University of Pretoria etd, Schmulian A (2006)

CHAPTER 1: INTRODUCTION

The Bonsmara is one of the few beef breeds in the world that was founded from the standpoint "Man must measure" as stated by one of the founders of the breed, Prof. Jan Bonsma (Bonsma, 1980). The development of the breed was based upon scientific breeding: selected not for any subjective aesthetic values but for functional efficiency, growth and adaptability. This was achieved through the application of principles of endocrinology, physiology together with performance and progeny testing. Founding animals were regularly photographed and their entire genealogy was painstakingly recorded (Bonsma, 1980).

The selection program resulted in the Bonsmara breed containing 5/8 Afrikaner and 3/8 British beef (Hereford and Shorthorn). It was important that animals bred for the subtropics, should be capable of overcoming tick and parasite infestation. Crosses from these different breeds were therefore compared and then selected for smooth coat and thick hide, with well-developed subcutaneous muscling. The effect of the climate on the different animals also player a major role in the choice of the foundation breeds. Out of all the crossbred animals at the Mara research station where the breed was being developed, animals which did not show symptoms of tropical degeneration, were identified as possible founders. Finally, bulls from exceptionally fertile cows were chosen and traits strictly selected for, included superior growth, milking ability and temperament (Bonsma, 1980). This strict attention to functional traits paid off with a beef breed that is adapted to its' environment.

In 1951 a Bonsmara bul-lending scheme was initiated by the National Department of Agriculture to evaluate the breed in farming regions across the country. In 1960 it became compulsory for these associated herds to take part in National Performance testing and in 1964 the breed association was established with Mr. Rex Ball at the helm. (Bonsma, 1980).

Lines within the breed were developed based on outstanding performance of certain bulls. The Edelheer line developed first, and is today still the most dominant line within the breed (Hennie Snyman, personal communication, 12 September 2001). Daughters of this bull were exceptionally fertile and adapted and bulls from this line also did well in performance tests. Others like the Wesselsvlei and Roodebos lines were also developed separately.

University of Pretoria etd, Schmulian A (2006)

The polled trait occurs spontaneously in several herds of Bonsmara cattle and can be ascribed to Shorthorn/Hereford forebears. In other herds the polled trait was infused by upgrading to Bonsmara stud status from Red Poll and Angus cows. Polledness however was not a trait selected for by breeders. The unconfirmed view, that polled animals were inferior to horned animals, was partly to blame for this (Hennie Snyman, personal communication, 12 September 2001). Recently however, this view was challenged when a group of stud breeders realized the advantages of marketing naturally polled stud animals and subsequently started to develop pedigreed, polled, Bonsmara herds.

Polled herds only came into being after the change of the millennium and they are bred mainly for export markets (Charl Uys, personal communication, 17 October, 2002). In many South American countries, beef cattle farming occurs under extensive farming conditions which makes the management of dehorning difficult. Ranchers in these countries therefore prefer polled animals due to a reduction in labor costs. Bonsmara breeders in South Africa obtain higher prices for polled animals exported to South America. Embryo transfer from top polled breeding cows had become a profitable enterprise. The prevalence of polled animals is on the increase in the national Bonsmara herd and homozygous polled animals are much sought-after.

The advantages of polled animals have been well documented in other countries. According to Stookey and Goonewaardene (1996), polled animals are less likely to bruise, take up less space in feed lots and feeding pens than horned animals. Dehorning in the feedlot also requires increased labor and this adds to the cost of production. Stress from this procedure of dehorning can lead to stock loss or growth retardation. Dehorning according to the above-mentioned authors, even though it is less traumatic requires time, labor and equipment. The solution therefore, is to breed animals polled.

In 1998 a group of Bonsmara breeders expressed their interest in having a molecular test developed so that they could differentiate between polled homozygote and heterozygote Bonsmara breeding stock. The current study for the investigation of a genetic test for the polled trait in Bonsmara cattle was then initiated by the Bonsmara Cattle Breeders' Association.

University of Pretoria etd, Schmulian A (2006)

There are two important reasons for applying a Marker Assisted Selection (MAS) test. Firstly, the status (polled or horned) of the animals can be identified at a much earlier stage without the use of test crossing or phenotypic examination. This in turn is useful because it would allow for selection of the polled trait in a much faster and more efficient manner. Secondly, the ability to label the animals as heterozygous or homozygous polled with a certificate of surety, would give a marketing advantage. This will allow for premium prices to be asked for homozygous breeding stock.

While isolation and sequencing of the gene itself would provide an excellent and accurate test in the Bonsmara, current resolution of the physical and genetic map did not allow for this route to be followed. The DNA sequence of the polled gene has not yet been published or identified as such. Linkage between the polled trait and several microsatellite markers has been shown and the position of polled, has been identified close to the centromere of cattle chromosome one. If whole families, in which the gene is segregating, are tested for linkage, these markers can subsequently be used to indicate the presence of the Polled gene for animals within that family (Taylor *et al*., 1998).

The exact position and sequence in base pairs of the polled gene remains unmapped, although breed and family specific tests have been developed. Schmutz (2004) reports a test on her website for Charolais, Limousin and Simmental cattle. The test works using three microsatellites of which the specific allele patterns for the above-mentioned breeds provide the results. Further it is stated that the same allele is not linked to the polled trait in all the families.(http://skyway.usask.ca/~schmutz/polled.htm#test, 2004)

Another test made available on the Internet just after this study was completed, is validated for use with Charolais, Gelbvieh, Hereford, Limousin, Salers and Simmental cattle. Although the accuracy of the test is high, the number of individuals that cannot be identified by this test is also high. The official figure as advertised on the website for inconclusive results is 10 to 15% (http://www.metamorphixine.com/faqtrupolled.html, 2004). A test that works across all families and breeds, all the time, still needs to be developed.

The pedigrees used in this study belonged to three Bonsmara stud breeders who had been selecting for the Polled trait for some years. Three to five generations of polled pedigrees existed in these

University of Pretoria etd, Schmulian A (2006)

herds for which breeders had accurately recorded phenotypic horned/polled status of the individual animals. Several of these pedigrees were used to test microsatellite markers on request of the Bonsmara breeders' Association.

In order to achieve the aim of developing a test to differentiate between homozygous polled animals and heterozygous Bonsmara animals, it would be necessary to first show that the gene action was the same in the Bonsmara as for other breeds tested. This could be verified by comparing the phenotypic data as well as linkage data in this breed, to the model currently held for polledness. The first objective (stated below) set for this study would also help to confirm this model from a genotypic perspective.

The aims of this study were as follows:

1) Determination of linkage between the polled gene and microsatellite markers found near the centromere of BTA1.

2) Determination of association between specific alleles at these microsatellite markers and the polled trait by tracing of these alleles in three pedigrees.

University of Pretoria etd, Schmulian A (2006)

CHAPTER 2: LITERATURE OVERVIEW

2.1 Polledness and animal production

Animal breeding focuses mainly on production traits such as reproduction fitness and growth. Below follows a discussion on the negative effects of dehorning and the associated benefits of breeding genetically polled animals. Polledness in cattle has become popular because it is a welfare friendly alternative to the painful process of dehorning and has positive management and cost implications for feedlot and extensive ranch management. There is however a perception among some breeders that polledness could be linked to reproductive abnormalities (Willie Snyman, personal communication, 2002), but as yet no conclusive data has been found to support these perceptions. Several studies, discussed below on polled versus horned beef cattle have found no conclusive, large-scale evidence for lower performance in polled animals.

Stookey & Goonewardene, (1996) and Goonewardene *et al.* (1999a), showed that de-horning adds to production costs in feedlots. This makes it necessary to de-horn cattle at a young age but somehow there is always a group of cattle that are only de-horned when they enter the feedlot. Occasionally such animals die, due to the removal of horns. At the least, the stress of de-horning leads to poor growth for the first period in the feedlot.

Data on selected production traits of polled and horned cattle was analyzed by Stookey & Goonewardene (1996), at a performance testing stations near Saskatchewan and Alberta, Canada. The cattle measured in the study were of the Charolais and Hereford breed and altogether 2438 bulls were tested spanning a period of eight years. The results indicated that there was no significant difference between polled and horned types for average daily gain, weight per day of age, adjusted scrotal circumference and adjusted yearling weight.

Even later studies by Goonewardene *et al.* (1999a) and Goonewardene *et al.* (1999b) on carcass growth and reproductive traits showed no difference between polled and horned animals. The reproductive traits that they studied, included pregnancy rates, calving and weaning weights, calf

University of Pretoria etd, Schmulian A (2006)

birth and weaning weights, calf pre-weaning average daily gains, dystocia score, cow weights, cow condition scores at calf birth and at calf weaning.

Preputial eversion and spiral deviation of the penis have been reported in literature as traits with a higher frequency in polled animals. In a study by Long *et al*. (1970) prepuces from 20 horned bulls and 10 polled were dissected after slaughtering and it was found that the retractor muscles were missing in the polled bulls. The muscle of the prepuce is responsible for stabilizing the preputial epithelium. These bulls were also viewed alive before the dissection. It was observed that the polled bulls everted preputial epithelium whereas only one horned bull had the condition before slaughter. Bulls from several different breeds were included in the study but the polled bulls examined were exclusively from the Aberdeen-Angus and Galloway breeds.

A higher prevalence of premature spiral deviation of the penis was found in a study done on polled bulls in Western Australia. A total of 1083 British breed beef bulls were tested using the serving capacity test and the prevalence of 16% among the polled bulls vs. 1% horned show a significant difference between the two types (Blockey & Taylor, 1984). However, the authors did not show conclusively that there was a genetic correlation between this anomaly and polledness.

Polledness was examined in an Australian state herd in which animals were selected on production and adaptive traits from 1970 to 1979 (Frisch *et al*., 1980). These animals in four closed, crossbred lines were tested on live weight, mortality and fertility. The base animals consisted of horned Hereford and polled Shorthorn cows. These were mated to horned Brahman, Hereford and Africander bulls as well as polled Shorthorn bulls and data was collected on their progeny. It was found that the frequency of the polled gene stayed constant in the herd indicating that no deleterious effect could be linked to it. These results are of interest to the current study due to the fact that the Bonsmara had similar breeds in its genesis.

Lange (1989) also reported that German polled Simmental animals exhibited the same level of performance as horned animals for all the production traits that they were tested for. These included growth, carcass yield and composition, and reproductive performance.

