al., 1991). The D-loop can be divided into three domains. The left (L) and right (R) domains are low in guanine content and contain the 5′ and 3′ D-loop ends, respectively (Saccone et al., 1987). They vary greatly in both sequence and length, but do contain conserved sequence blocks. The R domain contains the two promoters and the origin of replication making it the most important functional part of the regulatory region. The promoters have been thoroughly studied in the human and the mouse and it has been shown that their location with respect to other elements is maintained (Saccone et al., 1991). The L domain contains two conserved palindromic sequences that may be involved as a recognition site for the termination of H strand synthesis. The central domain, located between the L and R domains, is an area that is higher in guanine content. It is well preserved in mammals and contains an open reading frame of variable length. This is the region that behaves as a good molecular clock and that can thus be used to obtain estimates of the times of divergence of different species (Saccone et al., 1991).

In view of the autonomous nature of mtDNA replication and its physical isolation from nuclear DNA in eukaryotic cells, it is not really surprising that its inheritance differs from that of the nuclear genome. Early studies in non-mammalian species showed that mtDNA is maternally inherited. That is, the mitochondrial genome of an individual organism, detectable by the techniques available at the time, is derived solely from its

mother. Hutchinson et al. (1974) were the first to examine this same question in mammals. Using two common equine species, the horse and the donkey, and the progeny of reciprocal crosses between the two (mule and hinny), these researchers were able to show that mtDNA in mammals also appeared to be maternally inherited. They purified mtDNA from the liver and heart of individual animals, digested it with a restriction endonuclease (HaeIII) and analyzed the cleavage pattern using polyacrylamide gel electrophoresis. The DNA banding patterns on the different gels obtained were clearly distinguishable. The horse and donkey each produced a unique banding pattern and analysis of the mule and hinny gels showed that the banding displayed resembled that of the respective maternal parent. Furthermore, it was observed that the liver of a hinny female examined contained predominantly horse glucose 6 phosphate dehydrogenase but only donkey mtDNA. The gene for glucose 6 phosphate dehydrogenase is found on the X chromosome. Non-random inactivation of one of the X chromosomes is known to occur in hybrids and, in this case, the maternal X chromosome was inactivated. The same inactivation model thus could not be used to explain the observed maternal inheritance of mtDNA. The researchers stated that the sensitivity of their detection method was such that as little as 5% paternal DNA would have been detectable. Considering the fact that both the sperm and the egg contain mitochondria, although in vastly differing quantities, they proposed two mechanisms for the maternal inheritance they had

demonstrated: 1) an inability of the sperm mitochondria to replicated during embryo development after fertilization or 2) a dilution effect where the vast excess of maternal or egg mitochondria overwhelms the few paternal or sperm mitochondria in a fertilized ovum. They further postulated that paternal mitochondria may be partitioned into cells during blastocyst formation. If this were to occur before the start of mitochondrial replication, cells derived from originals lacking in paternal mitochondria would thus all be free of these organelles. Hutchinson and coworkers results were subsequently confirmed by several other individuals for higher eukaryotes (Hayashi et al., 1978; Giles et al., 1980). A definite mechanism of inheritance could not be determined, however.

In 1979 Brown et al. published a report detailing the fact that the rate of evolution of mtDNA was approximately ten times faster than that of the single copy fraction of the nuclear genome. They showed this by constructing a restriction endonuclease cleavage map of the mtDNA of various higher primates and determining the degree of nucleotide sequence divergence between the species. Plotting the degree of divergence against time then allowed them to calculate the rate of base substitution. Any mechanism of mtDNA inheritance proposed thus has to account not only for the seemingly maternal nature of its inheritance but also for this fast rate of evolution. Many different proposals have been made, but, as was suggested by Koehler et al. (1991), originally a mutation in the mtDNA would have had to occur in a

homoplasmic individual. This sequence variant in the now heterplasmic individual must then segregate in such a way as to become the only mtDNA in some future generation. That such segregation is extremely rapid was clearly shown by Laipis et al. (1988) and Ashley et al. (1989) using heteroplasmic Holstein cattle. In fact, it was shown in the latter paper that mitochondrial genotypes may segregate in opposite directions in maternal siblings and that homoplasmy may be re-achieved in as little as two generations. Koehler et al. (1991) have even shown that a particular sequence variant can be fixed within a single generation without there being a detectable heterplasmic intermediate. It has been suggested that there may thus be a differential amplification of one or a few germ-line mtDNA molecules to produce the mitochondrial genome of the next generation (Koehler et al., 1991).

In 1991, Gyllensten and colleagues reported findings that shed some new light on the possible manner of mitochondrial inheritance alluded to by Hutchinson and others. Having previously shown by less sensitive methods that on average an individual inherits less than one paternal mtDNA molecule in a thousand, Gyllensten et al. used the polymerase chain reaction to demonstrate that, in mice, sperm contribute mitochondria to the fertilized egg. They detected paternally inherited DNA molecules at a frequency of approximately 10^{-4}. While this number is extremely small, it does, as the authors suggest, provide a means of generating heteroplasmy and must be

considered in phylogenic studies since the effective population size for a species is consequently increased.

The use of mtDNA in the study of populations and their evolution has been encouraged by many of the factors mentioned earlier. As was stated by Cann et al. (1987), the variation present in a given gene pool is magnified in mtDNA due to the rapid rate of accumulation of mutations found there. In addition, because of the largely maternal haploid nature of inheritance of mtDNA and the general lack of recombination of mitochondrial genes in mammals, it is much easier to relate individuals to one another than it would be using nuclear DNA. Furthermore, its mode of inheritance makes mtDNA much more sensitive to drastic decreases in population size than its nuclear counterpart.

Using the results of mapping with various restriction enzymes, it has thus been possible to deduce many interesting things about different populations. In humans, for example, such studies have allowed researchers to produce genealogical trees relating nucleotide sequences in the D-loop for different human DNAs. Cann et al. (1987), using the mtDNA from people in five different geographic populations were able to show that all of the populations except the one from Africa had multiple origins and that all of the mtDNAs probably stem from one woman who lived in Africa approximately 200,000 years ago. It has since been suggested that this figure may need to be adjusted in view of the demonstration of some paternal

transmission of mtDNA. The existence of some paternal inheritance may mean that this figure is slightly exaggerated.

Further insight into the evolutionary history of populations is given by studies such as that of Di Rienzo and Wilson (1991). These researchers have used sequence analysis of the most variable part of the control region of the mtDNA of 117 individuals, mainly from Sardinia and the Middle East, to show that 88 different types could be identified in these people with differences between the types being found at 79 sites. They also showed that 12 of the identified types were shared by more than one individual. The geographic pattern of this sharing is consistent with the theory that Sardinians originated, by and large, from migrations of individuals from southern Europe that occurred 9000-6000 years ago and that Mediterranean and Middle Eastern populations stem from an ancestral population that lived in prehistoric times. Furthermore, the results of Di Rienzo and Wilson's research shows that the genealogical tree relating the different mitochondrial types exhibits a non-random pattern of branching. The majority of branching in the tree occurs approximately two-thirds of the way between the root and the tips of the tree in a narrow interval of sequence divergence. The distribution of pairwise sequence differences was, moreover, shown to be unimodal and resemble that of a Poisson distribution in these non-African groups, unlike the distribution seen for African populations which shows two or more modes. It is suggested that these results could thus reflect a

change in the probability of survival of mtDNA lineages due to a burst of population growth. This burst may be connected with the spread of an ancestral group of Africans to populate the rest of the world as was suggested by the work of Cann et al. described earlier.

A study done by Cunningham et al. (1994) also demonstrates the use of mtDNA analysis in the study of population evolution. It even sheds doubt on the commonly held belief that all of today's cattle stem from domestication events that occurred approximately 10,000 years ago in the Middle East. Using RFLP analysis coupled with direct sequence comparisons, these researchers compared cattle breeds from different continents. They found that the restriction endonucleases used detected 20 polymorphisms in the cattle representing 16 mitochondrial lineages. Two major branches representing *Bos taurus* and *Bos indicus* populations of cattle were detected. Cattle mitochondrial lineages were found to differ at about 1.6% of the sites examined. These results were supported by the results obtained from the sequence analysis of the D-loop region. The authors state that the differences seen between the two major lineages suggest that *Bos taurus* and *Bos indicus* cattle were domesticated separately, having diverged between 200,000 and one million years ago. This result refutes the theory that contends that *Bos indicus* represent artificially selected *Bos taurus* (Epstein, 1971).

Many other examples of the use of mtDNA polymorphisms in restriction endonuclease cleavage pattern and mitochondrial sequence

analysis to investigate the relationship between different breeds of farm animals and, indeed, species exist in the literature (Watanabe et al., 1985a,b; Watanabe et. al, 1986; Watanabe et. al, 1989; Bhat et al., 1990; Lan and Shi, 1993; Wu et. al, 1994). The papers by Watanabe et al. (1985b and 1986) are of particular interest. These researchers studied pigs belonging to four commercial breeds, Taiwan native breeds and miniature strains found in Japan as well as Japanese wild boars by this method. They found that pigs stem from two maternal origins- European and Asian wild boars. Furthermore, they showed that Large White pigs must stem from a combination of European and Asian swine. In addition, they demonstrated through the comparison of sequence data of porcine mtDNA at 237 positions with data from cattle, mice, rats and humans that the number of differences between cattle and pigs appears to be similar to the number of differences between pigs and mice. Pigs also appear to be more similar to rodents than cattle. When the number of replacements and transversions were examined, the difference between mice and rats was shown to be the smallest for the combinations studied and pigs and cattle appeared to be the species pair with the second smallest number of differences.

2.4 Genetic Distance and Phylogenetic Analysis

Genetic variation between populations can be the result of a number of factors including natural and artificial selection, mutation, migration, genetic

drift and non-random mating (Hedrick, 1975). Unfortunately, however, these differences don't always occur at sites that are available for genetic distance analysis. The approaches used to estimate genetic distance must, therefore, provide a means of ordering breeds and populations according to their level of phylogenetic distinction when such specific data is not available (Barker, 1994). In fact, in recent years, population genetics has been undergoing a resurgence in popularity and importance due to the increasing availability of information on molecular differences among alleles that serves as a valuable aid in the study of genetic differences and phylogeny.

The presence of genetic variation and the extent to which it exists between populations is one of the important factors that must be considered when deciding which breeds of animals should be conserved in any national or international domestic animal genetic resource conservation program. This is true because, as was mentioned earlier, genetic variation is the basis for all future genetic improvement. As was suggested by Barker (1994), one way to ensure the maintenance of the maximum possible genetic diversity would be to conserve the sub-set of all breeds in a species that show the largest amount of genetic differentiation and those that contain unique alleles and/or allelic combinations. Since it is impossible to easily characterize a breed in such a way so as to include all possible genetic traits in a manageable manner, measurements such as genetic distance can be used to provide an estimate of the diversity between breeds. Such genetic distances can then

often be used to construct phylogenetic trees to aid in decision making.

Genetic distance can be defined as "a measure of gene diversity between populations expressed as a function of genotype frequency" (Hoelzel and Bancroft, 1992), although in many ways this is an oversimplification. Numerous methods of estimating genetic distance have been proposed over the years, each with advantages and disadvantages and some better suited to certain types of data than others (Nei, 1987; Nei and Takezaki, 1994). These methods rely on statistical approaches of varying degrees of complexity. Two of the major types of data now being used to estimate genetic distances between breeds are DNA sequence data and microsatellite DNA data. Since it is too time-consuming and expensive to sequence extensive stretches of DNA, relatively short, highly variable regions such as the mtDNA D-loop have found favour for studying population differences. Microsatellite DNA, on the other hand, has the advantage over traditional serological and immunological markers of being useful for distinguishing between populations with close historical connections. These markers have a greater number of segregating alleles and as a result the extent of absolute genetic diversity within and between populations for these loci is much larger (Kimmel et al., 1996).

The genetic distance from DNA sequence data can be calculated as the proportion of nucleotide differences or the p-distance. The general formula for this distance is $p= n_d/n$, where n is the total number of nucleotides

studied and n_d is the number of nucleotide differences observed. As was discussed by Nei and Takezaki (1994), however, this simple distance measure does not take into account multiple nucleotide substitutions at the same site nor does it account for the higher rate of transitional nucleotide substitutions as compared to transversional substitutions. Additional procedures taking many of these factors into account have thus been developed (eg. Kimura, 1980).

DNA sequence data is most frequently used for phylogenetic analysis. Such an analysis of sequence data requires positional homology. This requires that alignments of sequences be made by inserting gaps corresponding to postulated insertions or deletions. In most instances, multiple alignments must be made. One strategy that is thus often employed and that is used by the popular computer program Clustal (Higgins et al., 1992) is to make pairwise alignments and then to add the sequences together by inserting additional gaps as needed. The order of the pairwise alignments is obtained from clusters in the initial tree produced from a matrix of distances across all pairwise alignments (Hillis et al., 1996). In order to infer phylogenies from such character data, methods based on the principle of maximum parsimony are widely employed. These methods function by selecting trees that tend to minimize the number of evolutionary steps needed to explain the data set being studied. Maximum likelihood methods are also often used to infer phylogenies. In contrast to the parsimony

methods, these methods strive to estimate the actual amount of evolutionary change according to a specific evolutionary model. Finally, pairwise distance methods such as were referred to earlier may be used to study phylogenies. The corrected distances obtained through such methods reflect the mean number of changes per site that have occurred between pairs of sequences since their divergence from a common ancestor (Swofford et al., 1996).

The genetic distance from allele frequency data such as that obtained from microsatellite analysis can also be calculated using many different formulae. Nei's standard genetic distance (Nei, 1972) is one measure often employed in studies examining the evolutionary relationship and distance between breeds using microsatellite data. Nei's identity for two taxa has been defined as: $id = J_{xy} / \sqrt{J_x J_y}$ where J_{xy}, J_x and J_y are the arithmetic means of j_{xy}, j_x and j_y ,respectively, over all loci. The latter three quantities are defined as follows: $j_x = \sum x_i^2$, $j_y = \sum y_i^2$, the probability of identity of two randomly chosen genes at a single locus in two populations, X and Y, and $j_{xy} = \sum x_i y_i$, the probability of identity of a gene at the same locus for two populations X and Y. The variables x_i and y_i are the frequencies of the ith alleles for the two populations, respectively . Nei's standard distance is then calculated as $D = -\ln(id)$. This distance is based on an infinite alleles model and assumes that the rate of gene substitution per locus is the same for all loci. Another

genetic distance measure that has been developed recently is the delta mu genetic distance for microsatellites (Goldstein et al., 1995a). This distance is closely related to the D1 or average squared distance measure developed by Goldstein et al. (1995b) which is based on a theory for wandering distributions developed by Moran (1975). Moran's theory deals with the distribution of alleles mutating under a strict stepwise mutation model in a population of finite and constant size with non-overlapping generations. He has shown that such a distribution will have a constant variance and it has been demonstrated that the average square difference in the number of repeats between alleles in two population will have a linear expectation with time. Due to this latter fact, the D1 distance has been shown to be superior for use with microsatellite data. Unfortunately, however, it has a large variance thus making it useful for phylogenetic reconstruction only when large amounts of data are available and for taxa that have been separated for long periods of time. The delta mu squared distance or Ddm was thus developed to improve some of the deficiencies of the D1 distance. Ddm is defined as the squared mean difference between alleles of two populations. This distance is also linear with time and has a much smaller variance than D1. Ddm has the additional advantage of being independent of population size. Finally, a third measure that has recently been adapted for use with microsatellites is the Rst distance developed by Slatkin (1995). This distance measure is analogous to Wright's Fst (Wright, 1969) but assumes a high rate stepwise mutation model

instead of a low rate infinite allele mutation model. The mutation process at microsatellite loci is not believed to conform with the latter model. Mutation rates at these loci have been shown to exceed 10^{-3} per generation. In addition, there is ample evidence that the size of a new mutant allele at a microsatellite locus depends on the size of the allele that mutated (Weber and Wong, 1993). The infinite alleles model meanwhile assumes that the mutation process erases any memory of the prior allelic state. Slatkins Rst is defined as follows: Rst = (Sbar - Sw)/Sbar, where Sw is the sum over all loci of twice the weighted mean of the within population variances V(A) and V(B) and Sbar is the sum over all loci of twice the variance V(A + B) of the combined population. This distance measure is also closely related to D1.

3. STUDY I: INBREEDING AND THE GENETIC BASE OF CANADIENNE, HOLSTEIN, JERSEY AND BROWN SWISS CATTLE BREEDS

3.0 OBJECTIVE

The purpose of this study was to identify and select the least related individuals within Canadian populations of four cattle breeds, namely Canadienne (CN), Holstein (HO), Jersey (JE) and Brown Swiss (BS), for blood sampling and to examine the degree and rate of inbreeding and the current genetic base of these breeds.

3.1 MATERIALS AND METHODS

3.1.1 Pedigree Data

Pedigree files for CN, JE and BS cattle were obtained from Canadian Livestock Records Corporation (CLRC) in Ottawa, Ontario. The files for the CN and BS breeds were combined as a fairly large amount of "upgrading" has been done in the CN breed using BS cattle and many of today's CN animals therefore have BS ancestors. These pedigree files included all computerized information available at CLRC. The file for CN and BS cattle consisted of 31,621 records while the file for JE cattle consisted of 256,945 records. Pedigree information for the HO breed was obtained from Holstein Canada in Brantford, Ontario. Due to the extremely large size of the HO population in this country, only an excerpt of the total pedigree file was analyzed. This file included all of the bulls presently found in the studs at Eastern Breeders Inc.,

United Breeders, Western Ontario Breeders and the Centre d'Insemmination Artificielle du Québec. Pedigree records of these bulls were traced back for 4 generations. This resulted in a file consisting of 5,539 records.

3.1.2 Computer Analysis

Inbreeding coefficients for each animal and the overall rate of inbreeding for the CN and BS breeds were determined from the additive relationship matrices. These matrices were generated using the tabular method for calculating relationship (Cruden, 1949; Emik and Terrill, 1949; Van Vleck, 1993). The same technique was used to analyze the pedigree file of the HO bulls. Unfortunately, due to the size of the JE file and computational requirements in excess of available computer RAM and hard disk space, it proved impossible to use this technique to evaluate the JE population as a whole.

In order to generate the additive relationship matrix, the data obtained from the various locations was first edited and correctly formatted through the use of a computer program written using the Statistical Analysis System language (SAS, 1990; eg. Appendix Table 1). A series of Fortran programs was developed and then used in order to perform the actual calculations (eg. Appendix Tables 2-5). Only the half matrix was stored. Nevertheless, the huge amount of output generated by even the smallest breed population examined necessitated the running of the program through the submission of

only small blocks of data at a time. Additional SAS and Fortran programs were, therefore, required to make this possible (eg. Appendix Tables 6-9). Animals in the base population were assumed to be unrelated. Regression analysis, carried out using SAS (eg. Appendix Table 10), was used to estimate the rate of increase in inbreeding based on the results obtained from the diagonal of the additive relationship matrix for the CN, BS and HO populations studied. Specifically, each year the mean value of the inbreeding coefficient was regressed on year to estimate the rate of inbreeding.

Wright (1931) developed a general formula to estimate the expected rate of inbreeding in a population based on effective numbers of males and females. This formula can be written as $\frac{1}{8Nm} + \frac{1}{8Nf}$, where Nm and Nf are the effective number of males and females, respectively. This method is often used in genetic resource conservation programs to obtain a rough estimate of the rate of inbreeding that can be expected in a given population. This method was used to derive estimates of the expected rate of inbreeding for the CN, BS and JE breeds (eg. Appendix Table 11). To obtain an estimate of the effective number of breeding animals, it was possible to use only those records for which the sire and dam was known. Records with missing values had to be excluded. This analysis was, therefore, restricted to the years 1983 to 1994 for which the most complete pedigree records were available.

3.1.3 Selection of Individual Animals for Blood Sampling

The additive relationships that were generated between animals within the CN and BS breeds were used to select the 20-21 least related individuals in each breed for blood sampling (eg. Appendix Table 12). As travel times and collection costs had to be considered, selected animals also had to be located within a reasonable distance from Ottawa. The collection area thus comprised Eastern Ontario and much of Québec. An excerpt of the pedigree records of JE cattle found on local farms was used for generating an additive relationship matrix in order to select those animals from this breed found in the region that were suitable for blood sampling.

