INDIVIDUAL IDENTIFICATION AND PARENTAGE ANALYSIS OF *STRUTHIO CAMELUS* (OSTRICH) USING MICROSATELLITE MARKERS

FATIMA ESSA

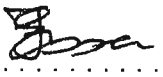
Submitted in fulfilment of the academic
requirements of the degree of
Master of Science
in the School of Biochemistry, Genetics, Microbiology and Plant Pathology,
University of KwaZulu-Natal

Pietermaritzburg 2005

PREFACE

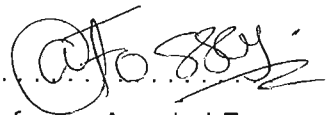
The experimental work described in this dissertation was conducted at the Institute for Animal Production at the Western Cape Department of Agriculture: Elsenburg, under the supervision of Professor Schalk Cloete, in collaboration with the University of KwaZulu-Natal, Pietermaritzburg, under the supervision of Professor Annabel Fossey.

The results have not been submitted in any other form to another University and except where the work of other is acknowledged in the text, are the results of my own investigation.



December 2005

I certify the above statement is correct.



Professor Annabel Fossey
Supervisor

ACKNOWLEDGEMENTS

I would like to acknowledge the contributions towards this project, which made it possible for me to start, and complete, my thesis.

I express my sincerest gratitude to my supervisors, Professor Annabel Fossey and Professor Schalk Cloete, whose help has been invaluable.

Thank you to the staff of the Oudtshoorn Experimental farm (KKADC) for the maintenance of the experimental ostrich population and the recording of data.
I would like to thank Zanell Brand for coordinating the collection of blood samples.
I am grateful for the technical information provided by Anel Engelbrecht and Dr Irek Malecki (University of Western Australia).

I appreciate the financial contributions made by the following organisations:

- The Klein Karoo Cooperation (for the use of the resource flock)
- THRIP programme of the South African Department of Trade and Industry.
- Western Cape Department of Agriculture – Young Professional Programme (Elsenburg)
- Western Cape Department of Agriculture – Institute for Animal Production (William Gertenbach)
- National Research Foundation Masters scholarship funding

I appreciate the help and support offered by my friends, Mariaan Viljoen, Ansie Scholtz and Pavarni Naidoo in proofreading and typing this thesis.

I am indebted to my parents for their continued support of my education.

Thanks to my husband, Bashier Amien, for his patience and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Ostrich (*Struthio camelus*) breeding is a well-developed industry in South Africa. However, successful genetic management has yet to be implemented. Parentage in colony breeding ostriches is unknown, where for a given offspring, a number of possible parents exist. Molecular markers have been extensively used in the livestock industry to resolve parentage issues and are only beginning to be utilized to address the issues of the ostrich industry. The aims of this investigation were to test known microsatellite markers developed for other ostrich subspecies in a South African Black ostrich population, and to further test these markers for their use in individual and parentage identification. DNA was extracted from venous blood obtained from two pair bred families and a colony of 97 individuals. Eleven polymorphic microsatellite markers were tested by PCR amplification of DNA samples followed by multiplexing on polyacrylamide gels to generate DNA fingerprints for each individual. Alleles were sized and quantified and used to create genotypes for each individual. Parentage analysis was performed using exclusion and likelihood methods. Pedigrees were constructed for the families by comparison of genotypes. Breeding statistics were calculated for the colony individuals. Three microsatellite markers did not amplify in this population and one marker was found to be monomorphic in this population. Four of the microsatellite markers that successfully amplified produced anonymous amplification products suggesting a second annealing site in the genome sequence of Blacks. All loci displayed low observed heterozygosities indicative of little genetic variation in this population. For the colony sample, four individuals were not assigned either parent and one female did not contribute any offspring. On average females produced 4.86 ± 2.71 fertile eggs during the sampling period with a coefficient of variation of 55.86%. A total of 79.2% of individuals were assigned paternity and 88.3% were assigned maternity. A greater number of loci are required to improve the power of parentage analysis within breeding flocks incorporating all eggs laid.

CHAPTER 1
INTRODUCTION

1.1 INTRODUCTION

The ostrich industry plays an important role in the South African economy, which started with the local demand for ostrich feathers. This sought after fashion accessory rapidly gained international popularity and, by 1913, worldwide demand for ostrich feathers reached its peak, with South Africa being the sole supplier of ostrich products (SAOBC 2004).

Changes in fashion trends, together with the worldwide economic depression at the time of the First World War, resulted in the supply of plumes exceeding demand. These market forces are what ultimately caused the initial collapse of the lucrative ostrich industry.

Currently the market for ostrich products includes: leather, meat, feathers and oil. Ostrich feathers, which initially spawned the industry, today only form a by-product of the ostrich industry. To sustain an operation based on feathers, the volume of production required combined with the labour-intensive nature of the process makes it a relatively unattractive economic prospect. The same applies to ostrich oils, considered to be therapeutic (Shanawany 1995).

Ostrich leather is the strongest commercially available leather in the world, and is valued for its distinctive quill pattern, softness, and suppleness. It is considered an exotic skin and is traded alongside the hides from reptiles and other wildlife. These exotic skins are used in the manufacture of garments, handbags, luggage and other small leather goods. Ostrich leather currently makes up the bulk of the exotic skins industry, in terms of volume traded (SAOBC 2004).

Whereas animal hides are generally produced as a by-product of the meat industry, the converse applies to the ostrich industry, with income from ostrich hides exceeding income from ostrich meat. Ostrich is a red meat, and has gained

widespread popularity over recent years, mainly due to its health benefits (Baronigg 2002). Ostrich meat is high in protein, but low in fat, calories and cholesterol, when compared to traditional red meats. Recent outbreaks of avian influenza have however somewhat curtailed the demand for ostrich meat products.

Ostrich products are widely accepted as luxury items, with in-elastic demand characteristics meaning there is no significant change in demand in response to a change in prices. There still appears to be a growing demand for ostrich products, with the major markets in order of size being as follows:

a) Leather products: Japan, USA, and Europe
b) Meat products: Europe (mainly Germany), North America and Japan.

Global awareness of the excellent economic potential of ostrich products has lured new industry participants such as China, Israel, United States, Australia and Europe into this industry in pursuit of this potential profitability. However, it would appear as if many of these countries have underestimated the costs of entering the market, with many of them struggling to produce acceptable economic results, and clearly being unable to compete with the established South African industry, which has benefited from strong economies (SAOBC 2004).

1.2 THE SOUTH AFRICAN OSTRICH INDUSTRY

South Africa, pioneers of the ostrich trade, revived the industry from its post World War II slump. In 1959 a one-channel co-operative marketing system, The Klein Karoo Cooperative, was established. In 1964 the first abattoir was built, while the first tannery was established in 1970. To date there are ten export-approved abattoirs, and 15 tanneries, throughout the country (SAOBC 2004).

The aim of the co-operative marketing system was to regulate the industry and set the quality standards for ostrich products. It also meant that South Africa was the only country to produce ostrich products on a commercial scale. This marketing system was eventually abandoned in 1993, opening up the markets to globalisation, thus allowing other countries to legally engage in commercial ostrich production.

The entrance of new countries into the market have resulted in a reduction of South Africa's share of global ostrich production from 82% in 1995, to 65% in 2000 (SAOBC 2004). Global ostrich production figures for 2004 (Figure 1.1) show that South Africa remains a world leader with 52% of the market (Stewart 2004). Despite the impact of increased competition from globalisation, there are approximately 600 registered breeding farms, with 450 of them registered for slaughter producing approximately 300 000 slaughter birds annually, and employing approximately 20 000 workers (SAOBC 2004).

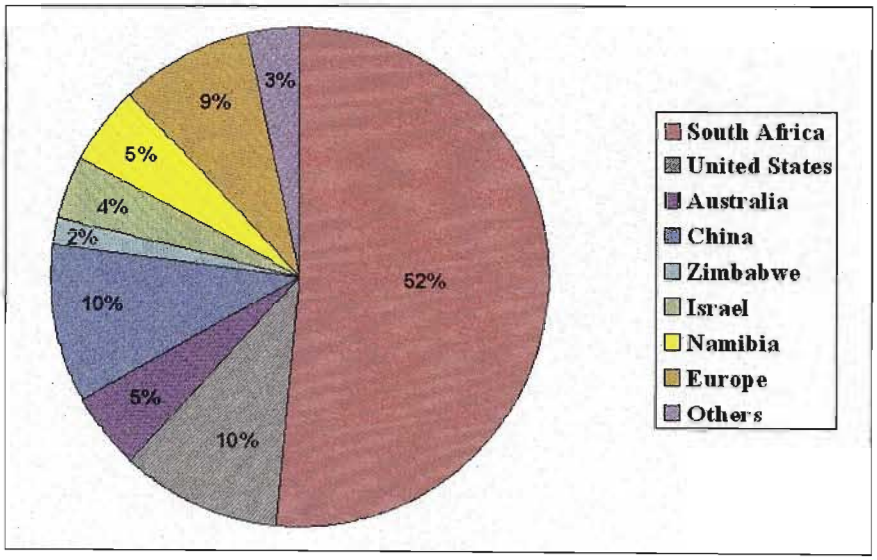


Figure 1.1 Global ostrich production for 2004 expressed as a percentage per country (Adapted from Stewart 2004).

The farms engaged in ostrich activities have spread from the Klein Karoo, in the Oudtshoorn area of the Western Cape to other provinces. The industry has been re-organised into an umbrella structure called the South African Ostrich Business Chamber (SAOBC), which is supported by two key players:

- The National Ostrich Processors of South Africa (NOPSA)
- The South African Ostrich producers Organisation (SAOO)

Total investment in direct ostrich activities exceeds R2,1 billion, with income from exports alone that amounts to approximately R1,2 billion annually which equates to 90% of all local meat and leather production. Typically, the value yielded by a

South African bird can be broken up as 45% skin, 45% meat and 10% feathers. This contrasts with foreign yield of approximately 75% meat, and 25% leather (SAOBC 2004).

South Africa enjoys a position of leadership in the ostrich industry owing to its long heritage and natural conditions that are ideal for the breeding of these unique birds.

1.3 THE OSTRICH

The ostrich (*Struthio camelus.*) is the largest living bird. Ostriches stand between 2.1 m and 2.5 m tall and can weigh over 100 kg at maturity. Their bodies are covered with soft brown feathers as chicks, which changes to black for males at maturity. The ostrich has a long flexible neck, long bare legs with developed thigh muscles and two toes. These strong legs allow them to run at a speed of up to 70 km/h. The great body size and reduced wing size of the ostrich renders this bird incapable of flying (Deeming 1999).

Ostriches are paleognathous birds belonging to the order *Struthioniformes* (Harlid & Arnason 1999) and are more commonly known as Ratites or flightless birds, which includes the emu, the cassowaries, the kiwi and the rheas. These birds have a flat breastbone and an archaic palate (Cooper *et al.* 1992; van Tuinen *et al.* 1998). The family *Struthionidae* has one species *Struthio camelus* with four currently recognized subspecies, which are all indigenous to Africa and a recently extinct subspecies *S. c. syriacus,* formerly found in the Middle East (Jarvis 1998). The subspecies are the North African strain *S. c. camelus*; the Kenyan 'Reds' *S. c. massaicus*; the Somali 'Blues' *S. c. molybdophanes* and the Zimbabwean 'Blues' *S. c. australis.* The 'Oudtshoorn Blacks', *S. c. domesticus*, are believed to be a hybrid between *S. c. camelus* and *S. c. australis* (Deeming 1999). These subspecies differ slightly in size and in the colour of the bare skin of the thighs and neck (Jarvis 1998; Kumari & Kemp 1998) as shown in figure 1.2.

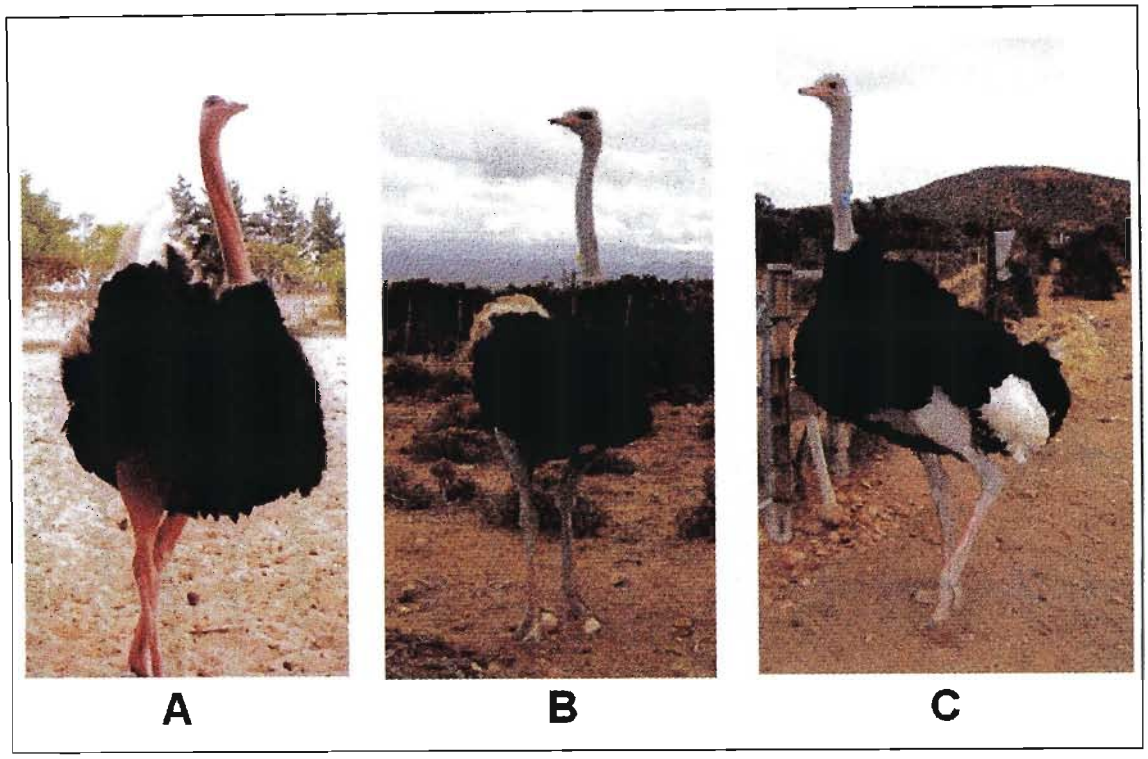


Figure 1.2 Phenotypes displaying three subspecies, Kenyan Red (A), Zimbabwean Blue (B) and South African Black (C).

Ostriches display a typical avian karyotype possessing macro- and microchromosomes. The diploid chromosome number for the ostrich is 2n=2x=80 (Takagi *et al.* 1972) with about six pairs that are distinguishable as macrochromosomes (Petitte & Davis 1999). The sex chromosomes, as with other avian species, are assigned ZZ for the male and ZW for the female. However, unlike other birds, ratite sex chromosomes are monomorphic meaning that the two chromosomes are indistinguishable in appearance as shown in the karyotype of figure 1.3 (Ogawa *et al.* 1998; Petitte & Davis 1999). The structural similarity of the sex chromosomes in this species is possibly reflected in the lack of sexual dimorphism in the juvenile ostriches (Takagi *et al.* 1972). Day-old chicks can be sexed fairly accurately by the visual inspection of the cloaca, but it is almost impossible to determine the gender of older chicks externally until the well-known dimorphic adult plumage is displayed at sexual maturation, at the age of approximately three years (Soley & Groenewald 1999). Despite the similarity between the sex chromosomes, rapid DNA-based methods for the sexing of

ostriches have been successfully developed (Bello & Sanchez 1999; Huynen *et al.* 2002; Malago *et al.* 2002; Mine *et al.* 2002; Hinckley *et al.* 2005).

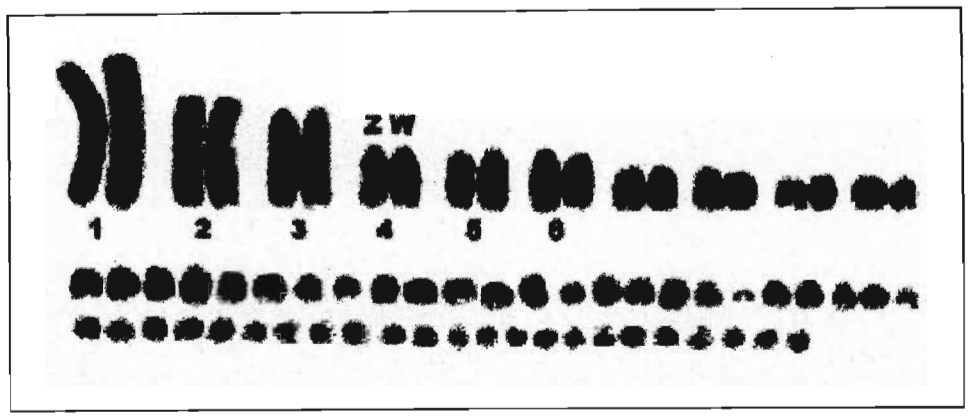


Figure 1.3 Karyotype of a female ostrich illustrating the macrochromosomes and microchromosomes. Sex chromosomes are labelled ZW (Adapted from Petitte & Davis 1999).

The ostrich is very adaptable and thrives under extreme conditions and very poorly vegetated desert areas. Ostriches are indigenous to the desert, semi-desert and savannah regions of Africa (Deeming 1999). Globalisation of the ostrich industry has found these birds to survive in colder climates such as Canada and Poland. In South Africa ostrich activities are centred around the Klein Karoo region in the Western Cape but has spread to the other provinces to include the Free State, Gauteng, Limpopo, Mpumalanga, North-West and North Cape regions (SAOBC 1994). Figure 1.4 shows the distribution map of all registered commercial ostrich farms within South Africa.

Generally the birds commercially farmed in South Africa are the Blacks. These hybrids came about through years of selection for feather production. This bird is characterized by its smaller stature and well-developed feather structure and its docile nature often lends it to being referred to as domesticated (Deeming 1999). In recent times Blues have been introduced into the breeding stock of some farms and attempts are being made to crossbreed the Blues and Blacks (Brand *et al.* 2005). However, these birds are larger and tend to be 'wilder' than the Blacks and very little is known about their performance in the South African conditions.

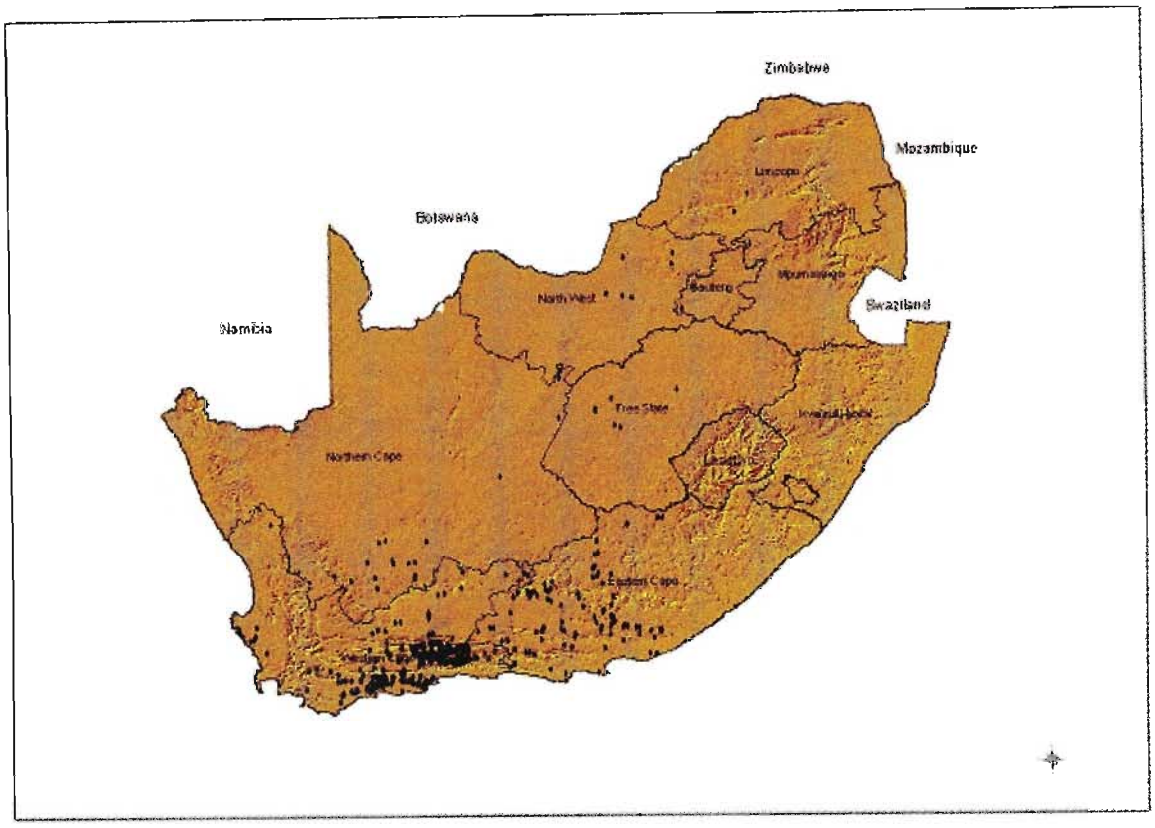


Figure 1.4 South African map showing the distribution of ostrich farms indicated by black dots.

1.4 EVOLUTIONARY AND DIVERSITY STUDIES

The mitochondrial DNA (mtDNA) of the ostrich has been sequenced and studied to determine the evolution of this bird as well as to assay genetic diversity between and within populations. The size of the complete mtDNA molecule of the ostrich is 16 591 nucleotides (Harlid *et al.* 1997).

There is much debate with regards to the origin of the ostriches and how these flightless birds came to be distributed across the southern continents. Van Tuinen *et al.* (1998) studied the phylogenetic relationships by assaying mitochondrial ribosomal genes and concluded that dispersal and vicariance (continental break-up) best explains the origin of these birds. In another study, Haddrath and Baker (2001), using phylogenetics concluded that dispersal best explains the present distribution of the ostrich. Ratites are recognized as paleognathous birds,

however, a study analysing the mtDNA suggests a neotonous origin of morphological characters of the ratites and that the ratites are descendants of flying neognathous ancestors (Harlid & Arnason 1999).

The mtDNA has been studied to assess the genetic diversity of populations. mtDNA was used to investigate the phylogeographic patterns of the wild subspecies of ostrich (Freitag & Robinson 1993). This analysis concurred with the currently accepted designations of subspecies. A later study by Bezuidenhout in 1999 used mtDNA to investigate the relationships between ostrich subspecies and to assess the genetic diversity between and within populations of the southern African ostrich *S. c. australis*, and found no genetic variability between the populations. A recent genetic diversity study that used nuclear DNA and microsatellites instead of mtDNA found the highest genetic variability in Blacks and the lowest in Reds in the population investigated. This study also showed that the greatest genetic distance exists between Blacks and Reds, as is shown by the neighbour-joining tree in figure 1.5, which indicates that the highest heterosis effect will be obtained by crossing the subspecies (Kawka 2005).

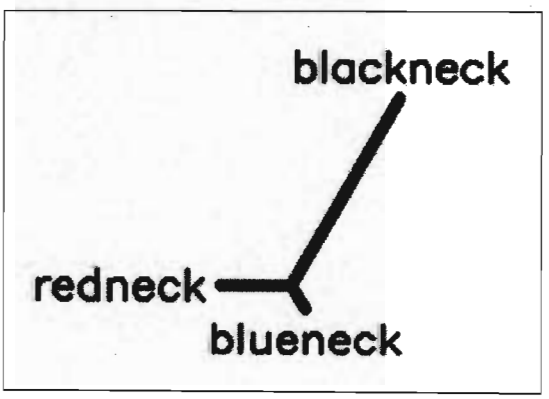


Figure 1.5 Neighbour-joining tree showing the genetic distance relationships between the three main subspecies (Kawka 2005).

1.5 BREEDING PRACTICES

There are three types of breeding systems commonly used within the ostrich industry. These are pair breeding, trio breeding and colony breeding.

Pair breeding is the monogamous type of mating where one male is mated with one female in a paddock. This type of breeding enables accurate pedigree recording. However, the management practice of repeatedly pairing the same males from year to year in the same breeding paddock confounds the potential random factors contributing to performance (Bunter 2002). Furthermore, the capital outlay required for the making of individual breeding paddocks, as is required for this type of breeding system, makes it prohibitive for the majority of ostrich farmers.

Trio-breeding, defined as one male mated with two females, leads to an improved data structure when compared to pair breeding. However, accurate recording of pedigrees is a problem unless eggs from the two females can be differentiated on egg size or shell structure (Cloete *et al.* 1998; Bunter 2002). This is feasible in practice, provided that there is some indication of a weight difference between the eggs of the females involved (Essa & Cloete 2004). The capital outlay required by this system is similar to that of the pair-breeding system making this system an unfavourable option to the majority of ostrich farmers.

Colony breeding or flock mating is a system that occurs naturally in wild populations, which is referred to as communal nesting. Within a controlled environment, such as a breeding farm, birds are mated in groups, usually at a ratio of six males to ten females (Cloete *et al.* 1998). In this system it is difficult for ostrich farmers to identify non-breeding birds within the flock because the parentage of eggs and chicks is unknown (Bunter 2002; Bunter & Cloete 2004). Due to the cost-effectiveness of this system it is the system practiced by the majority of ostrich farmers.

Since the growth of the industry was initially based on the international demand for feathers, birds were selected in terms of their phenotype for feather quality. Characteristics such as egg production and rate of gain were not considered at that stage (Petitte & Davis 1999). Despite the advances made in livestock breeding through selection, the majority of ostrich farmers continue to practice selection based on phenotypic characteristics with little progeny testing.

In an ongoing attempt to improve the breeding practices of ostrich farmers, artificial insemination (AI) technology and its application in ostriches has been explored. Semen collection methods for ostriches have been developed. Difficulties with the collection methods, with the semen preservation and storage, and inadequate knowledge of dosage rates that ensure ongoing fertility, have hindered its use in AI (Malecki & Martin 2002a; 2002b). Another complication for AI is that ostriches appear to be induced breeders that is, the presence of males acts as the stimulus for egg production in females (Cloete *et al.* 1998; Bunter & Cloete 2004). Although AI is a plausible option in most livestock breeding schemes, it is not routinely practiced in ostriches.

Commercial ostrich farming systems have characteristically low reproductive performances, a large variation in egg production and high chick mortalities when compared to small domestic poultry species (Van Schalkwyk *et al.* 1996; Cloete *et al.* 2001; Lambrechts 2004). A study undertaken in the United Kingdom (Deeming *et al.* 1993) on chick rearing showed a 67-78% survival rate to three months of age for ostrich chicks. This study also revealed that mortality was restricted mainly to the first four weeks of rearing. A study by Cloete *et al.* (2001) investigated the factors that related to the high levels of ostrich chick mortality in an intensive rearing system. In this study it was concluded that the high levels of mortality could be related to stress in chicks, resulting from an inability to adapt to the rearing environment. Soley and Groenewald (1999) suggested that a number of factors affected fertility. These include the use of immature males, periods of reproductive quiescence, nutrition, behavioural disorders, environmental stresses and diseases of the reproductive system. For the industry to move into a viable production phase, the average fertility rate of ostrich females needs to be improved significantly (Badley 1997).

The ostrich industry in South Africa practices no formal breeding strategy and lacks well-established breeding objectives (Petitte & Davis 1999). This is alarming when consideration is given to the marked advantages that were made with the breeding of domestic livestock over the past number of decades. These improvements have been particularly noticeable in the more intensive poultry, pig

and dairy cattle industries. Furthermore, crossbreeding is occurring in a random fashion without proper guidance and with little consideration to potential benefits such as heterosis (Cloete *et al.* 2002). A scientifically based breeding policy for farmed ostriches needs to be formulated but requires that production data first be obtained with linkage to pedigree information, to enable the derivation of genetic parameters and the estimation of breeding values (Van Schalkwyk *et al.* 1996). In recent times molecular technology has developed at a rapid rate. In the ostrich industry, knowledge about the molecular composition of individuals and populations of the different subspecies could facilitate an understanding of genetic diversity, which can contribute to marker-assisted selection in breeding and parentage analysis.

1.6 MOLECULAR GENETICS IN THE OSTRICH INDUSTRY

1.6.1 Introduction

Individuals have unique DNA. This DNA is susceptible to recombination and mutations, which are responsible for the generation of variation in DNA. In a population, these variations in the DNA are assessed in terms of polymorphism where, if more than one variant exists at a locus, the locus is termed polymorphic. The development of molecular markers based upon polymorphisms found in DNA has revolutionized areas of biology, including the estimation of genetic distances between populations, families and individuals.

1.6.2 Fingerprinting

The technique of DNA fingerprinting is based on the identification of polymorphic DNA that varies between individuals in a population. Molecular markers, the tools used to identify individuals and assess genetic variation, are based upon polymorphisms found in DNA (Cunningham & Meghen 2001). Molecular markers function by amplifying specific sequences of the DNA using short pieces of DNA called primers to produce amplification products. These products are then separated in a suitable matrix to produce a profile or fingerprint (Avise 2004).

Molecular markers can be divided into two classes, namely, random markers and specific markers. Random markers use arbitrary primers to amplify and do not require any prior knowledge of the DNA composition, thus making them easy and cost-effective to develop. Specific markers, however, require specific primers for amplification and require some prior knowledge of the DNA sequence to be amplified. It is, therefore, an expensive, labour-intensive process to develop these types of markers (Dodgson *et al.* 1997; Parker *et al.* 1998; Van Marle-Koster & Nel 2003). Molecular markers can be described as dominant or co-dominant depending on the type of output they yield. Dominant markers are indicated by the presence or the absence of a band and it is not possible to differentiate between a homozygous dominant individual and a heterozygous individual, both indicated by the presence of a band. The absence of a band indicates a homozygous recessive individual. Co-dominant markers, however, allow for the differentiation between homozygous and heterozygous individuals.

Random markers consist of random amplified polymorphic DNA (RAPD) and amplified fragment length polymorphism (AFLP). Polymorphisms generated by these markers display genetic dominance and are scored by the presence or absence of bands (Avise 2004).

The RAPD technique involves screening DNA for interpretable polymorphisms using short primers of 8-10 nucleotides of arbitrary sequence to amplify at random from anonymous genomic sequences (Avise 2004). No prior sequence information is required to utilise RAPDs or to produce a primer (Parker *et al.* 1998). The short length of the primers allows them to anneal to the genomic DNA of interest at multiple sites and amplification occurs if primers anneal in the proper orientation and at a suitable distance apart. This technique allows for the examination of multiple loci very rapidly since fragments are usually generated from different parts of the genome (Karp *et al.* 2001). RAPDs function as dominant markers meaning that it is not possible to differentiate between a homozygous dominant individual and a heterozygous individual. The absence of a band indicates a homozygous recessive individual whereas the presence of a band indicates either a homozygous dominant or heterozygous individual. RAPDs are a very powerful technique for screening populations for sequence diversity. The dominant property

of this marker makes it applicable to sex chromosome mapping (Dodgson *et al.* 1997). RAPDs have the advantage of being low-cost markers, however, they tend to have poor reproducibility (Dodgson *et al.* 1997; Parker *et al.* 1998; Avise 2004).