University of Pretoria etd, Schmulian A (2006)

Pressure from animal welfare groups concerning pain and stress of de-horned cattle as well as the greater management costs of this practice have prompted many breeders in Canada to opt for the alternative of breeding polled animals. The number of polled animals being registered in Canada has increased considerably over past few years. In 1989, 18% limousine and 14% Simmental registrations were polled. In 1999 these numbers increased to 49%, and 32% respectively. The greatest increase can be seen in the Charolais breed, with 38% registrations of polled animals in the period 1988-1990 jumping to 65.5% for the period 1998-2000 (http://www.usask.ca/wcvm/herdmed/applied-ethology/articles/dehorn.html, 2003). Breeders are selecting for the polled trait across breeds and no mention is made of any deleterious correlations with the trait.

Several polled breeds exist and polled animals can be found within almost every breed in the world. (www.ansi.okstate.edu/breeds/cattle, 2002) A brief discussion follows on the best-known examples of polled breeds and their popularity. Of interest to the current study is the fact that two of the founder breeds of the Bonsmara, the Hereford and the Shorthorn are well known for having polled lines. In South Africa several breeds are polled while others have started selecting for polled lines.

Probably the best-known polled breed of cattle is the Aberdeen-Angus. It originated in Scotland in the 16th century and in terms of numbers, the Angus breed is the largest breed in the world (Campher *et al.*, 1998). Three other polled breeds originating from the British Isles and are known as the Galloway and the Red Polled Norfolk and Suffolk (www.ansi.okstate.edu/breeds/cattle, 2003).The Red Poll originated in England when farmers crossed the Norfolk Red with the Suffolk Dun in the 19th century. The breed is well known for its red color and polledness. It was later introduced to South Africa by Cecil John Rhodes while the breeders Society was established in 1921. This breed is used for dairying in the United Kingdom but is bred primarily for its beef qualities in South Africa (Scholtz *et al.*, 2000).

The Shorthorn originated from England but was imported into the United States early on and genetically polled Shorthorns make up about 60% of the beef Shorthorns registered in the United States today. (www.bartleby.com/65/sh/Shorthor.html, 2001). Campher *et al.* (1998), states that the

University of Pretoria etd, Schmulian A (2006)

breed was imported to South Africa as early as 1836. It was also used in the foundation stock of the Bonsmara (Bonsma, 1980).

The Hereford developed from cattle kept for centuries in Herefordshire in England and the first herd book was established in 1846. Occasionally calves were born that did not develop horns and these mutants, bred by the farmers of Hereford County were used to develop polled Herefords. (http://www.cattle-today.com/poll_hereford.htm, 2003).

Sample groups of polled animals also occur in other important breeds as shown in the linkage studies mentioned later. Most notably the Charolais, Simmental, Pinzgaur and South Devon are well-known breeds in which the polled gene is segregating. Various composite breeds have developed in the last century, consisting of polled founding stock. In South Africa one such popular example of a polled composite is the Brangus, a cross between Angus and Brahman (Campher *et al*., 1998).

Maree & Casey (1993) stated that there is much variation in shape and horn size in the indigenous Southern African cattle breeds. They noted that some breeds like the Kuri and Watusi have very large horns while short horn breeds abound in the southern-most tip of Africa. The Tuli bred in Zimbabwe is probably the best example of a naturally polled breed from Southern Africa.

2.2 The inheritance pattern of the polled trait

The inheritance of polledness in cattle was described as early as 1927, when Auld was of opinion that in fact it is not hornlessness that needs explaining, but rather the horned character. Auld argued that before modern times all cattle were polled and horns developed as a mutation. This at least would explain why a recessive trait became universal over the dominant polled character and perhaps why polled animals are found to occur in so many different places all at the same time (Auld, 1927).

The main view held at the beginning of the century on the evolution of the polled trait, was that polledness was the result of selection for an ancestral hornless type that existed before the pliocene

University of Pretoria etd, Schmulian A (2006)

period. Clear examples of polled cattle in ancient Egypt; as well as ancient Greek and Roman artifacts indicate its' presence in early cattle types. Furthermore ancient drawings show that such animals occurred in France during the Palaeolithic period. An Assyrian artifact of a polled calf and a similar one found in Scotland indicate that the polled character was widespread long before our modern time. Polledness was a character specific to males and later transferred to females (Auld, 1927).

Watson (1921) also described the polled phenotype to be dominant to the horned phenotype. Four independent loci were proposed by White & Ibsen (1936) to affect the polled trait. This was the first comprehensive explanation of the inheritance pattern of the polled trait. Early exceptions found by Smith (1927) and Churchill (1927) gave rise to the idea that in some families horns were limited to males. Similar findings were made by Williamson & Williamson (1952) who indicated a sex-linked influence on the trait.

According to the model of White & Ibsen (1936), genes involved, were expressed independently but epistasis was present. They postulated that all horned cattle were homozygous for the allele for horns (*H*). The polled locus was epistatic to this locus, and had two alleles polled (*P*) and horned (*p*). They noted that scurs (*Sc*) were sex-influenced with the heterozygote (*Scsc*) scurred in males but not in females carrying a polled gene. They also documented the African horn locus, with one allele indicating its presence (*HA*) and another absence (*ha*). The African horn gene was epistatic to *P* in males with its inheritance uncertain in females. They found that it was present in their sample group, but at low frequency. Long & Gregory (1978) did a comprehensive investigation of the polled trait that supported the above-mentioned, four-locus model of White and Ibsen.

The mode of inheritance of polledness was later refined to a three locus model. This is the model which is widely accepted and was shown in the current study and it is summarized in Table 2.1. The assumption was that polled and horned were two alleles at the same locus instead of two different loci, with polled being the dominant allele at that locus. An additional two loci, which work epistatic to this locus, were also proposed. The scurs and African horn locus both sex-influenced, were described as epistatic to the polled locus (Georges *et al*., 1993).

University of Pretoria etd, Schmulian A (2006)

This model differed from the one proposed by Long & Gregory (1978) for scurred males. They stated that scurs are only expressed when the males have the genotype (*PpScsc*). He also suggested the possibility that the scur and African horn factors were different alleles of the same polled locus. Recent data however would indicate that the scurs gene is on a different chromosome altogether (Asai, 2001).

Horn development and morphology is described as a phenotypic characteristic with a high degree of polymorphism. The trait has a continuously distributed phenotype, yet is grouped into three distinct categories. Three major phenotypic groupings are: horned, visible as fully developed horns attached to the frontal bone which also differ in size and shape, scurs identified as horn-like outgrowths that were not fused to the frontal bone and polled, showing no attachments to the scull whatsoever (Georges *et al*., 1993).

The high frequency of horned offspring they observed from crosses between purebred polled *Bos taurus* breeds and horned *Bos indicus* breeds, and the occurrence of polled offspring from horned parents may be due to the African horn gene. Although the gene is present in *Bos taurus* and *Bos indicus,* there appeared to be a higher frequency of the *Ha* allele in *Bos indicus* (Georges *et al*., (1993).

Table 2.1 The genetic model of the inheritance of the polled trait (Georges *et al., 1993).*

a) Inheritance with the scurs gene

Genotype	**Phenotype:**	**Males**	**Females**
P_ScSc		Scurs	Scurs
P_Scsc		Scurs	Polled
P_scsc		Polled	Polled
pp___		Horned	Horned

b) Epistatic effect of the African horn gene

Genotype	**Phenotype:**	**Males**	**Females**
P_HaHa		Horned	Horned
P_Haha		Horned	Polled
P_haha		Polled	Polled
pp____		Horned	Horned

University of Pretoria etd, Schmulian A (2006)

The appearance of scurs is very much the same as that of horns. In a linkage study of the polled trait using a large, experimental herd of cattle, there was some difficulty in discerning the horned from the scurs phenotypes.To truly differentiate between horned and scurred animals, researchers in this study by Brenneman (1996) dissected animals after slaughter to determine their horn status. Additionally, animals were phenotyped for scurs/horns more than once during their lifetime, at birth, weaning and adulthood. Progeny used in the study were scored at one year of age and scull dissections were done at 20 months. First the horns were removed and the corneal diverticulum, the sinus cavity extending from the frontal sinus into the processus cornus or horn core was visible. The examination showed htat in the case of horns, there were cavities that extended into the horn for several centimeters. The skulls of the scurred animals on the other hand possessed large protuberances at the point of attachment. Scurs were filled with a cartilaginous material unlike horns. Externally such scurs would probably have been judged to be horns. Two backcross Brahman animals appeared to have attached horns but with no sinus cavities extended into the processus cornus. These animals were reclassified as scurred. There also appeared to e a continuous distribution of the type of scurs, from no attachment to complete attachment of scurs to the scull. Animals scored as heavily scurred at one year and then as horned, after slaughter showed that problems in classification could become an important factor in identification even under experimental conditions (Brenneman *et al*., 1996).

Long and Gregory (1978) made the observation that the presence and size of scurs was not independent of the shape of the poll. A flatter poll tended to predispose scurs while scurs occurred less frequently on any individual with an extremely peaked poll. They observed a misclassification of a bull used in the study and suggested the use of scurs free animals to increase the frequency of the polled gene in a herd.

The discernment of the phenotypic status of animals is therefore hampered by the presence of scurs. However, there are indications that the scurs gene is found on bovine chromosome 19 (Asai, 2001). If this were the case then it would make the gene for scurs genetically unlinked to the polled character. Scurs therefore, should have no chromosomal relationship with the markers linked to polledness and would not act as a confounding factor in an actual linkage test for polledness.

University of Pretoria etd, Schmulian A (2006)

2.3 Status of the Bovine genome with specific reference to polledness

Developments in the mapping of the Bovine genome have made it possible to study genes associated with several traits in livestock. In the past, genetic evaluation consisted primarily of phenotypic analysis, whereas today traits of economic importance can be tracked by the analysis of known DNA sequences. In the future, animal breeding may be dominated by both quantitative and molecular genetics (Beuzen *et al*., 2000).

The application of microsatellite markers resulted in the localization of the polled locus onto Bovine chromosome one (BTA1).Microsatellites are di-, tri and tetranucleotide DNA sequence repeats. Alleles refer to multiples fo these repeats and most microsatellites consist of the $(CA)_n$ repeat (Strachan & Read, 1999). There are certain characteristics that make microsatellite markers suitable in such linkage experiments. In general any markers used for purposes such as this are required to be highly polymorphic in other words; lines or individuals will potentially be more heterozygous at loci because it is likely that different alleles will occur at each locus. Markers should also occur abundantly across the genome and must be neutral in relation to the gene of interest and to reproductive fitness. Finally all possible allele combinations need to be identifiable and therefore the markers need to be co-dominant (Falconer & Mackey, 1997).