3.2 RESULTS

Tables 1-3 show the additive relationship between the 20-21 animals selected from the CN, BS and JE breeds for blood sampling. As is evident from these tables, on average the relationship among related animals was negligible. It should be noted that it was impossible to find a group of 20 animals that were totally unrelated given the constraints of breed size, location and collection cost. The animals chosen were the least related ones of the group available.

Figure 1 gives a graphical representation of the average inbreeding that was estimated for the years 1983 to 1994 in the three breeds studied (CN, BS

Table 1. The additive relationship matrix for all selected CN cattle.

Animal	1	2	3	4*	5*	6*	7*	8*	9*	10*
1	-	0.063	NR	NR	NR	NR	NR	NR	NR	NR
2		-	NR	0.078	NR	0.063	0.031	NR	NR	0.156
3			-	NR	NR	0.125	0.016	NR	NR	0.016
4				-	NR	0.063	0.008	NR	0.125	0.063
5					-	NR	0.063	0.031	NR	0.016
6						-	NR	NR	NR	NR
7							-	0.141	NR	0.074
8								-	NR	0.031
9									-	NR
10										-
11										
12										
13										
14										
15										
16										
17										
18										
19										
20										
21										

Table 1 (cont'd). The additive relationship matrix for all selected CN cattle.

Animal	11*	12*	13*	14	15	16	17*	18*	19	20*	21
1	NR	NR	NR	0.031	NR	NR	NR	0.031	NR	0.016	NR
2	NR	0.031	0.063	0.031	0.016	NR	NR	NR	NR	0.016	NR
3	NR	NR	0.031	NR	NR	NR	NR	NR	0.031	0.063	NR
4	NR	NR	0.250	0.063	0.016	NR	NR	NR	NR	NR	NR
5	0.031	NR	NR	0.031	0.031	0.250	NR	0.125	NR	0.125	0.031
6	NR	NR	0.063	NR	NR	NR	NR	NR	NR	0.031	NR
7	0.016	0.016	0.031	0.031	0.070	0.125	0.063	0.078	NR	0.024	0.016
8	0.063	NR	NR	0.063	0.016	NR	0.125	0.125	NR	NR	0.063
9	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
10	NR	0.063	0.039	0.094	0.039	NR	0.031	NR	0.008	0.031	0.031
11	-	NR	NR	NR	0.016	NR	NR	NR	NR	NR	0.063
12		-	NR	0.063	0.031	NR	NR	NR	NR	0.031	NR
13			-	0.031	NR	NR	NR	NR	0.063	0.016	NR
14				-	0.047	NR	NR	NR	NR	0.031	0.063
15					-	NR	NR	NR	NR	NR	0.016
16						-	NR	0.125	NR	0.125	NR
17							-	0.063	NR	NR	NR
18								-	NR	0.078	NR
19									-	NR	NR
20										-	NR
21											-

* Animals used for the PAUP computer analysis in Study III.

Table 2. The additive relationship matrix for all selected BS cattle.

Animal	1*	2	3	4	5	6*	7	8	9	10
1	-	NR	NR	NR	NR	0.125	0.078	NR	0.016	NR
2		-	NR	NR	0.188	NR	NR	NR	NR	NR
3			-	NR	0.125	NR	NR	NR	NR	0.125
4				-	NR	NR	NR	NR	NR	NR
5					-	0.125	NR	NR	NR	0.031
6						-	NR	0.031	0.008	NR
7							-	NR	NR	NR
8								-	NR	NR
9									-	NR
10										-
11										
12										
13										
14										
15										
16										
17										
18										
19										
20										

• Animals used for the PAUP computer analysis in Study III.

Table 2 (cont'd). The additive relationship matrix for all selected BS cattle.

Animal	11	12	13	14	15	16	17	18	19	20
1	NR	NR	0.016	NR	NR	0.063	NR	0.031	0.063	NR
2	NR	NR	0.063	NR	NR	NR	NR	NR	NR	NR
3	NR	NR	0.003	NR	NR	NR	NR	0.016	NR	NR
4	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
5	NR	NR	0.031	NR	NR	NR	NR	NR	0.063	0.047
6	NR	NR	0.004	NR	NR	0.031	NR	NR	0.156	NR
7	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
8	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
9	NR	NR	0.125	NR	NR	0.016	NR	NR	0.016	NR
10	0.063	0.063	NR	NR	NR	NR	0.031	NR	NR	NR
11	-	0.063	NR	NR	NR	NR	0.031	NR	NR	NR
12		-	NR	NR	NR	NR	0.031	NR	NR	NR
13			-	NR	NR	0.008	NR	0.016	0.008	NR
14				-	NR	NR	NR	NR	NR	NR
15					-	NR	NR	NR	NR	NR
16						-	NR	NR	0.063	NR
17							-	NR	NR	NR
18								-	NR	NR
19									-	0.016
20										-

Table 3. The additive relationship matrix for all selected JE cattle.

Animal	1	2	3*	4*	5	6	7	8	9	10
1	-	0.023	NR	0.008	0.023	0.016	0.008	0.020	0.031	0.023
2		-	NR	0.016	0.022	0.016	0.002	0.023	0.023	0.015
3			-	NR	NR	NR	NR	NR	NR	NR
4				-	0.008	NR	NR	0.008	0.008	0.006
5					-	0.016	NR	NR	NR	NR
6						-	0.031	0.090	NR	0.027
7							-	0.070	0.023	0.047
8								-	NR	NR
9									-	NR
10										-
11										
12										
13										
14										
15										
16										
17										
18										
19										
20										

• Animals used for the PAUP computer analysis in Study III

Table 3 (cont'd). The additive relationship matrix for all selected JE cattle.

Animal	11	12	13	14	15	16	17	18	19	20
1	0.031	0.004	0.031	0.023	0.016	0.008	0.023	0.063	0.043	0.012
2	0.078	0.008	0.027	0.006	0.004	0.002	0.007	0.029	0.003	0.003
3	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
4	0.047	0.031	0.016	NR	NR	NR	0.002	0.012	0.004	NR
5	0.023	0.012	0.023	0.012	0.016	0.004	0.016	0.023	0.014	0.010
6	0.063	NR	0.133	NR	NR	NR	0.020	0.016	NR	NR
7	0.039	0.008	0.023	0.012	0.016	0.066	0.020	0.008	0.014	0.010
8	0.031	0.004	0.051	NR	NR	0.063	0.009	0.020	NR	NR
9	0.063	0.004	0.051	0.023	0.016	0.008	0.020	0.031	0.012	0.012
10	0.027	0.019	0.020	NR	0.016	0.031	0.065	0.012	0.016	0.008
11	-	0.016	0.078	0.023	0.016	0.008	0.027	0.059	0.016	0.012
12		-	0.008	NR	0.141	NR	0.017	0.004	0.016	0.008
13			-	0.023	0.016	0.008	0.020	0.031	0.012	0.012
14				-	0.023	0.070	0.070	0.039	0.047	0.047
15					-	0.008	0.023	0.016	0.027	0.066
16						-	0.063	0.023	0.035	0.035
17							-	0.027	0.051	0.043
18								-	0.078	0.020
19									-	0.094
20										-

and HO) using the programs to compute the additive relationship matrix. Table 4 lists the rate of inbreeding estimated using Wright's formula for the same years for CN, BS and JE cattle. An examination of the latter values reveals that the predicted rate of inbreeding remained fairly constant during this time period. Average rates of inbreeding were found to be 0.61 $\pm$ 0.03% for the CN breed, 0.08 $\pm$ 0.01% for the BS breed and 0.032 $\pm$ 0.001% for the JE breed. Values estimated using the tabular method showed that there were no significant ($P>0.05$) increases in average inbreeding coefficients for the CN and BS breeds. A more substantial increase in the average inbreeding coefficients of the HO bull population was observed, however. Average inbreeding levels calculated using this method ranged from 0 to 1.3% for the CN breed, from 0.01 to 0.7% for the BS breed and from 0.9 to 4.8% for the HO bulls. The average rates of inbreeding per year that were determined for CN and BS cattle and for the HO bulls over this 12 year period using the additive relationship matrix were 0.11% $\pm$ 0.01%, 0.07% $\pm$ 0.01% and 0.37% $\pm$ 0.02%, respectively.

The effective number of male and female parents in the CN, BS and JE breeds from 1983 to 1994 and the number of animals registered in each of those years are shown in Table 4. The difference in the size of the breeding populations seen between the three breeds is quite striking. The average effective number of breeding males and females between the years 1983 and 1994 were 24 $\pm$ 1 and 184 $\pm$ 20, respectively, for the CN breed, 193 $\pm$ 9 and

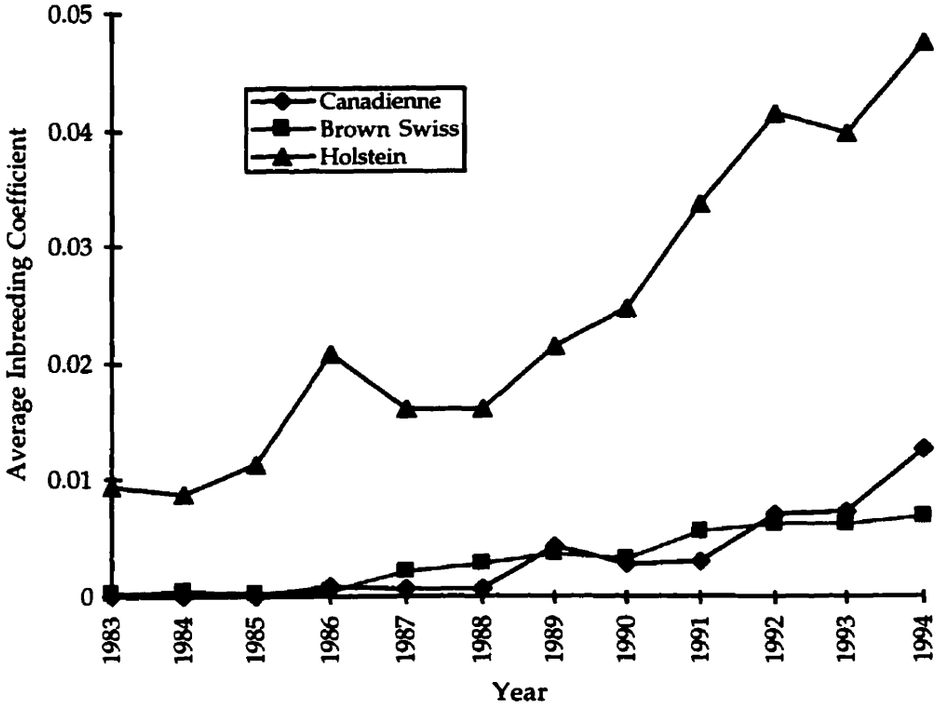


Figure 1. Average inbreeding coefficents of Canadienne, Brown Swiss and Holstein cattle obtained from the additive relationship matrix for each breed.

Table 4. Inbreeding estimates generated for 1983-1994 for Canadienne, Brown Swiss and Jersey breeds using Wright's formula and the effective number of males and females for these years.

Year	Rate of Inbreeding	N_m	N_f	# of Animals Registered
		Canadienne		
1983	0.00817	20	65	476
1984	0.00771	21	71	559
1985	0.00581	26	125	411
1986	0.00636	23	135	674
1987	0.00546	26	193	415
1988	0.00455	31	240	408
1989	0.00528	26	266	372
1990	0.00569	24	261	310
1991	0.00569	24	260	301
1992	0.00575	24	231	302
1993	0.00637	22	181	177
1994	0.00591	24	177	260
		Brown Swiss		
1983	0.00136	118	415	1524
1984	0.00100	162	536	1082
1985	0.00099	157	649	1375
1986	0.00083	178	961	1596
1987	0.00075	197	1072	1636
1988	0.00066	214	1604	1757
1989	0.00068	206	1751	1539
1990	0.00065	215	1808	2031
1991	0.00063	223	1759	1953
1992	0.00064	219	1801	2133
1993	0.00066	213	1724	2149
1994	0.00067	209	1774	2297

Table 4 (cont'd). Inbreeding estimates generated for 1980-1995 for Canadienne, Brown Swiss and Jersey breeds using Wright's formula and the effective number of males and females for these years.

Year	Rate of Inbreeding	N_m	N_f	# of Animals Registered
		Jersey		
1983	0.00027	489	6756	7032
1984	0.00027	490	6820	6725
1985	0.00028	481	6935	8098
1986	0.00031	427	6994	6907
1987	0.00030	448	7296	7362
1988	0.00029	455	6996	7084
1989	0.00029	460	6977	7155
1990	0.00033	403	7178	7126
1991	0.00035	373	6989	7333
1992	0.00038	347	6596	7049
1993	0.00040	331	6309	6564
1994	0.00041	318	6133	6447

1321 ± 160, respectively, for the BS breed and 419 ± 18 and 6832 ± 98, respectively, for the JE breed. The average effective population sizes for the CN, BS and JE breeds were 86, 672 and 1577 animals, respectively. Correspondingly, the average number of registrations annually were 389 ± 39, 1756 ± 105, and 7074 ± 124, respectively, during this 12 year span.

An examination of the ten most frequently sought sires between 1980 and 1995 for the CN, BS and JE breeds and the HO bulls reveals some interesting information. The ten most popular CN sires were found to have fathered approximately 61% of the CN population born during this time period. The same figures for the ten top BS and JE sires were 21% and 29%, respectively. Meanwhile, 41% of the HO bulls in the four studs were fathered by only ten sires.

3.3 DISCUSSION

For many types of genetic analysis it is advantageous to use animals that are as unrelated as possible as this approach gives one the greatest certainty of including all possible genotypes in a given sample. This is true for genetic distance studies done using microsatellites and mitochondrial D-loop sequence variation and is, in fact, common practice for such studies. The complete pedigree analysis of animal populations that are to be used in such experiments therefore becomes important. As is seen from the results of the

present study, such an analysis allows one to effectively choose the required animals and also generates some interesting additional information about the populations in question.

It is evident from Figure 1 that the degree of inbreeding found in the CN, BS and HO populations examined is low. It, therefore, appears that at this point in time the overall level of inbreeding is not of great concern in any of the breeds studied. This conclusion is supported by work done by Young et al., (1988) with registered Holsteins in the United States and a study done by Miglior et al. (1992) measuring inbreeding in Canadian Jersey cattle. Both of these groups found that the overall level of inbreeding in these dairy cattle populations had remained constant and fairly low over the years. Similar findings were also presented by Casanova et al. (1992) for Swiss Braunvieh cattle. These cattle are related to Brown Swiss cattle, having an average of 38.8% US Brown Swiss genes in 1984.

An examination of the increases in inbreeding that were seen in this study reveals that the rate of inbreeding in all of the populations was relatively low. The rate predicted using Wright's formula was greater what was actually seen in the CN and BS populations, respectively. Extrapolating from this, one can assume that the actual rate of inbreeding in the JE population is also less than the 0.032% average predicted with this method. The rate of inbreeding seen in the HO bull population in recent years, however, was more than three times that seen in the CN population and

approximately five times that seen in the BS population as a whole. This suggests that the HO bull population bears monitoring. Miglior and Burnside (1995) came to a similar conclusion after their very extensive study of Canadian Holsteins. The annual rate of increase in inbreeding of 0.21% they estimated for all Holstein bulls between 1987 and 1990 is lower than the rate of 0.37% found in the present study for 1983 to 1994. This suggests that inbreeding may be increasing more rapidly in A.I. bulls compared to all Holstein bulls in general. Miglior and Burnside's finding that the mean inbreeding of A.I. bulls approached 3% in 1990 is in agreement with the figures of 2.4% and 3.0% reported here for 1990 and 1991.

The results of this study show that as many as 61% of CN animals in recent years were fathered by just ten bulls and that the effective number of male parents used in the CN breed is extremely small. This suggests that inbreeding may become a concern in this breed in the near future. Due to the small size of the CN population and the relatively low number of sires that are being used (Table 4), it is impossible at this point in time to find a group of 20 animals that are completely unrelated. There are currently less than 50 Canadienne breeders left in the province of Québec and virtually none outside of the province. An examination of the herds of 6 of the largest breeders revealed that only 39 animals could be found that were not half or full-sibs. Furthermore, examination of the total number of registrations seen in the CN population each year and the effective number of breeding males

and females found in the population reveals that the breed meets or comes close to many of the criteria established by FAO and other conservancy organizations for endangered breeds. This poses a rather significant problem for the Société des Éleveurs de Bovins Canadiens and for all CN breeders. It is interesting to note that a phenomenon similar to that found in the CN breed in terms of animals sired by the top 10 bulls is present in the HO bull population studied. As was reported earlier, 41% of bulls in the four A.I. studs examined were sired by the same ten bulls.

The tremendous impact that the effective population size of a breed has on the amount of heterozygosity remaining after a given number of generations is illustrated by Table 5. The percentage of the heterozygosity remaining was calculated using the general formula $H_t = [1- (1/2N_e)]^t H_o$, where H_o is the initial heterozygosity and H_t is the heterozygosity after t generations and N_e is the effective population size (Weaver and Hedrick, 1997). From an examination of this table, it is clearly evident that a breed such as the CN, which was shown to have an effective population size of only 86, will lose approximately 44% of its heterozygosity within 100 generations. In contrast, the JE breed, which today has an effective population size of 1577, will have lost only about 3% of its present heterozygosity at that time.

A closer examination of the number of male and female parents in the CN breed shows that there is, on average, 1 male for every 8 females.

Table 5. Percentage of the present heterozygosity in the Canadienne (CN), Brown Swiss (BS) and Jersey (JE) populations remaining after 100 generations based on the effective population size of each breed.

Breed	Number of Generations			
	10	20	50	100
CN	94.3	89.0	74.7	55.8
BS	99.3	98.5	96.3	92.8
JE	99.7	99.4	98.4	96.9

Furthermore, one sire is present in the breed for every 16 animals registered. Similar calculations for BS cattle reveal that this breed has one breeding male for every seven breeding females and, on average, one sire for every nine animals registered. The JE breed has one male for every 16 breeding females and one sire for every 17 animals registered. As these figures indicate, the ratio of dams to sires in the CN and BS breeds is approximately half that in the more populous JE breed. This observation has implications for the future of these breeds in this country as it again reflects the much smaller genetic base that exists for these two breeds. It is important to remember that the breeding strategy that has been employed by the dairy industry in recent years involves mating less closely related animals than would occur at random. This fact has almost certainly helped keep the overall level of inbreeding in populations such as the ones studied here lower than would have been the case if such action had not been initiated. For the less populous breeds this strategy may not be enough in the future, however.

4. STUDY II: THE DETERMINATION OF GENETIC DISTANCE AMONG CATTLE BREEDS USING MICROSATELLITE ANALYSIS

4.0 OBJECTIVE

The objective of this study was to determine the genetic distance between the Canadienne (CN), Holstein (HO), Jersey (JE) and Brown Swiss (BS) cattle breeds found in Canada using 15 different, previously mapped microsatellite markers in order to draw conclusions about the relationship of these breeds, especially the CN, with each other.

4.1 MATERIALS AND METHODS

4.1.1 Animals

The animals sampled from three of the four breeds of cattle (CN, JE and BS) were chosen using the computer programs outlined in Study I (see Appendix Tables 1-11). Briefly, pedigree information was obtained from Canadian Livestock Records Corporation (CLRC) and Holstein Canada. The pedigree data obtained were first edited using a SAS computer program (SAS, 1990). The data were then analyzed using Fortran programs designed to calculate the additive relationship coefficients of the individual animals (Cruden, 1949; Emik and Terrill, 1949; Van Vleck, 1993). Animals were chosen for sampling based on their lack of or distance of relationship with each other and their location relative to Ottawa. Every attempt was made to sample those animals that were least related while at the same time being

located at a reasonable driving distance from the Centre. Tables 1-3 contain the relative relationships of all of the animals chosen for the study using this method. Holstein DNA samples were provided by Dr. Denis Petitclerc of Agriculture and Agri-Food Canada's Dairy and Swine Research and Development Centre in Lennoxville, Québec. These HO animals were completely unrelated for three generations.