AFLP is a technique based on the selective amplification of a subset of restriction fragments from a complex mixture of DNA fragments obtained after restriction endonuclease digestion of genomic DNA. Polymorphisms are detected by differences in the length of the amplified fragments by polyacrylamide gel electrophoresis (PAGE) (Karp *et al.* 2001). The process of generating AFLPs is much faster and provides more information than RAPDs. First the genomic DNA is digested with one or more restriction enzymes. Then double stranded oligonucleotide sequences or adapters are ligated to the restricted fragments. For this step to occur there has to be complementary base pairing between the restriction fragment overhang and the adapter overhang, therefore, knowledge of the restriction site is required to generate these primers. After ligation of the adapters to the fragments, PCR is carried out using primers complementary to the adapters. By manipulating the number of nucleotides in the adapters, the number of amplified fragments can be adjusted (Karp *et al.* 2001). Polymorphisms are detected as the absence or presence of bands due to a difference in restriction sites or insertions or deletions within the amplified restriction fragment. Hence AFLPs are dominant markers and no distinction can be made between a homozygous dominant individual and a heterozygous individual. The AFLP technique is used extensively in plant genome mapping but more recently has been applied in genetic studies of livestock (Van Marle-Koster & Nel 2003). The advantages of AFLP are that they are reliable and highly reproducible multi-locus markers and that the developmental costs are low. However, the running costs are high and extracted DNA must be of high quality and purity (Van Marle-Koster & Nel 2003).

Specific markers include restriction fragment length polymorphisms (RFLP) and simple sequence repeats (SSRs). These markers display co-dominance and are scored by the presence of one or two bands indicating homozygosity and heterozygosity, respectively (Parker *et al.* 1998).

The RFLP technique is based on the amplification of variable regions of the target genome, followed by sequence-specific cleavage with restriction enzymes (Van Marle-Koster & Nel 2003). Mutations can change the nucleotide sequences in restriction enzyme cleavage sites, thus prohibiting the enzymes from recognizing them, or they may create new restriction sites. These mutations then result in variations in the lengths of the DNA fragments produced by restriction enzyme digestion, hence the name restriction fragment length polymorphisms or RFLPs (Snustad & Simmons 2000; Karp *et al.* 2001; Avise 2004). RFLP variation can be visualized directly by ethidium bromide staining of an agarose gel following electrophoresis of the restricted DNA and this is commonly done for small molecules such as mitochondrial DNA (Parker *et al.* 1998; Snustad & Simmons 2000; Avise 2004). Alternatively, since RFLPs change the size and number of DNA fragments produced by restriction enzyme digestion, they can be detected by the Southern Blotting procedure. Briefly, this involves the hybridisation of labelled DNA probes near the restriction site allowing for the identification of the fragment position in the gel (Parker *et al.* 1998). RFLPs are co-dominant markers, meaning that the presence of both the dominant and recessive alleles can be detected in heterozygous individuals (Karp *et al.* 2001). Depending on the level of genetic variation in the population, this process can yield sufficient polymorphic loci to investigate questions within and among populations (Parker *et al.* 1998). RFLPs can be used in conjunction with the polymerase chain reaction (PCR) in a technique aptly named PCR-RFLPs. The idea behind this technique is first the amplification of a fragment of DNA under investigation followed by digestion with restriction enzymes. This practice enhances the chance of finding a polymorphism within the specific DNA fragment (Karp *et al.* 2001). PCR-RFLP technique was used in ostriches in a study of the population structure (Bezuidenhout 1999). RFLP is largely used in genome mapping and is an easy, reliable, cost-effective marker to use. However, it is labour-intensive to develop and to type and has a low polymorphic content compared to microsatellites (Dodgson *et al.* 1997).

SSRs are due to differences in the number of copies of short sequences that may be repeated many times in tandem at a particular locus in the genome. These can be present at different loci and may differ in the sequence and length of the repeating unit as well as in the number of tandem copies that occur in DNA

molecules in the population. SSRs can be divided into microsatellites or minisatellites depending on the length of the core repeating unit. Minisatellites have a longer core repeating unit than microsatellites and are typically 10-60 base pairs long and are distributed throughout the chromosomes of many eukaryotes. They are detected using Southern Blotting of agarose gels and can be single-locus or multilocus (Parker *et al.* 1998; Karp *et al.* 2001). When identified by unique sequence probes, a minisatellite polymorphism is called a variable number of tandem repeats (VNTR). The limitations of minisatellites are that they do not uniformly mark the genome and that the marker fragment is difficult to clone (Dodgson *et al.* 1997).

Microsatellites, also known as short tandem repeats (STRs), have a very short core repeating unit of 2-9 base pairs and are widely dispersed throughout eukaryotic genomes showing high polymorphism due to variation in the number of repeat units (Dodgson *et al.* 1997; Cunningham & Meghen 2001; Karp *et al.* 2001). It is the variable number of repeat units that confers polymorphism to microsatellites (Van Marle-Koster & Nel 2003). Microsatellite loci are analysed by amplifying the target region using specific primers that flank the repeated sequence (Parker *et al.* 1998). They have been used extensively in the livestock industry for genetic variability studies, parentage verification and genome mapping (Van Marle-Koster & Nel 2003). The initial cost of developing microsatellite markers is high and the process is labour-intensive, however, due to their high level of polymorphism, the markers developed for one species can sometimes be employed in a related species (Dodgson *et al.* 1997). Since the process lends itself to automation, microsatellite markers can be used to efficiently type large populations. These are co-dominant markers with two bands being produced for a heterozygote and one long band and one short band, respectively, for each homozygote. SSRs are highly polymorphic, and arise from non-coding DNA, therefore, are neutral to selection.

1.6.2 Applications of fingerprinting

Within the livestock industry, molecular markers are used extensively for parentage verification, individual identification, genome mapping, and evolutionary

and genetic biodiversity studies. The ostrich industry is following suit and molecular markers are beginning to make an appearance in ostrich studies.

Sexing of ostriches in their early stages is important for breeding farms for the early selection of breeding birds and for trading purposes (Mine *et al.* 2002). Ogawa *et al.* (1998) located and sequenced genes linked to the homomorphic Z and W sex chromosomes of the ostrich. Bello and Sanchez (1999) then used a RAPD assay to identify a sex-specific DNA marker in the ostrich and designed primers to perform a PCR diagnosis. Griffiths and Orr (1999) used AFLPs to isolate sex-specific markers. A sex-linked locus for ratites was found by Huynen *et al.* (2002), who likewise designed a PCR-based test to sex all species of ratites. The existing PCR-based methods were improved in 2002 by Malago *et al.* for large-scale sex-typing of ostriches using DNA extracted from feathers. The multiplex PCR method of Bello and Sanchez (1999) was evaluated on juvenile ostrich chicks by Mine *et al.* in 2002 and found to be successful under the Botswana farming conditions. More recently, specific sequence characterised amplified regions (SCARS) were developed by RAPD screening, for use in sex identification in the ostrich (Hinckley *et al.* 2005).

Multilocus DNA typing has been applied to ratites for individual identification and to evaluate genetic diversity. Petitte *et al.* (1996) used minisatellites for individual identification and in pedigree analysis of the ostrich and other ratite species. Their results suggest that a considerable amount of genetic variability was present in the population examined. A Polish study used minisatellites for individual identification in ostrich twins (Sacharczuk *et al.* 2002). A recent study used minisatellites to assess the genetic variability and genetic distance in commercial ratite populations in Poland (Kawka 2005).

Polymorphic single-locus microsatellite markers have been isolated for the ostrich and have been used in individual identification, parentage typing and to investigate relatedness between subspecies. Initially very few microsatellite markers were characterised for ostriches (Ward *et al.* 1994; Kimwele *et al.* 1998; Kumari & Kemp 1998; Ward *et al.* 1998) until recently when the need for more markers was met by Tang *et al.* (2003) who characterized 70 novel microsatellite markers. A recent

study made use of such markers to carry out a molecular genetic analysis of the communal nesting system of the ostrich, which included parentage identification (Kimwele & Graves 2003). Microsatellite markers were also used to characterise the genetics of the Polish ostrich population and express the amount of heterozygosity between subspecies (Kawka 2005).

DNA fingerprinting in the ostrich can assist in the design of breeding programmes aimed at maintaining genetic diversity and implementing successful selection strategies as well as to monitor levels of inbreeding within a population (Petitte *et al.* 1996; Kumari & Kemp 1998). It would also be useful in identifying individuals and in establishing parentage in the birds under colony breeding conditions to determine pedigrees. Such an approach would combine the advantage of knowledge of pedigree information (presently confined to pair-breeding systems) with an improved data structure (where the likelihood of confounded random effects is substantially reduced).

1.7 APPLICATIONS OF FINGERPRINTING IN GENETIC ANALYSES

Since its development, DNA fingerprinting has been successfully applied in the fields of biology, wildlife and conservation. The ability of molecular techniques to determine differences between individuals, populations and species has revolutionised livestock, wildlife, and conservation research. DNA fingerprinting has been applied to population-based studies to investigate structure, size and population-specific markers. However, the focus lies mainly on identifying individuals and determining parentage and kinship within populations (Avise 2004).

A direct relationship exists between the extent of genetic polymorphism required of the technique and the level of relatedness that can be addressed. Individual identification requires methods that reveal the highest level of variation possible (Parker *et al.* 1998). The assessment of parentage requires the same but also requires that Mendelian transmission genetics be taken into account (Avise 2004).

Parentage can be described at two levels of resolution, namely exclusion or assignment. An exclusion approach involves asking whether a particular individual could be the parent of the focal individual. Exclusion occurs when the putative parent's genotypes are incompatible with the genotypes of the juveniles under consideration. Assignment involves the identification of the two parents of the focal individuals by excluding with high levels of confidence all other possible parents in the population (Parker *et al.* 1998; Avise 2004).

To address the issue of relatedness and parentage, the use of maximum likelihood, a statistically based method, is efficient for the inference of relationships using genetic data obtained from fingerprints (Marshall *et* al. 1998; Luikart & England 1999). A number of software programmes has been developed that use likelihood as the principle for analysis of data (Luikart & England 1999). It is possible to calculate a likelihood ratio for each candidate parent (i.e. the likelihood of parentage of that candidate parent relative to the likelihood of parentage of an arbitrary unrelated candidate parent), and to compare the likelihood ratios of different candidate parents. Where available, software programmes make use of genetic information from the parent of the opposite sex to the one being tested. The likelihood ratio when one parent is known is different from the likelihood ratio when neither parent is known (Marshall *et al.* 1998).

1.8 AIMS

The developmental work of ascertaining pedigrees in farmed ostriches for production data is a prerequisite for the linkage of such pedigree information to production data in the broader South African industry. Individual performance levels for egg and chick production are also unknown under the more commonly practiced colony breeding systems. The use of molecular genetics specifically in individual identification and in parentage assignment is the first step forward to addressing these constraints to commercial ostrich production.

This investigation focused on a small part of the commercial ostrich population in South Africa to attempt to develop a system for ostrich identification and ostrich parentage identification. The ideal system for this purpose is microsatellite

markers, which offer ease of use and highly variable distinguishable loci with co-dominant alleles that can be unambiguously scored. Published microsatellite markers were considered as an alternative to the expensive, labour-intensive route of developing microsatellite markers for this investigation.

The two major aims of this investigation comprised:

- To test known microsatellite primers developed for other species of ostrich in the South African Blacks and,
- To test which of these microsatellite markers could be used for individual identification and, subsequently, parentage identification.

CHAPTER 2
MATERIALS AND METHODS

2.1 INTRODUCTION

In this investigation known microsatellite primers, some of which were developed for Kenyan Reds, were selected from literature to test their usability in the South African Black ostrich, *Struthio camelus domesticus.*

These primers were tested in two breeding systems, namely pair breeding (known parent combinations) and colony breeding (unknown parent combinations), to establish whether individual identification was possible.

The microsatellites that were suitable to identify individuals were then assessed for their use in the assignment of parentage.

The experimental population of South African Blacks was maintained at the Klein Karoo Agricultural Development Centre (KKADC) near Oudtshoorn, South Africa. This flock was developed as an experimental resource in the 1980's from the donation of commercial breeding birds. Further introductions were made in the 1990's giving rise to two predominant strains ('commercial' and 'feather') as founder parents in the flock, although the 'feather' strain was essentially developed from commercial animals displaying better feather quality characteristics. The flock was gradually expanded from the 1990's to 2005 with new breeders mostly obtained from within flock (Bunter & Cloete, 2004). The origin and management of the resource population is well documented (van Schalkwyk *et al.*, 1996; Bunter, 2002).

The protocol for all solutions used in this investigation have been taken up in Appendix A.

2.2 MATERIALS

Breeding birds and their progeny from a South African Black population were required for DNA sampling.

2.2.1 Sample Group

The major aim of this investigation was to test the suitability of published microsatellites to assign parentage. Therefore, two breeding pairs with known parent combinations and four known progeny of each pair were selected from the 2002-2003 breeding season to serve as the control group. These birds are maintained as part of the commercial flock at the KKADC.

One colony from the 2003-2004 breeding season, also maintained at the KKADC, was selected as the test sample. The production statistics of this colony are shown in Table 2.1. From this large colony all parents, consisting of six males and fourteen females, were sampled. A total of 77 progeny was included in the test sample.

Table 2.1 Production statistics of the selected colony of the 2003-2004 breeding season.

Eggs	n	%	No. Sampled
Total eggs produced	234		
Fertile eggs:	188	80.34	77
No. Eggs hatched	138	73.40*	66
No. Death-in-shell#	50	26.60*	11

* Calculated out of a total of 188 fertile eggs
# Chick died in shell during incubation, before hatching

A total number of 109 birds (12 pair, 97 colony) was sampled in this investigation.

2.2.2 DNA Source

Blood was used as a source of DNA, because it is obtained relatively easily and the DNA yield is high due to the nucleated erythrocytes of Aves. Using a sterile technique, blood was obtained from adult birds and day-old chicks. The jugular vein of the bird was located, the area of insertion first wiped with an alcohol swab, then pierced with a sterile syringe needle. Approximately 1-2 ml of blood was drawn and immediately expelled into Vacutainer™ EDTA tubes and shaken to prevent clotting. For death-in-shell chicks, the chicks were extracted from their eggs and dissected through the sternum. Blood was drawn directly from the heart using a syringe and needle and expelled into a Vacutainer™ EDTA tube. All blood samples were stored at 4 °C.

2.3 METHODS

The methods used in this investigation are detailed in three steps, namely, DNA extraction, verification and quantification of DNA yield, and generation of microsatellite fingerprints.

2.3.1 DNA Extraction

The non-mammalian whole blood protocol of PUREGENE® DNA purification kit (Gentra) was used to extract ostrich DNA. This protocol is known to be suitable for invertebrates, reptiles, fish and birds. The different steps of this extraction protocol entailed the following:

1. Stored blood samples were removed from the refrigerator and left at room temperature for at least one hour to thaw.
2. 4 µl whole blood was firstly added to a 1.5 ml microfuge tube containing 600 µl cell lysis solution.
3. Using a pipette, the blood containing solution was drawn up and down into the pipette tip for 3-5 times to lyse the cells until no cell clumps were visible.
4. 200 µl protein precipitation solution was then added to the cell lysate.

5. The tube was then vortexed vigorously at high speed for 20 seconds to mix the protein precipitation solution uniformly with the cell lysate.
6. The tube was then centrifuged at 13,000-16,000 × g *(12,300 rpm)* for 3 minutes. At this point, the proteins formed a tight dark brown pellet. If this protein pellet was not tightly logged, Step 4 was repeated followed by incubation on ice for 5 minutes. Thereafter Step 5 was repeated.
7. The supernatant containing the DNA (leaving behind the precipitated protein pellet) was then poured into a clean 1.5 ml microfuge tube containing 600 µl of 100% isopropanol (2-propanol).
8. This sample was then mixed by inverting the tube gently approximately 50 times.
9. The tube was then centrifuged at 13,000-16,000 × g *(12,300 rpm)* for 1 minute; the DNA was visible as a small white pellet.
10. The supernatant was then poured off and the tube drained on clean absorbent paper. 600 µl of 70% ethanol was then added and the tube was inverted to wash the DNA.
11. The tube was then centrifuged at 13,000-16,000 × g *(12,300 rpm)* for 1 minute. The ethanol was carefully poured off, whilst watching that the pellet did not dislodge.
12. The tube was inverted to drain on clean absorbent paper and the sample was allowed to air dry for 10-15 minutes.
13. To the air-dried tube, 200 µl of DNA Hydration Solution (100 µl will provide a concentration of 100 µg/ml for a yield of 10 µg DNA) was added.
14. DNA was rehydrated by incubating the sample for 1 hour at 65 °C and/or overnight at room temperature. The tube was tapped periodically to aid in dispersing the DNA.
15. The DNA solution was then stored at 4 °C for short-term storage and at –20 °C for long-term storage.

2.3.2 Verification and quantification of DNA yield

To verify the presence of DNA and to ensure that it was of high molecular weight, a diagnostic 0.8% agarose gel was run. The gel was prepared with 0.4 g agarose in 50 ml 1 X TAE and 2.5 µl of ethidium bromide (10 mg/ml). Each well was loaded with 10 µl of sample DNA and 2 µl of loading buffer. 4 µl of molecular weight marker III (Roche) (0.25 µg/µl) was included in the gel to verify the presence of DNA.

A uQuant plate-reader spectrophotometer (Biotek) was used to determine the concentration and purity of the DNA in each sample. Samples were blanked and diluted in 10 mM Tris, with a dilution factor of 10 times, consisting of 10 µl sample in 90 µl 10 mM Tris. The purity was calculated using the formula A_{260} / A_{280} and the concentration calculated with the formula A_{260} x dilution factor x 50 µg/ml.

2.3.3 Generation of microsatellite fingerprints

The generation of microsatellite fingerprints involved a number of sequential steps; the selection of microsatellite loci, the amplification of these loci and the subsequent electrophoresis in order to generate the fingerprints.

Selection of microsatellite loci

Published microsatellite loci were selected for this investigation. An attempt was made to select microsatellite loci that had primers with similar annealing temperatures and that displayed a high polymorphic content. Two rounds of selection were executed, the first during 2003 and the second during 2004. The first round yielded six loci published by Kumari and Kemp (1998) and Kimwele *et al.* (1998), all of which were developed for Kenyan Reds, a subspecies of *Struthio camelus*, which occur in central Africa. The second round of selection yielded a further five loci published by Tang *et al.* in (2003). These markers were developed for the two subspecies South African Blacks and Zimbabwean Blues. Table 2.2 lists the eleven selected microsatellite loci and their primer sequences.

Table 2.2 Microsatellite markers selected for this investigation.

Locus	Primer sequence (5' – 3')	Repeat unit	Reference
OSM 1	f: AATCTGCCTGCAAAGACCAG r: TCCCAGTCTTGAAGTCAGCA	$(CA)_{17}$	Kimwele *et al.* (1998)
OSM 2	f: AAGCCACGGCAATGAATAAG r: CCTCAACCATTCTGTGATTCTG	$(CA)_{22}$	Kimwele *et al.* (1998)
OSM 7	f: AGCATACACATGCAGACCCC r: TGTTTCCTGCCATTCTGTCA	$(CA)_{16}CT(CA)_5$	Kimwele *et al.* (1998)
LIST 005	f: ATGGTGCTTTCCAGTGGTGTGC r: CATTGACCCAGGCAAGAAATCC	$(TG)_2CG(TG)_{10}$	Kumari & Kemp (1998)
LIST 009	f: CATTGCAAACACTCTGCTGC r: TGAACGACAGGGTTATTGGC	$(CA)_{14}G(CA)_3CG(CA)_3$	Kumari & Kemp (1998)
CAU 14	f: ATTTAACTTCTCTAAGGCACTC r: GAGGAGCAATTCAGACAGAC	$(CA)_{16}$	Tang *et al.* (2003)
CAU 17	f: CGTAAACCCAGATAATCACAA r: AGTGGCATTGTAGCTCTTCA	$(CA)_{22}$	Tang *et al.* (2003)
CAU 40	f: ACGGGGAGACTCAAGGATG r: GCTTGCGTGTGCATGAGTAT	$(CA)_9$	Tang *et al.* (2003)
CAU 65	f: TGAGAGTCTCCCAGAAATGC r: CAGAGAAATATATGCCTGTAAAT	$(TA)_{12}(CA)_9$	Tang *et al.* (2003)
CAU 90	f: CCATCCAAAACATACCACACC r: TCCAGTCCCAACTGAGCTAAA	$(AC)_{20}$	Tang *et al.* (2003)
VIAS-OS 29	f: TTTTCGTCTTCCACCCACTG r: CTGCTTCTTCCGTGTGTGTC	$(AC)_{13}GG(AC)_6GG(AC)_4$	Ward *et al.* (1998)

f = forward primer; r = reverse primer

Microsatellite locus amplification

All microsatellite PCR amplifications used Go*Taq*® DNA Polymerase with green buffer (Promega) and PCR Nucleotide Mix (Promega). The PCR conditions proposed by Kimwele and Graves (2003) were employed to amplify all the selected microsatellite loci using the PCR reagents at concentrations listed in Table 2.3. Amplification was effected on the Geneamp PCR System 2700 (Applied Biosystems).

Table 2.3 PCR reagents concentrations.

Reagent	Concentration
PCR Buffer	1 x
$MgCl_2$	1.5 mM
dNTPs	0.2 mM each
Primer forward	12.5 pM
Primer reverse	12.5 pM
Taq	0.5 U
DNA	10 ng

A first round of amplification, using four samples and the reagent and amplification conditions proposed by Kimwele and Graves (2003), was executed to evaluate the performance of this method under conditions in this laboratory. Little or no amplification product was generated. This indicated the need for optimisation. Optimisation was then undertaken by testing a range of annealing temperatures for each primer independently. The presence of amplification product was verified through gel electrophoresis. A 6% non-denaturing polyacrylamide gel was prepared, run at 40-60 mA, stained with ethidium bromide and viewed under UV light.

Microsatellite fingerprints

The microsatellite fingerprints of all individuals were generated by separating the PCR amplification products in 6% non-denaturing polyacrylamide gels. These gels were run at 40-60 mA with two standard DNA ladders, 20 base pairs (ABGene) and 100 base pairs (Fermentas), for size referencing. Two types of fingerprints were generated, namely individual and multiplex fingerprints. Individual microsatellite loci fingerprints were produced by electrophoresing 10 µl of amplification product to determine whether the particular microsatellite locus was polymorphic in the selected South African sample population. Multiplex fingerprints were then generated by electrophoresing a sample that contained 10 µl of each individually amplified locus for a particular individual.

2.4 GENOTYPIC ANALYSIS

A genotypic analysis of each individual was performed by firstly identifying the alleles at a particular microsatellite locus, then determining the size of the alleles and, lastly, quantifying the alleles at each microsatellite locus.

2.4.1 Allele identification and sizing

Single and multiplex gel images were photographed, saved and loaded using UVIsoft image acquisition analysis software (Uvitec). The programme UVIDocMW was then used to detect the alleles at each locus for each individual, after which the size of each allele was determined using a 20 bp molecular weight marker (ABGene) as a reference.

Figure 2.1 provides the steps involved in the execution of the programme. Step 1 and step 2 involved, respectively, the selection of the type of analysis to be carried out and the definition of lanes on the gel loaded. Step 3 involved the detection of bands on the gel, which was used for the determination of alleles. Bands that were not product, for example, sample residue at the top of lanes, were manually deselected. Steps 4 and 5 involved the sizing of the bands by comparison with the known sizes of the bands of the 20 bp molecular weight marker DNA ladder, and the subsequent output of the table of band sizes.

The allele detection step, specifically the number of alleles at a particular locus, was used to determine the number of microsatellite loci that were either monomorphic or polymorphic. A locus that displayed more than one allele in the population under consideration was regarded polymorphic, while a locus that contained only one allele was considered monomorphic.

Allele sizes were used to name the different alleles of each locus to facilitate the construction of genotypes for each individual.

2.4.2 Construction of genotypes

Multiplex gels were used to determine the genotypes of all the individuals in this study. The allele sizes, which were used to name the different alleles at each

locus, were entered into an Excel spreadsheet sequentially, thereby constructing the genotype of each individual.

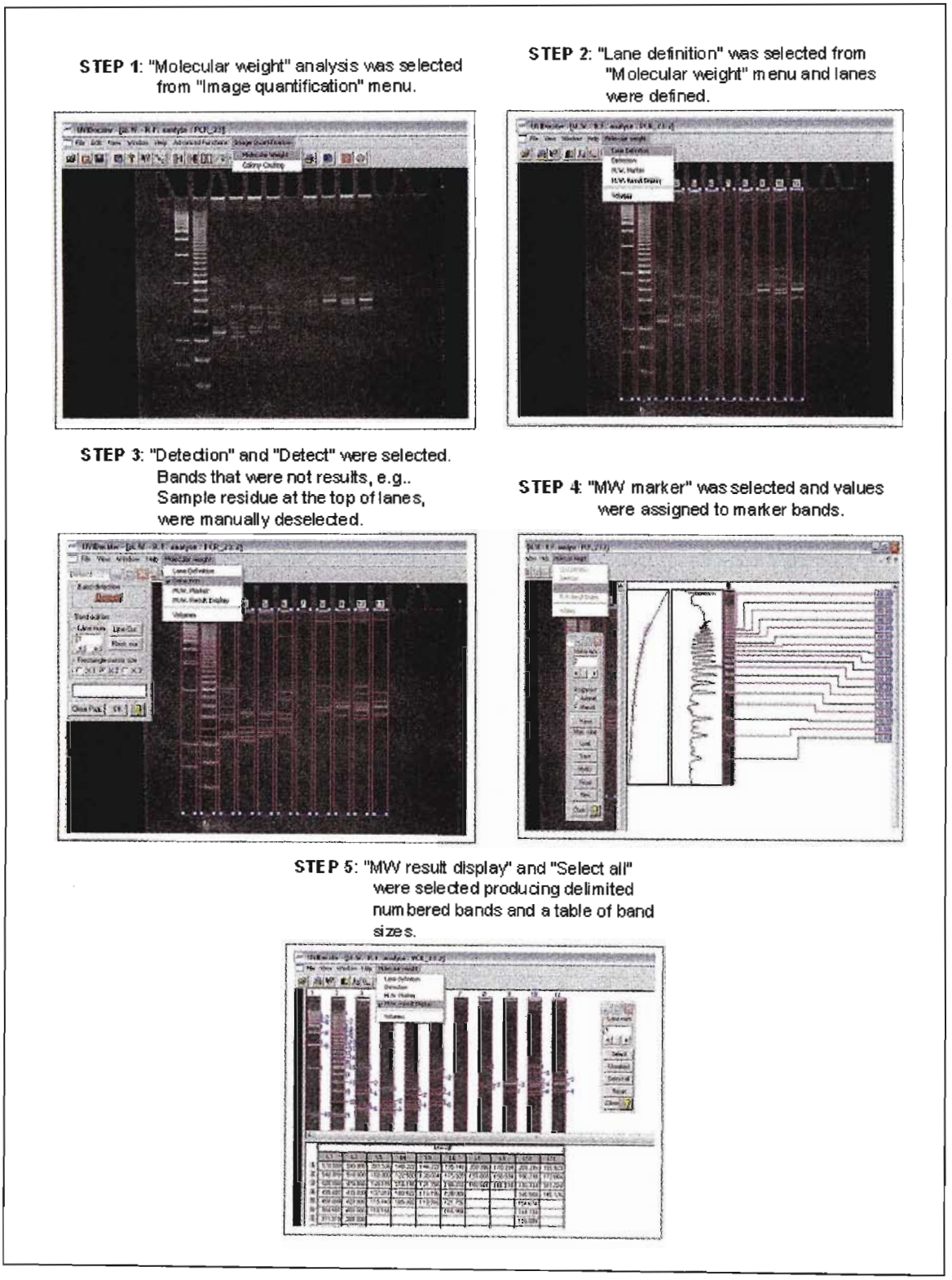


Figure 2.1 Allele detection and sizing in step-wise format using UVIDocMW programme.

2.4.3 Quantification of alleles

CERVUS Version 2.0 (Marshall *et al.*, 1998) was used to calculate the allele frequencies of the different alleles at the different microsatellite loci. These calculations were performed separately for each of the individuals from breeding pair families and colony individuals. These calculations required the construction of an input file containing the relevant data. The file that contained the genotypic information was named *GenotypesCorrected2.csv (BPGenotypes.csv* for breeding pair individuals) and was used as the input file. The options required by the programme regarding the input file were selected. Thereafter the "Allele Frequency Analysis" procedure (Figure 2.2) was initiated by selecting "Run".

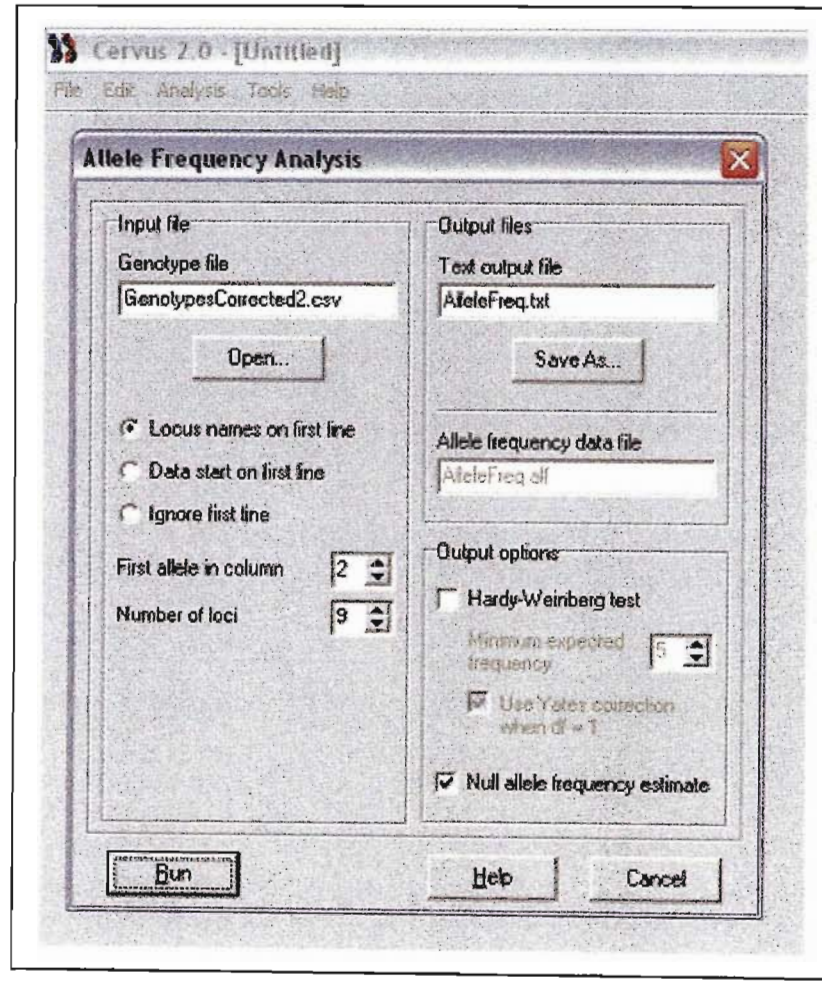


Figure 2.2 Allele frequency analysis setup screen in CERVUS.