Microsatellites are also useful in linkage studies due to the fact that they can be analyzed together in PCR multiplex. Microsatellites for which primers are created with non-overlapping allele sizes, can be amplified together and this represents a huge time saving as multiple markers can then be loaded in the same lane of an automated gel (Strachan & Read, 1999).

Much work had initially been done to determine several microsatellites on the bovine genome (Barendse *et al.,* 1994; Bishop *et al.,* 1994) and later by Vaiman *et al.* (1994) and Ferreti *et al.* (1997). Studies to find the exact positions of genes, including the exact sequence of the polled gene were hampered initially by the low resolution of maps. Medium density maps of the genome such as those of Ma *et al.* (1997) and Barendse *et al.* (1997) were also not comprehensive enough for the process of fine mapping genes associated with polledness. When compared, they could not be used to resolve differences in marker order found by different linkage studies.

University of Pretoria etd, Schmulian A (2006)

A cytogenetically anchored genetic map of BTA1 was later created that reduced the average interval between markers to two centimorgan (cM). The aim of positional cloning ie. To identify a gene previously genetically linked, by mapping efforts, was still not within reach. It was concluded that a very dense genetic map of at least one marker per cM was necessary in order to carry through positional cloning and such work was not yet possible in cattle or pigs due to the low resolution of the above map. (Viaman *et al*., 1997).

The first workshop on the genetic map of BTA1 was held in 1998 in order to resolve such issues and obtain consensus for the map of BTA1 (Taylor *et al*. 1998). This map was useful for obtaining consensus (between different mapping efforts) in certain areas, however it could not be used to determine microsatellite order near the horned and polled gene. The polled gene, located proximal to the centromere, could not be ordered relative to BM6438, TGLA49, IFNAR, KAP8, INRA212 and INRA117 and the order conflicted with other linkage studies done on the horned/polled trait (Schmutz *et al.* 1995; Brenneman *et al.* 1996; Harlizius *et al.* 1997). The markers SOD1 and SODMICRO2 were treated separately and not included on their consensus map. Four of the above markers were placed into a region spanning 12.2 cM as indicated in **Figure 1** and these markers were used in the current study.

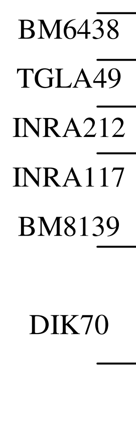

Figure 1 Section of the Sex-averaged linkage map of BTA1 (Taylor *et al*., 1998).

More recently, De Donato *et al.* (1999) presented the first report on the physical mapping of BAC clones to bovine chromosomes in which thirty-nine new microsatellites were assigned. The purpose of this study was to provide more physically anchored landmarks for the construction of contigs for positional cloning throughout the bovine genome as well as to integrate this physical map with the

University of Pretoria etd, Schmulian A (2006)

linkage maps. In a study by Rexroad III *et al*. (2000), a genetic map of BTA1 was created, using the method of radiation hybrid mapping. This map was specifically constructed to test a radiation hybrid panel using 18 markers linked to the polled gene. The study could not however produce a statistically ordered map of the region. The author concluded that a higher marker density and physical anchors would be necessary before this panel could be used efficiently for further mapping.

Work continues on the Bovine Genome. It is envisaged that full physical coverage of the bovine genome will soon be reached which will result in the determination and testing of the sequence of the gene responsible for polledness. Until that time, microsatellites can be applied to indicate the approximate position of the polled gene. At the time this study was initiated, the exact order of microsatellites around the polled gene was not known, nor had the position of polledness been pinpointed exactly.

Mapping initiatives have identified several microsatellites, which researchers tested in specific pedigrees and breeds to find linkage with the polled gene. Georges *et al*. (1993) were the first group to conclusively link the polled gene to two microsatellites placing it on on BTA1. Linkage between each of the microsatellite markers TGLA49, BM6438 and the polled gene in a Canadian population of the Charolais breed was shown by Schmutz *et al.* (1995). They were also able to trace a specific allele through a family to determine the status of a cow as homozygous polled. In a test used to determine if SOD1MICRO2 was linked, Marquess, *et al*. (1997) determined that the allele frequencies were polymorphic in one sample group but monomorphic when tested in another population. This could indicate the great differences in allele frequencies between families.

In a comprehensive study by Brenneman and colleagues (1996) at Texas A&M University, the polled locus was localized at 4.9 cM from TGLA49 in a large herd of Angus x Brahman crosses. They also concluded that the gene for polled had the same gene action in *Bos Indicus* breeds. Harlezius and co-workers (1997), tested four microsatellites for linkage with polledness in German Simmental and the Austrian Pinzgauer breeds. They were not able to order the polled locus relative to the markers due to the low number of individuals in the sample and resulting low number of recombinants. However, three markers in this study did show significant LOD scores: KAP 8 gene, INRA212 and BM6438.

University of Pretoria etd, Schmulian A (2006)

2.4 Concept overview of Linkage Analysis

Linkage analyses play an important role in the search for traits of economic importance by indicating a distance relationship between the trait and genetic markers situated close to it. Linkage allows for a trait to be placed relative to these markers, within a specific interval on the chromosome. Linkage is based on the determination of the recombination fraction (θ). This measure is used to test the hypotheses of free recombination versus linkage. If loci segregate independently and are on different chromosomes, they are unlinked and the free recombination fraction is θ =0.5. The process is illustrated in **figure 2**. In contrast, linked loci always have θ values smaller than 0.5. Very closely linked loci that rarely show recombination events will have θ values approaching zero (Ott, 1991).

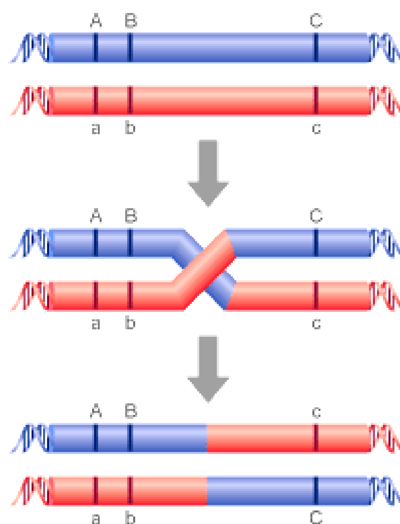

Figure 2 The process of recombination.

The maternal (top) and paternal (bottom) chromosomes are shown as the first two bars. Crossover takes place between them and this is passed on to the sperm or egg cell, which represents the next generation. Let A be the disease gene. B and C are genetic markers. Because A and B are so close together few recombination events will take place in this interval but the disease gene A can be mapped relative to the markers B and C due to recombination here.

(www.wellcome.ac.uk/en/genome/genesandbody/hg06b004.html., 2002)

University of Pretoria etd, Schmulian A (2006)

In this study, a disease gene is not being identified but a trait of economic importance for which phenotypes are available. The recombination fraction (θ) is the chance that a gamete from parent A is recombinant. This recombination fraction event would have taken place between the two parental alleles of parent A as indicated in **Figure 2.** Recombination events are passed from parents to children and these can be counted to give the recombination fraction.

As shown above on the chromosomes, the recombination fraction may be applied to define a distance measure between markers. This is not a simple relationship neither is it found to be constant across chromosomes and sections of chromosomes. There is a tendency for more recombination near the telomeres of chromosomes in males while more recombination in the centromeric region in females takes place.

The ratio of the physical distance to the genetic distance on a chromosome is therefore variable in nature. An indication of how well the recombination fraction estimates the physical distance on BTA1 was given in a study by De Donato *et al*. (1999). They stated a ratio of 1.24 Mb/cM (physical to genetic map distance) for this chromosome. Most bovine chromosomes have a physical/genetic relationship close to one except in the case of BTA 1 and BTA 2. Deviation from a 1 to 1 relationship of these chromosomes could be ascribed to the lower recombination rate on longer chromosomes, resulting in lower accuracy of the statistical methods used (De Donato *et al.,* 1999).

An important requirement for the estimation of linkage of genes according to Weir (1997), is accurate and complete pedigree data. A linkage analysis cannot be carried out on unrelated individuals and requires phenotypic information on family members in a pedigree. To score recombinants is often difficult, because individuals need to be informative. The chances that a mating will be informative for two genes are greater if that individual is a heterozygote at those genes (Ott, 1991). The pedigree should represent several generations as a recombination event can only be measured when it is passed from a parent to a child. The order of the genetic markers (associated with the trait of interest) and their distance relative to each other can further be calculated from the recombination fraction between them. The recombination fraction therefore indicates the strength of linkage between the trait and the genetic markers.

University of Pretoria etd, Schmulian A (2006)

Mapping functions have been developed to take these and other factors into consideration. Two of the most widely used functions are the Haldane and the Kosambi. The Haldane function is used if crossovers occur randomly and do not influence each other. The Kosambi function is used more widely because interference occurs and the formation of one chiasma inhibits the formation of another close-by (Strachan & Read, 1999). Results for the Bovine chromosomes mentioned above indicate a variable distance measure across chromosomes so aping functions and the statistical methods used to determine genetic maps need to be refined.

An important factor in the determination of linkage is the type of pedigree structure available. This will determine how easily the trait can be traced for an individual within the pedigree. Phase refers to the possibility of determining which Haplotype (combination of alleles) an individual received from which one of its grandparents. If the animals in the pedigree are informative, the alleles can be traced as they are passed from the parents to the individual. Such an individual is phase-known (Ott, 1991). This of course requires that the grandparental alleles must be known. In phase-unknown individuals recombinant and non-recombinant haplotypes cannot be distinguished unambiguously because we do not know which genes came from the grandparents. In such a case probabilities are allocated.