4.1.2 Extraction of Total Genomic DNA from Whole Blood Samples

Total genomic DNA was extracted from individual blood samples using the Super Quik-Gene DNA isolation kit (Analytical Genetic Testing Center, Inc., Denver, Colorado, USA). Briefly, blood samples were collected from individual animals in 8.5 ml ACD-Vaccutainer tubes (VWR Scientific, Ville Mont-Royal, Québec, Canada.; ACD: 22.0 g/l trisodium citrate, 8.0 g/l citric acid, and 24.5 g/l dextrose) and kept at 4 °C until the time of DNA extraction. Every effort was made to extract the DNA from the blood samples within a day or two of collection. Blood samples were first centrifuged at 2500 rpm for 20 min. in a Jouan CR 4 22 refrigerated centrifuge (Canberra Packard Canada, Mississauga, Ontario, Canada). They were then allowed to stand at 4°C for 2 hrs in order to allow the white blood cells to aggregate and make removal of the buffy coat more efficient. As much of the plasma as possible was then removed from each tube and the buffy coat and approximately 1 ml of the underlying red blood cells were transferred to a 15 ml polypropylene

Falcon tube (Fisher Scientific, Nepean, Ontario, Canada). RBC lysis buffer (1X) provided in the extraction kit was then added to each tube to a total volume of 13 ml. The mixture was mixed for 15 min. at room temperature and then centrifuged for 20 min. at 2500 rpm. After centrifugation, the supernatant was decanted and each tube was washed with 2-3 ml of RBC lysis buffer (1X), without disturbing the white blood cell pellet, in order to remove any excess stroma and supernatant binding to the tube. A total volume of 1.5 ml WBC lysis buffer was added to each tube and the pellet was broken up using a Pasteur pipet. The tubes were then incubated for 30 min. at 55 °C. After the incubation, 0.2ml of 10% SDS and 0.5 ml of Protein Precipitating Reagent were added to each tube and the mixture was shaken vigorously by hand for 30 sec. All tubes were then incubated for 15 min. in a 55 °C water bath. After the incubation period, the tubes were again centrifuged for 20 min. at 2500 rpm in order to pellet the precipitated protein. The clear supernatant was transferred to a new 15 ml polypropylene Falcon tube and 2 volumes of room temperature absolute ethanol were added. The tubes were inverted gently until the precipitated DNA could be seen as a mass of white, stringy fibres. Finally, the precipitated DNA was spooled on to a plastic inoculation loop and transferred to a 1.5 ml microcentrifuge tube containing Tris-EDTA (TE) buffer (10 mM Tris-base, 1 mM EDTA, pH 7.4).

4.1.3 PCR and Gel Electrophoresis

The samples of extracted DNA were used to perform PCR reactions in order to amplify specific microsatellite sequences. The loci studied and primers used in these reactions are shown in Table 6. Primers for these loci were provided by Dr. Craig Beattie, USDA, Clay Center, Nebraska and by Dr. James Derr, Texas A & M University, College Station, Texas. Primers were end-labelled with ^{32}P using $\gamma^{32}P$ ATP (Mandel Scientific Company Ltd., Guelph, Ontario, Canada) and T4 polynucleotide kinase (Promega; Fisher Scientific, Nepean, Ontario, Canada). Approximately 250 ng of genomic DNA was used per PCR reaction. These samples were amplified in a reaction volume of 30 µl containing 21.7 µl of deionized, distilled water, 3.0 µl of 10X PCR buffer, 2.0 µl of 25mM $MgCl_2$, 0.8 µl of a mixture of 10 mM dNTPs, 0.8 µl of DMSO, 1.0 µl of the labelled forward primer, 0.5 µl of the reverse primer and 0.2 ul of AmpliTaq DNA polymerase (Applied Biosystems Canada Inc., Mississauga, Ontario, Canada). A standard thermocycling protocol of 3 min. at 94 °C followed by 30 cycles of 1 min. at 94 °C, 30 sec at the annealing temperature and 1 min. at 72 °C and ending with an extension phase of 4 min. at 72 °C was used. An MJ Research DNA Engine was employed for all PCR runs. The amplified product was diluted with 20 µl of Stop Solution (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol FF) and stored at -20 °C for a maximum of 3 days before use.

A "T- ladder" reference marker for the subsequent gel electrophoresis was made using the Sequenase Version 2.0 DNA Sequencing Kit (Amersham Canada Ltd., Oakville, Ontario, Canada). Briefly, 1 µl of -40 primer, 2 µl of reaction buffer, 5 µl of M13mp18 single stranded DNA and 2 µl of distilled, deionized water were combined and warmed to 65 °C for 2 min. in a heat block. The mixture was then allowed to cool slowly to 35 °C over a period of approximately 30 min. To this annealed template/primer mixture, 1 µl of DTT (0.1M), 2 µl of diluted labelling mix, 0.7 µl of α^{35}S dATP and 2 µl of Sequenase were added and mixed. The mixture was incubated for 5 min. at room temperature. Following the incubation, 14 µl of the above reaction mixture were added to 10 µl of ddTTP and incubated for 5 min. at 37 °C. Finally, 16 µl of Stop solution was added and the resulting volume of labelled "T ladder" marker was frozen at -20 °C until use.

All amplified samples and the labelled "T ladder" reference marker were denatured by incubation at 94 °C for 5 min. and run on a 6% denaturing polyacrylamide gel (thickness 0.4 mm) in 1X TBE at 50 °C and 123 watts on a BIO RAD sequencing apparatus (BIO RAD Laboratories (Canada) Ltd., Mississauga, Ontario, Canada). The gel was then transferred to Whatman 3MM paper, covered in plastic wrap and dried for 2.5 hrs on a BIO RAD gel dryer (BIO RAD Laboratories (Canada) Ltd., Mississauga, Ontario, Canada). All dried gels were exposed overnight to Kodak BIOMAX film (Kodak,

Ottawa, Canada) and developed in an Ecomat 4000 automatic developer (eg Appendix Figure 1). Allele sizes were approximated by comparison to the "T ladder" reference marker and the different alleles present on each film were identified visually. Each animal was scored to establish which alleles it carried at each microsatellite locus.

4.1.4 Genetic Distance Analysis

Genetic distances were estimated between the four breeds using the allele frequency data collected by scoring the individual animals for the various microsatellites. To facilitate the analysis the MICROSAT 1.5b computer package (Minch, 1997) was used. Genetic distances were calculated using three different measures: Nei's standard genetic distance (D; Nei, 1972), delta mu squared (Ddm; Goldstein et al., 1995b) and Rst (Slatkin, 1995). Bootstrapping over 1000 replications was carried out in all cases. All loci were tested for divergence from Hardy-Weinberg equilibrium (HWE) using the GENEPOP (Version 3.0) computer program (Raymond and Rousset, 1995).

4.2 RESULTS

Fifteen previously described microsatellite loci (Bishop et al., 1994) were used in this study (Table 6). Loci on separate chromosomes were chosen so as to avoid linkage. All loci used were polymorphic in all breeds of cattle

studied. All loci, except BM2113 and BM6501, were found to be in HWE when tested across breeds using the GENEPOP program's "probability test" ($P>0.05$). When breed-locus combinations were tested individually, all breeds were found deviate from HWE at at least one locus. The CN breed's deviation occurred at locus ADCY2 ($P<0.05$). Meanwhile, BS cattle deviated from equilibrium at loci BM3413 ($P<0.05$) and BM203 ($P<0.01$), JE cattle at loci BM2113 ($P<0.01$), BM3413 ($P<0.05$) and BM6501 ($P<0.01$), and HO cattle deviated from HWE at loci CSSM47 ($P<0.05$), BM720 ($P<0.05$), BP20 ($P<0.05$), BMC8012 ($P<0.05$) and BM6501 ($P<0.05$).

Allele frequencies at the 15 loci for each breed of cattle are shown in Table 7. Allelic frequencies were found to be variable among breeds at most loci. In fact, only five of the loci, BM864, BM1237, BP20, BM3413 and BM226, were found to have the same predominant allele in each of the breeds.

Tables 8 and 9 show the average heterozygosity and the average variance in allele size across loci and across breeds. As can be seen, the average heterozygosity was high for all loci studied, ranging from 0.447 to 0.828. The average heterozygosity did not differ greatly for CN, BS and HO breeds. The JE breed tended to have a slightly lower heterozygosity, however. Allele size variances for the different loci ranged from 0.718 to 21.493. These variances were similar for all four breeds of cattle, although there was a tendency for BS cattle to have a slightly lower variance and JE cattle to have a slightly higher variance than the other breeds. The average number of alleles

Table 6. Data on the microsatellite loci examined in the study.

Chromosome	Locus	Primers (5′ -> 3′) (Forward; Reverse)	Annealing Temp. (o C)	# of known alleles	Size (bp)	Reference
1	BM 864	TGGTAGAGCAATATGAAGGCC GGAAATCCAAGAAAGAGGGG	58	14	214-274	Bishop et al., 1994
2	BM 2113	GCTGCCTTCTACCAAATACCC CTTCCTGAGAGAAGCAACACC	58	11	123-143	Bishop et al., 1994
7	BM6501	ACTAATAAGAAATTCTGCATGTGTG CCACCATGACTCAGAAGTAGTTC	58	8	87-107	Stone et al., 1995
8	CSSM47	TCTCTGTCTCTATCACTATATGGC CTGGGCACCTGAAACTATCATCAT	58	8	~150	Barendse et al., 1994
10	BM 1237	TCATCTTGGGCATAAGACAGG ATTGTTCCCAGCATCTTAGAGG	58	9	187-223	Bishop et al., 1994
13	BM 720	ACATCTCATTCTTGTGTCATGG GAAATTCAGTTTAGGGTTCCCC	56	14	210-240	Bishop et al., 1994
15	ADCY2	AAAGTGACACAACAGCTTCTCC ACAAGTGAGTGCGTAACAAAGG	54	7	185-205	Bishop et al., 1994
16	BM 121	TGGCATTGTGAAAAGAAGTAAA ACTAGCACTATCTGGCAAGCA	58	14	118-160	Bishop et al., 1994
19	BP 20	TCTGTGGGTGAACAAGCAAG GGCTCCCTAAAGACCCACTC	56	8	219-233	Bishop et al., 1994

Table 6. (cont'd) Data on the microsatellite loci examined in the study.

Chromosome	Locus	Primers (5' -> 3') (Forward; Reverse)	Annealing Temp. (o C)	# of known alleles	Size (bp)	Reference
21	BM 3413	TCCCTGGTAACCAATGAATTC CAATGGATTTGACCCTCCC	58	9	170-192	Bishop et al., 1994
24	BM 226	ATTGCCTTGTCCGTGTATCC CCGGCTGAATTGCTATAAGC	58	10	128-164	Bishop et al., 1994
25	BMC 8012	AATTCCATGCACAGAGGACC GATTCCAGAAAGTTCCCCCA	58	10	197-215	Bishop et al., 1994
27	BM 203	GGGTGTGACATTTTGTTCCC CTGCTCGCCACTAGTCCTTC	58	11	203-233	Bishop et al., 1994
28	BM 2515	GATTCCCTGACTCTCTGTCCC AGTATTGGCAAGTCAATGGAGG	58	5	132-148	Bishop et al., 1994
X	BM 6017	TCTTCTGTTTTCCTCCATCCC GGAAACTAGCTTATGCTGTGGG	58	10	112-139	Bishop et al., 1994

Table 7. Allelic frequencies for Canadienne(CN), Brown Swiss (BS), Jersey (JE) and Holstein (HO) cattle at 15 different microsatellite loci located on separate chromosomes.

Breed	Allele												
	1	2	3	4	5	6	7	8	9	10	11	12	13
Locus: BM864													
CN	0.167	0.048	0.000	0.024	0.095	0.000	0.167	0.071	0.119	0.048	0.262	0.000	0.000
BS	0.000	0.000	0.050	0.000	0.025	0.025	0.025	0.025	0.000	0.000	0.825	0.025	0.000
JE	0.000	0.000	0.000	0.000	0.000	0.000	0.100	0.025	0.050	0.225	0.600	0.000	0.000
HO	0.000	0.000	0.000	0.025	0.000	0.075	0.025	0.000	0.425	0.000	0.400	0.025	0.025
Locus: BM2113													
CN	0.286	0.310	0.143	0.095	0.000	0.000	0.095	0.071					
BS	0.225	0.000	0.075	0.175	0.000	0.000	0.225	0.300					
JE	0.000	0.000	0.475	0.350	0.025	0.125	0.000	0.025					
HO	0.316	0.263	0.237	0.026	0.000	0.000	0.105	0.053					
Locus: CSSM47													
CN	0.905	0.071	0.024	0.000	0.000	0.000							
BS	0.275	0.225	0.425	0.050	0.025	0.000							
JE	0.825	0.000	0.000	0.000	0.000	0.175							
HO	0.575	0.225	0.050	0.100	0.050	0.000							

Table 7 (cont'd). Allelic frequencies for Canadienne(CN), Brown Swiss (BS), Jersey (JE) and Holstein (HO) cattle at 15 different microsatellite loci located on separate chromosomes.

Breed	Allele												
	1	2	3	4	5	6	7	8	9	10	11	12	13
Locus: BM1237													
CN	0.024	0.024	0.167	0.429	0.048	0.238	0.071	0.000	0.000				
BS	0.075	0.075	0.075	0.575	0.000	0.125	0.050	0.025	0.000				
JE	0.050	0.025	0.000	0.400	0.000	0.250	0.225	0.025	0.025				
HO	0.158	0.105	0.158	0.342	0.053	0.079	0.079	0.000	0.026				
Locus: BM720													
CN	0.190	0.333	0.214	0.000	0.000	0.024	0.119	0.071	0.000	0.000	0.000	0.048	
BS	0.150	0.025	0.525	0.050	0.025	0.025	0.000	0.075	0.000	0.075	0.000	0.050	
JE	0.650	0.050	0.000	0.025	0.025	0.000	0.000	0.250	0.000	0.000	0.000	0.000	
HO	0.200	0.100	0.275	0.000	0.000	0.000	0.000	0.100	0.025	0.175	0.075	0.050	
Locus: ADCY2													
CN	0.357	0.190	0.333	0.000	0.000	0.119							
BS	0.000	0.200	0.300	0.000	0.225	0.275							
JE	0.500	0.000	0.075	0.000	0.425	0.000							
HO	0.000	0.111	0.528	0.028	0.222	0.111							

Table 7 (cont'd). Allelic frequencies for Canadienne(CN), Brown Swiss (BS), Jersey (JE) and Holstein (HO) cattle at 15 different microsatellite loci located on separate chromosomes.

Breed	Allele												
	1	2	3	4	5	6	7	8	9	10	11	12	13
Locus: BM121													
CN	0.048	0.119	0.190	0.024	0.310	0.119	0.190	0.000	0.000				
BS	0.000	0.350	0.275	0.250	0.050	0.025	0.000	0.000	0.050				
JE	0.150	0.000	0.100	0.000	0.400	0.000	0.100	0.150	0.100				
HO	0.150	0.225	0.250	0.000	0.225	0.025	0.100	0.025	0.000				
Locus: BP20													
CN	0.667	0.024	0.310	0.000	0.000	0.000							
BS	0.700	0.025	0.275	0.000	0.000	0.000							
JE	0.800	0.175	0.025	0.000	0.000	0.000							
HO	0.575	0.000	0.225	0.025	0.150	0.025							
Locus: BM3413													
CN	0.500	0.357	0.143	0.000									
BS	0.500	0.450	0.025	0.025									
JE	0.925	0.050	0.000	0.025									
HO	0.632	0.316	0.053	0.000									

Table 7 (cont'd). Allelic frequencies for Canadienne(CN), Brown Swiss (BS), Jersey (JE) and Holstein (HO) cattle at 15 different microsatellite loci located on separate chromosomes.

Breed	Allele												
	1	2	3	4	5	6	7	8	9	10	11	12	13
Locus: BM226													
CN	0.024	0.190	0.119	0.048	0.024	0.071	0.048	0.238	0.167	0.024	0.048		
BS	0.000	0.075	0.050	0.025	0.125	0.000	0.150	0.325	0.175	0.050	0.025		
JE	0.000	0.050	0.025	0.000	0.000	0.000	0.200	0.450	0.275	0.000	0.000		
HO	0.000	0.000	0.100	0.000	0.000	0.000	0.075	0.575	0.250	0.000	0.000		
Locus: BMC8012													
CN	0.000	0.000	0.048	0.167	0.286	0.310	0.167	0.024					
BS	0.100	0.150	0.000	0.250	0.250	0.200	0.025	0.025					
JE	0.100	0.125	0.450	0.175	0.050	0.100	0.000	0.000					
HO	0.000	0.079	0.079	0.711	0.105	0.026	0.000	0.000					
Locus: BM203													
CN	0.024	0.310	0.048	0.310	0.000	0.262	0.048	0.000	0.000	0.000	0.000	0.000	0.000
BS	0.050	0.375	0.250	0.075	0.150	0.000	0.075	0.000	0.000	0.000	0.000	0.000	0.025
JE	0.000	0.025	0.000	0.325	0.025	0.275	0.000	0.100	0.025	0.000	0.000	0.100	0.125
HO	0.000	0.325	0.050	0.225	0.175	0.025	0.000	0.000	0.025	0.075	0.075	0.025	0.000

Table 7 (cont'd). Allelic frequencies for Canadienne(CN), Brown Swiss (BS), Jersey (JE) and Holstein (HO) cattle at 15 different microsatellite loci located on separate chromosomes.

Breed	Allele												
	1	2	3	4	5	6	7	8	9	10	11	12	13
Locus: BM2515													
CN	0.024	0.571	0.095	0.286	0.000	0.024							
BS	0.050	0.525	0.025	0.150	0.250	0.000							
JE	0.025	0.575	0.200	0.200	0.000	0.000							
HO	0.000	0.375	0.175	0.425	0.000	0.025							
Locus: BM6017													
CN	0.071	0.048	0.167	0.119	0.357	0.048	0.119	0.000	0.024	0.000	0.024	0.024	
BS	0.000	0.025	0.050	0.075	0.275	0.275	0.025	0.050	0.200	0.000	0.000	0.025	
JE	0.000	0.100	0.100	0.150	0.275	0.075	0.075	0.075	0.150	0.000	0.000	0.000	
HO	0.118	0.000	0.118	0.088	0.059	0.059	0.235	0.088	0.176	0.059	0.000	0.000	
Locus: BM6501													
CN	0.310	0.071	0.024	0.190	0.095	0.262	0.048	0.000	0.000				
BS	0.425	0.075	0.125	0.100	0.000	0.175	0.050	0.050	0.000				
JE	0.475	0.350	0.000	0.000	0.025	0.000	0.150	0.000	0.000				
HO	0.250	0.075	0.000	0.200	0.125	0.275	0.050	0.000	0.025				

Table 8. Average heterozygosity and average allele size variance for each microsatellite locus studied.

Locus	Average Heterozygosity	Average Variance
BM 864	0.596	12.900
BM 2113	0.737	4.724
BM 6501	0.735	4.846
CSSM 47	0.429	8.115
BM 1237	0.721	21.493
BM 720	0.702	5.794
ADCY 2	0.666	4.494
BM 121	0.775	21.169
BP 20	0.458	1.779
BM 3413	0.447	0.718
BM 226	0.733	8.847
BMC8012	0.691	1.744
BM 203	0.765	9.766
BM 2515	0.614	2.894
BM 6017	0.828	9.317

Table 9. Average heterozygosity and average variance observed for each of four breeds of cattle studied.

Breed	Average Heterozygosity	Average Variance
Canadienne	0.690	7.907
Brown Swiss	0.675	7.761
Jersey	0.585	8.075
Holstein	0.690	7.884

was similar for CN, BS and HO cattle (Table 10). JE cattle, however, tended to have a lower average number of alleles ($P>0.05$).

The 15 microsatellite loci examined were analyzed for the occurrence of breed specific alleles. The results obtained are presented in Table 11. It is evident from this table that alleles that appear to be specific to each of the four breeds were identified. A total of six alleles at four different loci were found to be specific to the CN breed. Three alleles at three loci were found for the BS breed, four alleles at three loci for the JE breed and ten alleles at six loci were identified as being specific for the HO breed. The combined frequency of the breed specific alleles found at each locus was extremely small, varying from 6.9% to 12.5%.