2.5 PARENTAGE ASSIGNMENT

The multiplex fingerprints were used to identify the parents of the different progeny. Parentage analysis was performed separately for breeding pair families and for colony individuals in the same way. The software programme CERVUS Version 2.0 (Marshall *et al.*, 1998) was used to perform a parentage analysis. The programme uses exclusion and a likelihood-based approach to assign parentage. The first step was to run a simulation. The logarithm of the odds (LOD) scores was calculated in the simulation using the allele frequencies. The simulation generated criteria that permit the assignment of parentage to the most likely candidate parent and provides a level of statistical confidence for the particular assignment. A separate simulation was carried out for the male candidate parents and for the female candidate parents (Figure 2.3).

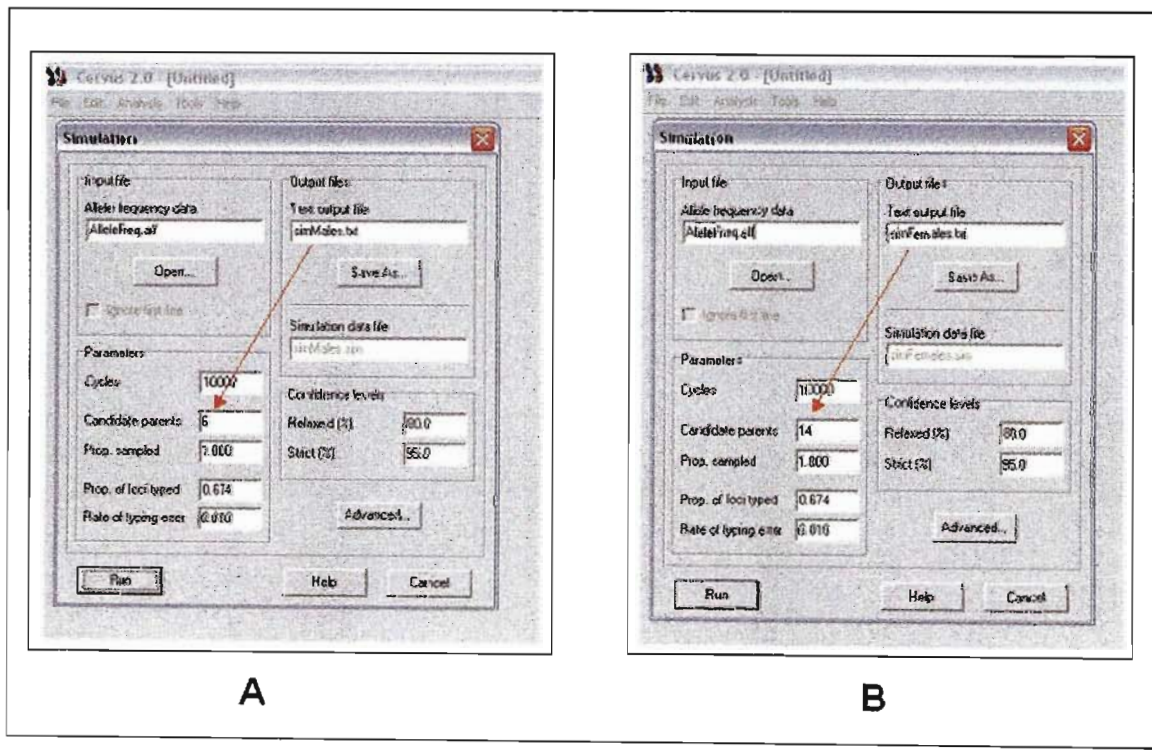


Figure 2.3 CERVUS screens showing simulation parameters for male parent assignment (A), and female parent assignment (B).

Once the simulations were conducted, the "Parentage Wizard" of the programme was employed to assign the parents to the different progeny. The different steps

followed are shown in Figure 2.4. Step 1 of the wizard required the input of relationship data, specifically the offspring genotypes. Step 2 required input of candidate parent genotypes. Step 3 required input of genetic data in the form of genotypes for all individuals. Step 4 required results files generated by CERVUS for the allele frequency analysis and the simulation. Step 5 required the naming of output files and selecting the option of "Most-likely parent" to be included for each offspring in the output parentage file. Candidate parent input files were created by extracting the relevant parent genotypic data from the individual genotype data file into separate data files.

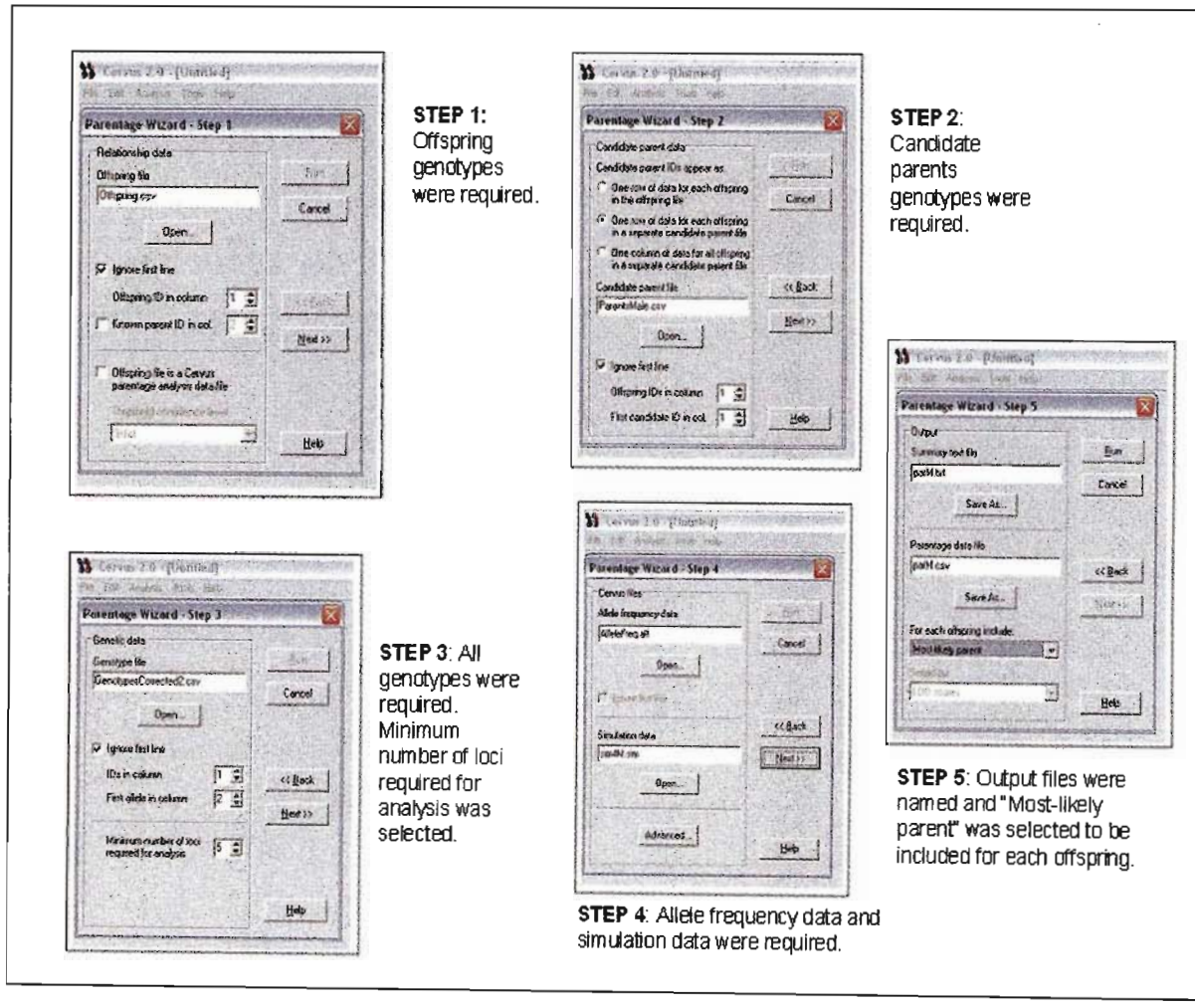


Figure 2.4 CERVUS screens displaying parentage wizard steps and required input files and parameters.

Because neither parent was known, CERVUS recommends a two-step analysis with the first step to run the group of parents with fewer candidates, males in this case,

and the second step to run the analysis with the females using the results of the first step. This two-step analysis was also carried out for the breeding pair families. For the female parentage analysis, step 1 of the wizard was modified such that the input offspring genotype file required was the output male parentage analysis file and the threshold confidence level was set to "Strict" (Figure 2.5).

Figure 2.5 CERVUS screen displaying step 1 of the parentage wizard for female analysis where the parentage analysis output file of the males was used as the input offspring genotype file to assign maternity.

The programme eliminates exclusion of parentage because of allelic mismatch, which could be due to actual allelic mismatch or more commonly to erroneous laboratory typing or the presence of mutations or null alleles.

2.6 CONSTRUCTION OF PEDIGREES

The microsatellite fingerprints were used to construct pedigrees. For the two breeding pair families, pedigrees were constructed by comparing the genotypes of the parents to the genotypes of the progeny. For the colony on the other hand, the results of the parentage analysis were used to create a mating table for the parents of the colony to identify assigned progeny of the respective matings.

2.7 DETERMINATION OF BREEDING STATISTICS

Breeding statistics were calculated for the production potential of females in the colony during the sampling period. From the parentage analyses, a count was done to determine the number of fertile eggs produced by individual females during the sampling period. Similarly, the number of eggs fertilized by individual males during the sampling period was determined. The mean and standard deviation were calculated from these counts, for the production of fertile eggs by females, and for the fertilization of eggs by males. The coefficient of variation for the production of fertile eggs by individual females during the sampling period was determined by:

Coefficient of Variation (%) = standard deviation / mean x 100

CHAPTER 3
RESULTS

3.1 INTRODUCTION

Known microsatellite primers, some of which were developed for other ostrich species, were selected to test for use in the South African Black ostrich. The selected microsatellite loci were then tested for their use in individual identification and parentage assignment.

Output files of the various computer analyses have been taken up in Appendix B.

3.2 DNA YIELD

DNA was extracted from whole blood using Gentra's PUREGENE® DNA purification kit and verified by running a diagnostic gel. Generally, DNA yield and purity are calculated using a spectrophotometer. However, the spectrophotometer available proved to have a technical problem rendering results unreliable, thus DNA was quantified on a gel as shown in Figure 3.1.

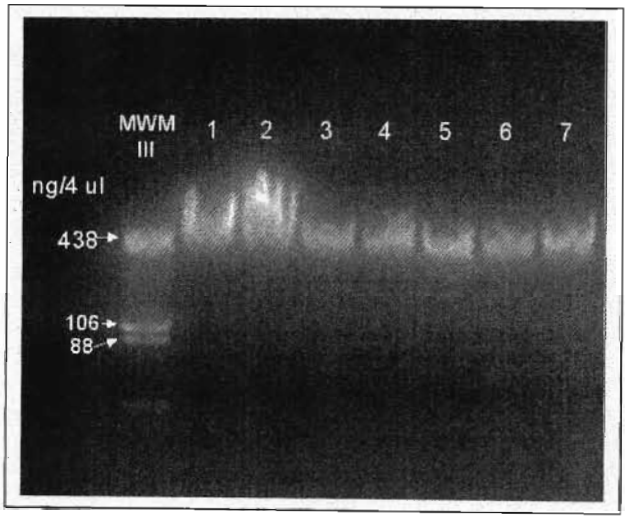


Figure 3.1 Diagnostic gel (0.8% agarose) used for visual DNA quantification by comparison of band intensity of DNA samples (lanes 1-7) to standard molecular weight marker (MWM) III.

The concentration of the DNA was determined in the following manner. The manufacturer's concentration of molecular weight marker (MWM) III was 0.25 µg/µl. However, only 4 µl of MWM III was loaded in a gel resulting in a band with the intensity produced by 1 µg. Since the intensity of the bands of the DNA samples loaded was approximately equal to the intensity of the bands produced by the marker, it could be concluded that there was approximately 1 µg of sample DNA in each lane. In this investigation however, 10 µl of sample DNA was loaded in a gel resulting in a concentration of 0.1 µg/µl. The DNA was quantified in this manner for all samples with an approximate yield of 20 µg per sample.

Purity of the DNA was determined by the colour of the pellet at the rehydration step of the extraction procedure, where a white-colourless pellet was indicative of relatively pure DNA. All DNA samples were sufficiently pure to enable amplification of microsatellite loci.

3.3 GENERATION OF MICROSATELLITE FINGERPRINTS

Microsatellite fingerprints were generated through the PCR amplification of the individual selected microsatellite loci. The first step in the optimisation of the PCR conditions involved the adjustment of cycling conditions, while the second step involved the modification of the annealing temperatures (T_A) of the markers. The reaction conditions, on the other hand, were maintained as proposed by Kimwele and Graves (2003).

In the first step the cycling conditions proposed by Kimwele and Graves (2003) were initially tested. These conditions, however, resulted in little or no amplification of the microsatellite loci. Once the number of cycles was increased to 30, satisfactory amplification was achieved. Figure 3.2 provides a summary of the cycling conditions used in this investigation.

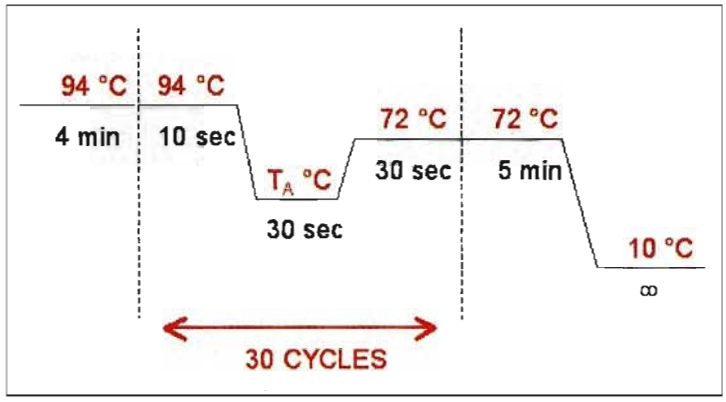


Figure 3.2 Cycling conditions used for all amplification reactions, where T_A indicates the specific annealing temperature for each microsatellite marker.

The second step in the optimisation of the amplification conditions required the adjustment of the individual annealing temperatures for the different microsatellite loci. The annealing temperatures reported in the literature (Kimwele *et al.* 1998; Kumari & Kemp, 1998; Ward *et al.* 1998; Tang *et al.* 2003) for each set of primers for each microsatellite locus, were initially tested. These temperatures were then optimised by implementing a series of 0.5°C increments and decrements. The annealing temperature that produced the most satisfactory amplification of each microsatellite locus was recorded and used throughout subsequent amplifications. It should, however, be mentioned that three of the eleven microsatellite loci that were tested did not produce an amplification product after extensive modification and testing of the annealing temperature. It was, therefore, concluded that the homology of the primers (developed for the Reds) for these microsatellite loci was too low to allow for optimisation by modification of the annealing temperature alone, and probably required intensive testing of all the different variables in the amplification process. These loci were thus excluded from all subsequent experimentation. Table 3.1 provides a list of the annealing temperatures of all primers tested.

Table 3.1 Optimised annealing temperatures of the different microsatellite loci.

Microsatellite name	Literature T_A (°C)	Modified T_A (°C)	Reference
OSM 1	57.0	57.5	Kimwele *et al.* (1998)
OSM 2	57.0	57.5	Kimwele *et al.* (1998)
OSM 7	58.0	57.5	Kimwele *et al.* (1998)
LIST 005	55.0	56.0	Kumari & Kemp (1998)
LIST 009	55.0	56.0	Kumari & Kemp (1998)
CAU 14	58.5	57.5	Tang *et al.* (2003)
CAU 17	58.5	57.5	Tang *et al.* (2003)
CAU 40	65.0	–	Tang *et al.* (2003)
CAU 65	58.5	57.5	Tang *et al.* (2003)
CAU 90	56.5	–	Tang *et al.* (2003)
VIAS-OS 29	55.0	–	Ward *et al.* (1998)

– No amplification

3.4 GENOTYPIC ANALYSIS

The genotypes of all the individuals that participated in this investigation were determined through the analysis of two different types of microsatellite fingerprints. The fingerprints were single-locus and multi-locus fingerprints. Single-locus fingerprints were used for the verification of amplification as well as for the determination of polymorphic loci. Multi-locus fingerprints were used to create the genotypes of all individuals sampled.

3.4.1 Microsatellite fingerprints

Single-locus fingerprints of the control sample, breeding pairs and progeny, were generated to verify amplification reactions, to confirm that allele sizes were within the expected range, and to determine which loci were polymorphic. Stutter bands, artefacts produced by DNA polymerase slippage, that reduce the resolution between alleles, were identified in these fingerprints. These were noted and

excluded from the sizing process of the amplification fragments. Examples of single-locus amplification fragment bands and stutter bands are shown in figure 3.3.

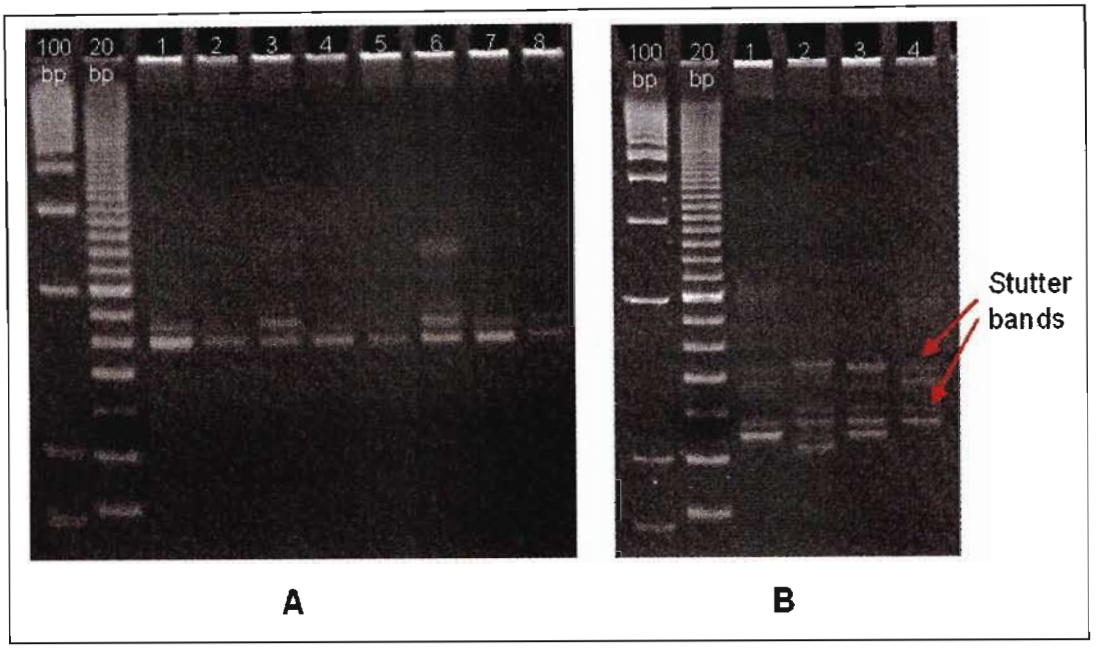


Figure 3.3 Polyacrylamide gels showing single-locus fingerprints of the CAU 17 locus (A) and OSM 1 locus (B). Stutter bands were produced when the OSM 1 locus was amplified (B).

Multi-locus fingerprints were generated to construct genotypes of all individuals sampled. These fingerprints were generated by poolling individual amplification products into a single sample for each individual and running on a gel, as shown by figure 3.4.

The expected size range of the amplification products for each microsatellite locus was obtained from the literature to identify the regions of amplification for each locus as shown in figure 3.4. Locus OSM 1 amplified in the size range of 100-120 base pairs and locus OSM 2 amplified in the size range of 121-141 base pairs. Locus CAU 14 amplified in the region of 142-160 base pairs and locus CAU 17 in the region of 161-180 base pairs. Locus CAU 65 produced amplification product of a size between 181-191 base pairs. The amplification product of locus LIST 005 was close in size to that of locus LIST 009 in the regions of 192-198 and 199-210

base pairs, respectively. Locus OSM 7 amplified in the size range of 210-230 base pairs.

Both single-locus and multi-locus fingerprints displayed amplification products that were outside the expected size range reported in the literature. These anonymous loci were named 'unknown locus 1' (ULoc1) and 'unknown locus 2' (ULoc2). The size range amplified by locus ULoc1 was 231-300 base pairs and for locus ULoc2 was 291-395 base pairs. From the single-locus fingerprints of the individual microsatellite loci, it was determined that loci OSM 2, OSM 7, LIST 005, and LIST 009 gave rise to these anonymous amplification products. However, it could not be determined which of these known loci gave rise to the specific amplification product bands of the anonymous loci.

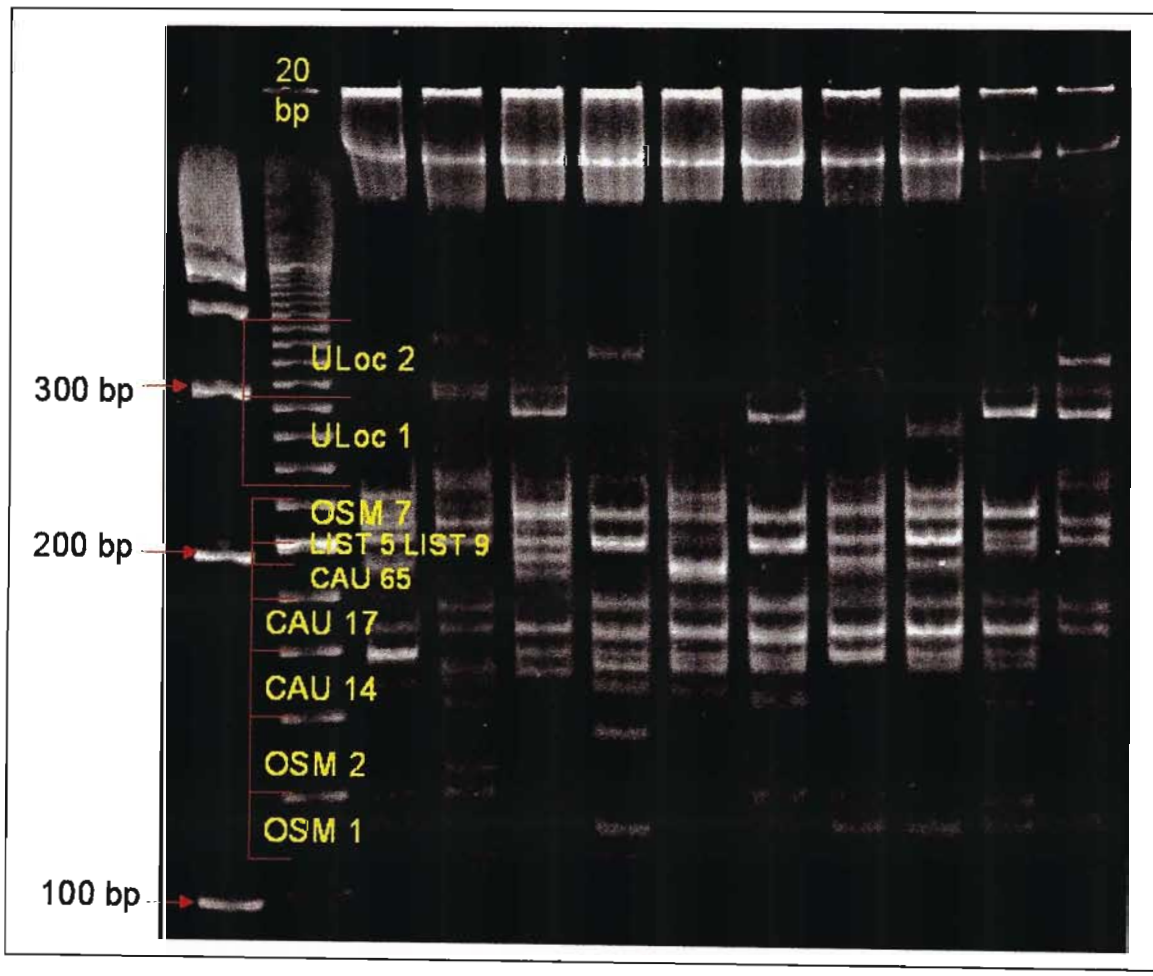


Figure 3.4 Multi-locus fingerprint showing the size ranges of the individual microsatellite loci amplification products.

3.4.2 Estimation of fragment sizes

The sizes of the amplification products were used to determine and name the different alleles at the different microsatellite loci. This was accomplished by capturing a fingerprint gel as an electronic photograph using UVISave hardware and UVISoft gel documentation software (Uvitec). This electronic version of the gel was then opened in the UVIDocMw programme (Uvitec). The molecular weights of the DNA molecular marker that ran alongside DNA samples in a gel were used as size references by the programme. These molecular weights were entered into the programme, after which the molecular weight analysis routine was selected. The microsatellite fragment bands on the electronic gel were then delimited automatically, based on band intensity. The fragment sizes were then determined by the programme using the sizes of the molecular weight marker as reference (Figure 3.5). Allelic size determination was determined separately for the pairs and colony. Records of the analyses were named according to the file contents.

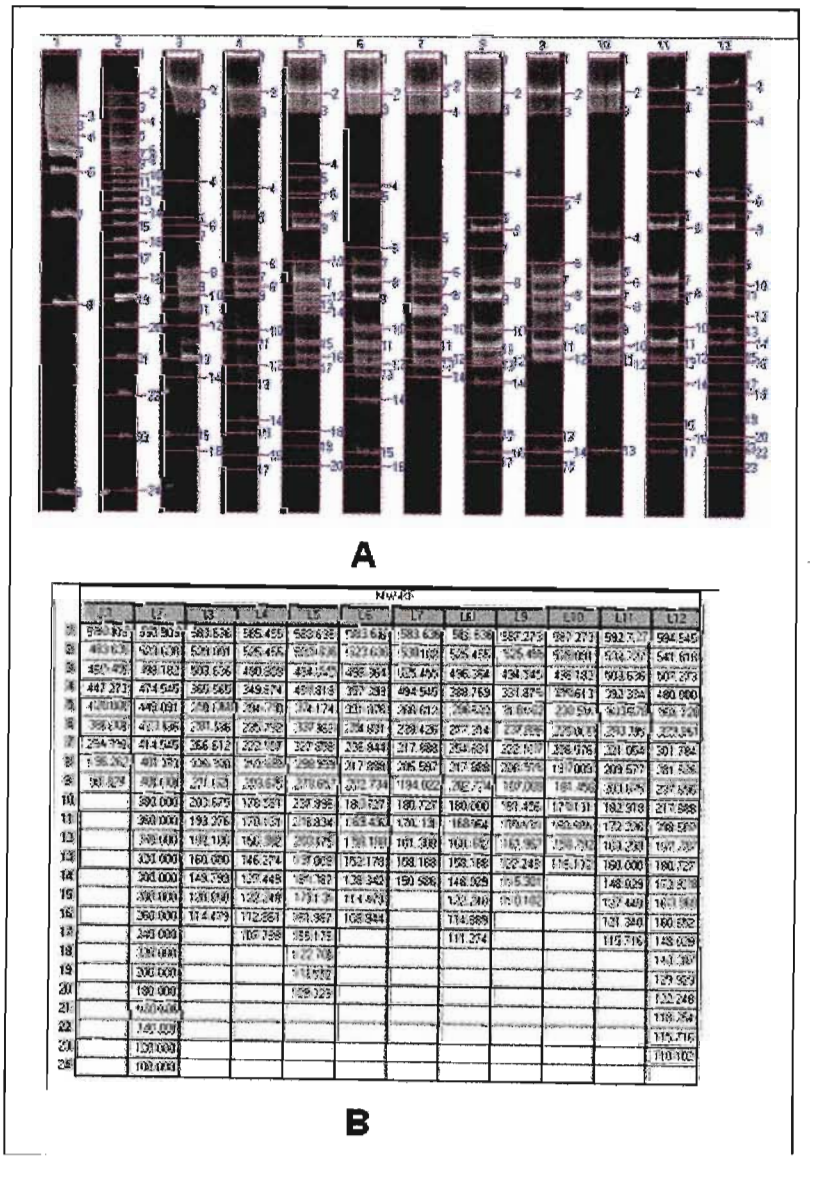
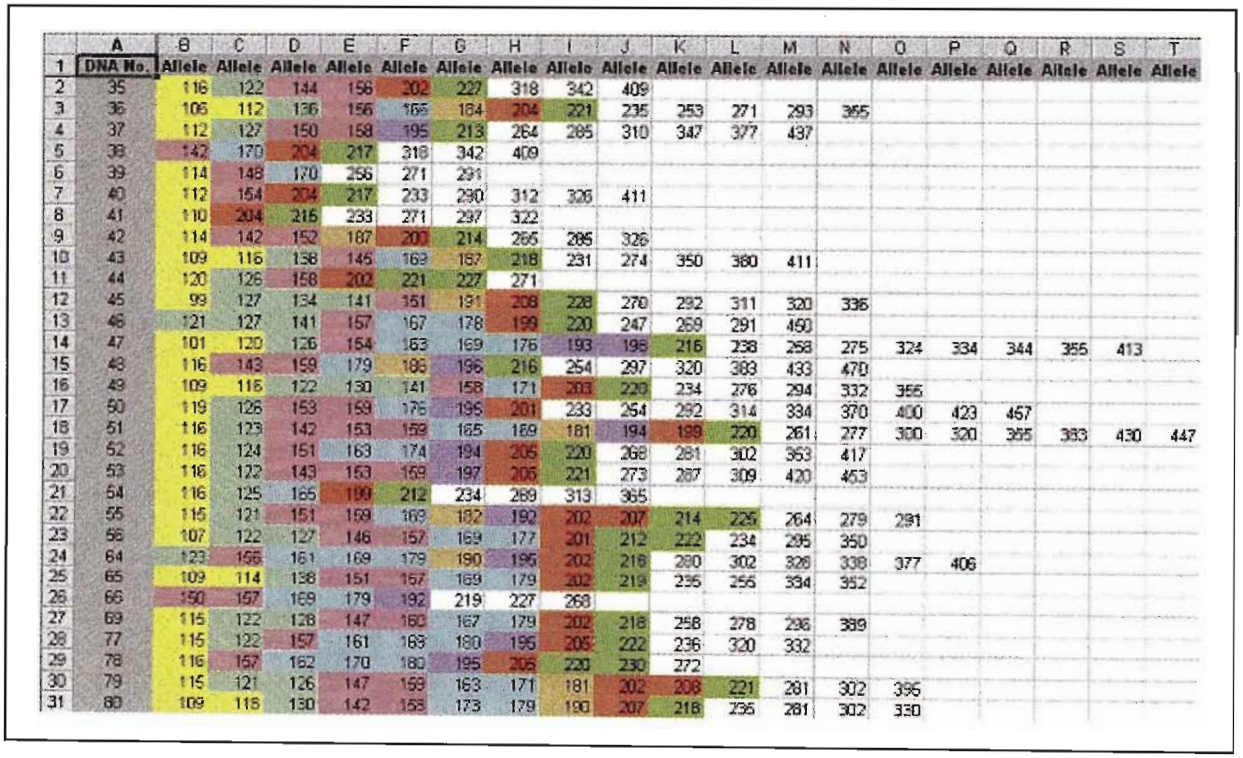


Figure 3.5 Delimitation of microsatellite fragment bands (A), and corresponding sizes of fragment bands (B).