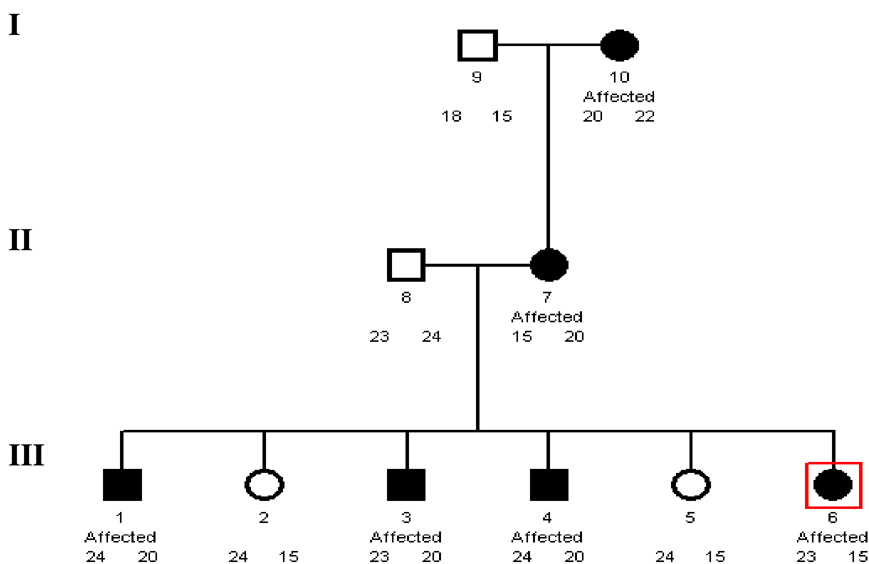

Figure 3 A phase-known family (Ott, 1991).

University of Pretoria etd, Schmulian A (2006)

This figure shows an example of a phase known family were Individuals suffering from an autosomal dominant disease are shown in solid black and marked as affected. The 20 allele of the marker B is inherited with the disease. Clearly individual 1 through 5 in generation III are non-recombinant while individual 6 (outlined in a square) is recombinant. If this family had been phase unknown, in other words, missing individuals 9 and 10 in generation I, then the affected mother, individual 7 in generation II, could have inherited either allele 15 or allele 20 with the trait therefore adding the probability that individuals 1 to 5 were recombinant and individual 6 non-recombinant.

The estimation of linkage requires rather complex and intensive computations in the case of phase unknown individuals within large families (Strachan & Read, 1999). Green *et al*. (1990), shows that if the individual is phase unknown, than the estimation of the recombination fraction can be calculated using maximum likelihood. The CRI-MAP software was specifically developed by him to calculate the recombination fraction for large pedigrees.

University of Pretoria etd, Schmulian A (2006)

CHAPTER 3: MATERIALS AND METHODS

3.1 Animal sample group

365 blood samples were collected from three registered Bonsmara herds of cattle in South Africa. The three herds are identified by the animal number prefixes: NFS, JPL, and MCU. Breeders were responsible for selecting phase-known families from these herds. The preferred pedigree structure as requested for the linkage study were test crossings such as a polled bull mated to horned cows, their parents and the offspring of the cross.

Blood was drawn from the jugular vein or from under the tail by the Bonsmara breeders. This was done under supervision of a veterinarian and a clean needle was used for every sample. 10 ml EDTA tubes were labeled and used for the collection of the blood. After the blood was drawn, the blood samples were kept on ice and then stored in a refrigerator at 4Cº. DNA extractions were done within three days after collection.

Strict selection for the polled trait had taken place for a number of years in the herds used in this study. As a result several horned cows had already been culled and few horned animals within the herds still existed. Furthermore, most of the base animals used to establish the polled trait in these herds had already been slaughtered due to advanced age. Fortunately the phenotypic status of these animals could still be traced using pedigree records.

Polled bulls bled for the study constituted the second generation of polled animals selected for within these herds. They had been used extensively among the three herds in order to increase the frequency of polled animals. Large numbers of their progeny were sampled for the study. Cows sampled, were very rarely mated to these bulls more than once and therefore ideal full-sib families could not be obtained. This is a well-known management practice within breeding herds in South Africa, specifically to prevent inbreeding.

University of Pretoria etd, Schmulian A (2006)

Several blood samples were obtained from horned animals unrelated to the above-mentioned bulls to act as a control for allele frequency determination. The identification numbers of the animals sampled, as well as their family relationships, sex and phenotype are indicated in **Addendum 1**.

Breeders kept careful records of the polled state and animals were phenotyped more than once. Three categories were used, namely, horned, polled and scurred. The presence of African horn had also been noted as polled offspring coming from horned parents. Calves were phenotyped when weighed at birth, phenotyped at de-horning (three months of age so that scurs could be distinguished from horns) and several times after this. The breeders discerned scurs as loose horns not attached to the scull of the animal. To increase the accuracy of the identification process, breeders confirmed the status of animals again when they were weaned at six months, then again at yearling age and as adult breeding cows after each calving.

3.2 DNA extraction and quantification

DNA extractions were done at the department of Animal and Wildlife Sciences, University of Pretoria. All the subsequent procedures were completed at the Institute National de′ la Recherche Agronomique′ (INRA) laboratory in Jouy en Josas, France.

The DNA extractions were done using the PUREGENE® DNA isolation kit (Gentra Systems, Minneapolis, USA) for buffy coat/whole blood. The protocol was adapted slightly due to the difficulty in removing haemoglobin from the samples. The amount of Red Blood Cell lysis was accordingly increased from 600µl to 900µl for each sample to remove heamoglobin. This step was repeated in order to obtain a pure white pellet. After obtaining a white cell pellet, cell lysis solution (500µl) was added. Proteinase k (20 mg/ml) to a final concentration of 100µg/ml was added and incubation of the samples took place at 37°C. Samples were left overnight to digest. This step was followed by protein precipitation, DNA precipitation in isopropanol, DNA washing in ethanol and DNA hydration to obtain the DNA samples.

DNA samples were quantified using spectrophotometry. The 100µl sample consisting of 96µl HQ and 4 µl extracted DNA was tested in a spectrometer and readings were taken for salt content, DNA

University of Pretoria etd, Schmulian A (2006)

concentration and the ratio of DNA to protein. This salt content determined from the 325nm wavelength reading had to be less than 0.5. An acceptable ratio of protein to DNA (260nm/280nm) was approximately 1.8. A ratio of two was assumed to be RNA contamination and readings lower than 1.0 was from protein contamination (GENTRA™ DNA EXTRACTION KIT, Gentra Systems, Minneapolis, USA). Samples with such readings were discarded and a re-extraction of the sample was done. Samples with a reading of between .01 and 1.0 for DNA concentration at the 260nm wavelength were considered pure enough to use in further analysis steps. To confirm the DNA concentrations, samples were run on agarose gels and diluted accordingly to a final concentration of 20 ng/µl. After dilution samples were kept at 4C° in the refrigerator.

On arrival at INRA laboratory Jouy an Josas, in France, DNA samples were transferred into 96 well dilution plates at 4C°. A test PCR with the primer INRA003 was performed to ensure that DNA had not degraded during the travel to France. The DNA was then transferred to dilution plates. The contents of the Hotstart PCR is shown in **Table 3.1** and consisted of the active mix and a general mix:

Table 3.1 Test PCR.

PCR mix per DNA sample	
5µl	DNA 20ng/µl
0.082µl	primer 100pMOL/µl
4.44µl	H_2O
0,1µl	TAQ POLIMERASE

General mix per 96 well plate	
7,14µl	PROMEGA dNTP
142,8µl	PROMEGA 10x BUFFER
200µl	Blue Sucrose*
85,7µl	$MgCl_2$
64,2µl	H_2O

*The blue sucrose consisted of Xylen Cyonol.. 2,5% and Sucrose 40% made up with water and filtered with a 0,22µm filter before commencing with the PCR.

The PCR program followed was: 5 minute Hotstart at 95 C°, 94 C° for 20 seconds, 55 C° for 30 seconds, 72 C° for 30 seconds for 34 cycles repeated twice. TAQ POLYMERASE was added

University of Pretoria etd, Schmulian A (2006)

during the final period of the Hotstart PCR. A large 1.5% agarose gel was used. Ethidium bromide was added to the agarose gel so that primers could later be visualized under ultraviolet light. The Blue Sucrose was added as a loading dye to the PCR product to determine how fast primers move.

15µl of this PCR product was loaded per well, and the gel was run at 80 amps/ 140 volts for 30 minutes. Once the run had completed the gels was then held under UV light and photographed. The DNA samples that did not amplify well and were not visible due to lack of PCR product were concentrated in a speed vacuum and the above process was repeated. In this way all DNA samples were brought to a suitable concentration of approximately 20ng/µl before commencing with the multiplex. Samples were then ready to be arranged into the dilution plates.

3.3 Plate design for large scale PCR

The dilution plates were designed to have animals in families grouped together. This design is attached as **Addendum 2**. The families consisted of a bull followed by the dam followed by the offspring of that mating. In the case of a dam with calves from many different bulls, the calves were grouped near the dam. Spaces were left in the plates corresponding to a loading scheme for the ABI gels. All samples were also assigned a new laboratory number.

Empty wells were left in every PCR plate for the control animal 14066. If offspring of a bull spanned more than one gel, a space was also left so that the bull sample could be loaded here in the final ABI-sequencer loading step. Also in the case were a dam was mated to more than one bull from a large bull family, the dam was loaded in both those gels, near its offspring. Finally DNA was transferred from the dilution plates in 5µl quantities to a duplicate set of PCR plates, which were frozen for later use.

University of Pretoria etd, Schmulian A (2006)

3.4 Choice of microsatellite markers for the multiplex

A total of 14 microsatellite markers were tested in two sets for the multiplex. SODMICRO2 was tested separately because it could not be unambiguously placed on the consensus map of BTA1 (Taylor *et al*. 1998). The forward and reverse primers for each microsatellite are shown in **Table 3.2**. Optimization tests were done to determine which microsatellites would work well in multiplex and the final primer mixes, are shown in **Table 3.3**

The following guidelines were used to find suitable microsatellite markers for the multiplex. Primers had been designed previously at INRA to all fall within the same Primer annealing temperature of 55 C°. Primers were designed in such a way that the amplified products were of different sizes and therefore did not overlap. In cases where overlapping alleles could not be avoided; on of the primers would be labeled suing a different fluorescent label. These microsatellites could therefore be added to the same primer mix and be differentiated in the data-formatting phase by their colour. The PCR product is run on a 377 ABI PRISM® automated sequencer (Applied Biosystems, CA, USA), which allows for the use of three possible fluorofores with each marker. These colors are Hex (yellow), Fam (blue) and Tet (Green). The labels for each primer used in the study are indicated in **Table 3.2.1.** The internal lane standard loaded in each well is labeled with a red fluorofore (Rox). A spacing of 15 base pairs (bp) in length between microsatellites with the same color fluorescent label allowed for unambiguous typing. This was done in case a new allele was found that had to be assigned to one of the two microsatellites.