Genetic distance values between cattle breeds were estimated using three different distance measures to analyze the data generated through microsatellite analysis. Bootstrapping over 1000 replications was used for each method. These values are shown in Table 12. As can be seen from the table, the three methods vary somewhat in the results they generate. Nei's D values ranged from 0.156 ± 0.057 to 0.427 ± 0.124. All genetic distances calculated using this method were significantly different from zero. Genetic distances did not differ significantly among breeds, however. The genetic distance between CN and HO cattle tended to be somewhat lower than those between CN and BS or JE cattle, whereas BS and JE breeds tended to be most distantly related. The Ddm genetic distance method generated distance

Table 10. The average number of alleles for Canadienne (CN), Brown Swiss (BS), Jersey (JE) and Holstein (HO) cattle over 15 microsatellite loci.

Breed	Average Number of Alleles
CN	6.2667 ± 0.6360
BS	6.2667 ± 0.4925
JE	4.9333 ± 0.4727
HO	6.1333 ± 0.4866

Table 11. Breed-specific microsatellite alleles observed in animals sampled from Canadienne (CN), Brown Swiss (BS), Jersey (JE) and Holstein (HO) cattle populations.

Locus	Allele (bp)*	Breed
BM 864	274 (1)	CN
	264 (3)	BS
	258 (2)	CN
	226 (13)	HO
BM 2113	134 (5)	JE
	132 (6)	JE
BM 6501	108 (9)	HO
	92 (8)	BS
CSSM 47	172 (6)	JE
BM 720	242 (9)	HO
	230 (11)	HO
	220 (7)	CN
BP 20	237 (6)	HO
	225 (4)	HO
	221 (5)	HO
BM 226	154 (1)	CN
	138 (6)	CN
BM 203	229 (8)	JE
	223 (10)	HO
	211 (11)	HO
BM 2515	134 (5)	BS
BM 6017	143 (11)	CN
	117 (10)	HO

* Numbers in brackets refer to the number assigned to that particular allele at that locus in Table 7.

Table 12. Genetic distance estimates (± SE) among Canadienne (CN), Brown Swiss (BS), Jersey (JE) and Holstein (HO) cattle derived using three different genetic distance measures: Nei's standard distance (D), the delta mu squared distance (Ddm) and Slatkin's Rst over 1000 bootstrap replications.

	D			Ddm			Rst		
	CN	BS	JE	CN	BS	JE	CN	BS	JE
BS	0.243 ± 0.078			3.385 ± 1.204			0.134 ± 0.040		
JE	0.235 ± 0.066	0.427 ± 0.124		3.717 ± 1.114	3.672 ± 1.492		0.202 ± 0.052	0.146 ± 0.042	
HO	0.156 ± 0.057	0.211 ± 0.047	0.320 ± 0.084	3.136 ± 1.166	1.301 ± 0.443	3.088 ± 1.158	0.142 ± 0.040	0.075 ± 0.030	0.129 ± 0.040

estimates ranging from 1.301±0.443 to 3.717±1.114. Again the individual values did not differ significantly from one another. The Ddm method, however, indicated a tendency for the CN breed to be most closely related to BS and HO breeds. BS and HO cattle tended to exhibit a smaller distance from each other than from the other breeds. Finally, the Rst method gave similar overall results to the Ddm method. Distance estimates based on the Rst method ranged from 0.075±0.030 to 0.202±0.052.

4.3 DISCUSSION

A population is said to be in HWE if the gene and genotype frequencies are constant from one generation to another. Factors such as genetic drift, selection, mutation and migration will all disrupt HWE in a population. It is interesting to see from the results that generally HWE was observed within the populations tested. The fact that some significant deviations were observed at particular loci agrees with results obtained by MacHugh et al. (1994) who similarly found some degree of heterogeneity among the European cattle populations they studied. Of particular interest is the fact that, within populations, HO cattle showed disequilibrium at the largest number of loci. This may be a result of the heavy selection pressure that has been applied to HO cattle. The use of A.I. is also especially widespread in the HO breed compared to the other three breeds and this may, therefore, be

having a more pronounced genetic influence in this population.

The three methods used to calculate genetic distance differed from each other in a few fundamental ways which naturally affect the results obtained from each method. Nei's genetic distance is an allele frequency based method that assumes that the infinite alleles model holds true. Ddm and Rst, on the other hand, are allele size and allele size variance based distance methods designed to fit the stepwise mutation model. While Rst represents the fraction of the total variance of allele size that is between populations, Ddm uses the squared mean difference between the alleles of two populations to estimate genetic distance. This latter distance has the advantage of being independent of population size. It has been shown and discussed widely in the literature (eg. Schlötterer and Tautz, 1992; Goldstein et al., 1995b; Slatkin, 1995) that microsatellite loci appear to adhere most closely to the stepwise mutation model, especially over relatively short time scales. Distances based on Nei's D method may, therefore, be less appropriate for microsatellites. However, since this distance has been used widely in previous studies, it was included in the present one for the sake of consistency. It is interesting to note that, knowing the fact that the CN breed has been crossbred with BS cattle in recent years, the two allele-size methods appear to generate more plausible results. Both show that the CN breed tends to be most closely related to the BS and HO breeds while with Nei's method there is a tendency for the HO breed to be more closely related to the CN breed than is the BS

breed. The apparent superiority of the Ddm and Rst methods over Nei's method in accounting for breed history tend to support the opinion, outlined above, that the stepwise mutation model is more suitable for microsatellite loci. In contrast, Forbes et al. (1995), in a study involving domestic and bighorn sheep, concluded that Nei's distance, among others, was more sensitive to population history than variance-based distances such as Rst.

The relatively more distant relationship that was found between the CN breed and the JE breed, particularly with the Ddm and Rst distance methods, is somewhat surprising. Historically, CN cattle are believed to have originated from cattle brought by the early settlers from the Normandy and Brittany regions of France. These cattle are believed to have similar origins to the Channel Island breeds such as the Jersey. In fact, CN cattle are fairly similar to JE cattle in overall appearance and have been referred to as Black Jerseys on occasion. Likewise, the fact that the HO breed was consistently shown to be quite closely related to the CN is surprising. From a strictly phenotypic standpoint there are very few similarities between the two breeds. Recent historical records, furthermore, do not indicate any direct link between these breeds. One contributing cause to their relatively close relationship may be associated with the introduction of BS into the CN breed. As can be seen from Table 12, BS and HO cattle appear to be very closely related.

Finally, an examination of the variations in allele frequencies found at

the different loci in the four breeds and the number of apparently breed specific allele raises the question whether such information could be used to characterize a breed. In fact, this has been shown to be possible by Buchanan et al. (1994) in a study involving various sheep breeds. These authors were able to use combined allele frequency differences to identify an individuals breed with relative certainty.

5. STUDY III: MITOCHONDRIAL DNA D-LOOP SEQUENCE VARIATION WITHIN AND AMONG CATTLE BREEDS.

5.0 OBJECTIVE

The objective of this study was to determine the extent of the mitochondrial DNA (mtDNA) D-loop sequence variation present in four Canadian cattle breeds: Canadienne (CN), Holstein (HO), Jersey (JE) and Brown Swiss (BS), to estimate the genetic distances among these breeds and to perform a phylogenetic analysis.

5.1 MATERIALS AND METHODS

5.1.1 Animals

The animals sampled from three of the four breeds of cattle (CN, JE and BS) were chosen using the computer program outlined in Study I (see also Appendix Tables 1-11). Briefly, pedigree information was obtained from Canadian Livestock Records Corporation (CLRC) and the Holstein Association of Canada. The pedigree data were edited using a SAS computer program (SAS, 1990) and the additive relationship coefficients among the individual animals (Cruden, 1949; Emik and Terrill, 1949; Van Vleck, 1993) were determined. Every attempt was made to sample those animals that were least closely related while at the same time being located at a reasonable driving distance from Ottawa. DNA samples of unrelated HO animals were provide by Dr. Denis Petitclerc of Agriculture and Agri-Food Canada's Dairy

and Swine Research and Development Centre in Lennoxville, Québec.

5.1.2 Extraction of Total Genomic DNA from Whole Blood Samples

Total genomic DNA was extracted from individual blood samples using the Super Quik-Gene DNA isolation kit (Analytical Genetic Testing Center, Inc., Denver, Colorado, USA) as was described in Study II. Briefly, blood samples were collected from individual animals in 8.5 ml ACD-Vaccutainer tubes (VWR Scientific, Ville Mont-Royal, Québec, Canada; ACD: 22.0 g/l trisodium citrate, 8.0 g/l citric acid, and 24.5 g/l dextrose) and kept at 4 °C until the time of DNA extraction. Every effort was made to extract the DNA from the blood samples within a day or two of collection. Blood samples were first centrifuged at 2500 rpm for 20 min. in a Jouan CR 4 22 refrigerated centrifuge (Canberra Packard Canada, Mississauga, Ontario, Canada). They were then allowed to stand at 4°C for 2 hrs in order to allow the white blood cells to aggregate and make removal of the buffy coat more efficient. As much of the plasma as possible was then removed from each tube and the buffy coat and approximately 1 ml of the underlying red blood cells were transferred to a 15 ml polypropylene Falcon tube (Fisher Scientific, Nepean, Ontario, Canada). RBC lysis buffer (1X) provided in the extraction kit was then added to each tube to a total volume of 13 ml. The mixture was mixed for 15 min. at room temperature and then centrifuged for 20 min. at 2500 rpm. After centrifugation, the supernatant was decanted and each tube

was washed with 2-3 ml of RBC lysis buffer (1X) without disturbing the WBC pellet to remove any excess stroma and supernatant binding to the tube. A total volume of 1.5 ml WBC lysis buffer was added to each tube and the pellet was broken up using a Pasteur pipet. The tubes were then incubated at 55 °C for 30 min. After the incubation, 0.2ml of 10% SDS and 0.5 ml of Protein Precipitating Reagent were added to each tube and the mixture was shaken vigorously by hand for 30 sec. All tubes were then incubated for 15 min in a 55 °C water bath. After the incubation period, the tubes were again centrifuged for 20 min. at 2500 rpm in order to pellet the precipitated protein. The clear supernatant was then transferred to a new 15 ml polypropylene Falcon tube and 2 volumes of room temperature absolute ethanol were added. The tubes were inverted gently until the precipitated DNA could be seen as a mass of white, stringy fibres. Finally, the precipitated DNA was spooled on to a plastic inoculation loop and transferred to a 1.5 ml microcentrifuge tube containing Tris-EDTA (TE) buffer (10 mM Tris-base, 1 mM EDTA, pH 7.4).

5.1.3 PCR Procedure

The D-loop of bovine mtDNA was amplified with a set of primers provided by Dr. James Derr of Texas A & M University. They generated a distinct PCR product of approximately 1100 bp in size. The set consisted of primers complimenting the bovine sequence in the conserved threonine

tRNA region of mtDNA (forward 5'- AGAGAAGGAGAACAACTAACCTCC-3'; position 15695) and the 12S rRNA gene (reverse 5'AACAGGAAGGC TGGGACC- 3'; position 457). All of the primer position numbers indicated correspond to Anderson et al. (1982) and the primers were designed to work under identical PCR conditions. The PCR was carried out in a 90 µl reaction volume using 66.5µl of sterile, deionized, distilled water, 9.0 µl of 10X PCR buffer, 6.0 µl of 25mM $MgCl_2$, 2.5 µl of a mixture containing 10mM of each of the dNTPs, 2.5 µl of DMSO, 1.5 µl of each of the two primers and 0.5 µl of AmpliTaq DNA polymerase (Perkin Elmer, Applied Biosystems Canada Inc., Mississauga, Ontario, Canada). A total of 35 cycles consisting of 94 °C for 1 min., 55 °C for 1 min. and 74 °C for 1 min. were used.

5.1.4 Purification of Amplified DNA

Before sequencing, the amplified DNA products were first purified from the primers, nucleotides and polymerases in the reaction mixture using the *QIAquick PCR Purification Kit* (QIAGEN Inc., Chatsworth, California, USA). Briefly, 5 volumes of *PB* buffer were added to 1 volume of the reaction mixture and mixed thoroughly. The sample was then loaded onto a *QIAquick* spin column in a 2 ml collection tube and centrifuged for 1 min. at maximum speed in a bench top microcentrifuge. The flowthrough fraction was drained and the column was washed with 0.75 ml of *PE* buffer by centrifuging for 1 min. at maximum speed. The *PE* buffer flowthrough was

again drained and the columns were spun once more. The *QIAquick* column was then placed in a clean 1.5 ml microcentrifuge tube and the DNA was eluted using 30-40 µl of sterile water.

Following the above purification, the purity and quantity of amplified DNA was estimated by running the samples with a ΦX174 standard on a 1.5% agarose gel. Occasionally, the low quantity of amplified product found in the some of the samples necessitated the reamplification and purification of the fragments under the same conditions described above.

5.1.5 Sequencing

Sequencing of all amplified D-loop fragments was performed using an ABI Prism 377 automatic sequencer (Perkin-Elmer, Forest City, California, USA; eg. Appendix Figure 2). The actual sequencing reaction was carried out using the *ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction Kit.* Four primers were used in the sequencing process: 1) THR (forward, 5′-GGAGAACAACTAACCTCC-3′), 2) Loftus et al. (forward, 5′-CTGCAG TCTCACCATCAACC-3′), 3) 1975 (reverse, 5′- CTGGACTTAACTGCA TCTTGAGC-3′) and 4) G-Jump (reverse, 5′- CAATTTAGCACTCCAA ACAAAGTC-3′). All were designed to work under identical conditions.

5.1.6 Sequence Processing and Phylogenetic Analysis

The sequences obtained through the automatic sequencing protocol described above were aligned using the Clustal V computer program (Higgins et al., 1992). Distances calculated were corrected for multiple hits by the two-parameter method of Kimura (1980) and sites representing gaps were excluded from the analysis. Trees were constructed under the criterion of maximum parsimony using the "branch and bound" search method in the PAUP computer package (Swofford, 1993). In the latter analysis, it was necessary to decrease the number of animals included in the data set as PAUP was unable to resolve the full set. All unique sequences found in CN cattle were thus included along with 2 unique sequences from each of the other 3 breeds (see Tables 1-3; for the HO breed animals 4 and 6 were used). An N'Dama and a Tharparkar sequence derived by Loftus et al. (1994a) were used as the outgroup for the analysis.

5.2 RESULTS

5.2.1 Mitochondrial DNA Variation

Mitochondrial DNA D-loop sequences consisting of 814 bp were determined for 20 CN, 9 BS , 10 JE and 10 HO animals. Among the 49 animals of the four breeds studied, 36 different mitochondrial haplotypes were identified. Of the 20 CN animals, 12 possessed haplotypes that differed from

all others in the breed. Similarly, 8 of the 9 BS animals, all 10 JE animals and 9 of the 10 HO animals carried haplotypes that were unique within their particular breed. When the populations were considered together, 11 of the CN, 6 of the BS, 9 of the JE and 8 of the HO cattle possessed haplotypes that were unique both within their own breed and within the general population. One haplotype was shared between the CN and BS breeds and one haplotype was shared between the BS, JE and HO breeds. Polymorphisms creating these 36 unique haplotypes were observed at 55 sites. The majority (56%) of the differences seen were the result of nucleotide substitution events. Most of these (19) were the result of transitions. Twelve transversion events, 11 deletions, 12 insertions and 1 inversion were observed. Insertions and deletions were of a single nucleotide in all cases.

Average pairwise sequence divergence estimates between the breeds and among animals within each breed were generated (Table 13). Intra-breed variability ranged from 0.29% to 0.73% while the inter-breed variability ranged from 0.43% to 0.80%. The intra-breed variability observed was generally of the same magnitude as the inter-breed variability. Only BS cattle were an exception to this rule. The variability seen within this breed of cattle was much lower than the variability observed between the BS breed and the other breeds studied.

The mtDNA D-loop sequences that were generated in this study were also compared to published D-loop sequences for one animal each of the

Table 13. Inter-breed and intra-breed (on diagonal) variability estimated from the D-loop sequences of four cattle breeds.

	CN	BS	JE	HO
CN	0.57	0.45	0.80	0.64
BS		0.29	0.56	0.43
JE			0.73	0.68
HO				0.61

CN= Canadienne, BS= Brown Swiss, JE= Jersey and HO= Holstein

Table 14. Average variability among Canadienne (CN), Brown Swiss (BS), Jersey (JE) and Holstein (HO) D-loop sequences and a representative published* sequence from six other breeds.

	FR	HE	BU	ND	HA	TP
CN	0.69	0.87	0.95	1.19	5.63	5.77
BS	0.48	0.65	0.77	1.04	5.48	5.66
JE	0.95	1.13	1.25	1.54	5.68	5.80
HO	0.74	0.92	1.02	1.28	5.69	5.84

CN= Canadienne, BS= Brown Swiss, JE= Jersey, HO= Holstein, FR= Friesian, HE= Hereford, BU= Butana, ND= N'Dama, HA= Hariana and TP= Tharparkar.
* Loftus et al., 1994a

Friesian (FR), Hereford (HE) , Butana (BU), N'Dama (ND), Hariana (HA) and Tharparkar (TP) breeds (Loftus et al., 1994a). The results of this comparison are shown in Table 14. As expected, all four Canadian populations were shown to be most closely related to the Friesian breed of dairy cattle. The beef breed Hereford was the next most closely related breed followed by the African Butana and N'Dama breeds. The greatest variability in D-loop sequence was seen between the four Canadian populations and the Indian Hariana and Tharparkar breeds.

5.2.2 *Phylogenetic Analysis*

The results of the phylogenetic analysis performed using all the unique CN mtDNA D-loop sequences and two unique sequences randomly selected from those of each of the other three breeds are given in Figures 2 and 3. In total, 264 possible maximum parsimony trees were generated using the "branch and bound" search method in PAUP. Relevant statistics for the trees were as follows: number of steps: 91, homoplasy and consistency indices: 0.143 and 0.857, respectively, homoplasy and consistency indices excluding uninformative characters: 0.325 and 0.675, respectively, and retention index: 0.723. Figures 2 and 3 give a sample of two of these trees while Figure 4 contains the strict consensus tree generated from the 264 possible trees. As can be seen, no distinct separation was observed between the CN breed and the other 3 Canadian cattle populations. It is interesting to note that Friesian

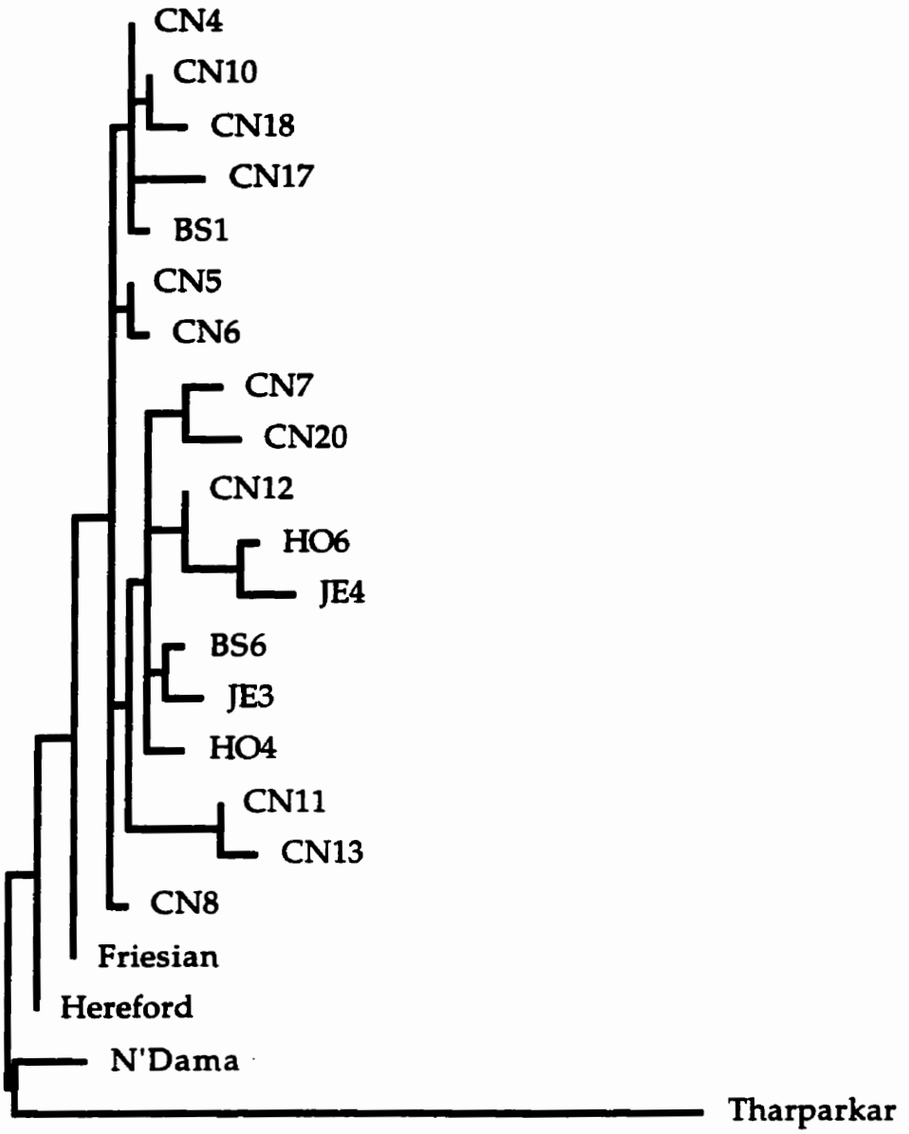


Figure 2. Phylogram depicting one of the 264 maximum parsimony trees generated using the "branch and bound" search in the PAUP program. Branch lengths depict the amount of inferred character change.