3.4.3 Genotypes

The genotypes of all the individuals in this investigation were compiled from the multi-locus fingerprints by arranging the alleles of each locus in sequential format. Figure 3.6 displays the genotypes of 30 colony individuals. Through careful inspection of the fingerprints the stutter bands were excluded from the final compilation of the genotypes. Each locus was represented in a different colour to facilitate the reading of the genotypes. The anonymous loci were not coloured. The monomorphic locus CAU 65 was excluded from the final genotypic compilation. This final genotypic compilation was named *GenotypesCorrected2.csv* (*BPGenotypes.csv* for pairs).



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	DNA No.	Allele	Allele	Allele	Allele	Allele	Allele	Allele	Allele	Allele	Allele	Allele	Allele	Allele	Allele	Allele	Allele	Allele	Allele	Allele
2	35	116	122	144	156	202	227	318	342	409										
3	36	106	112	136	156	166	184	204	221	235	253	271	293	365						
4	37	112	127	150	158	195	213	264	285	310	347	377	437							
5	38	142	170	204	217	316	342	409												
6	39	114	148	170	256	271	291													
7	40	112	154	204	217	233	280	312	326	411										
8	41	110	204	216	233	271	297	322												
9	42	114	142	152	187	200	214	266	285	326										
10	43	109	116	138	145	163	187	218	231	274	350	380	411							
11	44	120	126	158	202	221	227	271												
12	45	99	127	134	141	151	191	208	228	270	292	311	320	336						
13	46	121	127	141	157	167	178	199	220	247	269	291	450							
14	47	101	120	126	154	163	169	176	193	198	216	238	258	275	324	334	344	355	413	
15	48	116	143	159	179	188	196	216	254	297	320	383	433	479						
16	49	109	116	122	130	141	158	171	203	220	234	276	294	332	356					
17	50	119	126	153	159	176	195	201	233	254	292	314	334	370	400	423	457			
18	51	116	123	142	153	159	165	189	191	194	199	220	261	277	300	320	365	383	430	447
19	52	116	124	151	163	174	194	205	220	268	281	302	353	417						
20	53	116	122	143	153	169	197	205	221	273	287	309	420	453						
21	54	116	125	165	199	212	234	269	313	365										
22	55	115	121	151	159	169	182	192	202	207	214	226	264	279	291					
23	56	107	122	127	146	157	169	177	201	212	222	234	295	350						
24	64	123	166	181	189	179	190	196	202	218	280	302	326	338	377	406				
25	65	109	114	138	151	167	169	179	202	213	235	255	334	352						
26	66	150	157	169	179	192	219	227	268											
27	69	115	122	128	147	160	167	179	202	218	258	278	286	389						
28	77	115	122	157	161	169	180	196	205	222	236	320	332							
29	78	116	157	162	170	180	195	205	220	230	272									
30	79	115	121	126	147	159	163	171	181	202	208	221	281	302	396					
31	80	109	116	130	142	158	173	179	190	207	218	236	281	302	330					

Figure 3.6 Genotypes of 30 colony individuals. The numbers denote the different alleles and the colours the different loci. Alleles with sizes outside the expected range of the different microsatellite loci are uncoloured.

The genotypic compilation displaying the different alleles in terms of their size names was difficult to interpret and was therefore converted to a table where the allele names are provided as letters of the alphabet. None of the nine loci analysed were displayed in all individuals. The percentage of individuals that displayed the least number of loci was 1.03% while 6.19% individuals displayed the most number of loci. 12.37% individuals did not display anonymous loci. Table 3.2 was used to determine the within individual heterozygosity and the between individual, within locus heterozygosity for the colony individuals. The heterozygosity within individuals ranged from low to moderate (0-66.67%), while the heterozygosity between individuals, within loci ranged from 3.13-57.58%.

Table 3.2 Genotypic descriptions of colony individuals and measure of heterozygosity.

Individual	OSM1	OSM2	CAU14	CAU17	LIST5	LIST9	OSM7	ULoc1	ULoc2	No. Het.	% Het. (WI)
35	aa	bb	co	00	00	dd	oo	00	qG	2	22.22
36	bc	pp	oo	gg	00	ff	kt	AT	OO	3	33.33
37	cc	gg	iq	00	dd	00	cc	vM	jV	3	33.33
38	00	00	aa	kk	00	ff	gg	00	qG	1	11.11
39	dd	00	gg	kk	00	00	00	pA	aa	1	11.11
40	cc	00	mm	00	00	ff	gg	cQ	ku	2	22.22
41	ee	00	00	00	00	ff	es	AW	ss	2	22.22
42	dd	00	ak	00	00	bb	dd	wM	uu	2	22.22
43	fa	qq	dd	jj	00	00	hh	aD	KW	2	22.22
44	gg	ff	qq	00	00	dd	ko	AA	00	1	11.11
45	hh	gn	aj	00	00	jj	pp	zS	rD	4	44.44
46	00	ag	ap	hr	00	aa	jj	ky	00	4	44.44
47	ig	ff	mm	dj	bg	00	fu	qE	HO	6	66.67
48	aa	00	br	ss	ee	00	ff	nW	rX	3	33.33
49	fa	bj	aq	ll	00	ee	jj	FU	AO	5	55.56
50	jj	ff	lr	pp	dd	cc	rr	nS	mB	3	33.33
51	aa	cc	al	aj	cc	aa	jj	tG	bO	4	44.44
52	aa	dd	jj	dn	cc	gg	jj	xJ	dN	3	33.33
53	aa	bb	bl	aa	ff	gg	kk	CO	ii	2	22.22
54	aa	ee	00	ff	00	aa	bb	dP	IR	2	22.22
55	kk	aa	jr	jj	aa	di	dn	vR	00	4	44.44
56	ll	bg	ep	jq	00	cc	bl	dV	KK	5	55.56
64	00	cc	oo	bj	dd	dd	hh	ll	dV	2	22.22
65	fd	qq	jp	js	00	dd	ii	eo	BM	5	55.56
66	00	00	ip	js	aa	00	io	xx	00	3	33.33
69	kk	bh	fs	hs	00	dd	hh	qH	YY	4	44.44
77	kk	bb	pp	bj	dd	gg	ll	ff	rA	2	22.22

78	aa	00	pp	ck	dd	gg	jr	BB	00	2	22.22
79	kk	af	fr	dl	00	dj	kk	JJ	dZ	5	55.56
80	fm	jj	aq	ms	00	ii	hh	eJ	dy	5	55.56
81	nj	00	rr	ir	00	bb	bb	00	00	2	22.22
82	ng	mm	or	00	00	ii	dd	cV	00	3	33.33
83	lo	ag	ar	jr	dd	00	00	00	zl	5	55.56
84	jj	eo	rr	ir	00	ej	go	yy	00	4	44.44
86	cj	kk	kq	hp	00	gg	ar	GT	wN	6	66.67
91	om	ll	bn	aq	ee	hh	cc	EV	wP	5	55.56
92	np	oo	al	kk	00	00	gn	oo	FU	4	44.44
93	ee	ee	pp	go	00	jj	gg	ll	bb	1	11.11
94	qq	00	ko	gp	00	ah	cc	eX	00	4	44.44
95	00	00	lp	iq	bb	ii	ee	00	bb	2	22.22
96	rk	rr	ss	ss	00	hh	kk	gg	00	1	11.11
97	rk	00	cc	lr	cc	00	jj	DD	nn	2	22.22
98	dd	aa	ss	hr	00	00	00	mC	00	2	22.22
99	se	pp	cr	kk	dd	hh	00	00	SS	2	22.22
100	kk	ai	cc	ii	cc	ii	00	00	00	1	11.11
101	fa	ss	rr	jj	00	gg	ll	GW	00	2	22.22
102	aa	bb	00	ii	00	ee	kk	ee	00	0	0.00
103	fa	hp	00	jj	cc	00	00	ee	00	2	22.22
104	io	00	dr	00	00	gg	00	00	00	2	22.22
105	aa	ee	00	bb	00	00	00	00	00	0	0.00
106	ir	qq	00	00	cc	gg	00	lu	mJ	3	33.33
107	pp	ee	00	ll	00	bb	kk	00	00	0	0.00
108	te	00	00	00	ff	00	00	00	hh	1	11.11
109	mm	ee	kk	ff	00	cc	jj	tL	hT	2	22.22
110	dd	oo	qq	00	00	aa	aa	dK	fR	2	22.22
111	km	pp	aq	hp	gg	00	bb	bb	mC	3	33.33
112	kk	bs	hh	ee	00	00	aa	gg	LL	1	11.11
113	fp	00	00	00	ee	00	dd	br	00	2	22.22
114	fa	jp	00	ii	00	00	00	xx	00	2	22.22
115	ra	em	ss	00	00	00	00	00	00	2	22.22
116	ee	00	00	jj	00	ee	ii	00	00	0	0.00
117	ee	hh	00	00	00	bl	mm	OO	00	1	11.11
118	aj	00	00	00	00	ee	er	00	cq	3	33.33
119	ee	00	00	gg	00	ee	fr	zP	bk	3	33.33
120	cp	ho	ss	hh	00	ck	00	zO	pE	4	44.44
121	aa	io	00	bb	00	cc	pp	ww	gg	1	11.11
122	oa	pp	nn	fo	00	dd	en	xN	00	4	44.44
123	tf	00	qq	gg	00	00	00	vv	cc	1	11.11
124	00	ss	oo	en	ff	00	00	00	00	1	11.11
125	00	00	00	cc	bb	00	ee	iq	00	1	11.11
126	id	00	rr	ii	aa	00	00	00	qq	1	11.11
127	00	00	ar	ii	00	00	gg	zz	kk	1	11.11
128	kk	aa	gn	jj	00	00	ff	AS	jz	3	33.33
129	ra	00	qq	00	aa	gg	kk	xx	xx	1	11.11
130	tt	00	oo	fm	00	00	pp	jj	gg	2	22.22
131	rc	ej	pp	ee	dd	00	aa	uu	gg	2	22.22
132	rc	ll	js	kk	dd	jj	00	LL	ee	2	22.22
133	hq	dd	jj	gg	00	00	00	00	00	1	11.11

134	00	00	mm	en	00	00	nn	00	00	1	11.11
135	00	00	nn	ll	00	00	00	00	00	0	0.00
136	00	00	00	00	00	00	nn	00	sH	1	11.11
137	00	aj	cs	jj	00	dd	pp	PP	QQ	2	22.22
138	00	bb	00	00	00	cc	bn	KX	EE	1	11.11
139	00	af	00	hh	00	dd	00	zz	mt	2	22.22
140	mm	kp	nn	gq	00	cc	00	xO	jv	4	44.44
141	cj	qq	nn	gp	00	aa	00	zz	rr	2	22.22
142	nn	hh	ap	go	00	00	00	BO	qF	4	44.44
143	00	00	nn	ff	00	ll	qq	BB	qF	1	11.11
144	00	00	nn	ii	aa	00	00	00	00	0	0.00
145	ll	00	ll	ff	00	00	00	00	kC	1	11.11
146	00	00	00	00	00	00	00	AA	oo	0	0.00
147	00	gg	oo	hh	00	aa	kk	sE	dE	2	22.22
148	00	bb	kk	ds	00	ii	kk	hh	lB	2	22.22
149	00	dd	00	bk	00	00	ff	tt	oo	1	11.11
150	aa	00	rr	00	aa	00	kk	tX	EE	1	11.11
151	00	00	00	00	ff	00	kk	BB	KK	0	0.00
152	00	00	00	00	00	00	00	gg	00	0	0.00
% Het. (BI,WL)	45.33	30.16	43.24	44.59	3.13	10.17	21.43	56.00	57.58		

WI – Within individual heterozygosity.
BI, WL – Between individual, within locus heterozygosity.

3.4.4 Quantification of alleles

The frequencies of the different alleles of the different microsatellite loci were calculated from the input file generated with the genotypes, named *GenotypesCorrected2.csv* (*BPGenotypes.csv*). This input file of genotypes, was then imported into the programme CERVUS. The allele frequencies were determined by selecting the appropriate routine parameters. The programme generated an output allele frequency file, named *AlleleFreq.txt* (*AlleleFreqBP.txt*). The allele frequencies listed in this output file revealed that the Locus CAU 65 was monomorphic and was subsequently excluded from the input file to prevent an error message by the programme. Table 3.3 details the allelic variations of the nine ostrich loci tested in 97 colony individuals comprising 20 adults and 77 progeny. The number of alleles per locus ranged from 7-52 with an observed heterozygosity of 0.031-0.576 for the colony and from 5-15 alleles per locus with an observed heterozygosity of 0.125-0.778 for the breeding pairs. The programme calculated null allele frequencies for each microsatellite locus. CERVUS estimates

the frequency of any null allele segregating at each locus, using an iterative algorithm based on the difference between observed and expected frequency of homozygotes. A null allele occurs because of mutations in one or both primer binding sites, sufficient to prevent effective amplification of the microsatellite allele (Callen *et al.*, 1993). A locus with a large positive estimate of null allele frequency (large relative to other loci in the analysis) indicates an excess of homozygotes, but does not necessarily imply that a null allele is present. Furthermore, in the absence of known parent-offspring relationships, it is more difficult to identify a null allele with certainty (Marshall *et al.*, 1998).

Table 3.3 Allelic variations of the nine ostrich loci used in 97 colony individuals comprising 20 adults and 77 progeny.

Locus	No. Alleles	No. Individuals typed	Observed heterozygosity	Null allele frequency
OSM 1	20	75	0.453	0.3392
OSM 2	19	63	0.302	0.5091
CAU 14	19	74	0.432	0.3619
CAU 17	19	74	0.446	0.3523
List005	07	32	0.031	0.9269
List009	12	59	0.102	0.7966
OSM 7	21	70	0.214	0.6266
ULoc 1	50	75	0.560	0.2693
ULoc 2	52	66	0.576	0.2567

3.5 PARENTAGE ASSIGNMENT

The multiplex fingerprints were used to identify the parents of the different progeny. The software programme CERVUS was used to perform the parentage analyses by running simulations followed by the parentage wizard which assigns parentage to each offspring based on specific criteria.

The first step in the assignment of parents to progeny involved a simulation. The output file of the allele frequency analysis, *AlleleFreq.alf* (*AlleleFreqBP.alf*), was required as an input file for the simulation. Most of the default parameters were selected for the execution of the simulation, except for "Candidate parents" and "Prop. of loci typed". These parameters included the selection of the "Candidate parents" value, which was determined by the number of possible parents for the progeny sampled and the selection of the value for "Prop. of loci typed", which was obtained from the allele frequency output file. Separate simulations were carried out for the male candidate parents and for the female candidate parents of the colony generating the output files *sim4m.txt* and *sim4f.txt*. A single simulation was performed for breeding pairs since the number of candidate parents was equal, generating the output file *simBP.txt*. The delta criterion, the statistic used to assess the reliability of assigning parentage to the most likely candidate parent, calculated by the simulations for the colony, was 0.45 for assigning paternity with neither parent known, and 0.57 for assigning maternity with known paternity and 0.78 for assigning maternity with neither parent known. The delta criterion calculated by the simulation for the breeding pairs was 0.

The second step was to run the parentage wizard. The parentage wizard was run first for males then for females since both parents were unknown. Step 1 of the wizard required an offspring input file containing the genotypes of all the progeny. This input file, *Offspring.csv* (*offspringBP.csv*), was created by editing the genotypes file to include only the genotypes of the progeny. Step 2 of the wizard required a candidate parent data input file. This input file, *ParentsMale.csv* (*MalesBP.csv*), was created by editing the genotypes file to include only male candidate parent's genotypes. Step 3 required the genotype file, *GenotypesCorrected2.csv* (*BPGenotypes.csv*), as input. In this step, a value of 5 was selected for the minimum number of loci required for analysis. Step 4 required the allele frequency data and simulation data output files as input. Step 5 required the naming of the output files generated from the parentage analysis. "Most-likely parent" was the parameter selected to be included for each offspring in the output file *parM.csv* (*parMBP.csv*). The same procedure was followed for assigning maternity using the parentage wizard. The only difference occurred in step 1,

where the offspring input file was required. Here, the output file of the male parentage analysis, *parM.csv* (*parMBP.csv*), was used as the input file since paternity was already assigned i.e. one known parent. The output file generated from the maternal parentage analysis was named *parF.csv* (*parFBP.csv*).

Parentage assignments were made using the delta criterion at the 95% confidence interval (strict), the 80% confidence interval (relaxed) or not at all (unresolved). Individuals with a delta score of zero could not be assigned a parent.

For the breeding pairs, the assignment of males presented delta scores of zero for all individuals. However, most likely parents were suggested for each individual and seven out of the eight individuals were assigned paternity accurately. The assignment of females in the breeding pairs based on the paternity results presented two assignments at the 95% confidence interval, with the other six being unresolved. Here too the most likely parents were suggested and five out of the eight individuals were assigned maternity accurately. The mean observed error rate across loci for known parent-offspring mismatches amounted to 0.6658.

The results obtained from the parentage analyses of the colony were expressed in terms of number of assignments (Table 3.4). Of the 77 progeny typed for the colony, 61 were assigned fathers at the 80% confidence interval while 16 were unresolved with regards to paternity. 68 individuals were assigned mothers at the 80% confidence interval, of which 27 were assigned with an untyped known parent i.e. the known parent, adult males in this case, was typed at fewer than five loci. A total of nine individuals was unresolved as regards maternity. By comparison of unresolved individuals for paternity with those of maternity, it was conclusive that four individuals (sample ID: 112; 126; 144; 152) were unassigned to either a mother or a father. The mean observed error rate across loci for known parent-offspring mismatches amounted to 0.4833.

Table 3.4 Number of parentage assignments in the colony at respective confidence intervals.

Confidence interval	Assignments		
	Paternity	Maternity	
		KP* typed~	KP* untyped#
95 %	33	15	6
80 %	61	41	27
Unresolved	16	8	1

* KP Known parent, fathers in this case.
~ Typed at 5 or more loci.
# Typed at fewer than 5 loci.

3.6 PEDIGREES

Pedigrees were constructed using either the microsatellite fingerprints or the results of the parentage analyses. For the two breeding pair families, the microsatellite fingerprints were used because parentage was known. The pedigrees were constructed by comparing the genotypes of the progeny to the genotypes of the parents using manual inspection after computer analysis to increase the specificity. The pedigrees of the two breeding pair families are shown in Figure 3.7. The alleles of the progeny correspond to the alleles of their parents within 4 base pairs. This discrepancy is due to electronic gel scoring since a 4 base pair or smaller difference cannot be differentiated during band detection. Larger alleles that appear in these individuals are amplification products of the anonymous loci and were not used in the construction of these pedigrees.

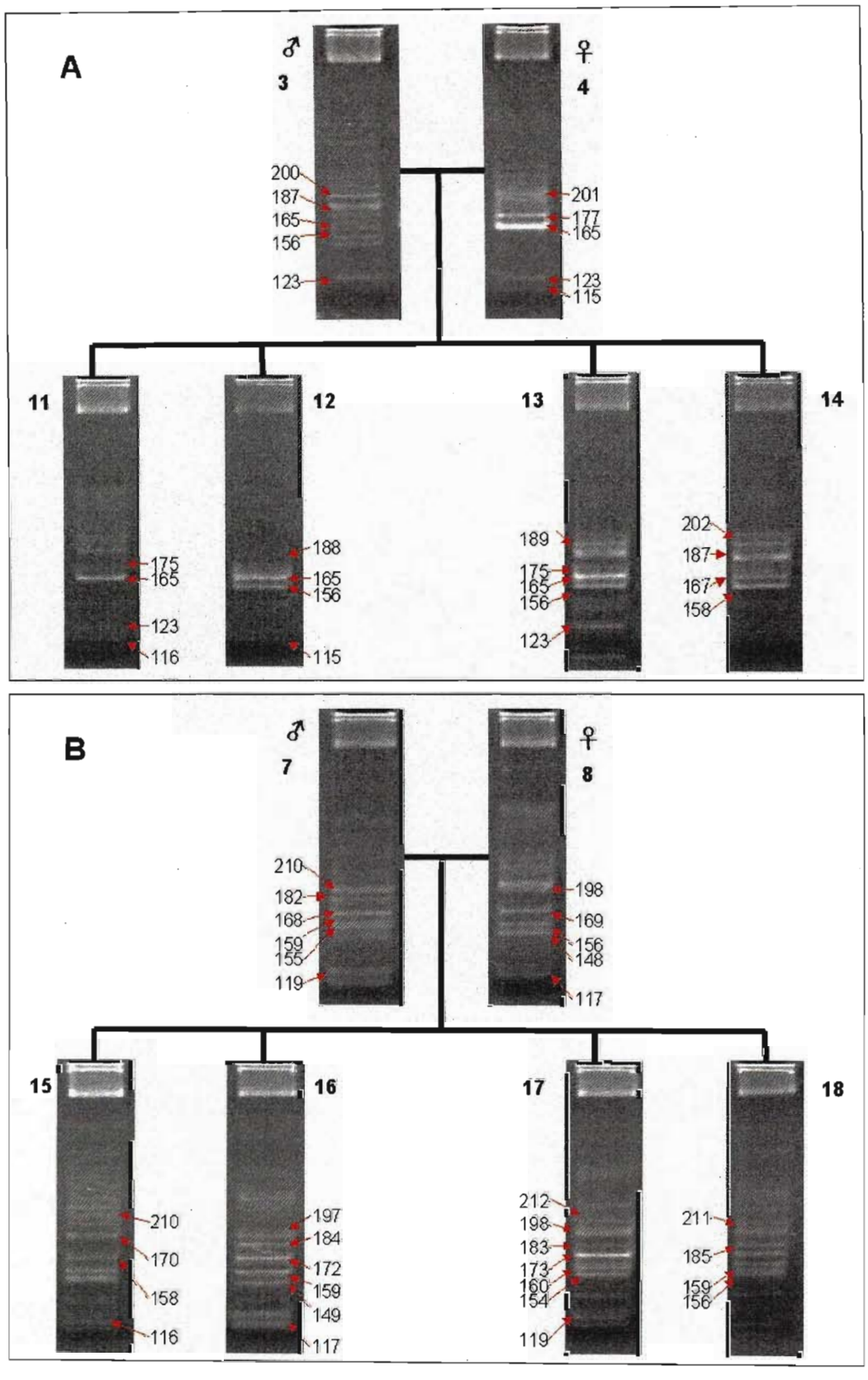


Figure 3.7 Pedigrees of the two breeding pair families A and B. Sample ID numbers (3,4,7,8, etc.) appear next to the individual multi-locus fingerprints. Sizes in base pairs correspond to bands indicated by red arrows.

For the colony, since parentage was unknown, the results of the parentage analyses, generated by CERVUS, were used. Table 3.5 illustrates the pedigrees of the colony in a male to female combination system, where the progeny sample ID number indicates products of each mating. The reliability of the assignments and death-in-shell progeny are included. 46.75% progeny individuals were not assigned both parents because they were either typed at too few loci to warrant assignment or the candidate parent was typed at too few loci to enable a comparison of loci.

Pedigree records of the colony parents identified three full sibs among the parents mated, namely sample ID: 40 and 49; 41 and 50; 52 and 54. Two full sib matings occurred with individuals 41 and 50 producing one chick and individuals 52 and 54 producing two chicks. Due to the limited number of crosses it was not possible to identify particular crosses or individuals that showed a predisposition to producing death-in-shell progeny.

Table 3.5 Mating combinations of colony parents and progeny produced based on 41 assignments.

♀ \ ♂	**48**	**49**	**50**	**51**	**53**	**54**
35	(69*)		138		102*;148*;150	
36		119	140		124; 133	
37			(56); 120; 132*		(77)	
38		80				
39	149		99			
40			(84)			
41			117*		108	
42						
43	(55*); 128	116*; 137		103*	101*	
44		123*; 146			96; 129*; 151	
45						
46			83*; 98*			
47			136			125; 134
52		114		97*	(78*); 104*	105; 109

* Assigned at 95% confidence interval.

() Death-in-shell eggs.

3.7 BREEDING STATISTICS

Breeding statistics were calculated for the production potential of the individual females in the colony. From the parentage analyses, the number of fertile eggs produced by individual females was derived as well as the number of eggs fertilized by individual males. This is depicted in the frequency distributions of figure 3.8 A and figure 3.8 B, respectively. It was interesting to note that one female (sample ID: 45) failed to produce any eggs during the sampling period. However, all males contributed to the fertilization of a minimum of five eggs during the sampling period. The mean and standard deviation were calculated for the production of fertile eggs by females and for the fertilization of eggs by males based on 68 and 61 assigned progeny, respectively. On average, females produced 4.86 ± 2.71 fertile eggs during the sampling period, while the males fertilized an average of 10.17 ± 3.25 eggs during the sampling period.

The coefficient of variation for the production of fertile eggs by individual females during the sampling period was calculated by taking the dividend of the standard deviation and the mean and expressing it as a percentage. This measure of variation in egg production amounted to 55.86%.

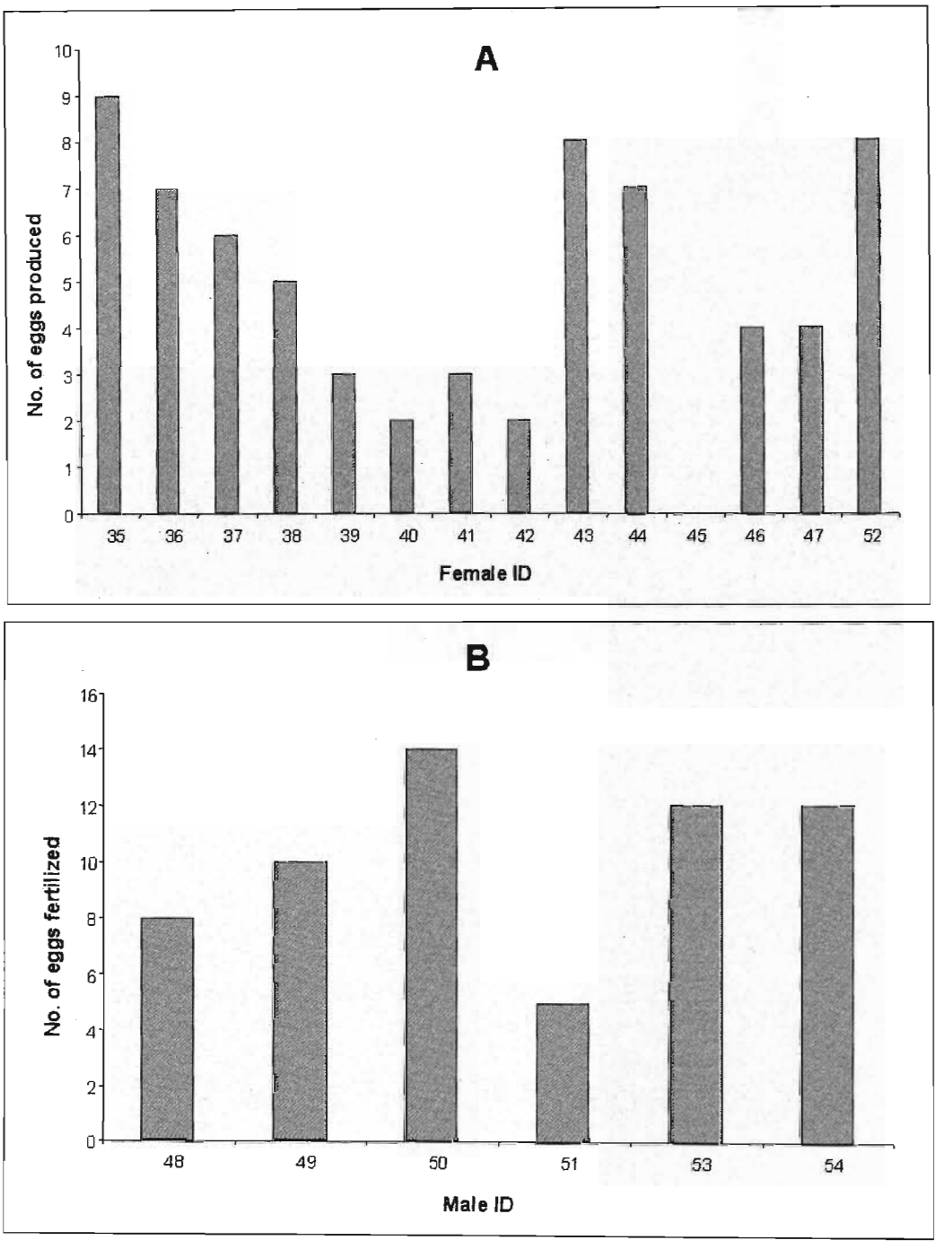


Figure 3.8 Number of eggs produced by individual colony females based on the genotypes of 68 progeny (A), and number of eggs fertilized by individual colony males based on the genotypes of 61 progeny (B).

CHAPTER 4
DISCUSSION AND CONCLUDING REMARKS

The importance of the ostrich industry in South Africa can be highlighted by the fact that income from ostrich products ranks as one of the top twenty income earners from agriculture related activities in this country. This income is derived mainly from exports of ostrich leather and ostrich meat. Both of these products are highly sought after in global markets: ostrich leather for its durability and suppleness, and ostrich meat as a healthy alternative to traditional red meat.

Ostriches are indigenous to Africa but are now commercially farmed throughout the world. The three subspecies most commonly used for commercial breeding are the Kenyan Reds, the Zimbabwean Blues and the South African Blacks. South Africa has the clear advantage in this increasingly competitive industry through the long genetic heritage of domesticated birds, which facilitate superior farm economics. In order for South Africa to maintain its leadership position in the industry, new breeding strategies need to be investigated.

Colony breeding ostriches share communal nests with the result that parentage of eggs and chicks is unknown. Within the South African ostrich industry about 80% of breeding birds are kept in colonies. This means that there is a lack of pedigree information, which hinders the assessment of production data and the reproduction potential of individual breeding birds. The development of molecular biological techniques, specifically the application of various DNA markers, has created new possibilities for the selection and genetic improvement of livestock (Van Marle-Koster & Nel 2003). The high variability of microsatellites makes them the most useful marker for use in genetic typing of individuals for parentage and kinship studies (Barker 2002).

Novel microsatellites are costly and labour-intensive to develop. When available for a species or related species, the known microsatellites are tested for informativeness. For ostriches, the few published microsatellite markers were developed for Kenyan Reds (Ward *et al.* 1994; Kimwele *et al.* 1998; Kumari &

Kemp 1998; Ward *et al.* 1998). However, 70 novel microsatellite markers were later developed for Blues and Blacks by Tang *et al.* (2003). This investigation addressed the applicability of a selection of these known microsatellite markers, developed for other ostrich species, in a South African Black ostrich population. The microsatellites were further tested for their ability to perform individual and parentage identification.

In testing the 11 selected microsatellite markers, the amplification reaction conditions of Kimwele and Graves (2003) were followed. However, optimisation of the protocol was necessary, since little or no amplification resulted, and hence the number of cycles was increased. This increase generated more products resulting in distinct banding after electrophoresis. The annealing temperatures of all the markers were optimised until an optimum product was obtained for each marker with minimum stutter bands. Three of the eleven markers did not amplify through a range of annealing temperatures indicating a low homology of these primers such that further optimisation with respect to other PCR variables is required. One of these three markers, VIAS-OS 29, amplified successfully in Blacks from a Polish ostrich population (Kawka 2005).

Additional amplification products were detected outside of the expected size range of products for four of the markers tested. These were markers developed for Reds and amplification may be attributed to an additional annealing site in the genome sequence of Blacks. This was not unexpected and provides indications of the genome diversity between these two subspecies.