3.5 Optimization

Microsatellite primers were firstly amplified in a simplex at the same concentrations and PCR conditions. If a primer amplified too strongly it was decreased by 0.2 µl and if it amplified only moderately it was decreased by 0.05µl. If, it amplified very weakly, the amount was increased by 0.2µl. Some markers amplified poorly regardless of any adjustment in the primer amounts and were identified as problem markers. The primers that amplified well were tested by grouping them according to the color of their labels. In these duplex/triplex reactions, some microsatellite primers amplified poorly because of primer-primer interactions. The type of interactions included:

University of Pretoria etd, Schmulian A (2006)

1) Interaction between unlabelled primers causing a decrease in the final product.
2) A new unexpected peak due to a reaction between a labeled primer and another unlabelled one.

Primers giving problems were re-designed and re-tested. Finally all the microsatellites that had amplified in the above tests and fell in non-overlapping allele size ranges were added together to form tow final multiplex mixes as illustrated in **Figure 4** and **Figure 5**

Table 3.2.1 Forward and reverse primers, labels and primer Tm for each microsatellite marker in SC1.

Marker	Label	Ann T.	Marker Sequence
90F5-24	-	58C°	CGAAGCGTAATTGCTTGACA
	Hex	62C°	CAGCACGGAGAGCTGTCTC
BMS574	Hex	60C°	ATGTTCTTTGACCACATGGATT
	-	64C°	GAACAAGCATTCTGACCATAGC
BM8139	-	60C°	TGCAAGGGGAAAGTCTTTATG
	Hex	66C°	GACCCAGCACACTAGAAGTGG
DIK70	Tet	60C°	CTCACTCTGCAGATCTGTAG
	-	60C°	CAGGGGGTTAAAGTCATGCA
INRA212	Tet	70C°	GCCATTGCCTTCTCCTAATAATGC
	-	72C°	CAGCAGTCCTTCCTAGTTAGAGTC
N90F5-9	Fam	60C°	GCCACAGAGAGCCGGATAT
		64C°	TTCTGGGATCTGGTAATAAGGG
AGLA017	Fam	62C°	CTGTTAGCAATGGGATGTGAAA
		64C°	CACTGCAGCTGATTCCTTACC

University of Pretoria etd, Schmulian A (2006)

Table 3.2.2 Forward and reverse primers, labels and primer Tm for each microsatellite marker in SC2.

Marker	Label	Ann T.	Marker Sequence
BMS1928	Fam	60 °C	TTTGCCAATCTCCAAAATTACG
	-	62 °C	CAGACATGACTTGGCACCTG
BM6438	Fam	64 °C	TTGAGCACAGACACAGACTGG
	-	60 °C	ACTGAATGCCTCCTTTGTGC
DVEPC42	Tet	60 °C	CAGAGACATGATTCCACTACT
	-	58 °C	ACTCCAGTAAAGAATCTGCC
DVEPC88	Tet	60 °C	CCTAAATGTTCATTGACAGATG
	-	62 °C	CTACCCAAGTTGCTACAATGG
INRA117	Hex	56 °C	GATCTCAGACAAGTTTTTAAC
	-	52 °C	AGGGAAGTCAGTTCATAG
INRA257	Tet	64 °C	AGCTGGAATTCACACTCTGGC
	-	66 °C	GTGTTTCCTGCATTGCAAGCAG
TGLA049	Fam	64 °C	GGCAGGACTTCACTCTTTTTCA
	-	64 °C	AGAAAAGGAATAATGAGACAGATTA

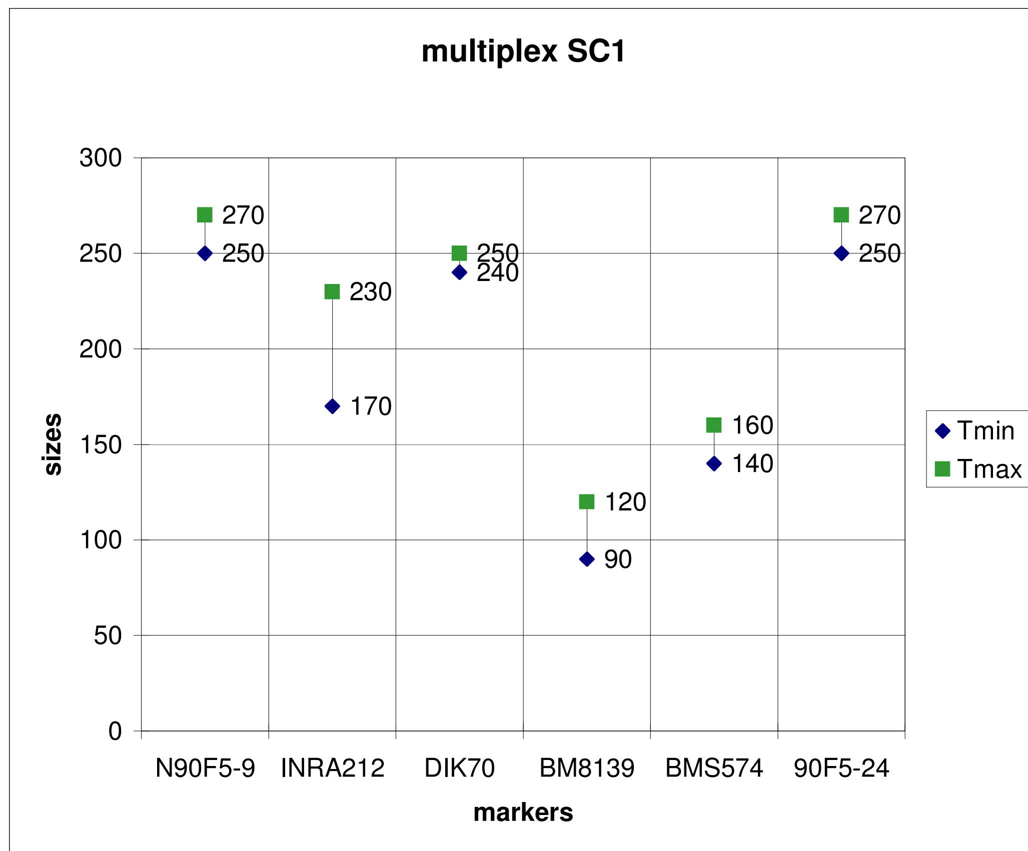

Figure 4 Size range determination for each marker in multiplex SC1.

University of Pretoria etd, Schmulian A (2006)

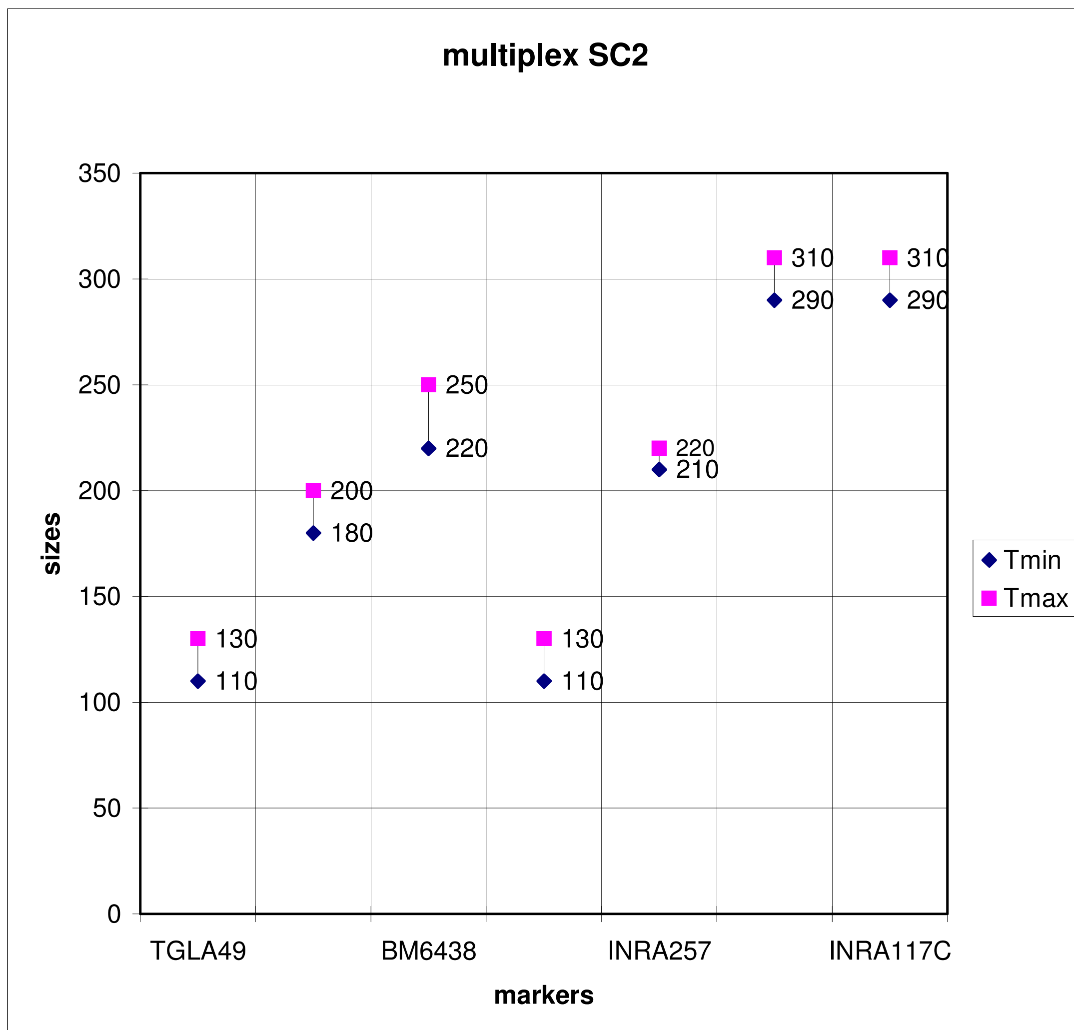

Figure 5 Initial size range determination for each marker in Multiplex SC2.

SOD1MICRO2 was tested on all samples in a separate simplex using the same method as described by Marquess *et al*. (1997). The marker primer sequences used were:

Forward: 5'-AGGGCTACAGTCCACGGGTTG-3'

Reverse: 5'-AGCGATTCACAGTCACCTCACCTA-3'

University of Pretoria etd, Schmulian A (2006)

The relative concentration of each primer in the primer mix as determined in the optimization step, are indicated in **Table 3.3**

Table 3.3 primer mix used in multiplex SC1 and SC2 at annealing temperature 55 °C.