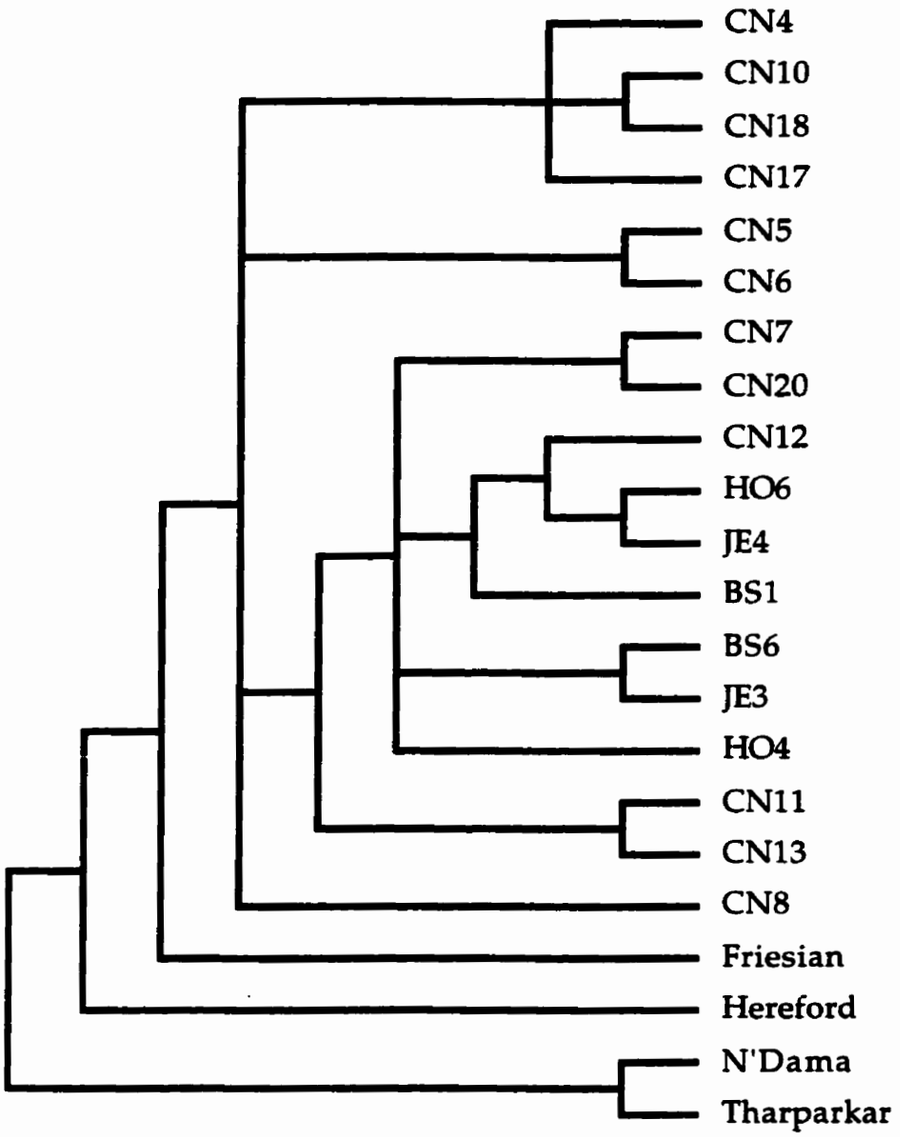


Figure 3. Rectangular cladogram depicting another of the 264 maximum parsimony trees generated using the "branch and bound" search in the PAUP program. Branch lengths are arbitrary.

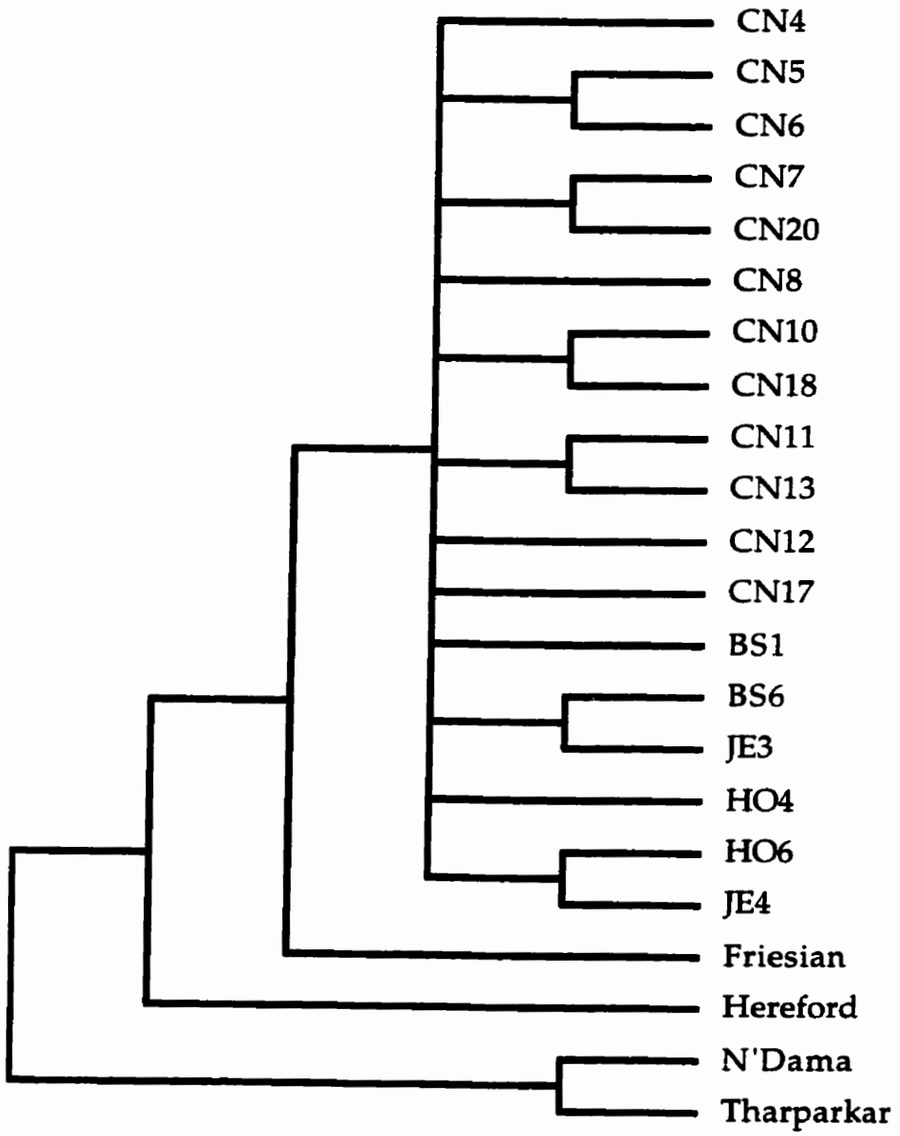


Figure 4. Strict consensus tree generated from the 264 maximum parsimony trees found using the "branch and bound" search of the PAUP program.

cattle appeared to be slightly separated from the four populations focused on in this study. As expected, N'Dama and Tharparkar cattle, which were used as the outgroup for this analysis, were most distinct from the other populations.

5.3 DISCUSSION

Previous studies examining differences in bovine mtDNA, with the exception of that by Loftus et al. (1994a), have traditionally used polymorphism in restriction enzyme cleavage pattern (Watanabe et al., 1985a; Watanabe et al., 1989; Bhat et al., 1990). These studies were successful in showing breed differences but, as RFLP analysis cannot identify variation that is present at sites other than those specific for restriction enzyme cleavage, a fair amount of variation was probably missed by these studies. While Loftus and coworkers failed to show distinct phylogenetic lineages within the European cattle breeds they studied, the pairwise distances they were able to calculate showed that variation does exist between these breeds. As a result, it was of interest to examine the variation that exists in the D-loop sequence of several Canadian populations of dairy cattle, one of which, the CN, is considered to be endangered.

The nature of the polymorphisms seen in the various sequences in the present study show that there was a transitional bias in nucleotide substitutions. This finding is in agreement with that reported by Loftus et al.

(1994a), although their data favoured transitions to an even greater extent, and has been found to be a characteristic of mammalian mitochondrial evolution (Wilson et al., 1985). The present data set also showed a larger number of insertion-deletion events than these authors reported finding. The significance of this observation is unclear.

As can be seen from Table 13, the inter-breed variability or distance between CN and BS cattle was found to be lower than between CN and JE or HO cattle. This finding is not unexpected given the fact that, beginning in the year 1969, CN cattle were heavily crossbred with BS. As a result, today many of the animals comprising the CN breed contain some BS blood. This is true also for the animals sampled. A lower amount of variation in the D-loop sequence between the two breeds is, therefore, likely to result from the introduction of BS maternal lineages into the CN breed.

Another interesting observation that can be made from Table 13 is that among the breeds, the most variability appears to occur between a particular breed and JE cattle. This is particularly noteworthy for the CN breed. Canadienne cattle are believed to have originated from animals brought to North America from the Normandy and Brittany regions of France. It has, therefore, been postulated that they come from the same general lineage as Jersey cattle, among others. It would, therefore, have been reasonable to expect the D-loop of CN cattle to be somewhat less variable when compared to JE cattle. In fact, the opposite appears to be true. Of the three breeds studied,

the CN shows the greatest amount of variation in its D-loop sequence in relation to the JE.

Table 14, which presents the average variability in D-loop sequence between the four Canadian breed populations and representative published sequences for FR, HE, BU, ND, HA and TP breeds, shows that the HO and FR breeds appear to have differentiated somewhat from one another at the mtDNA level. In fact, the variability between the D-loop of HO and FR cattle was greater than those between HO and CN, JE, or BS breeds. This result is again somewhat surprising since HO cattle were only recognized as a separate breed from the FR very recently.

Bhat et al. (1990) and Watanabe et al. (1985a) characterized variations in mtDNA restriction enzyme cleavage patterns within and between various European, North American, Indian and Asian cattle breeds. Similarly, Loftus et al. (1994a, b) have demonstrated, using both restriction fragment length polymorphism (RFLP) analysis of mtDNA and mtDNA D-loop sequence variation, that, while numerous differences between animals and between breeds can be identified, sequence variation between European and African *Bos taurus* animals and breeds is not significant enough to point to more than one major mitochondrial lineage. These results are confirmed by the present study. All of the CN, BS, JE and HO animals studied group together on the phylogenetic tree (Fig. 4). Only the *Bos indicus* TP breed shows any major separation. Furthermore, CN cattle do not appear to possess any

mitochondrial variation that would uniquely separate them from the remaining *Bos taurus* breeds in the form of a minor lineage.

The number of different mitochondrial haplotypes found in CN cattle give some reasons for concern about this breed. As was mentioned earlier, 57% of CN animals carried haplotypes that were unique within the breed. This compares with 89% of BS animals, 100% of JE animals and 90% of HOs. It is thus evident that the decrease in the size of the CN population may have resulted in a decrease in the percentage of animals carrying unique maternal lineages compared with the much more populous JE and HO breeds.

6. GENERAL CONCLUSIONS

The quantitative and molecular genetic methodologies employed to study the genetic distance among CN, BS, JE and HO populations in Canada and to examine intra-breed relationships and levels of inbreeding combine to provide us with an interesting picture of the breeds. Such information is vital for informed decision-making to take place in individual, national and international conservation programs.

Study I used pedigree analysis to select unrelated CN, BS and JE cattle and to estimate the inbreeding coefficients and rate of inbreeding of these cattle and HO bulls using the tabular method of calculating the additive relationship matrix and/or Wright's formula. Although inbreeding has been an important issue for many dairy cattle breeders for a number of years, it appears that at the present time the overall level of inbreeding is not high enough to be of concern. Signs of possible future danger do exist, however. The level of inbreeding of HO bulls in the A.I. studs has increased at an annual rate of 0.37% over the last few years. In addition, there are only an average of 24 breeding bulls available in the CN breed and the ten top bulls account for an alarmingly large proportion (61%) of animals born in recent years in the CN population and of HO bulls presently found in the various A.I. studs. The effective number of breeding parents in the CN breed and the number of registrations seen in recent years suggest that the Canadienne

breed can be classified as endangered.

In study II, microsatellite analysis at 15 loci was used to measure the genetic distance among the four cattle breeds. Genetic distances were estimated using three established methods for calculating such values. As would be expected considering the different assumptions associated with each of these genetic distance measures, the results obtained tended to vary between methods. All three generated genetic distances between the breeds that were significantly different from zero, however, confirming that genetic variation definitely exists among the breeds. Ranking of breeds in order of increasing genetic distance was not possible due to the large standard error associated with individual distance estimates. Nevertheless, CN cattle tended to be most closely related to BS and HO cattle. On the other hand, BS and HO breeds appeared to have the closest relationship of all of the breeds studied, In contrast, the JE breed tended to be most distantly related to all of the breed examined. Larger numbers of animals and/or loci may need to be sampled before a definitive ranking of breeds according to genetic distance is possible. The increase in precision of the estimates obtained versus the cost associated with sampling more animals and/or loci would have to be evaluated when such data becomes available to ascertain the benefit of such an exercise.

Analysis of the mtDNA D-loop region in study III indicated that CN cattle exhibited the most variability in D-loop sequence when compared to JE cattle and were found to be least variable in comparison to the BS breed.

Again HO cattle were found to be fairly similar to CN cattle, exhibiting approximately the same comparative amount of variation in D-loop sequence as was found between the Friesian and the HO. Phylogenetic analysis revealed that the strict consensus tree obtained from the sequence data did not show any significant separation among the breeds studied. Animals of all four breeds were grouped together on the tree indicating that the variation seen in their D-loop sequences was not evolutionarily significant.

The finding that CN cattle tend to be more closely related to HO cattle than to the JE breed raises some important questions that need to be addressed in the future. It would be interesting to determine whether this closeness of relationship is due largely to the presence of BS blood in the CN breed as a result of the crossbreeding that occurred after 1969 or whether some HO blood made its way into the breed during its development. BS cattle did, after all, tend to show the closest relationship to HO cattle of any of the breeds examined in both the microsatellite and mitochondrial D-loop analysis.

It is important to realize from the results of these studies, that making decisions about breed conservation will never be easy. As was stated above, genetic variation among breeds exists and concerns about inbreeding and the genetic base of at least some of these breeds are not unfounded. More specific research addressing the genetic uniqueness of a particular breed is, however, required. This may be possible through the identification of breed specific markers or alleles and requires that the cost of analysis be reduced

significantly. In the meantime we must consider the new scientific evidence we have in combination with historical data and the cultural significance of a breed when deciding whether conservation efforts are warranted.

7. REFERENCES

Anderson, S., Bankier, A.T., Barrell, B.G., deBruijn, M.H.L., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., Schreier, P.H., Smith, A.J.H., Staden, R., and Young, I.G., 1981. Sequence and organization of the human mitochondrial genome. Nature 290: 457-465.

Anderson, S., deBruijn, M.H.L., Coulson, A.R., Eperon, I.C., Sanger, F., and Young, I.B., 1982. Complete sequence of bovine mitochondrial DNA. Conserved features of the mammalian mitochondrial genome. J. Mol. Biol. 156: 683-717.

Arnason, U., and Johnsson, E., 1992. The complete mitochondrial DNA sequence of the harbor seal, *Phoca vitulina.* J. Mol. Evol. 34: 493-505.

Arnason, U., Gullberg, A., and Widegren, B., 1991. The complete nucleotide sequence of the mitochondrial DNA of the fin whale, *Balaenoptera physalus.* J. Mol. Evol. 33: 556-568.

Arnason, U., Gullberg, A., Johnsson, E., and Ledje, C., 1993. The nucleotide sequence of the mitochondrial DNA molecule of the grey seal, *Halichoerus grypus,* and a comparison with mitochondrial sequences of other true seals. J. Mol. Evol. 37: 323-330.

Ashley, M.V., Laipis, P.J., and Hauswirth, W.W., 1989. Rapid segregation of heteroplasmic bovine mitochondria. Nucleic Acids Res. 17: 7325-7331.

Barendse, W., Armitage, S.M., Kossarek, L.M., Shalom, A., Kirkpatrick, B.W., Ryan, A.M., Clayton, D., Li, L., Neibergs, H.L., Zhang, N., Grosse, W.M., Weiss, J., Creighton, P., McCarthy, F., Ron, M., Teale, A.J., Fries, R., McGraw, R.A., Moore, S.S., Georges, M., Soller, M., Womack, J.E., and Hetzel, D.J.S., 1994. A genetic linkage map of the bovine genome. Nature Genetics 6: 227-235.

Barker, J.S.F., 1994. A global protocol for determining genetic distances among domestic livestock breeds. Proceedings of the 5th World Congr. Genet. Appl. Livest. Prod. 21: 501-508.

Bhat, P.P., Mishra, B.P., and Bhat, P.N., 1990. Polymorphism of mitochondrial DNA (mtDNA) in Cattle and Buffaloes. Biochem. Genet. 28: 311-318.

Bernier, J.-G., 1995. Canadienne bovine breed report. Proc. Third Global Conf. on Conserv. Domest. Anim. Genet. Resour., Rare Breeds International, Kenilworth, UK. pp. 148-156.

Bibb, M.J., Van Etten, R.A., Wright, C.T., Walberg, M.W., and Clayton, D.A., 1981. Sequence and gene organization of mouse mitochondrial DNA. Cell 26: 167-180.

Bishop, M.D., Kappes, S.M., Keele, J.W., Stone, R.T., Sunden, S.L.F., Hawkins, G.A., Toldo, S.S., Fries, R., Grosz, M.D., Yoo, J., and Beattie, C.W., 1994. A genetic linkage map for cattle. Genetics 136: 619-639.

Bixby, D.E., Christman, C.J., Ehrman, C.J., and Sponenberg, D.P., 1994. Taking Stock. The North American Livestock Census. McDonald and Woodward Publishing Co., Blacksburg, USA, 182 pp.

Bonczek, R.R. and Young, C.W., 1980. Comparison of the ancestry and inbreeding of two lines of Holstein cattle and the Holstein breed. J. Dairy Sci. 63 (Suppl. 1): 107.

Brown, W.M., George Jr., M., and Wilson, A.C., 1979. Rapid evolution of animal mitochondrial DNA. Proc. Natl. Acad. Sci., USA 76: 1967-1971.

Buchanan, F.C., Adams, L.J., Littlejohn, R.P., Maddox, J.F., and Crawford, A.M., 1994. Determination of evolutionary relationships among sheep breeds using microsatellites. Genomics 22: 397-403.

Cann, R.L., Stoneking, M., and Wilson, A.C., 1987. Mitochondrial DNA and human evolution. Nature 325: 31-36.

Casanova, L., Hagger, C., Kuenzi, N., and Schneeberger, M., 1992. Inbreeding in Swiss Braunvieh and its influence on breeding values predicted from a repeatability animal model. J. Dairy Sci. 75: 1119-1126.

CLRC, 1995. Annual Report. Canadian Livestock Records Corporation, Ottawa, Canada, 68 pp.