It was interesting to note that five loci displayed a greater number of alleles than that reported in the literature (Kimwele *et al.* 1998; Tang *et al.* 2003) with about 19 alleles per locus. On the other hand, two loci displayed slightly fewer alleles than the reported number (Kumari & Kemp 1998). Locus CAU 65 was found to be monomorphic in this population despite literature citing it as polymorphic with six alleles (Tang *et al.* 2003). The relatively low observed heterozygosity at all loci suggests that there is little genetic variation in this population. A comparative study by Kimwele and Graves (2003) on a wild population of Reds using four of the same markers showed higher observed heterozygosity values. A high null allele

frequency was observed for all loci, however, it is difficult to identify a null allele with certainty in the absence of known parent-offspring relationships. Furthermore, a locus with a large positive estimate of null allele frequency indicates an excess of homozygotes but does not necessarily imply that a null allele is present (Hammond *et al.* 2002).

Analysis of the genotypes of the colony individuals revealed that 90% of individuals had a low number of heterozygous loci. Likewise, the heterozygosity within locus, between individuals, was reasonably low. Low heterozygosity suggests inbreeding in the population and it has yet to be investigated if this is the causative factor of the high chick mortality that is commonly observed with these birds. However, DeWoody and DeWoody (2005) have reported that genome-wide heterozygosity is poorly estimated by microsatellite loci and associations between phenotypes and heterozygosity should be established firmly on causative factors and not on simple correlations.

Using the control sample, the pair breeding families, this study found that eight markers were sufficient to assign parentage on comparison of genotypes. According to Dodds (2003), 10-13 markers are required to assign parentage at the 80-95% level without knowledge of sire-dam combinations using a co-dominant marker. On the other hand, Marshall *et al.* (1998) suggest that the number of loci required to resolve parentage with a given level of confidence depends on factors such as the level of variation at a locus (expected heterozygosity), the number of candidate parents, the proportion of candidate parents sampled and the availability of genetic data from a known parent.

The genotypic usefulness in parentage identification was evaluated using the success rates based on the control sample of known parentage. The success rate for assigning paternity correctly in the breeding pair samples was 87.5% and for assigning maternity correctly was 62.5%. However, it should be noted that the sample size of the breeding pair families was small. A total of 79.2% individuals was assigned paternity in the colony and 88.3% were assigned maternity. Four individuals (5.2%) could not be assigned either parent because they were typed at too few loci. Of the 68 individuals assigned maternity, 41 were assigned paternity

as well, thus some mating combinations could be resolved in the colony. Too few matings disabled the possibility of identifying a predisposition to death-in-shell chicks in some combinations. When the relationships between potential parents in the colony were considered, it was found that there were three sets of full sibs amongst the parents. This relatedness between parents can account for the low observed heterozygosity and high null allele frequencies at loci, due to full sibs having a higher probability of band sharing than other unrelated individuals. Although parentage could be assigned, the accuracy of these assignments is low but can be improved by increasing the number of polymorphic loci. One female parent was not assigned any progeny during the sampling period. The parentage technology could facilitate in the monitoring and evaluation of the breeding stock such that low or non-producers can be removed from the breeding system.

The assignment of parentage allowed for the allocation of fertile eggs to individual females. Infertile eggs were not sampled in this population owing to the difficulty in extracting DNA material from the eggs. A coefficient of variation of 55.8% was calculated for the production of fertile eggs by individual females during the sampling period. This measure of variation is about equivalent to variation in egg production in studies done on pair breeding females (51.8%: Bunter *et al.* 2001; 52.9%: Cloete *et al.* 2004), however these studies included infertile eggs.

The practical implication of this study was to assess the fertilization potential of individual males and, more importantly, the production potential of females within the colony mating system. In order to determine overall egg production, the parentage analysis will have to be carried out with a higher accuracy, incorporating all eggs laid i.e. fertile and infertile eggs. The high homozygosity and low genetic variation within this colony is indicative of some inbreeding. This can potentially be circumvented by assessing pedigrees of birds for relatedness before accepting them as replacements in colonies.

To conclude, microsatellite markers are an efficient, effective method for individual and parentage identification in colony breeding systems. The design of a robust set of microsatellite markers will need to address the issue of relatedness between birds to increase the effectiveness. Furthermore, to increase the accuracy with

which assignment is done, a greater number of loci are required in addition to the loci that were used. The further development of a robust DNA fingerprinting protocol as is available in conventional livestock breeding schemes should receive serious attention, if extrapolation to the broader industry is considered.

REFERENCES

Avise JC (2004) *Molecular markers, natural history, and evolution*, 2nd edn. Sinauer Associates, Massachusetts.

Badley AR (1997) Fertility, hatchability and incubation of ostrich (*Struthio camelus*) eggs. *Poultry and Avian Biology Reviews* 8, 53-76.

Barker GC (2002) Microsatellite DNA: a tool for population genetics. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 96, 21-24.

Baronnig R (2002) Ostrich meat market in EU. *Proceedings of World Ostrich Congress*, Warsaw, Poland, 26-29 September, 161-165.

Bello N, Sanchez A (1999) The identification of a sex-specific DNA marker in the ostrich using a random amplified polymorphic DNA (RAPD) assay. *Molecular Ecology* 8, 667-669.

Bezuidenhout CC (1999) Studies of the population structure and genetic diversity of domesticated and "wild" ostriches (*Struthio camelus*), PhD dissertation, Rhodes University, South Africa.

Brand MM, Cloete SWP, Hoffman LC, Muller M (2005) A comparison of liveweights, body measurements and reproductive traits in Zimbabwean Blue ostriches (*Struthio camelus*) and South African ostriches (*Struthio camelus australis*) and South African Black ostriches (*S. camelus* VAR. *Domesticus*). *Proceedings of the 3rd International Ratite Science Symposium and XII World Ostrich Congress*, Madrid, Spain, 14-16 October, 73-80.

Bunter KL (2002) The genetic analysis of reproduction and production traits recorded for farmed ostriches (*Struthio camelus*), PhD dissertation, University of New England, Australia.

Bunter KL, Cloete SWP (2004) Genetic parameters for egg-, chick- and live-weight traits recorded in farmed ostriches (*Struthio camelus*). *Livestock Production Science* 91, 9-22.

Bunter KL, Cloete SWP, Van Schalkwyk SJ, Graser H-U (2001) Factors affecting reproductive performance in farmed ostriches. *Proceeding of the Association for the Advancement of Animal Breeding and Genetics* 14, 43-46.

Callen DF, Thompson AD, Shen Y, *et al.* (1993) Incidence and origin of "null" alleles in the (AC)n microsatellite markers. *American Journal of Human Genetics* 52, 922-927.

Cloete SWP, Bunter KL, Brand Z, Lambrechts H (2004) (Co)variances for reproduction, egg weight and chick weight in ostriches. *South African Journal of Animal Science* 34, 17-19.

Cloete SWP, Lambrechts H, Punt K, Brand Z (2001) Factors related to high levels of ostrich chick mortality from hatching to 90 days of age in an intensive rearing system. *Journal of the South African Veterinary Association* 72, 197-202.

Cloete SWP, Van Schalkwyk SJ, Brand Z (1998) Ostrich breeding - Progress towards a scientifically based strategy. *Proceedings of the 2nd International Ratite Congress*, Oudtshoorn, South Africa, 21-25 September, 55-62.

Cloete SWP, Van Schalkwyk SJ, Bunter KL (2002) Progress in ostrich breeding research. *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production*, Montpellier, France, 19-23 August, 23-37.

Cooper A, Mourer-Chauvire C, Chambers GK, *et al.* (1992) Independent origins of moas and kiwis. *Proceedings of the National Academy of Sciences of the United States of America* 89, 8741-8744.

Cunningham EP, Meghen CM (2001) Biological identification systems: genetic markers. *Revue scientifique et technique Office international des Epizooties* 20, 491-499.

Deeming DC (1999) Introduction. In: *The Ostrich: Biology, Production and Health* (ed. Deeming DC), pp. 1-12. CABI Publishing, UK.

Deeming DC, Ayres L, Ayres FJ (1993) Observations on the commercial production of ostrich (*Struthio camelus*) in the United Kingdom: rearing of chicks. *Veterinary Record* 132, 627-631.

DeWoody YD, DeWoody JA (2005) On the estimation of genome-wide heterozygosity using molecular markers. *Journal of Heredity* 96, 85-88.

Dodds KG (2003) The number of markers required for parentage assignment. *Proceeding of the Association for the Advancement of Animal Breeding and Genetics* 15, 39-42.

Dodgson JB, Cheng HH, Okimoto R (1997) DNA Marker Technology: A Revolution in Animal Genetics. *Poultry Science* 76, 1108-1114.

Essa F, Cloete SWP (2004) Differentiation between females of ostrich breeding trios based on egg weights. *South African Journal of Animal Science* 34, 20-22.

Freitag S, Robinson TJ (1993) Phylogeographic patterns in mitochondrial DNA of the ostrich (*Struthio camelus*). *The Auk* 110, 614-622.

Griffiths R, Orr K (1999) The use of amplified fragment length polymorphism (AFLP) in the isolation of sex-specific markers. *Molecular Ecology* 8, 671-674.

Haddrath O, Baker AJ (2001) Complete mitochondrial DNA genome sequences of extinct birds: ratite phylogenetics and the vicariance biogeography hypothesis. *Proceedings of the Royal Society of London* 268, 939-945.

Hammond EL, Lymbery AJ, Martin GB, Groth D, Wetherall JD (2002) Microsatellite analysis of genetic diversity in wild and farmed emus (*Dromaius novaehollandiae*). *The Journal of Heredity* 93, 376-380.

Harlid A, Arnason U (1999) Analyses of mitochondrial DNA nest ratite birds with the Neognathae: supporting neotenous origin of ratite morphology. *Proceedings of the Royal Society of London* 266, 305-309.

Harlid A, Janke A, Arnason U (1997) The mtDNA sequence of the ostrich and the divergence between paleognathous and neognathous birds. *Molecular Biology and Evolution* 14, 754-761.

Hinckley JD, Park RL, Xion S, Anderson WR, Kooyman DL (2005) Identification and development of sex specific DNA markers in the ostrich using polymerase chain reaction. *International Journal of Poultry Science* 4, 663-669.

Huynen L, Millar CD, Lambert DM (2002) A DNA test to sex ratite birds. *Molecular Ecology* 11, 851-856.

Jarvis MJF (1998) The subspecies and races of ostriches and their present status in the wild. *Proceedings of the 2nd International Ratite Congress*, Oudtshoorn, South Africa, 21-25 September, 4-8.

Karp A, Isaac PG, Ingram DS (2001) *Molecular Tools for Screening Biodiversity* Kluwer Academic Publishers, Netherlands.

Kawka M (2005) Genetic characteristics of African ostrich (*Struthio camelus*) in Poland using molecular methods, PhD dissertation, Polish Academy of Sciences, Poland.

Kimwele CN, Graves JA (2003) A molecular genetic analysis of the communal nesting of the ostrich (*Struthio camelus*). *Molecular Ecology* 12, 229-236.

Kimwele CN, Graves JA, Burke T, Hanotte O (1998) Development of microsatellite markers for parentage typing of chicks in the ostrich *Struthio camelus*. *Molecular Ecology* 7, 247-255.

Kumari P, Kemp SJ (1998) Polymorphic microsatellite markers in the ostrich (*Struthio camelus*). *Molecular Ecology* 7, 133-140.

Lambrechts H (2004) Reproductive efficiency of ostriches (*Struthio camelus*), PhD dissertation, Universiy of the Free State, South Africa.

Luikart G, England PR (1999) Statistical analysis of microsatellite DNA data. *TREE* 14, 253-256.

Malago W, Franco HM, Matheucci E, Medaglia A, Henrique-Silva F (2002) Large scale sex typing of ostriches using DNA extracted from feathers. *BMC Biotechnology* 2, 19.

Malecki IA, Martin GB (2002a) Semen collection in the emu and ostrich. *Proceedings of World Ostrich Congress*, Warsaw, Poland, 26-29 September, 38-43.

Malecki IA, Martin GB (2002b) Egg fertilisation in the emu and ostrich - How much sperm do they need? *Proceedings of World Ostrich Congress*, Warsaw, Poland, 26-29 September, 44-49.

Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology* 7, 639-655.

Mine OM, Mochakana ME, Mpapho T, Motlhanka DTM, Kgwatalala P (2002) Application of a sex identification technique in juvenile ostriches and its potential application in Botswana. *South African Journal of Animal Science* 32, 160-163.

Ogawa A, Murata K, Mizuno S (1998) The location of Z- and W-linked marker genes and sequence on the homomorphic sex chromosomes of the ostrich and the emu. *Proceedings of the National Academy of Sciences of the United States of America* 95, 4415-4418.

Parker PG, Snow AA, Schug MD, Booton GC, Fuerst PA (1998) What molecules can tell us about populations: Choosing and using a molecular marker. *Ecology* 79, 361-382.

Petitte JM, Davis G (1999) Breeding and genetics. In: *The Ostrich: Biology, Production and Health* (ed. Deeming DC), 275-292. CABI Publishing, UK.

Petitte JN, Petitte JM, Scheideler SE (1996) Determination of genetic diversity in commercial ratite stocks using multilocus DNA fingerprinting. *Proceedings of the 1st International Ratite Conference*, Manchester, England, 27-29 March, 69-77.

Sacharczuk M, Horbanczuk JO, Jaszczak K (2002) Genetic studies of ostrich by DNA fingerprinting method. *Proceedings of World Ostrich Congress*, Warsaw, Poland, 26-29 September, 56-58.

SAOBC (2004) The South African Ostrich Industry. South African Ostrich Business Chamber.

Shanawany MM (1995) Recent developments in ostrich farming. *World Animal Review* 83, 3-8.

Snustad DP, Simmons MJ (2000) *Principles of Genetics*, 2nd edn. John Wiley & Sons, New York.

Soley JT, Groenewald HB (1999) Reproduction. In: *The Ostrich: Biology, Production and Health* (ed. Deeming DC), 129-158. CABI Publishing, UK.

Stewart S (2004) World Ostrich Association and World Ostrich Economics. In: *World Ostrich Association Conference*. World Ostrich Association, Xi'an, China.

Takagi N, Itoh M, Sasaki M (1972) Chromosome studies in four species of *Ratitae* (*Aves*). *Chromosoma* 36, 281-291.

Tang B, Huang YH, Lin L, *et al.* (2003) Isolation and characterization of 70 novel microsatellite markers from ostrich (*Struthio camelus*) genome. *Genome* 46, 833-840.

Van Marle-Koster E, Nel LH (2003) Genetic markers and their application in livestock breeding in South Africa: A review. *South African Journal of Animal Science* 33, 1-10.

Van Schalkwyk SJ, Cloete SWP, de Kock JA (1996) Repeatability and phenotypic correlations for body weight and reproduction in commercial ostrich breeding pairs. *British Poultry Science* 37, 953-962.

Van Tuinen M, Sibley CG, Hedges SB (1998) Phylogeny and biogeography of ratite birds inferred from DNA sequences of the mitochondrial ribosomal genes. *Molecular Biology and Evolution* 15, 370-376.

Ward WK, Matthews ME, Murray ND, Robinson NA (1994) An ostrich dinucleotide repeat polymorphism at the VIAS-OS2 locus. *Animal Genetics* 25, 291.

Ward WK, McPartlan HC, Matthews ME, Robinson NA (1998) Ostrich microsatellite polymorphisms at the VIAS-OS4, VIAS-OS8, VIAS-OS14, VIAS-OS22, and VIAS-OS29 loci. *Animal Genetics* 29, 331.

APPENDIX A

BUFFERS AND REAGENTS

---

5 x TBE (1 L)

- 54 g Tris
- 27.5 g Boric Acid
- 20 mL 0.5 M EDTA pH 8.0

Dissolve in distilled water and make up to 1 L. Autoclave.

Loading Buffer (1 mL)

- 600 µl Glycerol
- 280 µl sterile distilled water
- 120 µl 0.5 M EDTA
- 0.002 g Bromophenol Blue

Mix by vortexing and aliquot into tubes. Store at 4 °C.

40% Polyacrylamide Stock (5%C)

- 38 g Acrylamide
- 2 g Bisacrylamide

Make up to a total volume of 100 mL with distilled water. Heat slowly to dissolve.

NB: Solution expands. Store in a dark bottle at room temperature.

6% Polyacrylamide Gel (40 mL)

- 6 mL 40% Polyacrylamide stock
- 4.8 mL 5 x TBE
- 29.2 mL distilled water
- 120 µl 25% Ammonium Persulfate (APS)
- 40 µl TEMED

Add APS and TEMED last to set gel.

Ethidium Bromide Stain

- 1 L 1 x TBE
- 50 µl Ethidium Bromide (10 mg/mL)
- Destain in 1 L distilled water

APPENDIX B

OUTPUT FILES

BREEDING PAIR ANALYSES OUTPUT FILES

- **BPGenotypes.csv**

	OSM1a	OSM1b	OSM2a	OSM2b	CAU14a	CAU14b	CAU17a	CAU17b	LIST5a	LIST5b	LIST9a	LIST9b	OSM7a	OSM7b	UAll1a	UAll1b	UAll2a	UAll2b
3	100	100	123	123	150	156	165	178	0	0	200	200	211	223	231	280	302	322
4	115	115	123	123	156	156	165	177	195	195	201	201	214	223	239	268	0	0
7	99	119	0	0	145	155	168	168	193	193	204	210	0	0	0	0	0	0
8	117	117	122	141	148	156	169	169	196	198	0	0	220	229	235	235	322	322
11	116	116	123	123	143	143	165	175	195	195	209	209	226	226	277	283	0	0
12	115	115	121	141	156	156	165	165	0	0	0	0	0	0	0	0	0	0
13	104	104	123	137	145	156	165	175	192	192	201	201	0	0	268	268	0	0
14	0	0	141	141	145	158	167	167	0	0	202	202	214	214	0	0	0	0
15	112	116	133	133	152	158	171	171	195	195	210	210	221	221	233	247	305	305
16	117	117	124	134	149	159	172	172	197	197	0	0	222	222	249	255	0	0
17	114	119	135	137	154	160	164	173	198	198	0	0	212	212	238	251	291	291
18	109	114	137	137	156	159	165	176	0	0	201	201	211	211	231	259	338	338

- **AlleleFreqBP.txt**

```
Cervus 2.0... Allele Frequency Analysis
© Copyright Tristan Marshall 1998-2001


**** Summary statistics ****

Number of loci:                          9
Number of individuals:                   12

Locus       k   N     Hets  Homs  H(O)    H(E)    PIC     Excl(1)   Excl(2)
Null freq
OSM1        10  11    4     7     0.364   0.918   0.863   0.598     0.750
+0.4191
OSM2        9   11    5     6     0.455   0.853   0.792   0.470     0.644
+0.2972
CAU14       12  12    9     3     0.750   0.877   0.828   0.539     0.704
+0.0623
CAU17       12  12    6     6     0.500   0.902   0.854   0.587     0.741
+0.2480
LIST5       6   8     1     7     0.125   0.825   0.744   0.393     0.574
+0.0000
LIST9       6   8     1     7     0.125   0.825   0.744   0.393     0.574
+0.0000
OSM7        9   9     3     6     0.333   0.928   0.864   0.596     0.749
+0.0000
UAl11       15  9     7     2     0.778   0.980   0.921   0.738     0.849
+0.0000
UAl12       5   5     1     4     0.200   0.867   0.745   0.386     0.565
+0.0000

Mean number of alleles per locus:        9.33
```

```
Mean proportion of individuals typed:     0.787
Mean expected heterozygosity:             0.886
Mean PIC:                                 0.817
Total exclusionary power (first parent):  0.999024
Total exclusionary power (second parent): 0.999980


**** Files ****

Input
  Genotype data file:     BPgenotypes.csv
Output
  Summary text file:      AlleleFreqBP.txt
  Allele frequency file:  AlleleFreqBP.alf


**** Loci ****

1    OSM1
2    OSM2
3    CAU14
4    CAU17
5    LIST5
6    LIST9
7    OSM7
8    UA111
9    UA112


**** Locus OSM1 ****

Number of alleles:                  10
Number of individuals typed:        11
  Heterozygotes:                    4
  Homozygotes:                      7
Observed heterozygosity:            0.364

Allele  Count  Heterozygotes  Homozygotes  Frequency  Frequency with null
99      1      1              0            0.0455     0.0382
100     2      0              1            0.0909     0.0382
104     2      0              1            0.0909     0.0382
109     1      1              0            0.0455     0.0382
112     1      1              0            0.0455     0.0382
114     2      2              0            0.0909     0.0780
115     4      0              2            0.1818     0.0780
116     3      1              1            0.1364     0.0780
117     4      0              2            0.1818     0.0780
119     2      2              0            0.0909     0.0780

Expected heterozygosity:                0.918
Polymorphic information content (PIC):  0.863
Average exclusion probability (1):      0.598
Average exclusion probability (2):      0.750
Null allele frequency estimate:         0.4191


**** Locus OSM2 ****

Number of alleles:                  9
Number of individuals typed:        11
```

```
  Heterozygotes:                        5
  Homozygotes:                          6
Observed heterozygosity:                0.455

Allele  Count  Heterozygotes  Homozygotes  Frequency  Frequency with null
121     1      1              0            0.0455     0.0423
122     1      1              0            0.0455     0.0423
123     7      1              3            0.3182     0.1824
124     1      1              0            0.0455     0.0423
133     2      0              1            0.0909     0.0423
134     1      1              0            0.0455     0.0423
135     1      1              0            0.0455     0.0423
137     4      2              1            0.1818     0.1332
141     4      2              1            0.1818     0.1332

Expected heterozygosity:                0.853
Polymorphic information content (PIC): 0.792
Average exclusion probability (1):      0.470
Average exclusion probability (2):      0.644
Null allele frequency estimate:         0.2972


**** Locus CAU14 ****

Number of alleles:                      12
Number of individuals typed:            12
  Heterozygotes:                        9
  Homozygotes:                          3
Observed heterozygosity:                0.750

Allele  Count  Heterozygotes  Homozygotes  Frequency  Frequency with null
143     2      0              1            0.0833     0.0424
145     3      3              0            0.1250     0.1334
148     1      1              0            0.0417     0.0424
149     1      1              0            0.0417     0.0424
150     1      1              0            0.0417     0.0424
152     1      1              0            0.0417     0.0424
154     1      1              0            0.0417     0.0424
155     1      1              0            0.0417     0.0424
156     8      4              2            0.3333     0.2915
158     2      2              0            0.0833     0.0868
159     2      2              0            0.0833     0.0868
160     1      1              0            0.0417     0.0424

Expected heterozygosity:                0.877
Polymorphic information content (PIC): 0.828
Average exclusion probability (1):      0.539
Average exclusion probability (2):      0.704
Null allele frequency estimate:         0.0623


**** Locus CAU17 ****

Number of alleles:                      12
Number of individuals typed:            12
  Heterozygotes:                        6
  Homozygotes:                          6
Observed heterozygosity:                0.500

Allele  Count  Heterozygotes  Homozygotes  Frequency  Frequency with null
```

164	1	1	0	0.0417	0.0399
165	7	5	1	0.2917	0.2715
167	2	0	1	0.0833	0.0399
168	2	0	1	0.0833	0.0399
169	2	0	1	0.0833	0.0399
171	2	0	1	0.0833	0.0399
172	2	0	1	0.0833	0.0399
173	1	1	0	0.0417	0.0399
175	2	2	0	0.0833	0.0815
176	1	1	0	0.0417	0.0399
177	1	1	0	0.0417	0.0399
178	1	1	0	0.0417	0.0399

Expected heterozygosity: 0.902
Polymorphic information content (PIC): 0.854
Average exclusion probability (1): 0.587
Average exclusion probability (2): 0.741
Null allele frequency estimate: 0.2480

**** Locus LIST5 ****

Number of alleles: 6
Number of individuals typed: 8
  Heterozygotes: 1
  Homozygotes: 7
Observed heterozygosity: 0.125

Allele	Count	Heterozygotes	Homozygotes	Frequency	Frequency with null
192	2	0	1	0.1250	0.0399
193	2	0	1	0.1250	0.2715
195	6	0	3	0.3750	0.0399
196	1	1	0	0.0625	0.0399
197	2	0	1	0.1250	0.0399
198	3	1	1	0.1875	0.0399

Expected heterozygosity: 0.825
Polymorphic information content (PIC): 0.744
Average exclusion probability (1): 0.393
Average exclusion probability (2): 0.574
Null allele frequency estimate: Not done

**** Locus LIST9 ****

Number of alleles: 6
Number of individuals typed: 8
  Heterozygotes: 1
  Homozygotes: 7
Observed heterozygosity: 0.125

Allele	Count	Heterozygotes	Homozygotes	Frequency	Frequency with null
200	2	0	1	0.1250	0.0399
201	6	0	3	0.3750	0.2715
202	2	0	1	0.1250	0.0399
204	1	1	0	0.0625	0.0399
209	2	0	1	0.1250	0.0399
210	3	1	1	0.1875	0.0399

Expected heterozygosity: 0.825

Polymorphic information content (PIC): 0.744
Average exclusion probability (1): 0.393
Average exclusion probability (2): 0.574
Null allele frequency estimate: Not done

**** Locus OSM7 ****

Number of alleles: 9
Number of individuals typed: 9
  Heterozygotes: 3
  Homozygotes: 6
Observed heterozygosity: 0.333

Allele	Count	Heterozygotes	Homozygotes	Frequency	Frequency with null
211	3	1	1	0.1667	0.0399
212	2	0	1	0.1111	0.0815
214	3	1	1	0.1667	0.0399
220	1	1	0	0.0556	0.0399
221	2	0	1	0.1111	0.0399
222	2	0	1	0.1111	0.0399
223	2	2	0	0.1111	0.2715
226	2	0	1	0.1111	0.0399
229	1	1	0	0.0556	0.0399

Expected heterozygosity: 0.928
Polymorphic information content (PIC): 0.864
Average exclusion probability (1): 0.596
Average exclusion probability (2): 0.749
Null allele frequency estimate: Not done

**** Locus UA111 ****

Number of alleles: 15
Number of individuals typed: 9
  Heterozygotes: 7
  Homozygotes: 2
Observed heterozygosity: 0.778

Allele	Count	Heterozygotes	Homozygotes	Frequency	Frequency with null
231	2	2	0	0.1111	0.0399
233	1	1	0	0.0556	0.0399
235	2	0	1	0.1111	0.0399
238	1	1	0	0.0556	0.0000
239	1	1	0	0.0556	0.0399
247	1	1	0	0.0556	0.0399
249	1	1	0	0.0556	0.0399
251	1	1	0	0.0556	0.0000
255	1	1	0	0.0556	0.0399
259	1	1	0	0.0556	0.0000
268	1	1	0	0.0556	0.0399
277	1	1	0	0.0556	0.0399
280	1	1	0	0.0556	0.0815
283	1	1	0	0.0556	0.2715
288	2	0	1	0.1111	0.0399

Expected heterozygosity: 0.980
Polymorphic information content (PIC): 0.921
Average exclusion probability (1): 0.738

```
Average exclusion probability (2):      0.849
Null allele frequency estimate:         Not done


**** Locus UAll2 ****

Number of alleles:                      5
Number of individuals typed:            5
  Heterozygotes:                        1
  Homozygotes:                          4
Observed heterozygosity:                0.200

Allele  Count  Heterozygotes  Homozygotes  Frequency  Frequency with null
291     2      0              1            0.2000     0.0000
302     1      1              0            0.1000     0.0399
305     2      0              1            0.2000     0.0399
322     3      1              1            0.3000     0.0399
338     2      0              1            0.2000     0.0399

Expected heterozygosity:                0.867
Polymorphic information content (PIC): 0.745
Average exclusion probability (1):      0.386
Average exclusion probability (2):      0.565
Null allele frequency estimate:         Not done


*************************************************************************
```

- **simBP.txt**

```
Cervus 2.0... Simulation of parentage analysis
© Copyright Tristan Marshall 1998-2001


**** Summary statistics ****

Critical values and success rates (one parent known):

Level       Confidence(%)  Delta Criterion  Tests  Success Rate
Strict      95.00          0.00             9980   100%
Relaxed     80.00          0.00             9980   100%
Unresolved                                  20     0%

Critical values and success rates (neither parent known):

Level       Confidence(%)  Delta Criterion  Tests  Success Rate
Strict      95.00          0.00             9967   100%
Relaxed     80.00          0.00             9967   100%
Unresolved                                  33     0%


**** Files ****

Input
  Allele frequency file:  AlleleFreqBP.alf
    Number of loci:       9
Output
  Summary text file:      simBP.txt
  Simulation data file:   simBP.sim
```

```
**** Loci ****

1     OSM1
2     OSM2
3     CAU14
4     CAU17
5     LIST5
6     LIST9
7     OSM7
8     UAll1
9     UAll2


**** Simulation parameters ****

Input
  Cycles (number of offspring):                10000
  Number of candidate parents:                 2
  Proportion of candidate parents sampled:     1.000
  Proportion of loci typed:                    0.787
  Proportion of loci mistyped:                 0.010
Output
  Relaxed confidence level:                    80.00%
  Strict confidence level:                     95.00%


**** Delta distributions ****

One parent known:

Outcome                   N        Mean Delta   Standard Deviation
True parent most likely   9929     4.41         1.65
Non-parent most likely    51       0.64         0.67
No most likely parent     20
Total                     10000

Neither parent known:

Outcome                   N        Mean Delta   Standard Deviation
True parent most likely   9799     2.74         1.17
Non-parent most likely    168      0.58         0.51
No most likely parent     33
Total                     10000

*******************************************************************************
```

- **offspringBP.csv**

	OSM1a	OSM1b	OSM2a	OSM2b	CAU14a	CAU14b	CAU17a	CAU17b	LIST5a	LIST5b	LIST9a	LIST9b	OSM7a	OSM7b	UAll1a	UAll1b	UAl2a	UAll2b
11	116	116	123	123	143	143	165	175	195	195	209	209	226	226	277	283	0	0
12	115	115	121	141	156	156	165	165	0	0	0	0	0	0	0	0	0	0
13	104	104	123	137	145	156	165	175	192	192	201	201	0	0	288	283	0	0
14	0	0	141	141	145	158	167	167	0	0	202	202	214	214	0	0	0	0
15	112	116	133	133	152	158	171	171	195	195	210	210	221	221	233	247	305	305
16	117	117	124	134	149	159	172	172	197	197	0	0	222	222	249	255	0	0
17	114	119	135	137	154	160	164	173	198	198	0	0	212	212	238	251	291	291
18	109	114	137	137	156	159	165	176	0	0	201	201	211	211	231	259	338	338

- **MalesBP.csv**

	OSM1a	OSM1b	OSM2a	OSM2b	CAU14a	CAU14b	CAU17a	CAU17b	LIST5a	LIST5b	LIST9a	LIST9b	OSM7a	OSM7b	UAII1a	UAII1b	UAII2a	UAII2b
3	100	100	123	123	150	156	165	178	0	0	200	200	211	223	231	280	302	322
7	99	119	0	0	145	155	168	168	193	193	204	210	0	0	0	0	0	0

- **FemalesBP.csv**

	OSM1a	OSM1b	OSM2a	OSM2b	CAU14a	CAU14b	CAU17a	CAU17b	LIST5a	LIST5b	LIST9a	LIST9b	OSM7a	OSM7b	UAII1a	UAII1b	UAII2a	UAII2b
4	115	115	123	123	156	156	165	177	195	195	201	201	214	223	239	268	0	0
8	117	117	122	141	148	156	169	169	196	198	0	0	220	229	235	235	322	322