Multiplex SC1		**Multiplex SC2**	
Marker	**1 Sample**	**Marker**	**1 Sample**
N90F5-9	0.06µl	TGLA49	0.015µl
AGLA17	-	BMS1928	0.26µl
INRA212	0.14µl	BM6438	0.04µl
DIK70	0.09µl	*DVEPC88*	0.055µl
BM8139	0.03µl	*INRA257*	0.03µl
BMS574	0.05µl	*DVEPC42*	0.17µl
90F5-24	0.04µl	INRA117C	0.09µl
H2O qsp10µl	9.18µl	H2O qsp10µl	8.68µl

A multipipet was used to pipet 10µl of the primer mix into each well of the PCR plate, which was then placed into the thermocycler. 20µl mineral oil was then added to cover the mixture and prevent evaporation. The PCR program was as follows: 5 minute Hotstart activation step at 95C°, denaturation (94C° for 20 seconds), annealing (55C° for 30 seconds) and the extension step (72C° for 30 seconds). A total of 38 cycles were completed in order to obtain sufficient PCR product for further analysis. After the final extension step (72C° for 7 minute) the samples were cooled to 4C° and transferred to the refrigerator. The PCR or active mix is shown in **Table 3.4**

University of Pretoria etd, Schmulian A (2006)

Table 3.4 Mix used in PCR multiplex SC1 and SC2

Active ingredients	**1 Sample**
Stock Solution	8,5µl
DMSO	1,25µl
PROMEGA TAQ POLYMERASE	2,5µl
Primer Mix	10µl
Stock Solution	**1 Sample**
$MgCl_2$ 25mM	2µl
PROMEGA dNTP	0,625µl
PROMEGA 10x BUFFER	2,5µl
High Quality H_2O	3,375µl
Total	8,5µl

3.6 Fragment determination of DNA samples

PCR products were analyzed using a 377 ABI PRISM® automated sequencer (Applied Biosystems, CA, USA). Polyacrylamide gels were prepared according to the GENESCAN® manual version 2.1 (Applied Biosystems, CA, USA). 12cm ABI glass plates were used for the gels. Once the gel had polymerized, the plates were rinsed, especially the bottom third were the laser reading is taken. The plates were then clamped vertically in the ABI sequencer and the sharktooth comb was carefully inserted into the slot at the top of the gel. Care was taken not to push it in to deeply and damage the gel. A plate check was performed as in the user manual (Applied Biosystems, CA, USA). If there were no bubbles or warping in the gel, the upper and lower buffer trays were filled with 1.5 liter of 1x TBE and a pre-run was performed. During this time a sample sheet was prepared manually with the laboratory numbers of the samples to be placed in each of the 36 wells.

The PCR samples for gel loading consisted of 1.5µl PCR product mixed with 3µl of the TAMRA standard mixture which consisted of 2µl formamide, 0.6µl TAMRA Buffer and 0.40µl GENESCAN-500 which is the internal lane standard. The samples were denatured for two minutes in a denaturing oven at 100C° after which they were placed on ice. Just before loading, a syringe was used to circulate water in the slot of the gel prior to loading to remove any bubbles or ammonia

University of Pretoria etd, Schmulian A (2006)

that could have accumulated around the comb. The comb was removed and samples were then loaded into the 36 wells of the polyacrylamide gel. The average running time of gels was 1.5 hours. The gel run was recorded and the gel image transmitted to an Apple Macintosh for further analysis with GENESCAN® and GENOTYPER® software (Applied Biosystems, CA, USA).

3.7 Collection and formatting of genotypic data

Before any gel runs were initiated, the expected allele sizes for each microsatellite in the study were set up in GENOTYPER®. A minimum and a maximum allele size were added for each microsatellite so that any suitable peaks within this range could be auto-labeled. In the assignment of alleles a deviation of 0.6 base pairs for auto-labeling were allowed. Only in very rare cases were a mutation was found, was it necessary to update this list. A macro was used to filter and label all peaks within the size range and then open the peaks in a plot window. Every one of the labels was verified manually starting with the red label. If these peaks were satisfactory in signal strength, the other colored peaks could then be analyzed one at a time. Peaks for the alleles were read between 150 and 5000 units. Because of baseline noise, peaks below the threshold of 150 were not accepted as a result.

In some cases ghost peaks/stutter bands, the third hump of peak, as well as other unexpected anomalies were erroneously labeled by the auto-labeling procedure and these were removed manually. The accuracy of the manual checking method was greatly increased by the fact that pedigrees could be cross-referenced using the control lane. The parents and offspring loaded in adjacent lanes on the gels also simplified this checking process.

Anomalies or artifacts found were due to stutter band peaks, DNA contamination, overflow from one lane to another on the gel, spikes due to fluctuations in electricity supply and non specific fragments. Some microsatellites like the INRA257 marker and the BMS 1928 had non-standard peak shapes more like humps. In these cases a procedure was followed to always label in the same manner and these specific gels were also double-checked after all runs were completed to check for consistency in labelling method.

University of Pretoria etd, Schmulian A (2006)

After the manual analysis, allele sizes were exported to an EXCEL sheet by means of a second macro. This spreadsheet of allele sizes was in turn imported into GEMMA**,** an in-house software program developed at Toulouse, INRA, France.

Farm pedigree data sets were created by the breeders over a period of several years using the hern management program STUDMASTER®(Agricultural Data Control, Pretoria, South Africa). These data sets contained all the pedigree information and polled status for animals used in the study. Due to the fact that it is compulsory for Bonsmrara breeders to take part in the National Recording scheme, pedigree data had been captured for every animal in the National Herd and therefore the records of these breeders could be considered accurate. This pedigree data for samples was converted and consolidated for the three herds and each sample received a unique laboratory number. Also imported into GEMMA were the results of each gel run, while at the same time the relationships between sample animals were clear from the pedigree file. GEMMA was then used t filter this data on animals and markers to show inconsistencies and typing errors.

Parentage verification was also calculated using GEMMA. The program analyzed the data in such a way that it indicated whether an allele was inherited from the dam or the sire of a specific animal. If an incompatible match occurred, it was manually set to zero for the offspring animal in GEMMA. Some cases of mistaken parentage were found and these animals were then removed completely. Finally, the error free file was exported from GEMMA. At this point the data could be further analyzed using CRIMAP (Green *et al.* 1990).

University of Pretoria etd, Schmulian A (2006)

3.8 Statistical evaluation of the data

3.8.1 Linkage Analysis

The Linkage Analysis was conducted using CRI-MAP software (Green *et al*., 1990). All family data was organized into the CRI-MAP raw genotype data file with the *.gen* extension. Each family was named according to the stud prefix of the bull. Only animals with an anambiguous polled status were sued so as to show clear linkage with the polle trai (Poll). The CRI-MAP program options used were PREPARE and TWOPOINT as well as CHROMPIC and were performed as described in the CRI-MAP user manual. No changes were made to default parameters (Green *et al.,* 1990).

After running the option PREPARE, the output file with the *.loc* file extension was generated and used in order to determine that the number of meiosis per marker were sufficient for use in further options. All families were analyzed together and combined into one file as the LOD score is additive across families (Ott, 1991).

The resultant LOD scores were then calculated using the TWOPOINT option and were included as a final result set. The recombination fraction (θ) between Poll and each of the other microsatellites was found and expressed where the LOD score was highest for the range of scores calculated in CRI-MAP. Also included in the final result was the LOD score generated by the TWOPOINT option, for the interval INRA212-INRA257. This interval was of interest due to the fact that INRA257 was a new microsatellite and its distance relative to Poll was not yet known.

Because of the high levels of selection for the polled trait in the sample herds, the level of inbreeding was determine for each animal in the sample group. High levels of inbreeding may have been possible through repeated use of valuable polled bulls over several generations of cows. Inbreeding would result in inflated LOD score values. The Agricultural Research Council at Irene, Pretoria, administers the national Bonsmara database with all performance data and pedigree data of the National herd. They provided the co-efficients for the specific animals in this study and these values are attached as **Addendum 3**. Nearly all the animals, especially the sires, in the sample group, had values that fell in the acceptable level for an outbred population (Bernice Mostert,

University of Pretoria etd, Schmulian A (2006)

personal communication, 2003). From this it was assumed that there was no major effect on the LOD scores from inbreeding.

Finally the output from CHROMPIC was checked as a measure to determine data errors in the sample set. The chance of having a true double recombinant with markers that are five cM apart due to interference is 0.0025 according to Strachan and Read (1999). If this value is taken as accurate across all regions of all chromosomes, then the number for markers lying closer together should be comparable. If much higher values are obtained, this indicates errors in marker typing or a region with a higher that normal recombination rate.

3.8.2 Allelic association using the Transmission Disequilibrium Test.

The Chi-Square test may be applied as an association test by using a two-by-two contingency table. When the data is viewed as a single sample with two observed variables, the relationship expressed by the null hypothesis (H_o) is called statistical independence of the row variable and the column variable. Variables that are not independent are associated (Samuals, 1993).

For the polled trait the hypothesis (H_o) is as follows: The allele at the linked marker is transferred independently of Polledness. This study is to show that there is an association. Therefor the idea was to disprove the (H_o) conclusively and accept the alternative hypothesis (H_a). The (H_a) was directional because it was already assumed that the marker and therefore alleles at that marker, were linked to the polled trait. The (H_a). can therefore be stated as follows: The allele of the linked marker transmits more frequently together with the polled gene. Because a directional (H_a) was calculated, frequency tables were first set up to identify alleles which were found more often in polled animals that in horned. Once identified, the Chi-Square could be calculated for such an allele, and the resultant P-value halved.

In order to isolate an allele to test in the Chi-Square, frequency tables of the different genotypes per marker were first calculated. Base on these frequencies, on of the two alleles in a genotype could be shown to be the one linked to the polled trait. In the case of some of the markers a small frequency difference was found or the frequency of the genotype was low and no pattern was clear. In cases

University of Pretoria etd, Schmulian A (2006)

like these, the test for association was not done. Only the first two alleles found in the most frequent polled genotypes, were tested.