Couture, J.A., 1909. French Canadian cattle. Livre de généalogie du bétail canadien 1:13-18.

Cruden, D., 1949. The computation of inbreeding coefficients for closed populations. J. Hered. 40: 248-251.

Cunningham, P., Loftus, R.T., MacHugh, D.E., and Bradley, D.G., 1994. Molecular evolution of African, European and Asian Cattle. Proc. 5th World Congr. Genet. Appl. Livest. Prod. 21: 86-89.

Cunningham, P., 1996. Genetic diversity in domestic animals: Strategies for conservation and development. In: Beltsville Symposia in Agricultural Research XX- Biotechnology's Role in the Genetic Improvement of Farm Animals, American Society of Animal Science, USA, pp. 13-23.

Di Rienzo, A., and Wilson, A.C., 1991. Branching pattern in the evolutionary tree for human mitochondrial DNA. Proc. Natl. Acad. Sci., USA 88: 1597-1601.

Emik, L.O., and Terrill, C.E., 1949. Systematic procedures for calculation inbreeding coefficients. J. Hered. 40: 51-55.

Epstein, H., 1971. The Origin of the Domestic Animals of Africa. Africana Publishing Corporation, New York, USA.

F.A.O., 1993. World Watch List for Domestic Animal Diversity, 1st edition. R. Loftus and B. Scherf, editors. Rome, Italy.

Forbes, S.H., Hogg, J.T., Buchanan, F.C., Crawford, A.M., and Allendorf, F.W., 1995. Microsatellite evolution in congeneric mammals: domestic and bighorn sheep. Mol. Biol. Evol. 12: 1106-1113.

Fortin, L. de G., 1940. Histoire de la race bovine Canadienne. La Bonne Terre, Ecole Supérieure d'Agriculture, Ste-Anne de la Pocatière, Québec, Canada, 286 pp.

Falconer, D.S., 1989. Introduction to Quantitative Genetics, 3rd. edition. Longman Scientific and Technical, U.K. 438 pp.

Frankham, R., 1994. Conservation of genetic diversity for animal improvement. Proc. 5th World Congr. Genet. Appl. Livest. Prod. 21: 385-392.

Gadaleta, G., Pepe, G., De Candia, G., Quagliariello, C., Sibisa, E., and Saccone, C., 1989. The complete nucleotide sequence of the *Rattus norvegicus* mitochondrial genome: cryptic signals revealed by comparative analysis between vertebrates. J. Mol. Evol. 28: 497-516.

Giles, R.E., Blanc, H., Cann, H.M., and Wallace, D.C., 1980. Maternal inheritance of human mitochondrial DNA. Proc. Natl. Acad. Sci., USA 77: 6715-6719.

Goldstein, D.B., Ruis-Linares, A., Feldman, M., and Cavelli-Sforza, L.L., 1995a. An evaluation of genetic distances for use with microsatellite loci. Genetics 139: 463-471.

Goldstein, D.B., Ruis-Linares, A., Feldman, M., and Cavelli-Sforza, L.L., 1995b. Genetic absolute dating based on microsatellites and the origin of modern humans. Proc. Natl. Acad. Sci., USA 92: 6720-6727.

Gonzalez, P., Tunon, M.J., and Vallejo, M., 1988. Genetic diversity in Spanish cattle breeds inferred form the distribution of hemotypes. Biochem. Syst. Ecol. 16: 237-239.

Graml, R., Schmid, D.O., Erhard, L., Buchberger, J., Ohmayer, G., and Pirchner, F., 1986. Verwandschaft des Murnau-Werdenfelser Rindes zu anderen Rassen. Bay. Landw. Jahrb. 273-281.

Grisdale, J.H., 1909. French Canadian cattle. Livre de généalogie du bétail canadien 1:40-49

Gross, D.S., and Garrard, W.T., 1986. The ubiquitous potential Z-forming sequence of eukaryotes, $(dT\text{-}dG)_{nc}$ $(dC\text{-}dA)_{n}$, is not detectable in the genomes of eubacteria, archaebacteria, or mitochondria. Mol. Cell. Biol. 6: 3010-3013.

Gyllensten, U., Wharton, D., Josefsson, A., and Wilson, A.C., 1991. Paternal inheritance of mitochondrial DNA in mice. Nature 352: 255-257.

Hamada, H., Seidman, M., Howard, B.H., and Gorman, C.M., 1984. Enhanced gene expression by the poly(dT-dG) poly(dC-dA) sequence. Mol. Cell. Biol. 4: 2622-2630.

Hayashi, J.I., Yonekawa, H., Gotoh, O., Watanabe, J., and Tagashira, Y., 1978. Strictly maternal inheritance of rat mitochondrial DNA. Biochem. Biophys. Res. Commun. 83: 1032-1038.

Hedrick, P.W., 1975. Genetic similarity and distance: Comments and comparisons. Evolution 25: 276-280.

Hermas, S.A., Young, C.W., and Rust, J.W., 1987. Effects of mild inbreeding on productive and reproductive performance of Guernsey cattle. J. Dairy Sci. 70: 712-715.

Higgins, D.G., Bleasby, A.J., and Fuchs, R., 1992. Clustal V: Improved software for multiple sequence alignment. Comput. Appl. Biosci. 8: 189-191.

Hillis, D.M., Mable, B.K., Larson, A., Davis, S.K., and Zimmer, E.A., 1996. Nucleic acids IV: Sequencing and cloning. In: Molecular Systematics (D.M. Hillis, C. Moritz and B.K Mable, editors), Sinauer Associates Inc., Sunderland, Mass., USA.

Hoelzel, A.R. and Bancroft, D.R., 1992. Statistical analysis of variation, In: Molecular Genetic Analysis of Populations A Practical Approach, Oxford University Press, N.Y., pp. 297-304

Hudson, G.F.S. and Van Vleck, L.D., 1984. Effects of inbreeding on milk and fat production, stayability, and calving interval of registered Ayrshire cattle in the Northeastern United States. J. Dairy Sci. 67: 171-179.

Hutchinson, C.A., III, Newbold, J.E., Potter, S.S., and Edgell, M.H., 1974. Maternal inheritance of mammalian mitochondrial DNA. Nature 251: 536-538.

Kimmel, M., Chakraborty, R., Stivers, D.N., and Deka, R., 1996. Dynamics of repeat polymorphisms under a forward-backward mutation model: within- and between-population variability at microsatellite loci. Genetics 143: 549-555.

Kimura, M., 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16: 111-120.

Kimura, M., and Crow, J.F., 1964. The number of alleles that can be maintained in a finite population. Genetics 49: 725-738.

Kimura, M., and Ohta, T., 1978. Stepwise mutation model and distribution of allelic frequencies in a finite population. Proc. Natl. Acad. Sci., USA 75: 2868-2872.

Kit, S., 1961. Equilibrium centrifugation in density gradients of DNA preparations from animal tissues. J. Mol. Biol. 3: 711-716.

Koehler, C.M., Lindberg, G.L., Brown, D.R., Beitz, D.C., Freeman, A.E., Mayfield, J.E., and Myers, A.M., 1991. Replacement of bovine mitochondrial DNA by a sequence variant within one generation. Genetics 129: 247-255.

Laipis, P.J., Van de Walle, M.J., and Hauswirth, W.W., 1988. Unequal partitioning of bovine mitochondrial genotypes among siblings. Proc. Natl. Acad. Sci., USA 85: 8107-8110.

Lan, H., and Shi, L., 1993. The origin and genetic differentiation of native breeds of pigs in Southwest China: An approach from mitochondrial DNA polymorphism. Biochem. Genet. 31: 51-60.

Lewin, B., 1990. Genes IV, 4th. edition. Oxford University Press. 857 pp.

Loftus, R.T., MacHugh, D.E., Bradley, D.G., Sharpe, P.M., and Cunningham, E.P., 1994a. Evidence for two independent domestications of cattle. Proc. Natl. Acad. Sci., USA 91: 2757-2761.

Loftus, R.T., MacHugh, D.E., Ngere, L.O., Balain, D.S., Badi, A.M., Bradley, D.G., and Cunningham, E.P., 1994b. Mitochondrial genetic variation in European, African and Indian cattle populations. Animal Genetics 25: 265-271.

Lush, J.L., Holbert, J.C., and Willham, O.S., 1936. Genetic history of the Holstein-Fresian cattle in the United States. J. Hered. 27: 61.

MacHugh, D.E., Loftus, R.T., Bradley, D.G., Sharp, P.M., and Cunningham, P., 1994. Microsatellite DNA variation within and among European cattle breeds. Proc. R. Soc. Lond. B. 256: 25-31.

MacKay, S.L.D., Ghivizzani, S.C., Madsen, C.S., Tanhauser, S.M., Laipis, P.J., and Hauswirth, W.W., 1993. Transcribed heteroplasmic repeated sequences in the porcine mitochondrial DNA D-loop region. J. Mol. Evol. 37: 36-47.

Maijala, K., Cherekaev, A.V., Devillard, J.-M., Reklewski, Z., Rognoni, G., Simon, D.L., and Steane, D.E., 1984. Conservation of animal genetic resources in Europe. Final report of an EAAP working party. Livest. Prod. Sci. 11: 3-22.

Margulis, L., 1993. Symbiosis in Cell Evolution, 2nd. edition. W.H. Freeman and Company. 452 pp.

Medjugorac, I., Kustermann, W., Lazar, P., Russ, I., and Pirchner, F., 1994. Marker-derived phylogeny of European cattle supports demic expansion of agriculture. Animal Genetics 25 (Suppl. 1): 19-27.

Miglior, F., and Burnside, E.B., 1995. Inbreeding of Canadian Holstein cattle. J. Dairy Sci. 78: 1163-1167.

Miglior, F., Szkotnicki, B., and Burnside, E.B., 1992. Analysis of levels of inbreeding and inbreeding depression in Jersey cattle. J. Dairy Sci. 75: 1112-1118.

Miglior, F., Burnside, E.B., and Dekkers, J.C.M., 1995a. Nonadditive genetic effects and inbreeding depression for somatic cell counts of Holstein cattle. J. Dairy Sci. 78: 1168-1173.

Miglior, F., Burnside, E.B., and Kennedy, B.W., 1995b. Production traits of Holstein cattle: Estimation of nonadditive genetic variance components and inbreeding depression. J. Dairy Sci. 78: 1174-1180.

Minch, E., 1997. Microsat 1.5b, Stanford University, USA.

Moran, P.A.P., 1975. Wandering distributions and the electrophoretic profile. Theor. Pop. Bio. 8: 318-330.

Namikawa, T., Ito, S., and Amano, T., 1984. Genetic relationships and phylogeny of East and Southeast Asian cattle: genetic distance and principal component analysis. Z. Tierzüchtg. Züchtungsbiol. 101: 17-32.

Nei, M., 1972. Genetic distance between populations. Amer. Nat. 106: 283-292.

Nei, M., 1987. Molecular Evolutionary Genetics. Columbia University Press, New York.

Nei, M. and Takezaki, N., 1994. Estimation of genetic distance and phylogenetic trees from DNA analysis. Proc. 5th World Congr. Genet. Appl. Livest. Prod. 21: 405-412.

Pardue, M.L., Lowenhaupt, K., Rich, A., and Nordheim, A., 1987. $(dC\text{-}dA)_n$ $(dG\text{-}dT)_n$ sequences have evolutionarily conserved chromosomal locations in Drosophila with implications for roles in chromosome structure and function. EMBO J. 6: 1781-1789.

Quellar, D.C., Strassmann, J.E., and Hughes, C.R., 1993. Microsatellites and kinship. Trends Ecol. Evol. 8: 285-288.

Rafalski, J.A., and Tingey, S.V., 1993. Genetic diagonostics in plant breeding: RAPDS, microsatellites and machines. Trends in Genetics 9: 275-280.

Ramirez, V., Savoie, P., and Morais, R., 1993. Molecular characterization and Evolution of a duck mitochondrial genome. J. Mol. Evol. 37: 296-310.

Raymond, M., and Rousset, F., 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J. Heredity 86: 248-249.

Robertson, A., 1960. A theory of limits in artificial selection. Proc. Roy. Soc. London, B. 153: 234-249.

Saccone, C., Attimonelli, M., and Sbisa, E., 1987. Structural elements highly preserved during the evolution of the D-loop containing region in vertebrate mitochondrial DNA. J. Mol. Evol. 26: 205-211.

Saccone, C., Pesole, G., and Sbisa, E., 1991. The main regulatory region of mammalian mitochondrial DNA: Structure-function model and evolutionary pattern. J. Mol. Evol. 33: 83-91.

SAS Institute Inc., 1990. SAS/STAT® User's Guide. Version 6, Fouth Edition, Vol. 1 & 2, Cary, NC.

Schlötterer, C., and Tautz, D., 1992. Slippage synthesis of simple sequence DNA. Nucleic Acids Res. 20: 211-215.

Simon, D.L., 1984. Conservation of animal genetic resources- A review. Livest. Prod. Sci. 11: 23-36.

Slatkin, M., 1995. A measure of population subdivision based on microsatellite allele frequencies. Genetics 139: 457-462.

Smith, C., 1984. Genetic aspects of conservation in farm livestock. Livest. Prod. Sci. 11: 37-48.

Southern, S.O., Southern, P.J., and Dizon, A.E., 1988. Molecular characterization of a cloned dolphin mitochondrial genome. J. Mol. Evol. 28: 32-42.

Stallings, R.L., Ford, A.F., Nelson, D., Torney, D.C., Hildebrand, C.E., and Moyzis, R.K., 1991. Evolution and distribution of $(GT)_n$ repetitive sequences in mammalian genomes. Genomics 10: 807-815.

Stone, R.T., Pulido, J.C., Duyk, G.M., Kappes, S.M., Keele, J.W., and Beattie, C.W., 1995. A small-insert bovine genomic library highly enriched for microsatellite repeat sequences. Mammalian Genome 6: 714-724.

St-Pierre, A., 1936. Notes historiques sur la race bovine canadienne. Société des Éleveurs de Bovins Canadiens, 64 pp.

Swofford, D.L., 1993. PAUP: Phylogenetic analysis using Parsimony, Illinois Version 3.1.1. Formerly distributed by Illinois Natural History Survey, Champaign, IL.

Swofford, D.L., Olsen, G.J., Waddell, P.J., and Hillis, D.M., 1996. Phylogenetic inference. In: Molecular Systematics (D.M. Hillis, C. Moritz and B.K Mable, editors) Sinauer Associates Inc., Sunderland, Mass., USA.

Tautz, D., 1989. Hypervariability of simple sequences as a general source of polymorphic DNA markers. Nucleic Acids Res. 17: 6463-6471.

Tuggle, C.K., 1994. DNA Technology: Its use in animal breeding. Presented at a symposium on future genetics for the animal industry, St. Louis, MO, May 4.

Van Vleck, L.D., 1993. Selection index and introduction to mixed model methods. CRC Press Inc., Florida, USA, 481 pp.

Waring, M.J., and Britten, R.J., 1966. Nucleotide sequence repetition in a rapidly reassociating fraction of mouse DNA. Science 154: 791-794.

Watanabe, T., Hayashi, Y., Semba, R., and Ogasawara, N., 1985a. Bovine mitochondrial DNA polymorphism in restriction endonuclease cleavage pattern and the location of polymorphic sites. Biochem. Genet. 23: 947-957.

Watanabe, T., Hayashi, Y., Ogasawara, N., and Tomita, T., 1985b. Polymorphism of mitochondrial DNA in pigs based on restriction endonuclease cleavage patterns. Biochem. Genet. 23: 105-113.

Watanabe, T., Hayashi, Y., Kimura, J., Yasuda, Y., Saitou, N., Tomita, T., and Ogasawara, N., 1986. Pig mitochondrial DNA: Polymorphism, restriction map orientation, and sequence data. Biochem. Genet. 24: 385-396.

Watanabe, T., Masangkay, J.S., Wakana, S., Saitou, N., and Tomita, T., 1989. Mitochondrial DNA polymorphism in native Philippine cattle based on restriction endonuclease cleavage patterns. Biochem. Genet. 27: 431-438.

Watson, J.D., Gilman, M., Witkowski, J., and Zoller, M., 1992. Recombinant DNA, 2nd edition. Scientific American Books, New York, NY, USA. 626 pp.

Weaver, R.F., and Hedrick, P.W., 1997. Genetics, 3rd edition, Wm. C. Brown Publishers, USA.

Weber, J.L., and May, P.E., 1989. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am. J. Hum. Genet. 44: 388-396.

Weber, J.L., and Wong, C., 1993. Mutation of human short tandem repeats. Hum. Mol. Genet. 2: 1123-1128.

Wilson, A.C., Cann, R., Carr, S., George, M., Gyllensten, U.B., Helm-Bychowsky, K.M., Higuchi, R.G., Palumbi, S.R., Prager, E.M., Sage, R.D., and Stoneking, M., 1985. Mitochondrial DNA and two perspectives on evolutionary genetics. Biol. J. Linn. Soc. 26: 375-400.

Wright, S., 1931. Evolution in Mendelian populations. Genetics 16: 97-159.

Wright, S., 1969. Evolution and the Genetics of Populations. Vol. 2., University of Chicago Press, Chicago, pp. 295.

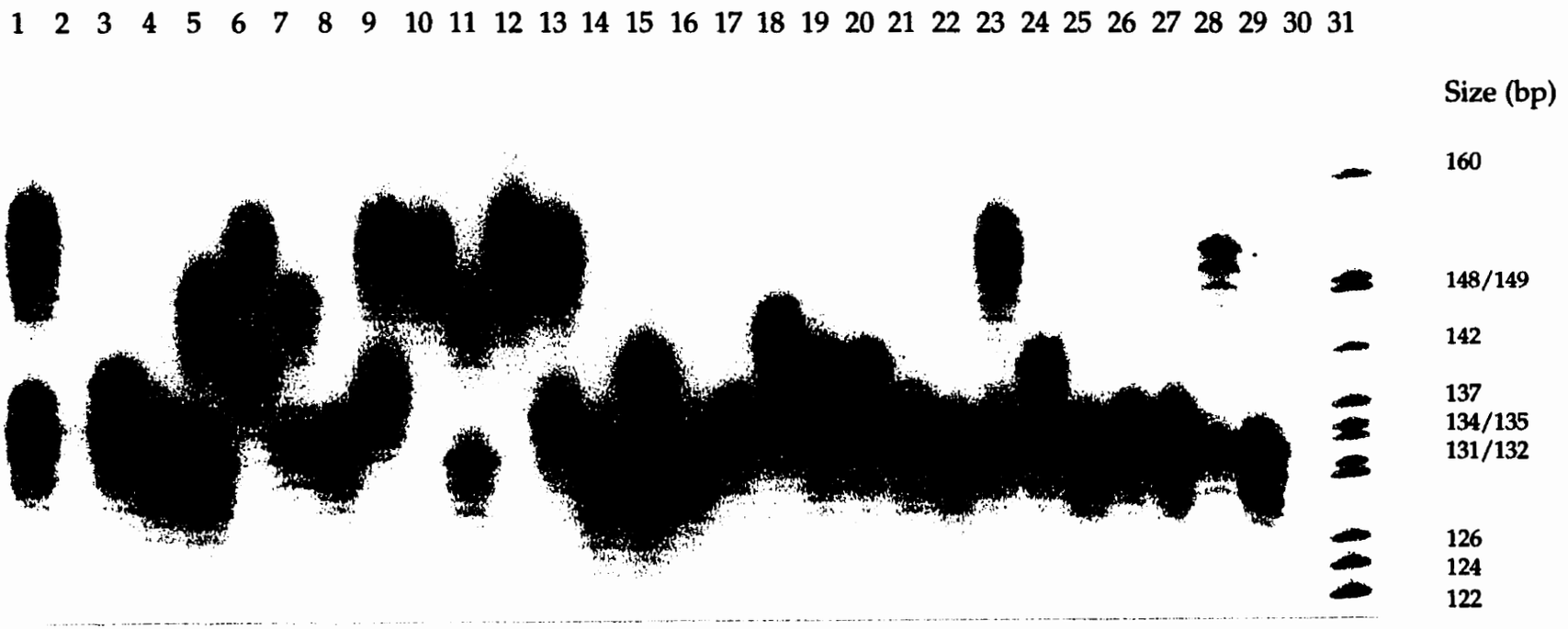
Wu, M.C., Shiue, Y.L., Liou, J.F., and Luoh, Y.S., 1994. Mitochondrial DNA polymorphism of native livestock breeds of Taiwan. The Third Global Conference on Conservation of Domestic Animal Genetic Resources Program and Abstracts. 50 pp.