- **parMBP.csv**

Offspring ID	O loci typed	Prob. non-exclusion	Candidate parent ID	CP loci typed	O-CP loci compared	O-CP loci mismatching	LOD	Delta	Confidence
11	8	8.33E-05	3	8	7	5	-1.03E+00	0.00E+00	
12	4	3.69E-02	3	8	4	2	-5.92E-01	0.00E+00	
13	7	1.68E-03	3	8	6	3	-1.30E+00	0.00E+00	
14	5	1.41E-03	7	5	3	2	-4.40E-01	0.00E+00	
15	9	7.05E-06	7	5	5	4	-8.90E-01	0.00E+00	
16	7	2.22E-05	7	5	4	4	-1.30E+00	0.00E+00	
17	8	1.83E-05	7	5	4	3	-4.90E-01	0.00E+00	
18	8	6.31E-04	7	5	4	4	-2.18E+00	0.00E+00	

- **parMBP.txt**

```
Cervus 2.0... Parentage analysis
© Copyright Tristan Marshall 1998-2001


**** Summary statistics ****

Offspring (total):                              8
  Tested:                                       8
    Known parent typed at 5 or more loci:       0
    Known parent typed at fewer than 5 loci:    8
  Not tested:                                   0
Candidate parents (total):                      2
  Tested (typed at 5 or more loci):             2
  Not tested:                                   0
    Typed at 0-4 loci:                          0
    Not typed:                                  0

Neither parent known:

Level       Confidence(%)  Delta Criterion  Tests       Success Rate
Strict      95.00          0.00             0 (8)       0% (100%)
Relaxed     80.00          0.00             0 (8)       0% (100%)
Unresolved                                  8 (0)       100% (0%)

(expected values predicted by the simulation are shown in brackets)
```

```
**** Files ****

Input
  Offspring file:          offspringBP.csv
  Candidate parent file:   MalesBP.csv
  Genotype data file:      BPgenotypes.csv
  Allele frequency file:   AlleleFreqBP.alf
  Simulation data file:    simBP.sim
Output
  Summary text file:       parMBP.txt
  Parentage data file:     parMBP.csv


**** Loci ****

1     OSM1
2     OSM2
3     CAU14
4     CAU17
5     LIST5
6     LIST9
7     OSM7
8     UAll1
9     UAll2


**** Simulation parameters ****

Input
  Cycles (number of offspring):                10000
  Number of candidate parents:                 2
  Proportion of candidate parents sampled:     1.000
  Proportion of loci typed:                    0.787
  Proportion of loci mistyped:                 0.010
Output
  Relaxed confidence level:                    80.00%
  Strict confidence level:                     95.00%


*******************************************************************************
```

- **parFBP.csv**

Offspring ID	O loci typed	Known parent ID	KP loci typed	KP class	O-KP loci compared	O-KP loci mismatching	Prob. non-exclusion	Candidate parent ID	CP loci typed	O-CP loci compared	O-CP loci mismatching	O-KP-CP loci compared	O-KP-CP loci mismatching	LOD	Delta	Conf
11	8	3	8	Typed	7	5	5.21E-02	4	8	8	5	7	6	8.48E-02	8.48E-02	*
12	4	3	8	Typed	4	2	2.77E-01	4	8	4	1	4	2	2.43E+00	2.43E+00	*
13	7	3	8	Typed	6	3	2.90E-03	4	8	7	3	6	6	-2.98E+00	0.00E+00	
14	5	7	5	Typed	3	2	1.51E-02	8	8	4	3	2	2	-6.55E-01	0.00E+00	
15	9	7	5	Typed	5	4	9.35E-04	4	8	8	7	5	5	-3.01E+00	0.00E+00	
16	7	7	5	Typed	4	4	7.64E-03	8	8	7	6	4	4	-1.36E-01	0.00E+00	
17	8	7	5	Typed	4	3	1.11E-03	8	8	8	7	4	4	-2.73E+00	0.00E+00	
18	8	7	5	Typed	4	4	1.11E-02	4	8	7	4	4	4	-1.52E+00	0.00E+00	

- **parFBP.txt**

```
Cervus 2.0... Parentage analysis
© Copyright Tristan Marshall 1998-2001


**** Summary statistics ****
```

```
Offspring (total):                              8
  Tested:                                       8
    Known parent typed at 5 or more loci:       8
    Known parent typed at fewer than 5 loci:    0
  Not tested:                                   0
Candidate parents (total):                      2
  Tested (typed at 5 or more loci):             2
  Not tested:                                   0
    Typed at 0-4 loci:                          0
    Not typed:                                  0

One parent known:

Level       Confidence(%)  Delta Criterion  Tests       Success Rate
Strict      95.00          0.00             2 (8)       25% (100%)
Relaxed     80.00          0.00             2 (8)       25% (100%)
Unresolved                                  6 (0)       75% (0%)

(expected values predicted by the simulation are shown in brackets)


**** Files ****

Input
  Offspring file:          parMBP.csv
  Candidate parent file:   FemalesBP.csv
  Genotype data file:      BPgenotypes.csv
  Allele frequency file:   AlleleFreqBP.alf
  Simulation data file:    simBP.sim
Output
  Summary text file:       parFBP.txt
  Parentage data file:     parFBP.csv


**** Loci ****

1    OSM1
2    OSM2
3    CAU14
4    CAU17
5    LIST5
6    LIST9
7    OSM7
8    UA111
9    UA112


**** Simulation parameters ****

Input
  Cycles (number of offspring):              10000
  Number of candidate parents:               2
  Proportion of candidate parents sampled:   1.000
  Proportion of loci typed:                  0.787
  Proportion of loci mistyped:               0.010
Output
  Relaxed confidence level:                  80.00%
  Strict confidence level:                   95.00%
```

**** Known parent-offspring mismatches ****

Locus name	Offspring ID	Genotype		Known parent ID	Genotype		Null
OSM1	11	116	116	3	100	100	Y
CAU14	11	143	143	3	150	156	
LIST9	11	209	209	3	200	200	Y
OSM7	11	226	226	3	211	223	
UA111	11	277	283	3	231	280	
OSM1	12	115	115	3	100	100	Y
OSM2	12	121	141	3	123	123	
OSM1	13	104	104	3	100	100	Y
LIST9	13	201	201	3	200	200	Y
UA111	13	288	288	3	231	280	
CAU17	14	167	167	7	168	168	Y
LIST9	14	202	202	7	204	210	
OSM1	15	112	116	7	99	119	
CAU14	15	152	158	7	145	155	
CAU17	15	171	171	7	168	168	Y
LIST5	15	195	195	7	193	193	Y
OSM1	16	117	117	7	99	119	
CAU14	16	149	159	7	145	155	
CAU17	16	172	172	7	168	168	Y
LIST5	16	197	197	7	193	193	Y
CAU14	17	154	160	7	145	155	
CAU17	17	164	173	7	168	168	
LIST5	17	198	198	7	193	193	Y
OSM1	18	109	114	7	99	119	
CAU14	18	156	159	7	145	155	
CAU17	18	165	176	7	168	168	
LIST9	18	201	201	7	204	210	

TOTAL: 27

**** Error rate analysis ****

Locus name	N compared	N mismatching	N null	Detection prob.	Est. error rate
OSM1	7	6	3	0.5978	0.7169
OSM2	3	1	0	0.4696	0.3549
CAU14	8	5	0	0.5388	0.5800
CAU17	8	5	3	0.5870	0.5324
LIST5	3	3	3	0.3929	1.2727
LIST9	5	4	2	0.3929	1.0182
OSM7	1	1	0	0.5955	0.8396
UA111	2	2	0	0.7377	0.6778
UA112	0	0	0	0.3858	0.0000

```
Mean observed error rate across loci: 0.6658
  (assumes all known parent-offspring pairs are equally independent)

********************************************************************************
```

COLONY ANALYSES OUTPUT FILES

- **GenotypesCorrected2.csv**

	OSM1a	OSM1b	OSM2a	OSM2b	CAU14a	CAU14b	CAU17a	CAU17b	LIST5a	LIST5b	LIST9a	LIST9b	OSM7a	OSM7b	UAII1a	UAII1b	UAII2a	UAII2b
35	116	116	122	122	144	156	0	0	0	0	202	202	227	227	0	0	318	342
36	106	112	136	136	156	156	186	166	0	0	204	204	221	235	271	293	355	355
37	112	112	127	127	150	158	0	0	195	195	0	0	213	213	264	285	310	377
38	0	0	0	0	142	142	170	170	0	0	204	204	217	217	0	0	318	342
39	114	114	0	0	148	148	170	170	0	0	0	0	0	0	256	271	291	291
40	112	112	0	0	154	154	0	0	0	0	204	204	217	217	233	290	312	326
41	110	110	0	0	0	0	0	0	0	0	204	204	215	233	271	297	322	322
42	114	114	0	0	142	152	0	0	0	0	200	200	214	214	265	285	326	326
43	109	116	138	138	145	145	169	169	0	0	0	0	218	218	231	274	350	380
44	120	120	126	126	158	158	0	0	0	0	202	202	221	227	271	271	0	0
45	99	99	127	134	142	151	0	0	0	0	208	208	228	228	270	292	320	336
46	0	0	121	127	142	157	167	178	0	0	199	199	220	220	247	269	0	0
47	101	120	126	126	154	154	163	169	193	198	0	0	216	238	258	275	344	355
48	116	116	0	0	143	159	179	179	196	196	0	0	216	216	254	297	320	383
49	109	116	122	130	142	158	171	171	0	0	203	203	220	220	276	294	332	355
50	119	119	126	126	153	159	176	176	195	195	201	201	230	230	254	292	314	334
51	116	116	123	123	142	153	160	169	194	194	199	199	220	220	261	277	300	355
52	116	116	124	124	151	151	163	174	194	194	205	205	220	220	268	281	302	353
53	116	116	122	122	143	153	160	160	197	197	205	205	221	221	273	287	309	309
54	116	116	125	125	0	0	165	165	0	0	199	199	212	212	234	289	313	365
55	115	115	121	121	151	159	169	169	192	192	202	207	214	225	264	291	0	0
56	107	107	122	127	146	157	169	177	0	0	201	201	212	222	234	295	350	350
64	0	0	123	123	156	156	161	169	195	195	202	202	218	218	280	280	302	377
65	109	114	138	138	151	157	169	179	0	0	202	202	219	219	235	255	334	352
66	0	0	0	0	150	157	169	179	192	192	0	0	219	227	268	268	0	0
69	115	115	122	128	147	160	167	179	0	0	202	202	218	218	258	278	389	389
77	115	115	122	122	157	157	161	169	195	195	205	205	222	222	236	236	320	332
78	116	116	0	0	157	157	162	170	195	195	205	205	220	230	272	272	0	0
79	115	115	121	126	147	159	163	171	0	0	202	208	221	221	281	281	302	395
80	109	118	130	130	142	158	173	179	0	0	207	207	218	218	235	281	302	330
81	113	119	0	0	159	159	168	178	0	0	200	200	212	212	0	0	0	0
82	113	120	133	133	156	159	0	0	0	0	207	207	214	214	233	295	0	0
83	107	111	121	127	142	159	169	178	195	195	0	0	0	0	0	0	331	346
84	119	119	125	135	159	159	168	178	0	0	203	208	217	227	269	269	0	0
86	112	119	131	131	152	158	167	176	0	0	205	205	211	230	277	293	328	353
91	111	118	132	132	143	155	160	177	196	196	206	206	213	213	275	295	328	357
92	113	117	135	135	142	153	170	170	0	0	0	0	217	225	255	255	340	368
93	110	110	125	125	157	157	166	175	0	0	208	208	217	217	280	280	300	300
94	105	105	0	0	152	158	166	176	0	0	199	208	213	213	235	296	0	0
95	0	0	0	0	153	157	168	177	193	193	207	207	215	215	0	0	300	300
96	108	116	140	140	160	160	179	179	0	0	206	206	221	221	237	237	0	0
97	108	115	0	0	144	144	171	178	194	194	0	0	220	220	274	274	315	315
98	114	114	121	121	160	160	167	178	0	0	0	0	0	0	250	273	0	0
99	102	110	136	136	144	159	170	170	195	195	206	206	0	0	0	0	366	366
100	115	115	121	129	144	144	168	168	194	194	207	207	0	0	0	0	0	0
101	109	116	141	141	159	159	169	169	0	0	205	205	222	222	277	297	0	0
102	116	116	122	122	0	0	168	168	0	0	203	203	221	221	235	235	0	0
103	109	116	128	136	0	0	169	169	194	194	0	0	0	0	235	235	0	0
104	101	111	0	0	145	159	0	0	0	0	205	205	0	0	0	0	0	0
105	116	116	125	125	0	0	161	161	0	0	0	0	0	0	0	0	0	0
106	101	108	138	138	0	0	0	0	194	194	205	205	0	0	248	263	314	349
107	117	117	125	125	0	0	171	171	0	0	200	200	221	221	0	0	0	0
108	104	110	0	0	0	0	0	0	197	197	0	0	0	0	0	0	307	307
109	118	118	125	125	152	152	165	165	0	0	201	201	220	220	261	283	307	367
110	114	114	135	135	158	158	0	0	0	0	199	199	211	211	234	282	305	365
111	115	118	136	136	142	158	167	176	198	198	0	0	212	212	232	232	314	335
112	115	115	122	141	149	149	164	164	0	0	0	0	211	211	237	237	351	351
113	109	117	0	0	0	0	0	0	196	196	0	0	214	214	232	259	0	0
114	109	116	130	136	0	0	168	168	0	0	0	0	0	0	268	268	0	0
115	108	116	125	133	160	160	0	0	0	0	0	0	0	0	0	0	0	0
116	110	110	0	0	0	0	169	169	0	0	203	203	219	219	0	0	0	0
117	110	110	128	128	0	0	0	0	0	0	200	210	224	224	287	287	0	0
118	116	119	0	0	0	0	0	0	0	0	203	203	215	230	0	0	301	318
119	110	110	0	0	0	0	166	166	0	0	203	203	216	230	270	289	300	312
120	112	117	128	135	160	160	167	167	0	0	201	209	0	0	270	287	317	338
121	116	116	129	135	0	0	161	161	0	0	201	201	228	228	265	265	306	306
122	111	116	136	136	155	155	165	175	0	0	202	202	215	225	268	286	0	0
123	104	109	0	0	158	158	166	166	0	0	0	0	0	0	264	264	301	301
124	0	0	141	141	156	156	164	174	197	197	0	0	0	0	0	0	0	0
125	0	0	0	0	0	0	162	162	193	193	0	0	215	215	241	258	0	0
126	101	114	0	0	159	159	168	168	192	192	0	0	0	0	0	0	318	318
127	0	0	0	0	142	159	168	168	0	0	0	0	217	217	270	270	312	312
128	115	115	121	121	148	155	169	169	0	0	0	0	216	216	271	292	310	331
129	108	116	0	0	158	158	0	0	192	192	205	205	221	221	268	268	329	329
130	104	104	0	0	156	156	165	173	0	0	0	0	228	228	245	245	306	306
131	108	112	125	130	157	157	164	164	195	195	0	0	211	211	263	263	306	306
132	108	112	132	132	151	160	170	170	195	195	208	208	0	0	283	283	304	304
133	99	105	124	124	151	151	166	166	0	0	0	0	0	0	0	0	0	0
134	0	0	0	0	154	154	164	174	0	0	0	0	225	225	0	0	0	0
135	0	0	0	0	155	155	171	171	0	0	0	0	0	0	0	0	0	0
136	0	0	0	0	0	0	0	0	0	0	0	0	225	225	0	0	322	344
137	0	0	121	130	144	160	169	169	0	0	202	202	228	228	289	289	360	360
138	0	0	122	122	0	0	0	0	0	0	201	201	212	225	282	298	338	338
139	0	0	121	126	0	0	167	167	0	0	202	202	215	0	270	270	314	324
140	118	118	131	136	155	155	166	177	0	0	201	201	0	0	268	287	310	327
141	112	119	138	138	155	155	166	176	0	0	199	199	0	0	270	270	320	320
142	113	113	128	128	142	157	166	175	0	0	0	0	0	0	272	287	318	340
143	0	0	0	0	155	155	165	165	0	0	210	210	229	229	272	272	318	340
144	0	0	0	0	155	155	168	168	192	192	0	0	0	0	0	0	0	0
145	107	107	0	0	153	153	165	165	0	0	0	0	0	0	0	0	312	335
146	0	0	0	0	0	0	0	0	0	0	0	0	0	0	271	271	316	316
147	0	0	127	127	156	156	167	167	0	0	199	199	221	221	260	275	302	338
148	0	0	122	122	152	152	163	179	0	0	207	207	221	221	240	240	313	334
149	0	0	124	124	0	0	161	170	0	0	0	0	216	216	261	261	316	316

- **AlleleFreq.txt**

```
Cervus 2.0... Allele Frequency Analysis
© Copyright Tristan Marshall 1998-2001


**** Summary statistics ****

Number of loci:                          9
Number of individuals:                   97

Locus       k   N     Hets  Homs  H(O)   H(E)   PIC    Excl(1)  Excl(2)
HW   Null freq
OSM1        20  75    34    41    0.453  0.917  0.906  0.703    0.825
NA   +0.3392
OSM2        19  63    19    44    0.302  0.936  0.924  0.747    0.855
NA   +0.5091
CAU14       19  74    32    42    0.432  0.930  0.919  0.732    0.846
NA   +0.3619
CAU17       19  74    33    41    0.446  0.932  0.921  0.740    0.851
NA   +0.3523
LIST5       7   32    1     31    0.031  0.831  0.794  0.466    0.641
NA   +0.9269
LIST9       12  59    6     53    0.102  0.903  0.885  0.647    0.787
NA   +0.7966
OSM7        21  70    15    55    0.214  0.936  0.926  0.753    0.859
NA   +0.6266
UAl11       50  75    42    33    0.560  0.977  0.969  0.887    0.940
NA   +0.2693
UAl12       52  66    38    28    0.576  0.981  0.972  0.898    0.946
NA   +0.2567

Mean number of alleles per locus:          24.33
Mean proportion of individuals typed:      0.674
Mean expected heterozygosity:              0.927
Mean PIC:                                  0.913
Total exclusionary power (first parent):   0.999997
Total exclusionary power (second parent):  1.000000


**** Files ****

Input
  Genotype data file:     GenotypesCorrected2.csv
Output
  Summary text file:      AlleleFreq.txt
  Allele frequency file:  AlleleFreq.alf


**** Loci ****

1     OSM1
2     OSM2
3     CAU14
4     CAU17
5     LIST5
6     LIST9
```

7 OSM7
8 UA111
9 UA112

**** Locus OSM1 ****

Number of alleles: 20
Number of individuals typed: 75
Heterozygotes: 34
Homozygotes: 41
Observed heterozygosity: 0.453

Allele	Count	Heterozygotes	Homozygotes	Frequency	Frequency with null
99	3	1	1	0.0200	0.0119
101	4	4	0	0.0267	0.0239
102	1	1	0	0.0067	0.0059
104	4	2	1	0.0267	0.0179
105	3	1	1	0.0200	0.0119
106	1	1	0	0.0067	0.0059
107	5	1	2	0.0333	0.0179
108	7	7	0	0.0467	0.0422
109	9	9	0	0.0600	0.0546
110	12	2	5	0.0800	0.0422
111	4	4	0	0.0267	0.0239
112	10	6	2	0.0667	0.0484
113	5	3	1	0.0333	0.0239
114	10	2	4	0.0667	0.0360
115	17	3	7	0.1133	0.0608
116	31	9	11	0.2067	0.1259
117	5	3	1	0.0333	0.0239
118	7	3	2	0.0467	0.0299
119	8	4	2	0.0533	0.0360
120	4	2	1	0.0267	0.0179

Expected heterozygosity: 0.917
Polymorphic information content (PIC): 0.906
Average exclusion probability (1): 0.703
Average exclusion probability (2): 0.825
Hardy-Weinberg equilibrium test: Not done
Null allele frequency estimate: 0.3392

**** Locus OSM2 ****

Number of alleles: 19
Number of individuals typed: 63
Heterozygotes: 19
Homozygotes: 44
Observed heterozygosity: 0.302

Allele	Count	Heterozygotes	Homozygotes	Frequency	Frequency with null
121	12	6	3	0.0952	0.0544
122	16	4	6	0.1270	0.0606
123	4	0	2	0.0317	0.0118
124	6	0	3	0.0476	0.0178
125	13	3	5	0.1032	0.0482
126	8	2	3	0.0635	0.0298
127	8	4	2	0.0635	0.0359
128	7	3	2	0.0556	0.0298

129	2	2	0	0.0159	0.0118
130	6	4	1	0.0476	0.0298
131	3	1	1	0.0238	0.0118
132	4	0	2	0.0317	0.0118
133	3	1	1	0.0238	0.0118
134	1	1	0	0.0079	0.0059
135	7	3	2	0.0556	0.0298
136	11	3	4	0.0873	0.0420
138	8	0	4	0.0635	0.0238
140	2	0	1	0.0159	0.0059
141	5	1	2	0.0397	0.0178

Expected heterozygosity: 0.936
Polymorphic information content (PIC): 0.924
Average exclusion probability (1): 0.747
Average exclusion probability (2): 0.855
Hardy-Weinberg equilibrium test: Not done
Null allele frequency estimate: 0.5091

**** Locus CAU14 ****

Number of alleles: 19
Number of individuals typed: 74
  Heterozygotes: 32
  Homozygotes: 42
Observed heterozygosity: 0.432

Allele	Count	Heterozygotes	Homozygotes	Frequency	Frequency with null
142	13	11	1	0.0878	0.0731
143	3	3	0	0.0203	0.0178
144	7	3	2	0.0473	0.0298
145	3	1	1	0.0203	0.0118
146	1	1	0	0.0068	0.0059
147	2	2	0	0.0135	0.0118
148	3	1	1	0.0203	0.0118
149	2	0	1	0.0135	0.0059
150	2	2	0	0.0135	0.0118
151	8	4	2	0.0541	0.0359
152	7	3	2	0.0473	0.0298
153	7	5	1	0.0473	0.0359
154	6	0	3	0.0405	0.0178
155	14	2	6	0.0946	0.0481
156	13	3	5	0.0878	0.0481
157	14	6	4	0.0946	0.0605
158	13	5	4	0.0878	0.0543
159	19	9	5	0.1284	0.0859
160	11	3	4	0.0743	0.0420

Expected heterozygosity: 0.930
Polymorphic information content (PIC): 0.919
Average exclusion probability (1): 0.732
Average exclusion probability (2): 0.846
Hardy-Weinberg equilibrium test: Not done
Null allele frequency estimate: 0.3619

**** Locus CAU17 ****

Number of alleles: 19

Number of individuals typed: 74
  Heterozygotes: 33
  Homozygotes: 41
Observed heterozygosity: 0.446

Allele	Count	Heterozygotes	Homozygotes	Frequency	Frequency with null
160	4	2	1	0.0270	0.0179
161	7	3	2	0.0473	0.0300
162	3	1	1	0.0203	0.0119
163	4	4	0	0.0270	0.0240
164	6	2	2	0.0405	0.0240
165	10	2	4	0.0676	0.0362
166	13	5	4	0.0878	0.0548
167	11	5	3	0.0743	0.0485
168	15	3	6	0.1014	0.0548
169	22	8	7	0.1486	0.0931
170	12	2	5	0.0811	0.0423
171	8	2	3	0.0541	0.0300
173	2	2	0	0.0135	0.0119
174	3	3	0	0.0203	0.0179
175	3	3	0	0.0203	0.0179
176	6	4	1	0.0405	0.0300
177	4	4	0	0.0270	0.0240
178	6	6	0	0.0405	0.0362
179	9	5	2	0.0608	0.0423

Expected heterozygosity: 0.932
Polymorphic information content (PIC): 0.921
Average exclusion probability (1): 0.740
Average exclusion probability (2): 0.851
Hardy-Weinberg equilibrium test: Not done
Null allele frequency estimate: 0.3523

**** Locus LIST5 ****

Number of alleles: 7
Number of individuals typed: 32
  Heterozygotes: 1
  Homozygotes: 31
Observed heterozygosity: 0.031

Allele	Count	Heterozygotes	Homozygotes	Frequency	Frequency with null
192	12	0	6	0.1875	0.0133
193	5	1	2	0.0781	0.0066
194	12	0	6	0.1875	0.0133
195	18	0	9	0.2813	0.0200
196	6	0	3	0.0938	0.0066
197	8	0	4	0.1250	0.0088
198	3	1	1	0.0469	0.0044

Expected heterozygosity: 0.831
Polymorphic information content (PIC): 0.794
Average exclusion probability (1): 0.466
Average exclusion probability (2): 0.641
Hardy-Weinberg equilibrium test: Not done
Null allele frequency estimate: 0.9269

**** Locus LIST9 ****

Number of alleles:	12
Number of individuals typed:	59
Heterozygotes:	6
Homozygotes:	53
Observed heterozygosity:	0.102

Allele	Count	Heterozygotes	Homozygotes	Frequency	Frequency with null
199	13	1	6	0.1102	0.0219
200	7	1	3	0.0593	0.0125
201	13	1	6	0.1102	0.0219
202	18	2	8	0.1525	0.0315
203	11	1	5	0.0932	0.0188
204	8	0	4	0.0678	0.0125
205	18	0	9	0.1525	0.0283
206	7	1	3	0.0593	0.0125
207	11	1	5	0.0932	0.0188
208	8	2	3	0.0678	0.0156
209	1	1	0	0.0085	0.0031
210	3	1	1	0.0254	0.0062

Expected heterozygosity:	0.903
Polymorphic information content (PIC):	0.885
Average exclusion probability (1):	0.647
Average exclusion probability (2):	0.787
Hardy-Weinberg equilibrium test:	Not done
Null allele frequency estimate:	0.7966

**** Locus OSM7 ****

Number of alleles:	21
Number of individuals typed:	70
Heterozygotes:	15
Homozygotes:	55
Observed heterozygosity:	0.214

Allele	Count	Heterozygotes	Homozygotes	Frequency	Frequency with null
211	7	1	3	0.0500	0.0175
212	8	2	3	0.0571	0.0219
213	6	0	3	0.0429	0.0131
214	7	1	3	0.0500	0.0175
215	7	3	2	0.0500	0.0219
216	8	2	3	0.0571	0.0219
217	10	2	4	0.0714	0.0264
218	8	0	4	0.0571	0.0175
219	5	1	2	0.0357	0.0131
220	13	1	6	0.0929	0.0308
221	22	2	10	0.1571	0.0535
222	5	1	2	0.0357	0.0131
224	2	0	1	0.0143	0.0043
225	8	4	2	0.0571	0.0264
227	5	3	1	0.0357	0.0175
228	8	0	4	0.0571	0.0175
229	2	0	1	0.0143	0.0043
230	6	4	1	0.0429	0.0219
233	1	1	0	0.0071	0.0043
235	1	1	0	0.0071	0.0043
238	1	1	0	0.0071	0.0043

Expected heterozygosity: 0.936
Polymorphic information content (PIC): 0.926
Average exclusion probability (1): 0.753
Average exclusion probability (2): 0.859
Hardy-Weinberg equilibrium test: Not done
Null allele frequency estimate: 0.6266

**** Locus UAl11 ****

Number of alleles: 50
Number of individuals typed: 75
  Heterozygotes: 42
  Homozygotes: 33
Observed heterozygosity: 0.560

Allele	Count	Heterozygotes	Homozygotes	Frequency	Frequency with null
231	1	1	0	0.0067	0.0062
232	3	1	1	0.0200	0.0124
233	2	2	0	0.0133	0.0124
234	3	3	0	0.0200	0.0187
235	7	3	2	0.0467	0.0314
236	2	0	1	0.0133	0.0062
237	6	0	3	0.0400	0.0187
240	2	0	1	0.0133	0.0062
241	1	1	0	0.0067	0.0062
245	2	0	1	0.0133	0.0062
247	1	1	0	0.0067	0.0062
248	1	1	0	0.0067	0.0062
250	1	1	0	0.0067	0.0062
254	2	2	0	0.0133	0.0124
255	3	1	1	0.0200	0.0124
256	1	1	0	0.0067	0.0062
258	3	3	0	0.0200	0.0187
259	1	1	0	0.0067	0.0062
260	1	1	0	0.0067	0.0062
261	5	3	1	0.0333	0.0250
263	3	1	1	0.0200	0.0124
264	4	2	1	0.0267	0.0187
265	3	1	1	0.0200	0.0124
268	9	3	3	0.0600	0.0378
269	3	1	1	0.0200	0.0124
270	9	3	3	0.0600	0.0378
271	8	4	2	0.0533	0.0378
272	7	1	3	0.0467	0.0250
273	2	2	0	0.0133	0.0124
274	3	1	1	0.0200	0.0124
275	3	3	0	0.0200	0.0187
276	1	1	0	0.0067	0.0062
277	3	3	0	0.0200	0.0187
278	1	1	0	0.0067	0.0062
280	4	0	2	0.0267	0.0124
281	4	2	1	0.0267	0.0187
282	2	2	0	0.0133	0.0124
283	3	1	1	0.0200	0.0124
285	2	2	0	0.0133	0.0124
286	1	1	0	0.0067	0.0062
287	6	4	1	0.0400	0.0314
289	4	2	1	0.0267	0.0187
290	1	1	0	0.0067	0.0062

291	1	1	0	0.0067	0.0062
292	3	3	0	0.0200	0.0187
293	2	2	0	0.0133	0.0124
294	1	1	0	0.0067	0.0062
295	3	3	0	0.0200	0.0187
297	3	3	0	0.0200	0.0187
298	3	3	0	0.0200	0.0187

Expected heterozygosity: 0.977
Polymorphic information content (PIC): 0.969
Average exclusion probability (1): 0.887
Average exclusion probability (2): 0.940
Hardy-Weinberg equilibrium test: Not done
Null allele frequency estimate: 0.2693

**** Locus UAll2 ****

Number of alleles: 52
Number of individuals typed: 66
  Heterozygotes: 38
  Homozygotes: 28
Observed heterozygosity: 0.576

Allele	Count	Heterozygotes	Homozygotes	Frequency	Frequency with null
291	2	0	1	0.0152	0.0071
300	6	2	2	0.0455	0.0287
301	3	1	1	0.0227	0.0143
302	5	5	0	0.0379	0.0360
304	2	0	1	0.0152	0.0071
305	1	1	0	0.0076	0.0071
306	6	0	3	0.0455	0.0215
307	3	1	1	0.0227	0.0143
309	2	0	1	0.0152	0.0071
310	3	3	0	0.0227	0.0215
312	5	3	1	0.0379	0.0287
313	2	2	0	0.0152	0.0143
314	4	4	0	0.0303	0.0287
315	2	0	1	0.0152	0.0071
316	4	0	2	0.0303	0.0143
317	1	1	0	0.0076	0.0071
318	7	5	1	0.0530	0.0434
320	5	3	1	0.0379	0.0287
322	3	1	1	0.0227	0.0143
324	1	1	0	0.0076	0.0071
326	3	1	1	0.0227	0.0143
327	1	1	0	0.0076	0.0071
328	2	2	0	0.0152	0.0143
329	2	0	1	0.0152	0.0071
330	1	1	0	0.0076	0.0071
331	2	2	0	0.0152	0.0143
332	2	2	0	0.0152	0.0143
334	3	3	0	0.0227	0.0215
335	2	2	0	0.0152	0.0143
336	1	1	0	0.0076	0.0071
338	6	2	2	0.0455	0.0287
340	3	3	0	0.0227	0.0215
342	2	2	0	0.0152	0.0143
344	2	2	0	0.0152	0.0143
346	1	1	0	0.0076	0.0071