Once the alleles from these frequently occurring polled genotypes, had been isolated, they were considered as M1 and tested for association as described by Spielman *et al.* (1993) and illustrated below. The Transmission Disequilibrium Test (TDT) is a refinement of the Chi-Square to show transmission within the generational family unit. This makes it possible to count transmissions of the allele. The test compares the number of parents who transmitted M1 to their polled offspring with the number who transmitted their other allele M2. Parents had to be heterozygous at marker M to be considered. Therefore only individuals within a family of this structure were counted: both the parents and the polled offspring. The final hypotheses for this directional Chi-square test is only altered slightly: Allele M1 is *transmitted* independently of the *transmission* of the polled trait, to the offspring.

Table 3.5 Transmission of alleles in a two-by-two contingency table (Spielman *et al*,. 1993).

Marker M	**M1**	**M2**	**Totals**
Transmitted	a	b	a+b
Not Transmitted	c	d	c+d
Totals	a+c	b+d	4n

A sum is then taken over these families in which at marker locus M there is a total of 4n parental alleles that can be transmitted. 2n are transmitted and 2n alleles are not transmitted. The Chi-Square test of significance with one degree of freedom is expressed as follows:

$$\chi_s^2 = \sum [(O-E)^2/E]$$

Where E = ([(Row total) x (Column total)]/Grand total) and E is the expected frequencies and O the observed frequencies. The Chi-Square $_{(0.01)}$ is calculated at 1 degree of freedom with the P-value halved and the threshold value indicated in **Table 3.6**. In order to be very certain of the result a very stringent allowable fault level was taken as shown below.

University of Pretoria etd, Schmulian A (2006)

Table 3.6 Critical values of the Chi-Square distribution (Samuals, 1993).

Probability	99%	99.90%
Chi-Square at 1 degree of freedom(df)	6.63	10.83

After the results of the Chi-Square are known it is necessary to trace these alleles, visually, in three generation, phase-known families. Three of the largest families in the sample group which had been used in the CRI-MAP run were drawn using CYRILLIC® pedigree drawing software. Unlike the CRI-MAP run, all typed individuals were included not just those that could be shown as heterozygote polled. These family trees are found in **Addendum 5**

University of Pretoria etd, Schmulian A (2006)

CHAPTER 4: RESULTS

A total of 14 microsatellite markers were typed on 365 animals from the Bonsmara families included in the study and final results were obtained for 11 of the markers. The initial linkage analysis in CRI-MAP gave an indication of which markers were suitable with a high number of meioses. The markers DVEPC88 and DVEPC42 were not analyzed in the final CRI-MAP run due to the small number of meiosis for them. The numbers of meiosis per marker are listed in **Table 4.1.**

Table 4.1 Output of the *loc* file Indicating the number of informative meioses per microsatellite marker.

MicrosatelliteMarker	Phase-Unknown	Phase - known
Poll	64	0
INRA257	42	1
INRA212	32	2
DVEPC88	7	0
SOD1MICRO2	34	0
90F5-24	51	1
BMS1928	51	1
INRA117C	38	0
BM8139	26	0
DIK70	51	0
DVEPC42	4	1
BMS574	43	0

The final CRI-MAP run for the markers using the TWOPOINT option was completed and the results are shown for the intervals in **Table 4.2**. CRI-MAP calculates. LOD scores for a range of recombination fractions, the most likely score being the highest one in the range. This is the point at which the LOD score is maximized, and the value is denoted as Z_{max}. A value for Z_{max} above three indicates linkage with a high certainty level (Ott, 1991). This measure therefore provides a built in measure of the significance of the results obtained.

The results show that the tested markers are closely linked to the polled trait (Poll). It was thought that a recombination hotspot occurred in this region and current mapping functions could not

University of Pretoria etd, Schmulian A (2006)

account for this. Distance measures were therefore not included in the results. The calculated recombination fractions (θ) were small, indicating that not only were the markers linked, but closely too, however the BUILD and ALL options responsible for MULTIPOINT LOD scores could not be used on the data due to the low number of phase-known meioses. An ordering of the loci into a genetic map was therefore not possible using CRI-MAP.

Table 4.2 Estimated recombination fractions and associated maximum LOD scores.

Marker Interval	θ	Z_{max}
Poll-INRA257	0.14	3.97
Poll-INRA212	0.15	3.22
INRA257-INRA212	0.00	6.32
Poll-SOD1MICRO2	0.10	4.15
Poll-90F5-24	0.18	3.46
Poll-BMS1928	0.14	5.07
Poll-INRA117C	0.11	3.40
Poll-BM8139	0.18	3.28
Poll-DIK 70	0.10	4.20
Poll-BMS574	0.15	5.28

Linkage was confirmed so the second aim of the study, namely the identification of specific alleles in association, could be tested. Numbers and percentages of each genotype for the different markers were calculated. The frequency of a genotype higher in polled animals than in horned was assumed as first indication that one of the alleles could be associated. Only animals with heterozygote polled genotypes were counted, therefore it was clear which allele was associated with the horned gene and which one with the polled gene. Horned genotypes added even more information at this stage because it is assumed that both alleles in a horned animal are obviously associated with the horned gene. Once an allele had been selected by following these assumptions, it was tested for association.

In **Table 4.3** the genotypes observed for microsatellite INRA257 are shown. The first column indicates the genotype, the next column the amount of typings observed and the next column its' frequency in polled animals. This count is also done on the horned animals in the sample group. The

University of Pretoria etd, Schmulian A (2006)

final column is the difference in percentage between the polled and horned typings. Were the difference values were the greatest, alleles were tested for association.

Table 4.3 Frequency estimates for genotypes of marker INRA 257

INRA 257	**Genotype**	**Polled**	**Frequency**	**Horned**	**Frequency**	**Difference**
	20/22	1	0.36	0	0.00	0.36
	20/29	1	0.36	0	0.00	**0.36**
	21/29	5	1.78	0	0.00	**1.78**
	22/29	40	14.23	2	0.71	**13.52**
	23/29	23	8.19	3	1.07	**7.12**
	24/29	2	0.71	0	0.00	**0.71**
	25/29	10	3.56	0	0.00	**3.56**
	27/29	2	0.71	0	0.00	**0.71**
	28/29	10	3.56	1	0.36	**3.20**
	29/29	6	2.14	0	0.00	**2.14**
	22/22	30	10.68	25	8.90	1.78
	22/23	12	4.27	14	4.98	-0.71
	22/25	0	0.00	3	1.07	-1.07
	21/21	3	1.07	1	0.36	0.71
	21/22	8	2.85	0	0.00	2.85
	21/23	3	1.07	1	0.36	0.71
	23/24	4	1.42	2	0.71	**0.71**
	22/24	5	1.78	0	0.00	**1.78**
	25/24	1	0.36	0	0.00	**0.36**
	23/23	6	2.14	24	8.54	-6.41
	23/25	0	0.00	2	0.71	-0.71
	25/22	2	0.71	0	0.00	0.71
	23/27	1	0.36	0	0.00	0.36
	22/27	2	0.71	0	0.00	0.71
	28/28	1	0.36	0	0.00	0.36
	27/28	0	0.00	1	0.36	-0.36
	20/28	1	0.36	0	0.00	0.36
	25/28	0	0.00	1	0.36	-0.36
	22/28	9	3.20	4	1.42	1.78
	23/28	2	0.71	6	2.14	-1.42
	25/25	1	0.36	0	0.00	0.36
Totals		**191**	**67.97**	**90**	**32.03**	

A total of 281 typings **(Polled plus Horned)** were obtained for INRA257. 191 typings were from heterozygote polled animals and 90 from horned animals. The frequency of the genotype 22/29 was

University of Pretoria etd, Schmulian A (2006)

13.52% higher in polled animals that in horned animals and genotypes 23/29 was 7.12% higher in polled animals. Allele 22 and 23 occurred frequently in horned genotypes, while allele 29 occurred in less than 1% of the horned genotypes. All genotypes containing the allele 29 had a higher frequency for polled animals. This allele was compared in the Chi-Square test by the time-consuming process of counting all transmissions of the allele from parents to child.

Using the same process of elimination, allele 24 of INRA257 was also tested. There was no significance in the Chi-Square result for this allele and therefore other possible associated alleles at low frequency were not tested. It must be stressed here that there could have been more than one allele per markers associated with the polled trait. As the aim of this study was to identify polled animals through associated alleles in a MAS test, such alleles had to be relevant for almost all animals in the Bonsmara population. This is why the very stringent significance level for the Chi-Square at 99% was used. In other words at this level, the chances are 1% of erroneously assuming association. Clearly the significance of association for the other tested alleles was low. The allele 29 of INRA257 was the only allele found through the TDT test to be very strongly associated with the polled trait. The results for the other markers and their frequency tables can be found in **Addendum 4.**

It was important to superimpose data onto family trees in CYRILLIC®. This step ensured that inheritance patterns within families could be verified. This is a necessary step in practice, due to recombination. The pedigree genotype data is necessary to confirm that a rare recombination event has not occurred in any individual, swapping the allele and transmitting it with a copy of the horned gene in progeny. It is also possible to find rare alleles or new mutations segregating in specific families. If allele 29 of INRA257 is to be used to test animals in the future it will be necessary to first check for these possibilities in the pedigrees.

Additionally, the transmission had to be shown in a three-generation, phase-known family without missing genotype and phenotype data. It was found that the family trees were not phase-known, however these trees can be useful in determining the polled status of specific animals in future generations as data is added. The three largest CYRILLIC® family trees are shown in **Addendum 5**.

University of Pretoria etd, Schmulian A (2006)

The frequency tables and Chi-Square results which are contained in **Addendum 4** show that none of the alleles tested for other markers showed a very strong association. Not all alleles were tested, only those with the greatest percentage difference between horned and polled animals. If the first allele at the highest percentage difference did not show a significant association, other associated alleles at lower frequencies were not tested.

Table 4.4 Results of the TDT.

Observed Frequencies	M1	M2	Row Total
Transmitted	40	11	51
Not Transmitted	40	49	89
Column Total	80	60	140

Expected Frequencies	M1	M2	Row Total
Transmitted	29	22	51
Not Transmitted	51	38	89
Column Total	80	60	140

As can be seen from the expected values of transmission for the 29 allele (M1) group and the other allele (M2) group, there is a frequency difference in the direction of the polled group. The results were then halved and compared to the critical Chi-Square value, which indicated a highly significant result.

P value: 14.8
P/2: 7.4 > 6.63 critical value

From Table 3.6, taking the significance at 99%, the critical Chi-Square value is 6.63. Therefore the P/2 of 7.4 is greater. There is a 1% chance that (H_0) is mistakenly rejected, and so the data provides sufficient evidence in favor of (H_A). The allele 29 of marker INRA257 is associated and transmitted with the polled gene from polled parent to polled offspring.