Young, C.W., 1973. Inbreeding and ancestry of registered Holstein-Fresian females born in 1970. J. Dairy Sci. 56: 659.

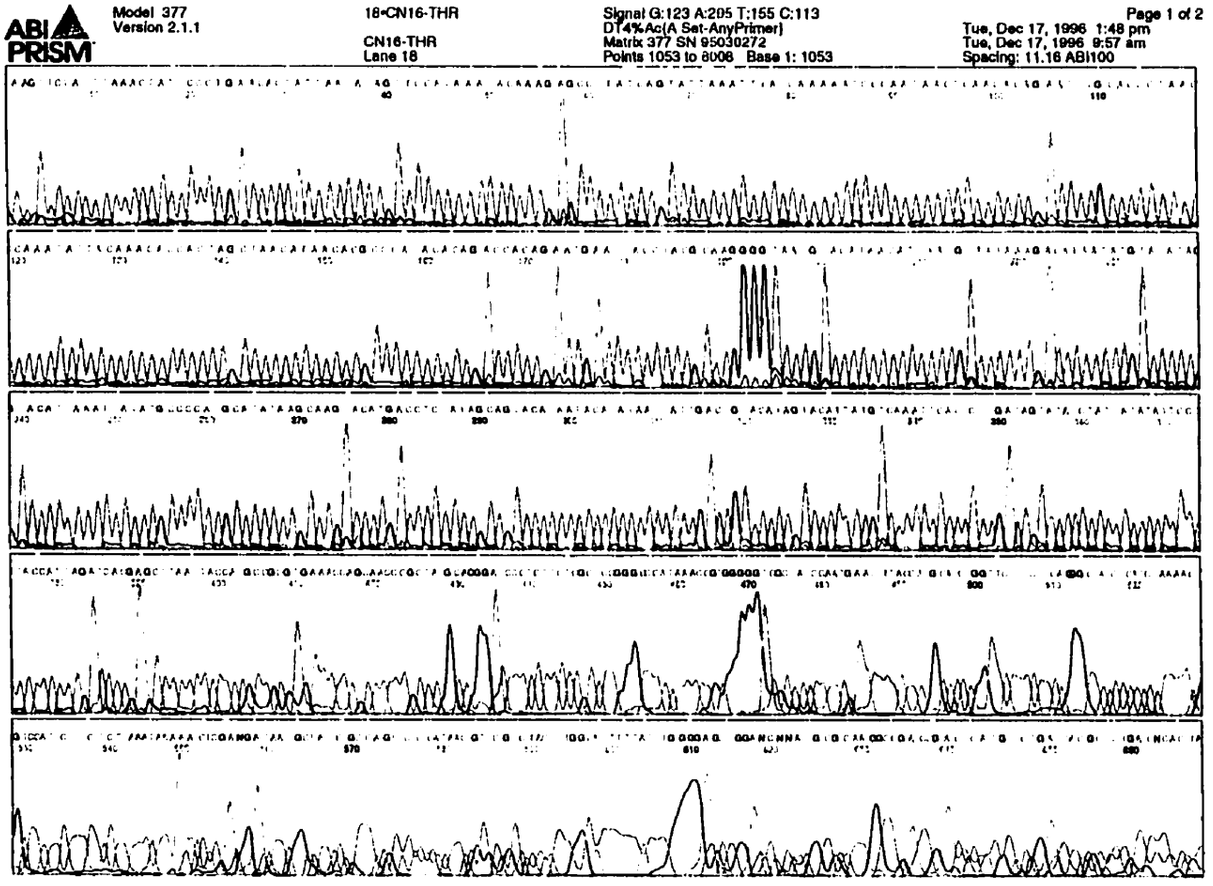
Young, C.W., 1984. Inbreeding and the gene pool. J. Dairy Sci. 67: 472-477.

Young, C.W., Bonczek, R.R., and Johnson, D.G., 1988. Inbreeding and relationship among registered Holsteins. J. Dairy Sci. 71: 1659-1666.

8. APPENDIX FIGURES



Appendix Figure 1. Example of an autoradiograph of a polyacrylamide gel for microsatellite locus BM226. Lanes represent the following animals: 1) CN4, 2) Blank, 3) CN6, 4) CN7, 5) CN8, 6) CN9, 7) CN10, 8) CN11, 9) CN12, 10) CN13, 11) CN14, 12) CN15, 13) BS4, 14) BS5, 15) BS6, 16) BS7, 17) BS8, 18) BS9, 19) BS10, 20) BS11, 21) BS12, 22) BS13, 23) BS14, 24) BS15, 25) JE4, 26) JE5, 27) JE6, 28) JE7, 29) JE8, 30) Blank, 31) Marker.



Appendix Figure 2. Example of the output generated by the ABI Prism 377 automatic sequencer. This particular sequence is for animal CN16 using sequencing primer THR.

9. APPENDIX TABLES

Appendix Table 1. Example of the SAS program for extracting and formatting the pedigree data

---

ANINDEX.HO

```
FILENAME DATA1 '[AG150BLUP.JNS]HERDBOOK.DAT';
FILENAME DATA2 '[AG150BLUP.HANSEN]ANIMAL.HOL';
DATA T1;
INFILE DATA1;
OPTION LINESIZE=78 ERRORS=0;
INPUT ANIMAL $ 1-18 @121 BDATE YYMMDD6. SIRE $ 153-170 DAM $ 210-227;
IF BDATE=. THEN BDATE=0;
DATA T2;
SET T1;
BDATE=0;
ANIMAL=SIRE;
IF ANIMAL='                ' THEN DELETE;
IF ANIMAL='HO              ' THEN DELETE;
SIRE='999999999999999999';
DAM='999999999999999999';
PROC SORT DATA=T2; BY ANIMAL BDATE;
* UNIQUE ANIMAL NUMBERS;
DATA S2;
SET T2;
IF LAST.ANIMAL;
BY ANIMAL;
DATA T3;
SET T1;
BDATE=0;
ANIMAL=DAM;
IF ANIMAL='                ' THEN DELETE;
IF ANIMAL='HO              ' THEN DELETE;
SIRE='999999999999999999';
DAM='999999999999999999';
PROC SORT DATA=T3; BY ANIMAL BDATE;
* UNIQUE ANIMAL NUMBERS;
DATA S3;
SET T3;
IF LAST.ANIMAL;
BY ANIMAL;
DATA FIRST;
SET T1 S2 S3;
IF SIRE='                  '
THEN SIRE='999999999999999999';
IF DAM='                  '
THEN DAM='999999999999999999';
IF BDATE NE 0 THEN BDATE= JULDATE(BDATE);
PROC SORT DATA=FIRST; BY ANIMAL BDATE;
* UNIQUE ANIMAL NUMBERS;
```

Appendix Table 1 (cont'd). Example of the SAS program for extracting and formatting the pedigree data.

```
DATA SECOND;
SET FIRST;
IF LAST.ANIMAL;
BY ANIMAL;
PROC SORT DATA=SECOND: BY BDATE;
* RENUMBER ANIMALS;
DATA INDEX (KEEP=ANIMAL NEWID);
SET SECOND;
RETAIN NEWID 0;
NEWID=NEWID+1;
* NEWID FOR ANIMAL;
DATA THIRD;
MERGE SECOND (IN=INA) INDEX (IN=INB);
IF INA;
* NEWID FOR SIRE;
DATA INDEXS (KEEP=SIRE SIREID);
SET INDEX;
SIRE=ANIMAL;
SIREID=NEWID;
PROC SORT; BY SIRE;
PROC SORT DATA=THIRD; BY SIRE;
DATA FOURTH;
MERGE THIRD (IN=INA) INDEXS (IN=INB); BY SIRE;
IF INA;
*NEWID FOR DAM;
DATA INDEXD (KEEP=DAM DAMID);
SET INDEX;
DAM=ANIMAL;
DAMID=NEWID;
PROC SORT; BY DAM;
PROC SORT DATA=FOURTH; BY DAM;
DATA FIFTH;
MERGE FOURTH (IN=INA) INDEXD (IN=INB); BY DAM;
IF INA;
PROC SORT; BY NEWID;
DATA SIXTH;
SET FIFTH;
IF SIREID=. THEN SIREID=0;
IF DAMID=. THEN DAMID=0;
IF ANIMAL='          ' THEN DELETE;
IF ANIMAL='HO          ' THEN DELETE;
NEWID=NEWID-2;
IF SIREID NE 0 THEN SIREID=SIREID-2;
IF DAMID NE 0 THEN DAMID=DAMID-2;
DATA _NULL_;
FILE DATA2;
```

Appendix Table 1 (cont'd). Example of the SAS program for extracting and formatting the pedigree data.

```
SET SIXTH;
PUT ANIMAL 1-18 NEWID 20-24 BDATE 26-31 SIRE 33-50
SIREID 52-56 DAM 58-75 DAMID 77-81;
```

Appendix Table 2. Example of the control statements for initial submission of the Fortran program to compute additive relationship.

REL-HOL.COM

```
$ SET DEF USERB7:[AG150BLUP.HANSEN]
$ FOR REL-HOL.FOR
$ LIN REL-HOL
$ ASS USERB7:[AG150BLUP.HANSEN]ANIMAL.HOL    FOR003
$ ASS USERB7:[AG150BLUP.HANSEN]FOROUT.HOL    FOR004
$ ASS USERB7:[AG150BLUP.HANSEN]DIAGOUT.HOL   FOR005
$ RUN USERB7:[AG150BLUP.HANSEN]REL-HOL
$ DEASS FOR003
$ DEASS FOR004
$ DEASS FOR005
$ DEL REL-HOL.OBJ;*,.EXE;*,.MAP;*
```

Appendix Table 3. Example of the Fortran program used to generate the additive relationship matrix that was used for the initial batch submission.

REL-HOL.FOR

```
C       USE NN IN THE DIMENSION
        REAL*4 R(5539), DIAG(6000), RX(400000), RRX(500)
        INTEGER ID(5539,3), NEWID, SIREID, DAMID,
        1IX(400000), JX(400000), IIX(500), JJX(500)
C       COUNTER TO STOP RUN
        ISTOP=0
C       COUNTER TO WRITE ROW AND COL RELATIONSHIP INTO IIX, JJX, RRX
C       PRINTER ARRAY
        JCOUNT=0
C       SETTING PRINTER ARRAY TO NULL
        DO 100 I=1,500
        IIX(I)=0
        JJX(I)=0
 100    RRX(I)=0.
C       NUMBER OF ELEMENTS IN THE RX(ICOUNT) ARRAY
C       IN SUBSEQUENT ROUNDS ICOUNT=DIAGNOL ELEMENTS
        ICOUNT=1
C       CONTINUE COMPUTING FROM IROW
        IROW=1
C       NUMBER OF ANIMALS USED TO COMPUTE RELATIONSHIP
        NN=5539
C       MAXIMUM CAPACITY OF ELEMENTS STORED FOR SUBSEQUENT RUNS
        NX=400000
C       SETTING DIAG TO NULL
        DO 1 I=1,NN
  1     DIAG(I)=0.
        DO 2 I=1,NX
        IX(I)=0
        JX(I)=0
  2     RX(I)=0.
C       READING DATA IN IDIAG(ND) AND DIAG(ND) ARRAY FROM ROUND N
C       DO 106 I=1,ND
C       READ (6,21) ILOC, RVAL
C       IDIAG(I)=ILOC
C       DIAG(I)=RVAL
C 106   CONTINUE
C       CLOSE (6)
C       COUNTER FOR READING DATA FROM IROW ONWARDS
C       READING DATA INTO RX(ICOUNT) ARRAY FROM ROUND 1
C       DO 107 I=1,ICOUNT
C       READ (2,18) N1, N2, R1
C       KCOUNT=KCOUNT+1
C       IX(KCOUNT)=N1
C       JX(KCOUNT)=N2
```

Appendix Table 3 (cont'd). Example of the Fortran program used to generate the additive relationship matrix that was used for the initial batch submission.

```
C       RX(KCOUNT)=R1
C 107   CONTINUE
C       CLOSE (2)
C       READING DATA INTO ID(NN,3) ARRAY
        DO 3 I=1,NN
        READ (3,17) ID(I,1), ID(I,2), ID(I,3)
  3     CONTINUE
        CLOSE (3)
C       READ EACH ROW AS I, COMPUTE RELATIONSHIP FOR DIAGONAL ELEMENTS
C       CHANGE START OF SOLUTION FOR I AS IROW
        DO 16 I=IROW,NN
C       SET A VALUE OF 1 FOR DIAGONAL
        R(I)=1.
        INEWID=ID(I,1)
        ISIRE=ID(I,2)
        IDAM=ID(I,3)
        IF (ISIRE.EQ.0.OR.IDAM.EQ.0) GO TO 4
        R(I)= R(I)+0.5*DIAG(I)
C       READ EACH COL AS J, COMPUTE RELATIONSHIP FOR OFFDIAGONAL
C       ELEMENTS
  4     DO 10 J=1,NN
        IF (I.EQ.J) GO TO 8
C       IF J LESS THAN I
        IF (J.GT.I) GO TO 6
        DO 5 IND=1,ICOUNT
C       COPYING ELEMENTS FROM UPPER TRIANGULAR MATRIX FOR SOLUTION
        IF (J.EQ.IX(IND).AND.I.EQ.JX(IND)) GO TO 500
        GO TO 5
  500   R(J)=RX(IND)
        GO TO 7
  5     CONTINUE
        GO TO 7
C       IF J GREATER THAN I
  6     JSIRE=ID(J,2)
        JDAM=ID(J,3)
        IF (JSIRE.EQ.0.AND.JDAM.EQ.0) GO TO 10
        IF (R(JDAM).EQ.0..AND.R(JSIRE).EQ.0.) GO TO 10
        IF (JSIRE.EQ.0) R(J)=R(J)+0.5*R(JDAM)
        IF (JDAM.EQ.0) R(J)=R(J)+0.5*R(JSIRE)
        IF (JSIRE.EQ.0.OR.JDAM.EQ.0) GO TO 7
        R(J)= R(J)+0.5*(R(JSIRE)+R(JDAM))
C       SKIP PRINTING ZERO VALUE
  7     IF(R(J).EQ.0.) GO TO 10
        IF (I.GT.J) GO TO 111
C       PRINT ROW COL AND RELATIONSHIP
```

Appendix Table 3 (cont'd). Example of the Fortran program used to generate the additive relationship matrix that was used for the initial batch submission.

```
  8     JCOUNT=JCOUNT+1
        ISTOP=ISTOP+1
        IIX(JCOUNT)=I
        JJX(JCOUNT)=J
        RRX(JCOUNT)=R(J)
C       GO TO PRINTER BEFORE STOPPING THE ROUND
        IF (ICOUNT.GE.NX) GO TO 108
        IF (ISTOP.GE.NX) GO TO 108
        IF (JCOUNT.EQ.1) GO TO 108
        GO TO 110
C       WRITING RX() FOR NEXT ROUND
 108    DO 109 JP=1,JCOUNT
        WRITE (4,18) IIX(JP), JJX(JP), RRX(JP)
 109    CONTINUE
C       TO STOP THE ROUND
        IF (ISTOP.GE.NX) GO TO 111
C       RESET COUNTER OF PRINTER
        JCOUNT=0
 110    CONTINUE
        IF(I.EQ.J) GO TO 10
        IF (ICOUNT.GE.NX) GO TO 19
        IF (ISTOP.GE.NX) GO TO 19
        IX(ICOUNT)=I
        JX(ICOUNT)=J
        RX(ICOUNT)=R(J)
        ICOUNT=ICOUNT+1
C       STORING ELEMENTS IN THE DIAGONAL
 111    DO 9 JK=IROW,NN
        KNEW=ID(JK,1)
        KSIRE=ID(JK,2)
        KDAM=ID(JK,3)
        IF (KSIRE.LT.1.OR.KDAM.LT.1) GO TO 9
        IF  (KSIRE.EQ.INEWID.AND.KDAM.EQ.J)  DIAG(KNEW)=R(J)
        IF (DIAG(KNEW).GT.0.) GO TO 9
        IF  (KDAM.EQ.INEWID.AND.KSIRE.EQ.J)  DIAG(KNEW)=R(J)
  9     CONTINUE
        IF (ICOUNT.GE.NX) GO TO 19
        IF (ISTOP.GE.NX) GO TO 19
 10     CONTINUE
 14     IF (I.EQ.NN) GO TO 16
C       CLEARING R(NN) ARRAY FOR NEXT ROUND
        DO 15 L=1,NN
 15     R(L)=0.
 16     CONTINUE
 17     FORMAT (20X, I4, 28X, I4, 21X, I4)
```

Appendix Table 3 (cont'd). Example of the Fortran program used to generate the additive relationship matrix that was used for the initial batch submission.

```
   18   FORMAT (I4,I4,F6.4)
C       WRITING NON ZERO ELEMENTS IN DIAG() FOR NEXT ROUND
   19   DO 20 I=1,NN
        IF (DIAG(I).EQ.0.) GO TO 20
        WRITE (5,21) I, DIAG(I)
   20   CONTINUE
   21   FORMAT (I4,F6.4)
        END
```

Appendix Table 4. Example of the control statements needed for subsequent submissions of the Fortran program to compute additive relationship.

REL-HOL.COMFY

```
$ SET DEF USERB7:[AG150BLUP.HANSEN]
$ FOR REL-HOL.FY
$ LIN REL-HOL
$ ASS USERB7:[AG150BLUP.HANSEN]DIAGIN.HOL   FOR006
$ ASS USERB7:[AG150BLUP.HANSEN]FORIN.HOL    FOR002
$ ASS USERB7:[AG150BLUP.HANSEN]ANIMAL.HOL   FOR003
$ ASS USERB7:[AG150BLUP.HANSEN]FOROUT.HOLY  FOR004
$ ASS USERB7:[AG150BLUP.HANSEN]DIAGOUT.HOLY FOR005
$ RUN USERB7:[AG150BLUP.HANSEN]REL-HOL
$ DEASS FOR002
$ DEASS FOR003
$ DEASS FOR004
$ DEASS FOR005
$ DEASS FOR006
$ DEL REL-HOL.OBJ;*,.EXE;*,.MAP;*
```

Appendix Table 5. Example of the Fortran program used to generate the additive relationship matrix that was used for all subsequent submissions.