349	1	1	0	0.0076	0.0071
350	5	1	2	0.0379	0.0215
351	2	0	1	0.0152	0.0071
352	1	1	0	0.0076	0.0071
353	2	2	0	0.0152	0.0143
355	5	3	1	0.0379	0.0287
357	1	1	0	0.0076	0.0071
360	2	0	1	0.0152	0.0071
365	2	2	0	0.0152	0.0143
366	2	0	1	0.0152	0.0071
367	1	1	0	0.0076	0.0071
368	1	1	0	0.0076	0.0071
377	2	2	0	0.0152	0.0143
380	1	1	0	0.0076	0.0071
383	1	1	0	0.0076	0.0071
389	2	0	1	0.0152	0.0071
395	1	1	0	0.0076	0.0071

```
Expected heterozygosity:                0.981
Polymorphic information content (PIC): 0.972
Average exclusion probability (1):      0.898
Average exclusion probability (2):      0.946
Hardy-Weinberg equilibrium test:        Not done
Null allele frequency estimate:         0.2567
```

****

- **sim4M.txt**

```
Cervus 2.0... Simulation of parentage analysis
© Copyright Tristan Marshall 1998-2001


**** Summary statistics ****

Critical values and success rates (one parent known):

Level       Confidence(%)  Delta Criterion  Tests  Success Rate
Strict      95.00          0.25             9341   93%
Relaxed     80.00          0.00             9978   100%
Unresolved                                  22     0%

Critical values and success rates (neither parent known):

Level       Confidence(%)  Delta Criterion  Tests  Success Rate
Strict      95.00          0.45             7525   75%
Relaxed     80.00          0.00             9981   100%
Unresolved                                  19     0%


**** Files ****

Input
  Allele frequency file:  af4.alf
    Number of loci:       9
Output
  Summary text file:      sim4M.txt
```

```
  Simulation data file:    sim4M.sim


**** Loci ****

1     OSM1
2     OSM2
3     CAU14
4     CAU17
5     LIST5
6     LIST9
7     OSM7
8     UA111
9     UA112


**** Simulation parameters ****

Input
  Cycles (number of offspring):                10000
  Number of candidate parents:                 6
  Proportion of candidate parents sampled:     1.000
  Proportion of loci typed:                    0.674
  Proportion of loci mistyped:                 0.010
Output
  Relaxed confidence level:                    80.00%
  Strict confidence level:                     95.00%


**** Delta distributions ****

One parent known:

Outcome                  N        Mean Delta  Standard Deviation
True parent most likely  9201     2.11        1.23
Non-parent most likely   778      0.45        0.40
No most likely parent    21
Total                    10000

Neither parent known:

Outcome                  N        Mean Delta  Standard Deviation
True parent most likely  8662     1.22        0.78
Non-parent most likely   1319     0.34        0.31
No most likely parent    19
Total                    10000


*******************************************************************************
```

- **sim4F.txt**

```
Cervus 2.0... Simulation of parentage analysis
© Copyright Tristan Marshall 1998-2001

**** Summary statistics ****

Critical values and success rates (one parent known):
```

Level	Confidence(%)	Delta Criterion	Tests	Success Rate
Strict	95.00	0.57	7572	76%
Relaxed	80.00	0.00	9994	100%
Unresolved			6	0%

Critical values and success rates (neither parent known):

Level	Confidence(%)	Delta Criterion	Tests	Success Rate
Strict	95.00	0.78	4406	44%
Relaxed	80.00	0.11	8889	89%
Unresolved			1111	11%

**** Files ****

Input
  Allele frequency file: af4.alf
    Number of loci: 9
Output
  Summary text file: sim4F.txt
  Simulation data file: sim4F.sim

**** Loci ****

1 OSM1
2 OSM2
3 CAU14
4 CAU17
5 LIST5
6 LIST9
7 OSM7
8 UA111
9 UA112

**** Simulation parameters ****

Input
  Cycles (number of offspring): 10000
  Number of candidate parents: 14
  Proportion of candidate parents sampled: 1.000
  Proportion of loci typed: 0.674
  Proportion of loci mistyped: 0.010
Output
  Relaxed confidence level: 80.00%
  Strict confidence level: 95.00%

**** Delta distributions ****

One parent known:

Outcome	N	Mean Delta	Standard Deviation
True parent most likely	8535	1.77	1.15
Non-parent most likely	1459	0.43	0.43
No most likely parent	6		
Total	10000		

Neither parent known:

Outcome	N	Mean Delta	Standard Deviation
True parent most likely	7525	0.99	0.70
Non-parent most likely	2473	0.32	0.31
No most likely parent	2		
Total	10000		

****

- **Offspring.csv**

	OSM1a	OSM1b	OSM2a	OSM2b	CAU14a	CAU14b	CAU17a	CAU17b	LIST5a	LIST5b	LIST9a	LIST9b	OSM7a	OSM7b	UAII1a	UAII1b	UAII2a	UAII2b
55	115	115	121	121	151	159	169	169	192	192	202	207	214	225	264	291	0	0
56	107	107	122	127	146	157	169	177	0	0	201	201	212	222	234	295	350	350
64	0	0	123	123	156	156	161	169	195	195	202	202	218	218	280	280	302	377
65	109	114	138	138	151	157	169	179	0	0	202	202	219	219	235	255	334	352
66	0	0	0	0	150	157	169	179	192	192	0	0	219	227	268	268	0	0
69	115	115	122	128	147	160	167	179	0	0	202	202	218	218	258	278	389	389
77	115	115	122	122	157	157	161	169	195	195	205	205	222	222	236	236	320	332
78	116	116	0	0	157	157	162	170	195	195	205	205	220	230	272	272	0	0
79	115	115	121	126	147	159	163	171	0	0	202	208	221	221	281	281	302	395
80	109	118	130	130	142	158	173	179	0	0	207	207	218	218	235	281	302	330
81	113	119	0	0	159	159	168	178	0	0	200	200	212	212	0	0	0	0
82	113	120	133	133	156	159	0	0	0	0	207	207	214	214	233	295	0	0
83	107	111	121	127	142	159	169	178	195	195	0	0	0	0	0	0	331	346
84	119	119	125	135	159	159	168	178	0	0	203	208	217	227	269	269	0	0
86	112	119	131	131	152	158	167	176	0	0	205	205	211	230	277	293	326	353
91	111	118	132	132	143	155	160	177	196	196	206	206	213	213	275	295	328	357
92	113	117	135	135	142	163	170	170	0	0	0	0	217	225	255	255	340	368
93	110	110	125	125	157	157	166	175	0	0	208	208	217	217	280	280	300	300
94	105	105	0	0	152	156	166	176	0	0	199	206	213	213	235	298	0	0
95	0	0	0	0	153	157	168	177	193	193	207	207	215	215	0	0	300	300
96	108	115	140	140	160	160	179	179	0	0	206	206	221	221	237	237	0	0
97	108	115	0	0	144	144	171	178	194	194	0	0	220	220	274	274	315	315
98	114	114	121	121	160	160	167	178	0	0	0	0	0	0	250	273	0	0
99	102	110	136	136	144	159	170	170	195	195	206	206	0	0	0	0	366	366
100	115	115	121	129	144	144	168	168	194	194	207	207	0	0	0	0	0	0
101	109	116	141	141	159	159	169	169	0	0	205	205	222	222	277	297	0	0
102	116	116	122	122	0	0	168	168	0	0	203	203	221	221	235	235	0	0
103	109	116	128	136	0	0	169	169	194	194	0	0	0	0	235	235	0	0
104	101	111	0	0	145	159	0	0	0	0	205	205	0	0	0	0	0	0
105	116	116	125	125	0	0	161	161	0	0	0	0	0	0	0	0	0	0
106	101	108	138	138	0	0	0	0	194	194	205	205	0	0	248	263	314	349
107	117	117	125	125	0	0	171	171	0	0	200	200	221	221	0	0	0	0
108	104	110	0	0	0	0	0	0	197	197	0	0	0	0	0	0	307	307
109	118	118	125	125	152	152	165	165	0	0	201	201	220	220	261	283	307	367
110	114	114	135	135	158	158	0	0	0	0	199	199	211	211	234	282	305	365
111	115	118	136	136	142	158	167	176	198	198	0	0	212	212	232	232	314	335
112	115	115	122	141	149	149	164	164	0	0	0	0	211	211	237	237	351	351
113	109	117	0	0	0	0	0	0	196	196	0	0	214	214	232	259	0	0
114	109	116	130	136	0	0	168	168	0	0	0	0	0	0	268	268	0	0
115	108	116	125	133	160	160	0	0	0	0	0	0	0	0	0	0	0	0
116	110	110	0	0	0	0	169	169	0	0	203	203	219	219	0	0	0	0
117	110	110	128	128	0	0	0	0	0	0	200	210	224	224	287	287	0	0
118	116	119	0	0	0	0	0	0	0	0	203	203	215	230	0	0	301	318
119	110	110	0	0	0	0	166	166	0	0	203	203	216	230	270	289	300	312
120	112	117	128	135	160	160	167	167	0	0	201	209	0	0	270	287	317	338
121	116	116	129	135	0	0	161	161	0	0	201	201	228	228	265	265	306	306
122	111	116	136	136	155	155	165	175	0	0	202	202	215	225	268	286	0	0
123	104	109	0	0	158	158	166	166	0	0	0	0	0	0	264	264	301	301
124	0	0	141	141	156	156	164	174	197	197	0	0	0	0	0	0	0	0
125	0	0	0	0	0	0	162	162	193	193	0	0	215	215	241	258	0	0
126	101	114	0	0	159	159	168	168	192	192	0	0	0	0	0	0	318	318
127	0	0	0	0	142	159	168	168	0	0	0	0	217	217	270	270	312	312
128	115	115	121	121	148	155	169	169	0	0	0	0	216	216	271	292	310	331
129	108	116	0	0	158	158	0	0	192	192	205	205	221	221	268	268	329	329
130	104	104	0	0	156	156	165	173	0	0	0	0	228	228	245	245	306	306
131	108	112	125	130	157	157	164	164	195	195	0	0	211	211	263	263	306	306
132	108	112	132	132	151	160	170	170	195	195	208	208	0	0	283	283	304	304
133	99	105	124	124	151	151	166	166	0	0	0	0	0	0	0	0	0	0
134	0	0	0	0	154	154	164	174	0	0	0	0	225	225	0	0	0	0
135	0	0	0	0	155	155	171	171	0	0	0	0	0	0	0	0	0	0
136	0	0	0	0	0	0	0	0	0	0	0	0	225	225	0	0	322	344
137	0	0	121	130	144	160	169	169	0	0	202	202	228	228	289	289	360	360
138	0	0	122	122	0	0	0	0	0	0	201	201	212	225	282	298	338	338
139	0	0	121	126	0	0	167	167	0	0	202	202	215	0	270	270	314	324
140	118	118	131	136	155	155	166	177	0	0	201	201	0	0	268	287	310	327
141	112	119	138	138	155	155	166	176	0	0	199	199	0	0	270	270	320	320
142	113	113	128	128	142	157	166	175	0	0	0	0	0	0	272	287	318	340
143	0	0	0	0	155	155	165	165	0	0	210	210	229	229	272	272	318	340
144	0	0	0	0	155	155	168	168	192	192	0	0	0	0	0	0	0	0
145	107	107	0	0	153	153	165	165	0	0	0	0	0	0	0	0	312	335
146	0	0	0	0	0	0	0	0	0	0	0	0	0	0	271	271	316	316
147	0	0	127	127	156	156	167	167	0	0	199	199	221	221	260	275	302	338
148	0	0	122	122	152	152	163	179	0	0	207	207	221	221	240	240	313	334
149	0	0	124	124	0	0	161	170	0	0	0	0	216	216	261	261	316	316
150	116	116	0	0	159	159	0	0	192	192	0	0	221	221	261	298	338	338
151	0	0	0	0	0	0	0	0	197	197	0	0	221	221	272	272	350	350
152	0	0	0	0	0	0	0	0	0	0	0	0	0	0	237	237	0	0

- **ParentsMale.csv**

	OSM1a	OSM1b	OSM2a	OSM2b	CAU14a	CAU14b	CAU17a	CAU17b	LIST5a	LIST5b	LIST9a	LIST9b	OSM7a	OSM7b	UAII1a	UAII1b	UAII2a	UAII2b
48	116	116	0	0	143	159	179	179	196	196	0	0	216	216	254	297	320	383
49	109	116	122	130	142	158	171	171	0	0	203	203	220	220	276	294	332	355
50	119	119	126	126	153	159	176	176	195	195	201	201	230	230	254	292	314	334
51	116	116	123	123	142	153	160	169	194	194	199	199	220	220	261	277	300	355
53	116	116	122	122	143	153	160	160	197	197	205	205	221	221	273	287	309	309
54	116	116	125	125	0	0	165	165	0	0	199	199	212	212	234	289	313	365

- **ParentsFemale.csv**

	OSM1a	OSM1b	OSM2a	OSM2b	CAU14a	CAU14b	CAU17a	CAU17b	LIST5a	LIST5b	LIST9a	LIST9b	OSM7a	OSM7b	UAII1a	UAII1b	UAII2a	UAII2b
35	116	116	122	122	144	156	0	0	0	0	202	202	227	227	0	0	318	342
36	106	112	136	136	156	156	166	166	0	0	204	204	221	235	271	293	355	355
37	112	112	127	127	150	158	0	0	195	195	0	0	213	213	264	285	310	377
38	0	0	0	0	142	142	170	170	0	0	204	204	217	217	0	0	318	342
39	114	114	0	0	148	148	170	170	0	0	0	0	0	0	256	271	291	291
40	112	112	0	0	154	154	0	0	0	0	204	204	217	217	233	290	312	326
41	110	110	0	0	0	0	0	0	0	0	204	204	215	233	271	297	322	322
42	114	114	0	0	142	152	0	0	0	0	200	200	214	214	265	285	326	326
43	109	116	138	138	145	145	169	169	0	0	0	0	218	218	231	274	350	380
44	120	120	126	126	158	158	0	0	0	0	202	202	221	227	271	271	0	0
45	99	99	127	134	142	151	0	0	0	0	208	208	228	228	270	292	320	336
46	0	0	121	127	142	157	167	178	0	0	199	199	220	220	247	269	0	0
47	101	120	126	126	154	154	163	169	193	198	0	0	216	238	258	275	344	355
52	116	116	124	124	151	151	163	174	194	194	205	205	220	220	268	281	302	353

- **parM.txt**

```
Cervus 2.0... Parentage analysis
© Copyright Tristan Marshall 1998-2001


**** Summary statistics ****

Offspring (total):                              77
  Tested:                                       77
    Known parent typed at 5 or more loci:       0
    Known parent typed at fewer than 5 loci:    77
  Not tested:                                   0
Candidate parents (total):                      6
  Tested (typed at 5 or more loci):             6
  Not tested:                                   0
    Typed at 0-4 loci:                          0
    Not typed:                                  0

Neither parent known:

Level       Confidence(%)  Delta Criterion  Tests        Success Rate
Strict      95.00          0.45             33 (58)      43% (75%)
Relaxed     80.00          0.00             61 (77)      79% (100%)
Unresolved                                  16 (0)       21% (0%)

(expected values predicted by the simulation are shown in brackets)


**** Files ****

Input
  Offspring file:          Offspring.csv
  Candidate parent file:   ParentsMale.csv
  Genotype data file:      GenotypesCorrected2.csv
  Allele frequency file:   af4.alf
  Simulation data file:    sim4M.sim
Output
  Summary text file:       parM.txt
  Parentage data file:     parM.csv
```

```
**** Loci ****

1     OSM1
2     OSM2
3     CAU14
4     CAU17
5     LIST5
6     LIST9
7     OSM7
8     UA111
9     UA112


**** Simulation parameters ****

Input
  Cycles (number of offspring):                10000
  Number of candidate parents:                 6
  Proportion of candidate parents sampled:     1.000
  Proportion of loci typed:                    0.674
  Proportion of loci mistyped:                 0.010
Output
  Relaxed confidence level:                    80.00%
  Strict confidence level:                     95.00%


*******************************************************************************
```

- **parM.csv**

Offspring ID	O loci typed	Prob. non-exclusion	Candidate parent ID	CP loci typed	O-CP loci compared	O-CP loci mismatching	LOD	Delta	Confidence
55	8	6.89E-06	48	7	6	5	-1.18E+00	0.00E+00	
56	8	2.99E-07	50	9	8	7	8.19E-01	8.19E-01	*
64	8	3.05E-07	50	9	8	7	1.88E-01	1.88E-01	+
65	8	4.58E-07	48	7	6	5	1.40E-01	1.40E-01	+
66	5	4.18E-04	48	7	5	4	2.24E-01	2.24E-01	+
69	8	1.49E-07	48	7	6	5	-7.56E-02	0.00E+00	
77	9	8.47E-08	53	9	9	7	1.01E+00	1.01E+00	*
78	7	4.04E-05	53	9	7	5	1.30E+00	6.73E-01	*
79	8	1.36E-06	50	9	8	6	1.06E-01	1.06E-01	+
80	8	2.13E-07	49	8	8	5	1.86E+00	1.86E+00	*
81	5	1.34E-04	50	9	5	3	8.19E-01	6.73E-01	*
82	6	2.40E-06	48	7	4	3	1.97E-02	1.97E-02	+
83	6	9.76E-05	50	9	6	4	1.12E+00	1.12E+00	*
84	7	4.57E-06	50	9	7	5	6.56E-01	5.91E-01	*
86	8	1.13E-07	50	9	8	5	6.44E-01	1.19E-01	+
91	9	1.22E-09	48	7	7	5	1.60E+00	1.60E+00	*
92	7	3.10E-07	50	9	7	6	4.77E-01	1.95E-01	+
93	8	9.35E-08	54	7	7	6	4.43E-01	4.43E-01	+
94	6	8.04E-06	54	7	5	4	5.14E-01	3.31E-01	+
95	6	1.46E-05	51	9	6	4	1.03E-01	1.03E-01	+
96	7	4.09E-07	53	9	7	6	4.72E-01	4.72E-01	*
97	7	3.51E-07	51	9	7	5	1.63E+00	6.37E-01	*
98	5	2.87E-05	50	9	5	5	-4.83E-01	0.00E+00	
99	7	2.31E-06	50	9	7	5	9.87E-01	9.87E-01	*
100	6	4.82E-05	51	9	6	5	3.21E-02	3.21E-02	+
101	7	3.69E-06	53	9	7	6	1.58E+00	6.94E-01	*
102	6	7.96E-05	53	9	6	3	3.40E+00	1.51E+00	*
103	5	1.05E-03	51	9	5	2	2.36E+00	1.67E+00	*
104	3	8.05E-03	53	9	3	2	1.15E+00	1.15E+00	*
105	3	6.70E-03	54	7	3	1	2.39E+00	1.18E+00	*
106	6	6.51E-06	51	9	6	5	5.92E-01	2.92E-01	+
107	5	4.50E-05	54	7	5	4	6.92E-01	5.98E-02	+
108	3	2.13E-03	53	9	3	2	1.00E+00	1.00E+00	*
109	8	4.92E-08	54	7	7	5	1.40E+00	6.03E-01	*
110	7	1.40E-07	54	7	6	3	8.33E-01	2.37E-01	+
111	8	1.22E-07	49	8	7	6	6.39E-02	6.39E-02	+
112	7	3.20E-08	53	9	7	6	-1.38E-01	0.00E+00	
113	4	1.63E-04	48	7	4	3	8.17E-01	4.20E-01	+
114	4	2.61E-03	49	8	4	2	1.76E+00	1.32E+00	*
115	3	1.51E-02	54	7	2	0	1.06E+00	6.26E-01	*
116	4	5.27E-04	49	8	4	3	5.82E-01	5.82E-01	*
117	5	5.99E-06	50	9	5	5	-3.02E-01	0.00E+00	
118	4	2.08E-03	49	8	4	2	1.05E+00	1.55E-01	+
119	6	2.31E-05	49	8	6	5	5.09E-01	5.09E-01	*
120	7	3.58E-06	50	9	7	6	-2.10E-01	0.00E+00	
121	7	3.83E-07	48	7	5	4	1.13E+00	2.33E-01	+
122	7	1.54E-05	48	7	5	4	1.40E-01	1.40E-01	+
123	5	1.11E-05	49	8	5	3	1.13E+00	1.13E+00	*
124	4	3.61E-04	53	9	4	3	1.11E+00	1.11E+00	*
125	4	3.09E-05	54	7	3	3	-6.44E-02	0.00E+00	
126	5	2.89E-04	48	7	5	4	-3.65E-01	0.00E+00	
127	5	8.84E-05	49	8	5	4	2.68E-01	1.17E-01	+
128	7	2.69E-06	48	7	6	5	-2.31E-01	0.00E+00	
129	7	7.22E-06	53	9	7	4	2.44E+00	2.34E+00	*
130	6	3.59E-07	54	7	5	4	1.41E-01	1.41E-01	+
131	8	1.42E-07	50	9	8	7	4.93E-01	2.87E-01	+
132	8	3.77E-08	50	9	8	7	4.63E-01	4.63E-01	*
133	4	1.29E-04	53	9	4	4	-2.47E-01	0.00E+00	
134	3	1.04E-03	54	7	2	2	-1.06E-01	0.00E+00	
135	2	1.90E-02	49	8	2	1	2.98E-01	2.98E-01	+
136	2	8.25E-03	50	9	2	2	-4.60E-02	0.00E+00	
137	7	8.26E-07	49	8	7	6	-2.24E-01	0.00E+00	
138	5	6.22E-05	50	9	5	4	9.47E-01	4.81E-01	*
139	6	9.94E-06	50	9	5	3	2.21E-01	2.21E-01	+
140	7	1.76E-06	50	9	7	6	7.37E-01	7.37E-01	*
141	7	2.16E-06	54	7	6	5	4.92E-01	1.85E-01	+
142	6	1.16E-05	49	8	6	5	2.35E-01	1.28E-01	+
143	6	4.45E-07	54	7	5	4	7.85E-01	7.85E-01	*
144	3	1.18E-02	54	7	1	1	-1.71E-01	0.00E+00	
145	4	8.16E-05	54	7	3	2	7.61E-01	1.73E-01	+
146	2	6.20E-03	49	8	2	2	-1.05E-02	0.00E+00	
147	7	1.50E-06	54	7	6	5	7.47E-01	2.14E-01	+
148	7	3.74E-07	53	9	7	5	2.15E+00	2.02E+00	*
149	5	9.70E-06	48	7	4	3	5.53E-01	5.53E-01	*
150	6	8.08E-05	53	9	6	4	1.81E+00	7.29E-01	*
151	4	4.60E-04	53	9	4	2	2.56E+00	2.56E+00	*
152	1	7.84E-02	49	8	1	1	-1.19E-03	0.00E+00	

- **parF.txt**

```
Cervus 2.0... Parentage analysis
© Copyright Tristan Marshall 1998-2001


**** Summary statistics ****

Offspring (total):                                77
  Tested:                                         77
    Known parent typed at 5 or more loci:         49
    Known parent typed at fewer than 5 loci:      28
  Not tested:                                     0
Candidate parents (total):                        14
  Tested (typed at 5 or more loci):               14
  Not tested:                                     0
    Typed at 0-4 loci:                            0
    Not typed:                                    0

One parent known:

Level       Confidence(%)  Delta Criterion  Tests       Success Rate
Strict      95.00          0.57             15 (37)     31% (76%)
Relaxed     80.00          0.00             41 (49)     84% (100%)
Unresolved                                  8 (0)       16% (0%)

Neither parent known:

Level       Confidence(%)  Delta Criterion  Tests       Success Rate
Strict      95.00          0.78             6 (12)      21% (44%)
Relaxed     80.00          0.11             21 (25)     75% (89%)
Unresolved                                  7 (3)       25% (11%)

(expected values predicted by the simulation are shown in brackets)


**** Files ****

Input
  Offspring file:          parM.csv
  Candidate parent file:   ParentsFemale.csv
  Genotype data file:      GenotypesCorrected2.csv
  Allele frequency file:   af4.alf
  Simulation data file:    sim4F.sim
Output
  Summary text file:       parF.txt
  Parentage data file:     parF.csv


**** Loci ****

1     OSM1
2     OSM2
3     CAU14
4     CAU17
5     LIST5
6     LIST9
7     OSM7
```

8 UA111
9 UA112

**** Simulation parameters ****

Input
Cycles (number of offspring): 10000
Number of candidate parents: 14
Proportion of candidate parents sampled: 1.000
Proportion of loci typed: 0.674
Proportion of loci mistyped: 0.010
Output
Relaxed confidence level: 80.00%
Strict confidence level: 95.00%

**** Known parent-offspring mismatches ****

Locus name	Offspring ID	Genotype		Known parent ID	Genotype	
Null						
OSM1	55	115	115	48	116	116
Y						
CAU17	55	169	169	48	179	179
Y						
LIST5	55	192	192	48	196	196
Y						
OSM7	55	214	225	48	216	216
UA111	55	264	291	48	254	297
OSM1	56	107	107	50	119	119
Y						
OSM2	56	122	127	50	126	126
CAU14	56	146	157	50	153	159
CAU17	56	169	177	50	176	176
OSM7	56	212	222	50	230	230
UA111	56	234	295	50	254	292
UA112	56	350	350	50	314	334
OSM1	69	115	115	48	116	116
Y						
CAU14	69	147	160	48	143	159
OSM7	69	218	218	48	216	216
Y						
UA111	69	258	278	48	254	297
UA112	69	389	389	48	320	383
OSM1	77	115	115	53	116	116
Y						
CAU14	77	157	157	53	143	153
CAU17	77	161	169	53	160	160
LIST5	77	195	195	53	197	197
Y						
OSM7	77	222	222	53	221	221
Y						
UA111	77	236	236	53	273	287
UA112	77	320	332	53	309	309
CAU14	78	157	157	53	143	153
CAU17	78	162	170	53	160	160
LIST5	78	195	195	53	197	197
Y						
OSM7	78	220	230	53	221	221
UA111	78	272	272	53	273	287

CAU17	80	173	179	49	171	171
LIST9	80	207	207	49	203	203
Y						
OSM7	80	218	218	49	220	220
Y						
UA111	80	235	281	49	276	294
UA112	80	302	330	49	332	355
CAU17	81	168	178	50	176	176
LIST9	81	200	200	50	201	201
Y						
OSM7	81	212	212	50	230	230
Y						
OSM1	83	107	111	50	119	119
OSM2	83	121	127	50	126	126
CAU17	83	169	178	50	176	176
UA112	83	331	346	50	314	334
OSM2	84	125	135	50	126	126
CAU17	84	168	178	50	176	176
LIST9	84	203	208	50	201	201
OSM7	84	217	227	50	230	230
UA111	84	269	269	50	254	292
OSM1	91	111	118	48	116	116
CAU17	91	160	177	48	179	179
OSM7	91	213	213	48	216	216
Y						
UA111	91	275	295	48	254	297
UA112	91	328	357	48	320	383
OSM1	96	108	115	53	116	116
OSM2	96	140	140	53	122	122
Y						
CAU14	96	160	160	53	143	153
CAU17	96	179	179	53	160	160
Y						
LIST9	96	206	206	53	205	205
Y						
UA111	96	237	237	53	273	287
OSM1	97	108	115	51	116	116
CAU14	97	144	144	51	142	153
CAU17	97	171	178	51	160	169
UA111	97	274	274	51	261	277
UA112	97	315	315	51	300	355
OSM1	98	114	114	50	119	119
Y						
OSM2	98	121	121	50	126	126
Y						
CAU14	98	160	160	50	153	159
CAU17	98	167	178	50	176	176
UA111	98	250	273	50	254	292
OSM1	99	102	110	50	119	119
OSM2	99	136	136	50	126	126
Y						
CAU17	99	170	170	50	176	176
Y						
LIST9	99	206	206	50	201	201
Y						
UA112	99	366	366	50	314	334
OSM2	101	141	141	53	122	122
Y						
CAU14	101	159	159	53	143	153

CAU17	101	169	169	53	160	160
Y						
OSM7	101	222	222	53	221	221
Y						
UA111	101	277	297	53	273	287
CAU17	102	168	168	53	160	160
Y						
LIST9	102	203	203	53	205	205
Y						
UA111	102	235	235	53	273	287
OSM2	103	128	136	51	123	123
UA111	103	235	235	51	261	277
OSM1	104	101	111	53	116	116
CAU14	104	145	159	53	143	153
CAU17	105	161	161	54	165	165
Y						
OSM1	108	104	110	53	116	116
UA112	108	307	307	53	309	309
Y						
OSM1	109	118	118	54	116	116
Y						
LIST9	109	201	201	54	199	199
Y						
OSM7	109	220	220	54	212	212
Y						
UA111	109	261	283	54	234	289
UA112	109	307	367	54	313	365
OSM1	112	115	115	53	116	116
Y						
CAU14	112	149	149	53	143	153
CAU17	112	164	164	53	160	160
Y						
OSM7	112	211	211	53	221	221
Y						
UA111	112	237	237	53	273	287
UA112	112	351	351	53	309	309
Y						
CAU17	114	168	168	49	171	171
Y						
UA111	114	268	268	49	276	294
OSM1	116	110	110	49	109	116
CAU17	116	169	169	49	171	171
Y						
OSM7	116	219	219	49	220	220
Y						
OSM1	117	110	110	50	119	119
Y						
OSM2	117	128	128	50	126	126
Y						
LIST9	117	200	210	50	201	201
OSM7	117	224	224	50	230	230
Y						
UA111	117	287	287	50	254	292
OSM1	119	110	110	49	109	116
CAU17	119	166	166	49	171	171
Y						
OSM7	119	216	230	49	220	220
UA111	119	270	289	49	276	294
UA112	119	300	312	49	332	355
OSM1	120	112	117	50	119	119

OSM2	120	128	135	50	126	126
CAU14	120	160	160	50	153	159
CAU17 Y	120	167	167	50	176	176
UA111	120	270	287	50	254	292
UA112	120	317	338	50	314	334
CAU17 Y	123	166	166	49	171	171
UA111	123	264	264	49	276	294
UA112	123	301	301	49	332	355
OSM2 Y	124	141	141	53	122	122
CAU14	124	156	156	53	143	153
CAU17	124	164	174	53	160	160
CAU17 Y	125	162	162	54	165	165
OSM7 Y	125	215	215	54	212	212
UA111	125	241	258	54	234	289
OSM1	126	101	114	48	116	116
CAU17 Y	126	168	168	48	179	179
LIST5 Y	126	192	192	48	196	196
UA112	126	318	318	48	320	383
OSM1 Y	128	115	115	48	116	116
CAU14	128	148	155	48	143	159
CAU17 Y	128	169	169	48	179	179
UA111	128	271	292	48	254	297
UA112	128	310	331	48	320	383
CAU14	129	158	158	53	143	153
LIST5 Y	129	192	192	53	197	197
UA111	129	268	268	53	273	287
UA112 Y	129	329	329	53	309	309
OSM1	132	108	112	50	119	119
OSM2 Y	132	132	132	50	126	126
CAU14	132	151	160	50	153	159
CAU17 Y	132	170	170	50	176	176
LIST9 Y	132	208	208	50	201	201
UA111	132	283	283	50	254	292
UA112	132	304	304	50	314	334
OSM1	133	99	105	53	116	116
OSM2 Y	133	124	124	53	122	122
CAU14	133	151	151	53	143	153
CAU17 Y	133	166	166	53	160	160
CAU17	134	164	174	54	165	165
OSM7 Y	134	225	225	54	212	212
OSM7 Y	136	225	225	50	230	230
UA112	136	322	344	50	314	334