University of Pretoria etd, Schmulian A (2006)

CHAPTER 5: DISCUSSION

The results of the microsatellite marker analyses have shown linkage in the Bonsmara families tested. The markers where chosen for their proximity to the polled gene. Results in this study were obtained from the analyses of 11 microsatellites tested on 365 Bonsmara cattle. The size of the sample allowed for conclusive LOD scores to be determined in the linkage analysis. The point at which the LOD score was maximized (Z_{max}) in the range generated by maximum likelihood, for each marker was well above the critical threshold of three, as is shown in **Table 4.2.**

From the preceding literature on other breeds tested for linkage for the polled trait, it was evident that the gene controlling the trait is located on BTA1. The Bonsmara breed is unique in that it is a composite of British *Bos taurus* and the Afrikaner which is indigenous to Southern Africa. The question of whether the gene action in the Bonsmara was the same could be answered by the linkage results and phenotypic observations. The location of polled gene close to the BTA1 markers was confirmed in the current linkage study, for the Bonsmara breed, indicating a similar genetic model as that found in other studies.

According to Weir (1997) calculating the recombination fraction through maximum likelihood, provides an inaccurate result when distances under one cM are considered. Therefore refining the position of the gene by using the linkage data is not possible and was not the aim here, especially since the published maps did not at the time of this study show any intervals for physical anchor markers under one cM.

Results of the CHROMPIC option for the data showed a high number of double recombinants. This points to possible typing errors or a very high recombination rate. Band *et al.* (1997) isolated microsatellites from a YAC clone that were linked to polled gene. He and his associates reported an abnormally high recombination rate present in the region, as the genetic distance between microsatellites observed by them was five cM. This is much greater that the total physical distance of a entire YAC clone; the YAC clone they used was less than one cM long. The indication is that the region around polled has an unusually high recombination rate which supports the results of our

University of Pretoria etd, Schmulian A (2006)

study. The CRI-MAP, CHROMPIC option used to indicate recombinants, therefore reflected this higher recombination rate (Green , 1990).

A higher than expected recombination rate would also cause a bias in any distance calculation for a genetic map. It is of critical importance in future to determine mapping functions for specific regions on a chromosome. More research is therefore required to determine the exact relationship between the physical and genetic distances in the region of the centromere of BTA1. This is still very difficult because sufficient resolution of the physical map is necessary. Until such a time at the region around the polled gene is fine mapped by physical means, its exact location remains difficult to pinpoint by constructing genetic linkage maps.

Due to the low numbers of phase-known meioses calculated by CRI-MAP (**Table 4.1**), a genetic map could not be created using the BUILD function. The grandparental genotypes of the heterozygote polled bulls used in the study were missing and therefore their phase was unknown. This was the main reason for the small number of phase-known meioses. A MULTIPOINT LOD score could not be generated, and the most likely order of the microsatellites was not calculated. This made it impossible to place the markers in order relative to each other and the polled trait.

There was no loss of power in the determination of the LOD score and a TWOPOINT LOD score is considered accurate and robust. In a TWOPOINT analysis, a bias in calculated recombination fraction is not serious for detecting linkage. This is not true in a multipoint analysis (Ott, 1991). The sample size in this study was large and the linkage result is cumulative over families. If a Z_{max} of above the value of three is found, this is significant regardless of the structure of the data set. At this level the significance equates to a thousand to one that you will be making a mistake to accept linkage. The pair wise linkage result calculated in the TWOPOINT procedure of this study show high LOD scores.

The pair wise linkage result calculated in the TWOPOINT LOD score for INRA212 with INRA257 was also included in the result set. It was hoped that INRA 257 could be placed in order on the map but due to the fact that a MULTIPOINT LOD score could not be calculated, this was not possible.

University of Pretoria etd, Schmulian A (2006)

Stone *et al*. (1995) estimated that the frequency of $(TG)_n$ microsatellites occur every 15000 to 44000 base pairs. The optimization procedure with this multiplex was difficult and several microsatellite primers had to be re-designed in order to amplify together. The multiplex was also done in two steps. A high level of background noise was found across all gels typed in the study. These two factors as well as the recombination rate in the region would indicate a structural difference on that part of the chromosome. This could be ascribed to a high frequency of repeat sequences over a long distance, causing non-specific annealing.

Green *et al*. (1990), the author of CRI-MAP notes that allele frequencies are usually estimated in a population different from that population from which the disease family is taken and that allele frequencies vary dramatically between populations. CRI-MAP does not use allele frequencies in the calculation of likelihood for missing individuals but in a full likelihood analysis the population allele frequencies are used to assign probabilities to various possible genotypes for a missing “original parent”. The other genotype information in that family is rather utilized by CRI-MAP for an allele in a child of that parent. A LOD score is therefore calculated for a family regardless of population allele frequencies.

The linage test in CRI-MAP does not differentiate on the level of specific alleles linked to the polled trait and therefore a separate association test had to be performed in order to determine specific alleles associated with the trait. With the use of the TDT, association at allelic level was shown. Allele 29 of INRA 257 was found to be associated with the polled animals by this method with a high level of significance. This allele may contribute to MAS test which researchers can use in future to determine homozygote polled animals.

Family data remains critical to the success of a TDT test because associated alleles can differ between families and breeds. The pedigree data must be complete, and the tested individual needs to be phase-known which implies a three-generation family tree. The breed and population from which the individual comes must therefore be known. There is the added possibility that a rare recombination between the associated allele and the polled trait might occur, swapping it to horned and the only way to determine this would be to trace the trait and allele visually in the pedigree tree.

University of Pretoria etd, Schmulian A (2006)

Visual inspection of all the genotypic and phenotypic data for a family must be done in order to allow for the identification of a rare allele associated with the polled trait or a new mutation. The CYRILLIC® family drawing tool was used in this study and allowed for very large families to be drawn and visualized as can be seen in **Addendum 5.**

Families must always be considered independently at this analysis stage if allelic association is the method used to determine polled status. Phenotypic data on polled status remains very important in identification efforts in the national herd. Data collection should be very accurate also so that errors in discerning horns from scurs, are not made. Scurred animals were not included in the linkage analysis. It is also suggested that scurred animals be culled from herds so that the polled gene is not masked phenotypically as in the case of heterozygote scurred males and homozygous scurred females and males. The gene action for this effect is indicated in **Table 2.1** above.

Few Bonsmara families were complete and no experimental herds were available for the breed. Few herds had an accurate selection program in place for polled animals and therefore some difficulty was experienced in finding phase-known families. However the breeders of the herds used in this study had been keeping very strict record of the polled trait. Identification of animals was also strictly regulated among them so that these breeders would maintain their reputation of supplying animals for top breeding and export purposes. Phenotypic data obtained on families in this study was therefore very accurate.

It is suggested that a national effort should be made to keep records of the polled trait, and more technicians must be trained to judge the phenotypic status of animals. Regular observations will increase the accuracy of the identification process and must include examination at calving (cow and calf), three months of age, weaning, yearling and at slaughter age.

The methods used by these breeders were verified by the Breeders' Association. There were several cases of polled animals that spontaneously occurred in other Bonsmara herds, which were not included in this study, due to the fact that those breeders did not have a polled selection program. As stated previously, few phase-known families were available for this study because in a natural population older animals are culled and this was the case for the grandparental generation of the

University of Pretoria etd, Schmulian A (2006)

bulls sampled. Some test crossings (*Ph x hh*) were however still available but not in the three generation structure.

Any variation in the trait, such as double horns, a single horn and the size and quality of horns should be noted in future. Occasional oddities were found by Bonsmara breeders, such as a cow with double horns on each side of the head. In the studies by Georges *et al*. (1993) and Brenneman *et al*. (1996) brief mention is made of the variability amongst the categories of horned and scurred animals. The exact interactions between these phenotypes need to be studied further and the model for polledness refined to include such examples.

In this study polledness was considered the major effect. Scurs has recently been mapped to a different chromosome and the frequency of African horn gene is low in the herds tested. A clear example for the African horn gene was found, confirming the presence of this effect in the Bonsmara. This animal was initially phenotyped as polled and later developed scurs.

The results of the current study indicated that nine of the markers were linked to the polled trait. This result together with the phenotypic data collected showed that the gene action in the Bonsmara matches that of the current model for polledness. These markers are therefore all postulated to lie close to the polled gene. The association test was successful for one allele on the marker INRA257 that could be used to test for polled status in future pedigrees. Breeders should continue to phenotype animals as accurately as possible so that this data as well as DNA samples can be submitted to add to the information of a central genetic bank for the Bonsmara breed. This could allow for the accurate verification of polled status across the entire breed in future.

University of Pretoria etd, Schmulian A (2006)

CHAPTER 6: CONCLUSION

This study has been the first step in the compilation of a MAS program focusing on the polled trait in the Bonsmara breed of cattle. At the current level of resolution of the Bovine genome map, a microsatellite linkage study was considered to be a suitable methodology to use for this purpose. From this study, it has been shown that the polled gene does occur in the same interval on BTA1, in the Bonsmara breed as in other breeds tested previously. These results also indicate that the model held by George *et al.* (1993) can be applied to the Bonsmara breed. Linkage was found with nine of the initial eleven microsatellites, although an ordering of the microsatellites into a genetic map was not possible. The first aim of the study was satisfied by proving linkage.

Linkage was determined using CRI-MAP but a TDT was used to find specific allelic association. The presupposition of linkage, allowed for the determination of association between allele 29 of marker INRA257 and the polled gene. A very stringent significance level was chosen so that this test could be used uniformly for the whole Bonsmara population. This allele can be used to test for polled status in future, if complete phase-known family data is also made available together with the DNA sample. The aim of showing association was therefore satisfied.

The genotypes of animals were superimposed on to CYRILLIC® trees and traced. It was however not possible to verify status of animals in these family trees because animals were not phase known. There are two very important reasons for the need to always visually inspect pedigrees. Firstly, as stated, a rare allele or a mutation forming a new allele can also occur in families in the Bonsmara population. If family trees are to be used, the parental generation must be phase-known.

The genotypic and phenotypic pedigree data from this study can be used in further microsatellite tests on future generations from these same families, as a tool to determine polled heterozygotes from homozygotes. This is because the next generation will be phase-known. The success of this MAS test therefore depends on knowledge of the breed and family as well as on correct phenotypic data and complete DNA sampling of these families.