REL-HOL.FY

```
C     USE NN IN THE DIMENSION
      REAL*4 R(5539), DIAG(2000), RX(567900), RRX(500)
      INTEGER ID(5539,3), NEWID, SIREID, DAMID,
      1IX(567900), JX(567900), IIX(500), JJX(500),
     2 IDIAG(2000)
      IFLAG=1
      JFLAG=0
C     COUNTER FOR DIAGNOL
      KDIAG=2000
C     COUNTER TO STOP RUN
      ISTOP=0
      JSTOP=50000
C     COUNTER TO WRITE ROW AND COL RELATIONSHIP INTO IIX, JJX, RRX
C     PRINTER ARRAY
      JCOUNT=0
C     SETTING PRINTER ARRAY TO NULL
      DO 100 I=1,500
      IIX(I)=0
      JJX(I)=0
  100 RRX(I)=0.
C     NUMBER OF ELEMENTS IN THE RX(ICOUNT) ARRAY
C     IN SUBSEQUENT ROUNDS ICOUNT=DIAGNOL ELEMENTS
      ICOUNT=517860
      NFLAG=ICOUNT+1
C     ICOUNT=6917303
C     CONTINUE COMPUTING FROM IROW
      IROW=5400
      KFLAG=IROW
C     NUMBER OF ELEMENTS IN THE DIAG(ND) ARRAY
      ND=140
C     NUMBER OF ANIMALS USED TO COMPUTE RELATIONSHIP
      NN=5539
C     MAXIMUM CAPACITY OF ELEMENTS STORED FOR SUBSEQUENT RUNS
      NX=567900
C     SETTING DIAG TO NULL
      DO 1 I=1,KDIAG
      IDIAG(I)=0
    1 DIAG(I)=0.
      DO 2 I=1,NX
      IX(I)=0
      JX(I)=0
    2 RX(I)=0.
```

Appendix Table 5 (cont'd). Example of the Fortran program used to generate the additive relationship matrix that was used for all subsequent submissions.

```
C*******************************************************
C   MAKE SURE DIAG.IN IS SORTED IN DIAG.SAS
C*******************************************************
C     READING DATA IN IDIAG(ND) AND DIAG(ND) ARRAY FROM ROUND N
      DO 106 I=1,ND
      READ (6,21) ILOC, RVAL
      IDIAG(I)=ILOC
      DIAG(I)=RVAL
 106  CONTINUE
      CLOSE (6)
C     COUNTER FOR READING DATA FROM IROW ONWARDS
      KCOUNT=1
C*******************************************************
C     MAKE SURE FORIN.IN IS SORTED IN TEMP.SAS
C*******************************************************
C     READING DATA INTO RX(ICOUNT) ARRAY FROM ROUND 1
      DO 107 I=1,ICOUNT
      READ (2,18) N1, N2, R1
      IX(KCOUNT)=N1
      JX(KCOUNT)=N2
      RX(KCOUNT)=R1
      KCOUNT=KCOUNT+1
 107  CONTINUE
      CLOSE (2)
C
C     READING DATA INTO ID(NN,3) ARRAY
      DO 3 I=1,NN
      READ (3,17) ID(I,1), ID(I,2), ID(I,3)
  3   CONTINUE
      CLOSE (3)
C     READ EACH ROW AS I, COMPUTE RELATIONSHIP FOR DIAGONAL ELEMENTS
C     CHANGE START OF SOLUTION FOR I AS IROW
      DO 16 I=IROW,NN
C     SET A VALUE OF 1 FOR DIAGONAL
      R(I)=1.
      INEWID=ID(I,1)
      ISIRE=ID(I,2)
      IDAM=ID(I,3)
      IF (ISIRE.EQ.0.OR.IDAM.EQ.0) GO TO 4
C
      DO 1000 ILL=1,ND
      IF (IDIAG(ILL).EQ.I) R(I)= R(I)+0.5*DIAG(ILL)
 1000 CONTINUE
C
   4  ICOL=1
 999  CONTINUE
```

Appendix Table 5 (cont'd). Example of the Fortran program used to generate the additive relationship matrix that was used for all subsequent submissions.

```
C     READ EACH COL AS J, COMPUTE RELATIONSHIP FOR OFFDIAGONAL ELEMENTS
      DO 10 J=ICOL,NN
C
      IF (J.EQ.JFLAG) GO TO 7
      IF (I.EQ.J) GO TO 8
C     IF J LESS THAN I
      IF (J.GT.I) GO TO 6
C
      IF (I.EQ.JX(IFLAG)) GO TO 499
C
      IF (I.EQ.KFLAG) GO TO 777
      ICOL=I
      GO TO 999
C
 777  DO 5 IND=NFLAG,KCOUNT
C     COPYING ELEMENTS FROM UPPER TRIANGULAR MATRIX FOR SOLUTION
      IF (J.EQ.IX(IND).AND.I.EQ.JX(IND)) GO TO 500
      GO TO 5
C
 500  R(J)=RX(IND)
      GO TO 7
 5    CONTINUE
      GO TO 7
C     SET NEW J VALUE
 499  ICOL=IX(IFLAG)
      JFLAG=ICOL
       R(ICOL)=RX(IFLAG)
      KFLAG=JX(IFLAG)
      IFLAG=IFLAG+1
      GO TO 999
C
C     GO TO 7
C
C     IF J GREATER THAN I
 6    JSIRE=ID(J,2)
      JDAM=ID(J,3)
      IF (JSIRE.EQ.0.AND.JDAM.EQ.0) GO TO 10
      IF (R(JDAM).EQ.0..AND.R(JSIRE).EQ.0.) GO TO 10
      IF (JSIRE.EQ.0) R(J)=R(J)+0.5*R(JDAM)
      IF (JDAM.EQ.0) R(J)=R(J)+0.5*R(JSIRE)
      IF (JSIRE.EQ.0.OR.JDAM.EQ.0) GO TO 7
      R(J)= R(J)+0.5*(R(JSIRE)+R(JDAM))
C     SKIP PRINTING ZERO VALUE
 7    IF(R(J).EQ.0.) GO TO 10
      IF (I.GT.J) GO TO 111
C     PRINT ROW COL AND RELATIONSHIP
```

Appendix Table 5 (cont'd). Example of the Fortran program used to generate the additive relationship matrix that was used for all subsequent submissions.

```
 8    JCOUNT=JCOUNT+1
      ISTOP=ISTOP+1
      IIX(JCOUNT)=I
      JJX(JCOUNT)=J
      RRX(JCOUNT)=R(J)
C     GO TO PRINTER BEFORE STOPPING THE ROUND
      IF (ICOUNT.GE.NX) GO TO 108
      IF (ISTOP.GE.JSTOP) GO TO 108
      IF (JCOUNT.EQ.1) GO TO 108
      IF (ND.GE.KDIAG) GO TO 108
      GO TO 110
C     WRITING RX() FOR NEXT ROUND
 108  DO 109 JP=1,JCOUNT
      WRITE (4,18) IIX(JP), JJX(JP), RRX(JP)
 109  CONTINUE
C     TO STOP THE ROUND
      F (ISTOP.GE.JSTOP) GO TO 111
C     RESET COUNTER OF PRINTER
      JCOUNT=0
 110  CONTINUE
      IF(I.EQ.J) GO TO 10
      IF (ICOUNT.GE.NX) GO TO 19
      IX(KCOUNT)=I
      JX(KCOUNT)=J
      RX(KCOUNT)=R(J)
      ICOUNT=ICOUNT+1
      KCOUNT=KCOUNT+1
C     STORING ELEMENTS IN THE DIAGONAL
 111  DO 9 JK=IROW,NN
      KNEW=ID(JK,1)
      KSIRE=ID(JK,2)
      KDAM=ID(JK,3)
      IF (ND.GE.KDIAG) GO TO 19
      IF (KSIRE.LT.1.OR.KDAM.LT.1) GO TO 9
      IF (KSIRE.EQ.INEWID.AND.KDAM.EQ.J) GO TO 8000
      IF (KDAM.EQ.INEWID.AND.KSIRE.EQ.J) GO TO 8001
      GO TO 9
8000  ND=ND+1
      IDIAG(ND)=KNEW
      DIAG(ND)=R(J)
      GO TO 9
8001  ND=ND+1
      IDIAG(ND)=KNEW
      DIAG(ND)=R(J)
 9    CONTINUE
      IF (ISTOP.GE.JSTOP) GO TO 19
```

Appendix Table 5 (cont'd). Example of the Fortran program used to generate the additive relationship matrix that was used for all subsequent submissions.

```
 10   CONTINUE
C     CLEARING R(NN) ARRAY FOR NEXT ROUND
      DO 15 L=1,NN
15    R(L)=0.
16    CONTINUE
17    FORMAT (20X, I4, 28X, I4, 21X, I4)
18    FORMAT (I4,I4,F6.4)
C     WRITING NON ZERO ELEMENTS IN DIAG() FOR NEXT ROUND
19    IF (ND.EQ.0) WRITE (5,22)
      DO 20 I=1,ND
      WRITE (5,21) IDIAG(I), DIAG(I)
20    CONTINUE
21    FORMAT (I4,F6.4)
22    FORMAT (16H***NO DIAGONAL***)
      END
```

Appendix Table 6. Example of the control statements used to submit the program to compress the data files generated.

COMPRESS.COM

```
$ SET DEF USERB7:[AG150BLUP.HANSEN]
$ FOR COMPRESS.FOR
$ LIN COMPRESS
$ ASS USERB7:[AG150BLUP.HANSEN]FOROUT.SJER    FOR008
$ ASS USERB7:[AG150BLUP.HANSEN]FOROUT.SJERX   FOR009
$ ASS USERB7:[AG150BLUP.HANSEN]COMPRESS.LIST  FOR006
$ RUN USERB7:[AG150BLUP.HANSEN]COMPRESS
$ DEASS FOR008
$ DEASS FOR009
$ DEASS FOR006
$ DEL COMPRESS.OBJ;*,.EXE;*,.MAP;*
```

Appendix Table 7. Example of the Fortran program used to compress the data files generated.

COMPRESS.FOR

```
      LOGICAL SW/.FALSE./
      REAL*4 JVALUE(100000),VALUE
      INTEGER JCOL(100000)
C
      OPEN(UNIT=8,STATUS='OLD',NAME='USERB7:[AG150BLUP.HANSEN]
     1FOROUT.SJER')
C
      OPEN(UNIT=9,CARRIAGECONTROL='NONE',RECL=18,STATUS='NEW',
     1  INITIALSIZE=11000,NAME='USERB7:[AG150BLUP.HANSEN]FOROUT.SJERX')
C
C
      OPEN(UNIT=6,CARRIAGECONTROL='NONE',RECL=100,STATUS='NEW',
     1  INITIALSIZE=11000,NAME='USERB7:[AG150BLUP.HANSEN]COMPRESS.LIST')
C
C**************************************************************************
C
C     OPEN(UNIT=9,CARRIAGECONTROL='NONE',RECL=11,STATUS='NEW',
C     1  INITIALSIZE=22000,NAME='USERB3:[AG150BLUP.JNS]
C 2   FORIN.CANBS1X',
C 3   ACCESS='SEQUENTIAL',FORM='UNFORMATTED',RECORDTYPE='VARIABLE')
C
      M2 = 0
      M3 = 0
C
      K = 0
C*****************************************
C     JROW IS THE ID OF THE FIRST ANIMAL
C     PROCESSED
C*****************************************
      JROW = 1
C
C*****************************************
C     100,000 IS THE MAXIMUM LIMIT FOR EACH
C     ANIMAL AND 20,000,000 IS THE TOTAL
C     NUMBER OF RECORDS THAT CAN BE PROCESSED
C*****************************************
      DO 10 M = 1,20000000
      READ(8,1,END=4)  IROW,ICOL,VALUE
      IF(IROW .NE. JROW) GO TO 5
      K = K + 1
      JCOL(K) = ICOL
      JVALUE(K) = VALUE
      GO TO 10
 5    CONTINUE
```

Appendix Table 7 (cont'd). Example of the Fortran program used to compress the data files generated.

```
      GO TO 6
  4   SW = .TRUE.
  6   M2 = M2 + 1
      WRITE(9,2)  JROW,K
C     WRITE(9)  JROW,K
      DO 9 L = 1,K
      M3 = M3 + 1
      WRITE(9,3) JCOL(L),JVALUE(L)
C     WRITE(9) JCOL(L),JVALUE(L)
  9   CONTINUE
      IF (SW) GO TO 20
      K = 1
      JCOL(K) = ICOL
      JVALUE(K) = VALUE
      JROW = IROW
C
 10   CONTINUE
 20   M1 = M - 1
      WRITE(6,30) M1, M2, M3
 30   FORMAT(' THERE ARE ',I10,' INPUT RECORDS IN FORIN.CANBSX'/
      1I10, 5X, 1H+, 5X, I10, '  OUTPUT RECORDS')
      STOP
C
  1   FORMAT(2I6,F6.4)
  2   FORMAT(2I6)
  3   FORMAT(I6,F6.4)
      END
```

Appendix Table 8. Example of the SAS program used to extract a subset of data for sequential processing by the relationship program from the main body of the data.

TEMP.SAS

```
FILENAME DATA25 '[AG150BLUP.HANSEN]FOROUT.CANBSY';
FILENAME DATA29 '[AG150BLUP.HANSEN]FOROUT.CANBSALL';
DATA F25;
INFILE DATA25;
OPTION LINESIZE=78 ERRORS=0;
INPUT ID 1-5 JD 6-10 RVAL 11-16 .4;
* DELETE ROWS AND COLS. LESS THAN T;
IF ID LT 30818 AND JD LT 30818 THEN DELETE;
* DELETE DIAGNOL ELEMENTS;
IF ID EQ JD THEN DELETE;
* DELETE ROWS LESS THAN T, COLS. GREATER THAN T+1;
IF ID LT 30818 AND JD GT 31100 THEN DELETE;
DATA _NULL_;
FILE DATA29;
SET F25;
PUT ID 1-5 JD 6-10 RVAL 11-16 .4;
```

Appendix Table 9. Example of the SAS program used to process the generated diagonal elements of the relationship matrix for resubmission.

DIAG.SAS

```
FILENAME DATA1 '[AG150BLUP.HANSEN]DIAGIN.CANBS';
FILENAME DATA2 '[AG150BLUP.HANSEN]DIAGIN.CANBS';
DATA SECOND;
INFILE DATA1;
OPTION LINESIZE=78 ERRORS=0;
INPUT ID 1-5 RVAL 6-11 .4;
PROC SORT; BY ID;
DATA THIRD;
SET SECOND;
IF FIRST.ID;
BY ID;
DATA _NULL_;
FILE DATA2;
SET THIRD;
PUT ID 1-5 RVAL 6-11 .4;
```

Appendix Table 10. Example of the SAS program used for the regression analysis of the inbreeding data.

REGCANBS.SAS

```
FILENAME NFLD 'USERB5:[AG150HONG.JNS]ANIMAL.CANBSX';
DATA T1;
INFILE NFLD;
OPTION LINESIZE=78 ERRORS=0;
INPUT BR $ 1-2 ANIMAL $ 1-22 NEWID 25-29 YR 32-33 BDATE 31-36
SIRE $ 38-59 SIREID 62-66 DAM $ 68-89 DAMID 92-96
RX 98-103 .4;
IF BDATE=. THEN DELETE;
IF YR=. THEN DELETE;
IF YR LT 83 THEN DELETE;
IF YR=95 THEN DELETE;
PROC SORT; BY BR;
DATA T2;
SET T1;
PROC REG; BY BR;
MODEL RX=YR;
```

Appendix Table 11. Example of the SAS program for calculating inbreeding using Wright's formula.

FREQCN.SAS

```
FILENAME NFLD 'USERB5:[AG150HONG.JNS]ANIMAL.CANBSX';
FILENAME DATA2 'USERB7:[AG150BLUP.HANSEN]CN.INBR';
DATA T1;
INFILE NFLD;
OPTION LINESIZE=78 ERRORS=0;
INPUT BR $ 1-2 ANIMAL $ 1-22 NEWID 25-29 YR 32-33 BDATE 31-36
SIRE $ 38-59 SIREID 62-66 DAM $ 68-89
DAMID 92-96 RX 98-103 .4;
IF BDATE=. THEN DELETE;
IF YR=. THEN DELETE;
IF BR='BS' THEN DELETE;
DATA TX;
SET T1;
IF SIRE='99999999999999999999' THEN DELETE;
IF SIREID=0 THEN DELETE;
YRSIRE=YR*100000 + SIREID;
PROC SORT; BY YRSIRE;
DATA T2;
SET TX;
If Last.YRSIRE;
By YRSIRE;
DATA T3;
SET T2;
PROC SORT; BY YR;
PROC MEANS; BY YR;
VAR YRSIRE;
OUTPUT OUT=DATAX N=NYRSIRE;
DATA T4;
SET DATAX;
PUT YR 1-2 NYRSIRE 4-7;
DATA T5;
SET TX;
PROC FREQ; BY YR;
TABLES SIREID;
DATA T6;
SET T1;
IF DAM='99999999999999999999' THEN DELETE;
IF DAMID=0 THEN DELETE;
YRDAM=YR*100000 + DAMID;
PROC SORT; BY YRDAM;
DATA T7;
SET T6;
If Last.YRDAM;
By YRDAM;
```

Appendix Table 11 (cont'd). Example of the SAS program used for calculating inbreeding using Wright's formula.

```
DATA T8;
SET T7;
PROC SORT; BY YR;
PROC MEANS; BY YR;
VAR YRDAM;
OUTPUT OUT=DATAY N=NYRDAM;
DATA T9;
SET T6;
PROC FREQ; BY YR;
TABLES DAMID;
DATA T10;
MERGE DATAX DATAY;
BY YR;
DATA T11;
SET T10;
INB=((1/(8*NYRSIRE)) + (1/(8*NYRDAM)));
DATA T12;
FILE DATA2;
SET T11;
PUT YR 1-2 INB 4-10 .5 NYRSIRE 12-18 NYRDAM 19-26;
```

Appendix Table 12. Example of the control statements used to submit the program to select unrelated CN, BS and JE cattle.

SELECT.COM

```
$ SET DEF USERB7:[AG150BLUP.HANSEN]
$ FOR SELECT.FOR
$ LIN SELECT
$ ASS USERB7:[AG150BLUP.HANSEN]CAN.AN          FOR021
$ ASS USERB7:[AG150BLUP.HANSEN]FORIN.CANBS6X   FOR001
$ ASS USERB7:[AG150BLUP.HANSEN]FORIN.CANBS10X  FOR002
$ ASS USERB7:[AG150BLUP.HANSEN]SELECT.CANLIST  FOR022
$ ASS USERB7:[AG150BLUP.HANSEN]SELECT.LIST     FOR023
$ RUN USERB7:[AG150BLUP.HANSEN]SELECT
$ DEASS FOR001
$ DEASS FOR002
$ DEASS FOR021
$ DEASS FOR022
$ DEASS FOR023
$ DEL SELECT.OBJ;*,.EXE;*,.MAP;*
```

Appendix Table 13. Example of the Fortran program used to select unrelated CN, BS and JE cattle.

SELECT.FOR

```
      REAL*4 VALUE(35000),KVALUE(35000)
      INTEGER JCOL(35000),KCOL(35000),JD(100),KANIM0(35000)
      INTEGER ID(100,2),IANIM,JANIM,KANIM(35000)
C
      NO=0
      IFLAG=0
C
      DO 999 INDEX=1,200000
      READ (21,888,END=333) JD(INDEX),ID(INDEX,1),ID(INDEX,2)
  888  FORMAT(A3,I4,1X,I5)
      NO=NO+1
  999 CONTINUE
C
  333 IFILE=2
C
      DO 666 IT=1,IFILE
C
      M1=0
      M2=0
      M4=0
      IN=0
C
      DO 5 I=1,3000000
      READ(IT,1,END=17) IROW, KLIMIT
    1   FORMAT (2I5)
      M1=M1+1
C*********************************************************
C     SELECTION BEGINS HERE
C*********************************************************
C
      DO 100 IR=1,NO
      IANIM=ID(IR,2)
      IF(IROW.EQ.IANIM)IFLAG=1
      IF(IROW.EQ.IANIM)IDENT=ID(IR,1)
      IF(IROW.EQ.IANIM)IDENT0=JD(IR)
  100 CONTINUE
C
      DO 3 K = 1,KLIMIT
      READ(IT,2,END=17) JCOL(K), VALUE(K)
    2   FORMAT(I5,F6.4)
      M2=M2+1
    3   CONTINUE
C
      DO 4 L=1,KLIMIT
```

Appendix Table 13 (cont'd). Example of the Fortran program used to select unrelated CN, BS and JE cattle.

```
      JC=JCOL(L)
C*************************************************************
C     SELECTION BEGINS HERE
C*************************************************************
      DO 101 JR=1,NO
      JDENT0=JD(JR)
       JDENT=ID(JR,1)
      JANIM=ID(JR,2)
      IF(JC.EQ.JANIM) GO TO 9
 101  CONTINUE
      GO TO 4
C
  9   IN=IN+1
      KANIM0(IN)=JDENT0
      KANIM(IN)=JDENT
      KCOL(IN)=JCOL(L)
      KVALUE(IN)=VALUE(L)
  4   CONTINUE
C
      IF (IFLAG.EQ.0) GO TO 777
      DO 8 IX=1,IN
      WRITE(22,60)  IDENT0,IDENT,IROW,KANIM0(IX),
     1KANIM(IX),KCOL(IX),KVALUE(IX)
 60   FORMAT  (A3,I4,1X,I5,1X,A3,I4,1X,I5,1X,F6.4)
      M4=M4+1
  8   CONTINUE
      IFLAG=0
 777  CONTINUE
C
      IN=0
C
C
  5   CONTINUE
C
 17   WRITE(23,20) M1,M2,M4
 20   FORMAT(5X,' THERE ARE ROWS', I10,'  +  COLS',I10,
     1'   INPUT',/, 5X,' THERE ARE RECORDS',
     2I10,' OUTPUT IN SELECT.FOR')
      CLOSE(IT)
 666  CONTINUE
      STOP
      END
```