CAU14	137	144	160	49	142	158
CAU17	137	169	169	49	171	171
Y						
LIST9	137	202	202	49	203	203
Y						
OSM7	137	228	228	49	220	220
Y						
UA111	137	289	289	49	276	294
UA112	137	360	360	49	332	355
OSM2	138	122	122	50	126	126
Y						
OSM7	138	212	225	50	230	230
UA111	138	282	298	50	254	292
UA112	138	338	338	50	314	334
OSM1	140	118	118	50	119	119
Y						
OSM2	140	131	136	50	126	126
CAU14	140	155	155	50	153	159
CAU17	140	166	177	50	176	176
UA111	140	268	287	50	254	292
UA112	140	310	327	50	314	334
LIST9	143	210	210	54	199	199
Y						
OSM7	143	229	229	54	212	212
Y						
UA111	143	272	272	54	234	289
UA112	143	318	340	54	313	365
CAU17	144	168	168	54	165	165
Y						
UA111	146	271	271	49	276	294
UA112	146	316	316	49	332	355
CAU14	148	152	152	53	143	153
CAU17	148	163	179	53	160	160
LIST9	148	207	207	53	205	205
Y						
UA111	148	240	240	53	273	287
UA112	148	313	334	53	309	309
CAU17	149	161	170	48	179	179
UA111	149	261	261	48	254	297
UA112	149	316	316	48	320	383
CAU14	150	159	159	53	143	153
LIST5	150	192	192	53	197	197
Y						
UA111	150	261	298	53	273	287
UA112	150	338	338	53	309	309
Y						
UA111	151	272	272	53	273	287
UA112	151	350	350	53	309	309
Y						
UA111	152	237	237	49	276	294

TOTAL: 194

**** Error rate analysis ****

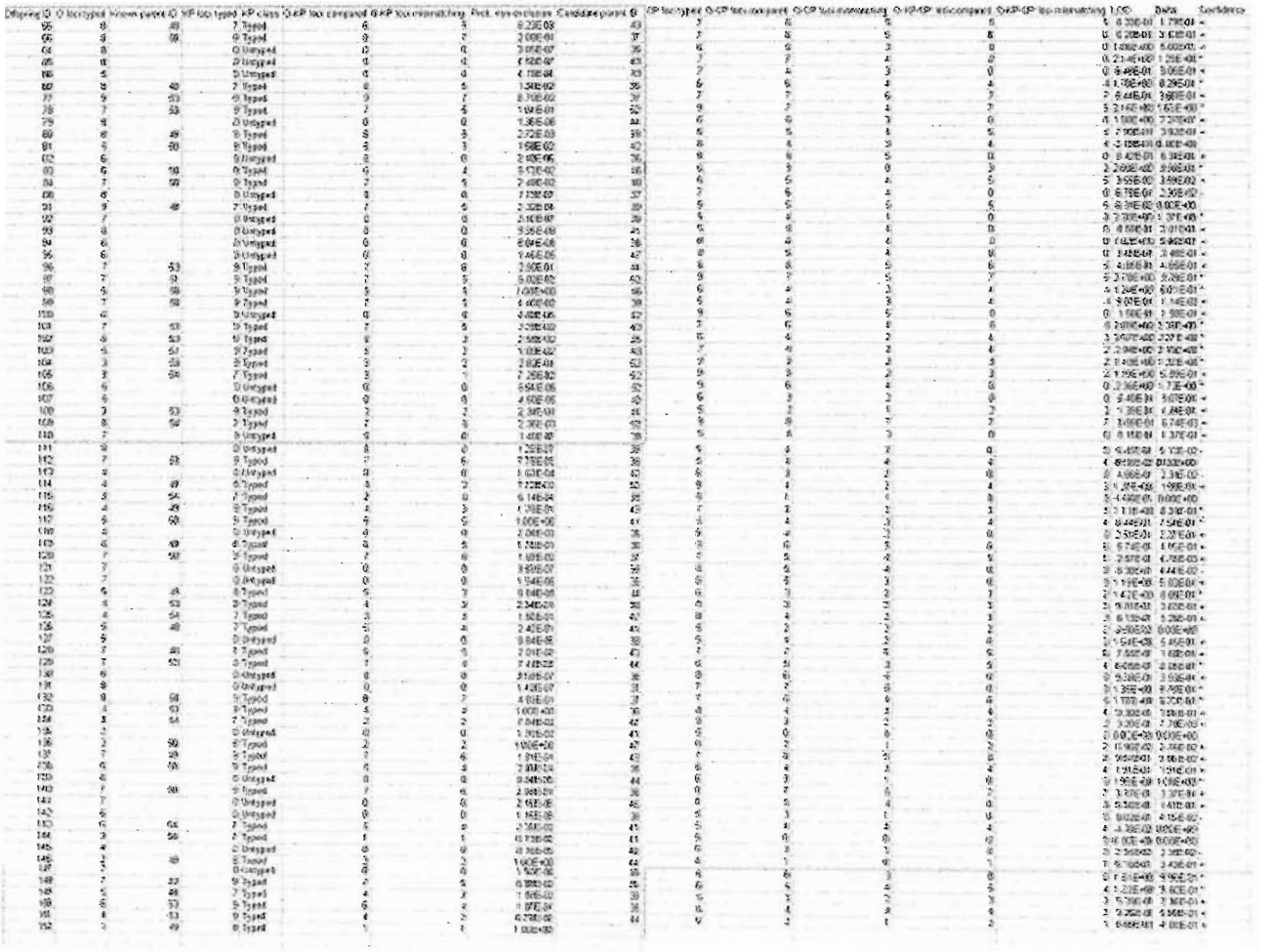
Locus name	N compared	N mismatching	N null	Detection prob.	Est. error rate
OSM1	36	23	10	0.7032	0.4543
OSM2	25	15	9	0.7469	0.4017

```
CAU14       29          20          0       0.7323          0.4709
CAU17       38          34          19      0.7402          0.6044
LIST5       15          6           6       0.4657          0.4295
LIST9       23          12          10      0.6466          0.4035
OSM7        30          21          15      0.7529          0.4649
UA111       35          35          0       0.8872          0.5636
UA112       28          28          5       0.8976          0.5571

Mean observed error rate across loci: 0.4833
  (assumes all known parent-offspring pairs are equally independent)

*****************************************************************************
```

- **parF.csv**



APPENDIX C
ARTICLES PUBLISHED

2004

Conference: The 2nd Joint Congress of the Grassland Society of Southern Africa and the South African Society of Animal Science, 28 June-1 July 2004, Goudini, South Africa.

Title: Differentiation between females of ostrich breeding trios based on egg weights

The article was peer-reviewed and appears in the South African Journal of Animal Science 2004, 34 (Supplement 2). The paper was presented at the conference and was well received.

2005

Conference: The 3rd International Ratite Science Symposium and the XII World Ostrich Congress, 14-16 October 2005, Madrid, Spain.

Title: Parentage determination of ostriches in breeding flocks using microsatellite markers

The article was peer-reviewed and appears in the Proceedings of the 3rd International Ratite Science Symposium and XII World Ostrich Congress, 14-16 October 2005, Madrid, Spain. The paper was presented at the conference and was well received.

All conference attendance was funded by the Young Professional Programme of the Western Cape Department of Agriculture, Elsenburg.

South African Journal of Animal Science 2004, 34 (Supplement 2)

Peer-reviewed paper: Joint South African Society for Animal Science/Grassland Society of Southern Africa Congress

Differentiation between females of ostrich breeding trios based on egg weights

F. Essa[1,2] and S.W.P. Cloete[2#]

[1]University of KwaZulu-Natal, Private Bag X01, Scottsville 3209, South Africa
[2]Elsenburg Agricultural Development Centre, Private Bag X1, Elsenburg 7607, South Africa

Abstract

Data, collected during the 1997-1998 breeding season from 14 ostrich trios, were used to determine whether two females in a trio could be differentiated based on egg size. For eight of the trios a difference between egg weights could be discerned and eggs could be assigned to one or to the other female. An average egg weight of approximately 1.5 kg and an average chick weight of approximately 0.9 kg was recorded. Coefficients of variation ranged from 9.8% for chick weight to 10.7% for egg weight. Overall hatchability was 53%. The repeatability of service sire observations for egg weight and chick weight were 0.32 and 0.28 respectively, whereas a higher repeatability of 0.82 and 0.67, respectively, was recorded for these traits when based on individual females. Four of the eight trios showed differences in the assumed hatchability between the females. These results suggest a preference of certain males for specific females. This study shows that the process of differentiating between females and their eggs based on the egg weight is feasible and could contribute to better data structures for ostrich breeders during genetic evaluation.

Keywords: Pedigree, egg weight, differentiation, hatchability, repeatability
[#]Corresponding author. E-mail: schalkc@elsenburg.com

Introduction

Within the ostrich industry there is a lack of basic pedigree and performance data recording, since about 80% of breeding birds are kept in colonies (usually at a male:female ratio of 6:10). It is impossible to obtain data for genetic analysis under such conditions. The industry thus lacks pedigree information linked to production data, which is the basis of livestock performance recording and evaluation schemes. Moreover, it is impossible to assess the reproduction potential of individual hens, which is known to be extremely variable. While breeding-pair mating systems enable accurate pedigree recording, the management practice of repeatedly pairing the same mates from year to year in the same breeding paddock means that potential random factors affecting performance will be confounded (Bunter, 2002). In order to improve data quality, reallocation of mates to different companions, and paddocks, from year to year can be done (Cloete *et al.*, 2002; Bunter, 2002). Trio breeding, consisting of one male with two females, leads to an improved data structure but accurate recording of pedigrees remains a problem unless eggs from the two females can be differentiated on egg size or shell structure (Cloete *et al.*, 1998; Bunter, 2002). High repeatability figures for egg weight from known breeding pairs reported by Cloete *et al.* (1998) and derived from variance ratios presented by Bunter (2002) support the reasoning that egg size could be used for differentiation. However, there is a suggestion that hatchability may be compromised under trio breeding conditions (Lambrechts *et al.*, 2002). Conventional wisdom in the ostrich industry suggests that all males are not necessarily compatible with all females, which could lead to lower hatchability under trio mating circumstances. This study investigates production records of trios to determine whether different females can be differentiated according to egg weights using historical data. Simultaneously hatchability was compared for the two females comprising a trio in cases where it was possible to differentiate on the basis of egg size.

Materials and Methods

Data were obtained from 14 trios of the 1997-1998 breeding season. These trios were maintained at the Klein Karoo Agricultural Development Centre, near Oudtshoorn, South Africa. For each trio, data were first sorted according to egg production date. Then egg weight was plotted against production date to determine if the eggs of the two females in a trio could be differentiated, based on egg weight. Wherever this was possible with a trio, the eggs of higher weight were assigned to one female and the eggs of lower weight to the other female. Nine eggs were excluded from the study as they were broken prior to weighing, and thus could not be assigned to a specific female.

The ASREML program (Gilmour *et al.*, 1999) was used to fit random effects of service sire initially, and then of dam to get an indication of the repeatability of egg weight and chick weight. Variance

Peer-reviewed paper: Joint South African Society for Animal Science/Grassland Society of Southern Africa Congress

components derived in this way were used to obtain repeatability estimates based on either sires or dams. It was reasoned that correspondence of these estimates with the literature would provide an indication of the success achieved with the allocation of eggs to individual females.

Hatchability was calculated using only the trios that could be differentiated. A chi-squared two-by-two contingency table was used to test if there were significant differences in hatchability between the two females in a trio (Siegel, 1956).

Results and Discussion

It was possible to discern between females based on egg weight differences for eight out of the 14 trios (Figure 1). In cases with very similar egg weights, it was not possible to make a reliable distinction between females in the trio (Figure 2).

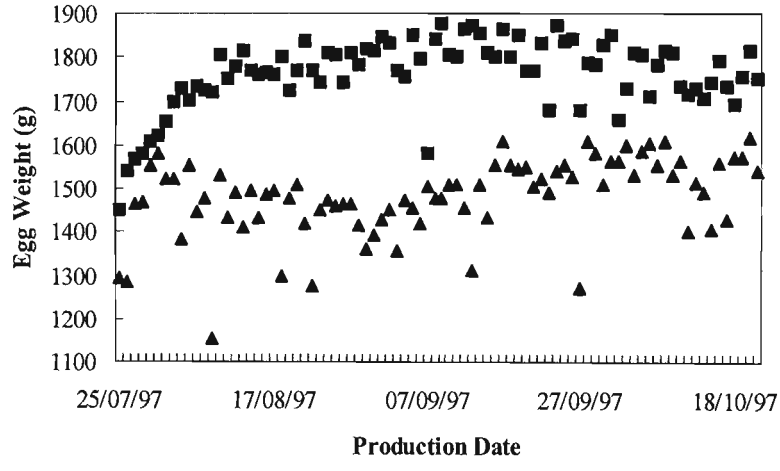


Figure 1 Egg weights plotted against production date for a trio where clear differences were evident between the two females, denoted by squares and triangles respectively

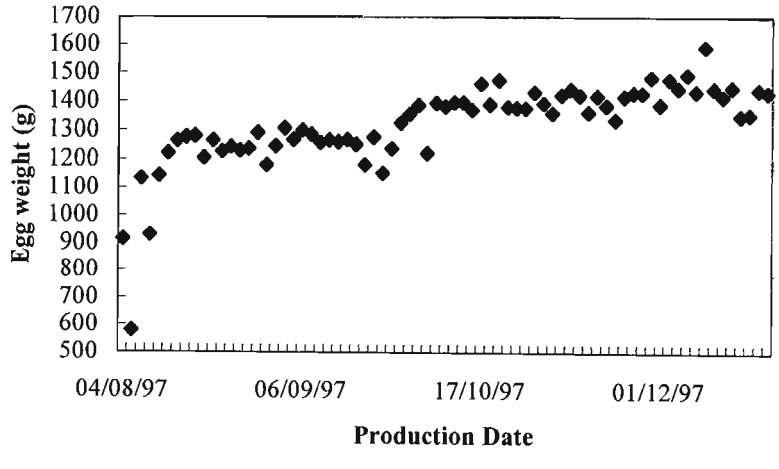


Figure 2 Egg weights plotted against production date for a trio where it was impossible to discern between the two females

Mean egg and chick weights were approximately 1.5 kg and 0.9 kg respectively (Table 1). Coefficients of variation ranged from 9.8% for chick weight to 10.7% for egg weight. These results are consistent with previous reports (Cloete *et al.*, 1998; Bunter, 2002). An overall hatchability of approximately 53% was computed from figures in Table 1, which is slightly higher than the average of 46% reported by Cloete *et al.* (1998). The repeatability based on service sire amounted to approximately 0.30 for both traits. This figure is substantially lower than previous figures ranging from 0.61 to 0.87 for egg weight

Peer-reviewed paper: Joint South African Society for Animal Science/Grassland Society of Southern Africa Congress

and from 0.56 to 0.73 for chick weight (Cloete *et al.*, 1998; Bunter, 2002). Estimates within these ranges (0.82 and 0.67 respectively) were obtained when the random source of variation was based on individual females.

In half of the trios under investigation, no difference could be discerned between the hatchability of eggs from the two females (Table 2). However, in the other four trios, there were indications of differences between the two members ($P = 0.11$ to $P < 0.01$). This indirect evidence support arguments in favour of varying levels of preference of certain males for specific females. These results also give substance to the reduction in hatchability of trios reported by Lambrechts *et al.* (2002).

Table 1 Descriptive statistics for egg and chick weights

Parameter	Egg weight (g)	Chick weight (g)
Number of observations	791	419
Overall mean	1502	892
Standard deviation	160	87
Repeatability		
Based on the male variance	0.32 ± 0.12	0.28 ± 0.11
Based on the female variance	0.82 ± 0.06	0.67 ± 0.09

Table 2 Egg numbers, hatchability figures and statistical information for individual females per trio

Sire ID	Hypothesized Dam ID	Hatchability (Dam A)	Hatchability (Dam B)	χ^2	Significance
56	149 / 150	23 / 65 = 0.35	23 / 56 = 0.41	0.207	P > 0.25
93	147 / 148	53 / 84 = 0.63	54 / 84 = 0.64	0.000	P > 0.25
101	151 / 152	42 / 68 = 0.62	5 / 51 = 0.10	30.788	P < 0.01
113	177 / 178	42 / 48 = 0.88	11 / 19 = 0.58	5.538	P < 0.05
131	179 / 180	40 / 57 = 0.70	12 / 22 = 0.55	1.099	P >0.25
159	159 / 160	16 / 30 = 0.53	36 / 49 = 0.73	2.518	P = 0.11
162	165 / 166	24 / 63 = 0.38	3 / 21 = 0.14	3.075	P = 0.08
168	173 / 174	34 / 57 = 0.60	10 / 17 = 0.59	0.049	P > 0.25

Conclusions

The differentiating between females in a trio based on egg size seems feasible, provided that there is some indication of a weight difference between the eggs of those females involved. The method should be particularly effective if previous knowledge of egg size, -shape, -colour and -structure of individual females could be obtained from birds previously subjected to pair breeding. This could contribute to better data structures without resorting to more expensive techniques, for example, DNA fingerprinting, which is being developed. The latter procedures are however required for verification of the accuracy of this technique.

References

Bunter, K.L., 2002. The genetic analysis of reproduction and production traits recorded for farmed ostriches (*Struthio camelus*). PhD dissertation, University of New England, Australia.

Cloete, S.W.P., Bunter, K.L. & van Schalkwyk, S.J., 2002. Progress towards a scientific breeding strategy for ostriches. Proc. 7th World Congr. Gen. Appl. Livest. Prod. 30, 561-568. Montpellier, France.

Cloete, S.W.P., van Schalkwyk, S.J. & Brand, Z., 1998. Ostrich breeding – progress towards a scientifically based strategy. Proc. 2nd International Ratite Congr., Ed. F.W. Huchzermeyer, Oudtshoorn, South Africa, 21-25 September 1998. pp. 55-62.

Gilmour, A.R., Cullis, B.R., Welham, S.J. & Thompson, R., 1999. ASREML – Reference manual. NSW Agriculture Biometric Bull. No. 3. NSW Agriculture, Orange Agricultural Institute, Forest Road, Orange 2800, NSW, Australia.

Lambrechts, H., Swart, D., Cloete, S.W.P., van Schalkwyk, S.J. & Greyling, J.P.C., 2002. Egg production and fertility of breeding ostriches as influenced by stocking density and male : female ratio. Proc. World Ostrich Congr., Ed. Horbanczuk, J.O., Warsaw, Poland, 26-29 September 2002. pp. 211-213.

Siegel, S., 1956. Nonparametric statistics for the behavioural sciences. McGraw-Hill Book Co. (Inc), New York.

PAPER 2

PARENTAGE DETERMINATION OF OSTRICHES IN BREEDING FLOCKS USING MICROSATELLITE MARKERS

F. Essa[1,3#], S.W.P. Cloete[2,3] & A. Fossey[1]

[1]University of KwaZulu-Natal, Private Bag X01, Scottsville 3209, South Africa
[2]Department of Animal Sciences, University of Stellenbosch, Private Bag X1, Matieland, 7602, South Africa
[3]Elsenburg Agricultural Development Centre, Private Bag X1, Elsenburg 7607, South Africa

Abstract

Parentage in colony breeding ostriches is generally unknown, where for a given offspring a number of possible parents exist. The aim of this study was to fingerprint and genotype parents and progeny in a breeding colony in order to assign parentage. DNA was extracted from blood samples and used together with microsatellite markers in a polymerase chain reaction to generate fingerprints of each individual. Multiplexing on a polyacrylamide gel was used to view the results of the amplification. Eleven polymorphic microsatellite DNA markers were randomly selected and tested. Three of these markers did not amplify in this population. The remaining eight markers were used to genotype this flock consisting of 14 females, 6 males, and 77 day-old or death-in-shell chicks. Parentage analysis was carried out on the basis of exclusion and likelihood-based methods. All loci displayed low observed heterozygosities. Paternity could be assigned with a 43% success rate at a 95% confidence level and 79% at an 80% confidence level while 21% of cases could not be resolved. Maternity could be assigned with 31% and 84% success rates at the 80% and 95% confidence levels respectively, while 16% of cases were unresolved. Four individuals could not be assigned either parent while one female did not contribute any offspring to this sample. On average, females produced 4.86 ± 2.71 fertile eggs, ranging from 0-9 during the sampling period. Males fertilized 10.17 ± 3.25 eggs on average, ranging from 5-14. A greater number of loci are required to improve the power of parentage analysis within breeding flocks.

Keywords: parentage; ostrich; microsatellite; production potential

#Corresponding author. E-mail: fatimae@elsenburg.com

Introduction

Colony breeding ostriches share communal nests with the result that parentage of eggs and chicks is unknown. Within the South African ostrich industry about 80% of breeding birds are kept in colonies. This means that there is a lack of pedigree information, which hinders the assessment of production data and reproduction potential of individual breeding birds. A number of studies indicated marked variation in the individual egg or chick production of ostrich females maintained in pairs (Bunter *et al.*, 2001; Cloete *et al.*, 2004). The development of molecular biological techniques, specifically the application of various DNA markers, has created new possibilities for the selection and genetic improvement of livestock (Van Marle-Koster & Nel, 2003). The high variability of microsatellites makes them the most useful molecular marker for use in genetic typing of individuals for parentage and kinship studies (Barker, 2002). Until recently, very few microsatellite markers were characterized for ostriches (Ward *et al.*, 1994; Kimwele *et al.*, 1998; Kumari & Kemp, 1998; Ward *et al.*, 1998). The need for more markers was met by Tang *et al.* (2003) when they characterized 70 novel microsatellite markers. Ostrich microsatellite markers have been extensively used in sex typing of the birds (Bello & Sanchez, 1999; Huynen *et al.*, 2002; Malago *et al.*, 2002; Mine *et al.*, 2002) and in an analysis of the communal nesting system (Kimwele & Graves, 2003). This study makes use of the available microsatellite markers to attempt a large-scale parentage analysis. Individual male fertilization and female egg production records were also derived from the analysis.

Materials and Methods

Sampling

This study was carried out on South African Black ostriches (*Struthio camelus*) maintained at the Klein Karoo Agricultural Development Centre in Oudtshoorn. The origin, background and husbandry of the commercial population at the Centre are described in the literature (van Schalkwyk *et al.*, 1996; Bunter, 2002). Ten breeding pairs from the 2002 – 2003 breeding sea-

son and a sample of their progeny were used initially. These birds served as a control sample of known parentage and kinship to develop the set of markers to be used. A colony from the 2003 – 2004 breeding season, consisting of 14 females and 6 males, was used as the test sample. A total of 77 progeny of this colony was sampled, comprising of 66 day-old chicks and 11 death-in-shell chicks. Blood was obtained from adult birds, day-old chicks, and death-in-shell chicks. Blood was taken by syringe from the jugular vein of adult birds and day-old chicks. For death-in-shell chicks, the chicks were extracted from their eggs and dissected through the sternum. Blood was taken by puncturing the heart. Samples were also obtained from eggs that did not show any macroscopic sign of embryonic development (termed as infertile eggs), but no DNA was extracted from them. All blood samples were stored in Vacutainer™ EDTA tubes and kept at 4 °C until needed.

DNA Extraction
DNA was extracted from blood samples using the PUREGENE® DNA Purification Kit (Gentra – Adcock Ingram). The non-mammalian whole blood protocol was followed and scaled up for 4 µl of blood.

PCR Amplification
Eleven polymorphic microsatellite markers were randomly selected from literature. These markers are detailed in Table 1.

Table 1. Microsatellite markers tested for genotyping

Primer Name	Reference
OSM 1	Kimwele *et al.* (1998)
OSM 2	Kimwele *et al.* (1998)
OSM 7	Kimwele *et al.* (1998)
LIST 005	Kumari & Kemp (1998)
LIST 009	Kumari & Kemp (1998)
CAU 14	Tang *et al.* (2003)
CAU 17	Tang *et al.* (2003)
CAU 40	Tang *et al.* (2003)
CAU 65	Tang *et al.* (2003)
CAU 90	Tang *et al.* (2003)
VIAS-OS 29	Ward *et al.* (1998)

All PCR amplifications used Go*Taq*® DNA Polymerase with green buffer (Promega) and PCR Nucleotide Mix (Promega). The cycling conditions were as described by Kimwele & Graves (2003) but the number of cycles was increased to 30. Amplification was effected on the Geneamp PCR System 2700 (Applied Biosystems). The annealing temperature was adapted to obtain optimal reactions.

Gel Analysis
PCR products were analysed by electrophoresis in 6% non-denaturing polyacrylamide gels. Multiplexing on the gel was used to generate the fingerprints using 10 µl of each sample per primer. Gels were run between 40 – 60 mA. A 20 bp ladder (ABGene) and 100 bp ladder (Fermentas) were used to size the alleles. The amplified loci were visualized by ethidium bromide UV fluorescent staining. Images were captured on the UVIsave gel documentation system.

Allele Scoring
Alleles were typed by length across the loci using UVIsoft Analysis software (Uvitec).

Parentage analysis
CERVUS Version 2.0 (Marshall *et al.*, 1998) was used to perform an allele frequency analysis and a parentage analysis. The programme uses exclusion and a likelihood-based approach to assign parentage. The logarithm of the odds (LOD) scores are calculated from a simulation using the allele frequencies. The simulation generates criteria that permit the assignment of parentage to the most likely candidate parent and to give a level of statistical confidence for this assignment. A separate simulation was carried out for the male candidate parents and for the female candidate parents. The programme eliminates exclusion of parentage because of allelic mismatch, which could be due to actual allelic mismatch or more commonly to erroneous laboratory typing or the presence of mutations or null alleles. Because neither parent was known, CERVUS recommends a two-step analysis with the first step to run the group of parents with fewer candidates, males in this case, and the second step to run the analysis with the females using the results of the first step. Pedigrees that could be obtained at the 80% confidence interval were used to derive the production of fertile eggs by individual female parents in the colony. The number of eggs fertilized by individual males was determined accordingly.

Results
Two breeding pairs and four of their progeny each were used as the known parentage sample to test the microsatellite markers. Three of the eleven markers viz. VIAS-OS 29, CAU 40, and CAU 90, did not amplify in this population, and were excluded from further usage. The remaining eight markers were used to type 20 adults and 77 progeny from the experimental breeding colony. Overall, production records for this colony show that 234 eggs were produced during this breeding season. Of these eggs, 188 could have been sampled by blood, being either live chicks or death-in-shell chicks. From this potential number, 77 (or 41%) were used as the progeny sample.

Locus	No. Alleles	No. Individuals typed	Observed heterozygosity
OSM 1	20	75	0.453
OSM 2	19	63	0.302
CAU 14	19	74	0.432
CAU 17	19	74	0.446
List005	7	32	0.031
List009	12	59	0.102
OSM 7	21	70	0.214
ULoc 1	50	75	0.560
ULoc 2	52	66	0.576

Table 2. Allelic variation of the 9 ostrich loci used in 97 individuals comprising 20 adults and 77 progeny

One locus, CAU 65, displayed no heterozygosity for any of its alleles for all individuals typed at this locus. This locus was therefore excluded from the CERVUS analyses. Scoring of bands revealed amplification for most individuals at loci not within the product size range of specified markers used. To prevent information loss, these loci were named Unknown Locus (ULoc) one and two, and were included in the CERVUS analyses. For the 97 individuals typed, 7-52 alleles per locus were detected with an observed heterozygosity of 0.031-0.576 (Table 2). All loci displayed positive high null allele frequencies.

The total exclusion probabilities for first and second parents were 0.999 and 1.000 respectively. All 20 candidate parents, consisting of 6 males and 14 females, were sampled. Paternity was identified at either the 95% confidence interval (33 assignments), at the 80% confidence interval (61 assignments) or not at all (16). Maternity was then identified with 15 assignments at the 95% confidence interval, 41 assignments at the 80% confidence interval, and 8 progeny could not be assigned. Four individuals were unassigned to either a mother or a father. One female was not assigned any progeny. The mean observed error rate across loci for known parent-offspring mismatches amounted to 0.4833.

The production of fertile eggs by individual females and the number of eggs fertilized by individual males are depicted in the frequency distributions of Figure 1 and Figure 2, respectively. No eggs were assigned to one of the females, and it was assumed that she failed to produce any eggs during the period of sampling. On average, females produced 4.86 ± 2.71 fertile eggs, with a range of 2-9 fertile eggs. Individual males fertilized 10.17 ± 3.25 eggs on average, with a range of 5-14. All males were represented in the potential offspring that were sampled.

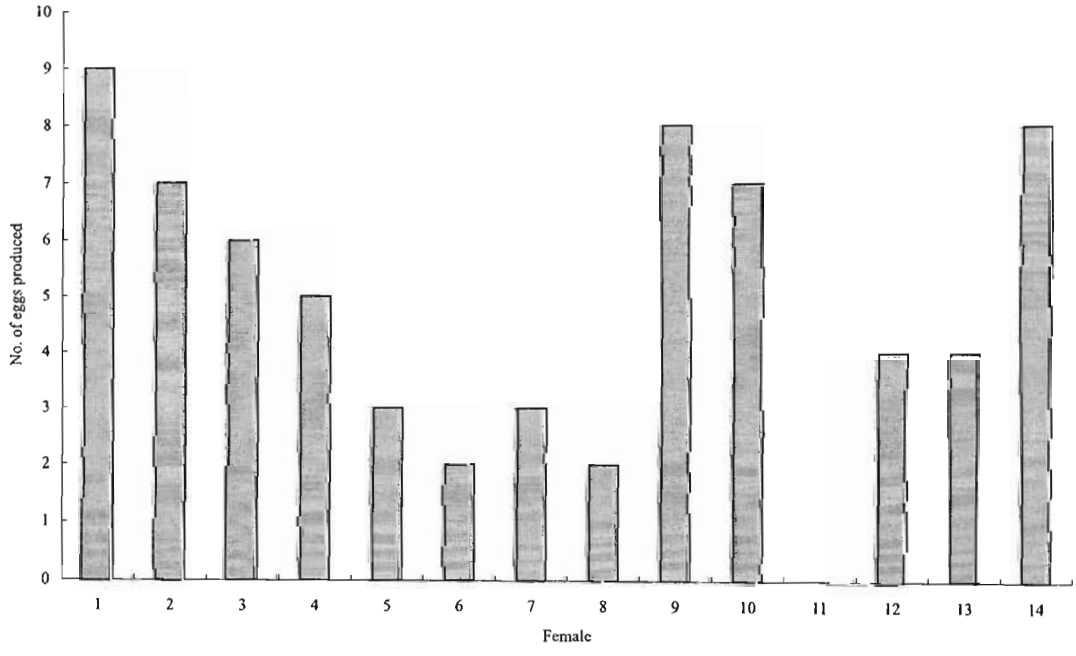


Figure 1. Number of eggs produced by individual females during the sampling period, as based on the genotypes of 68 progeny

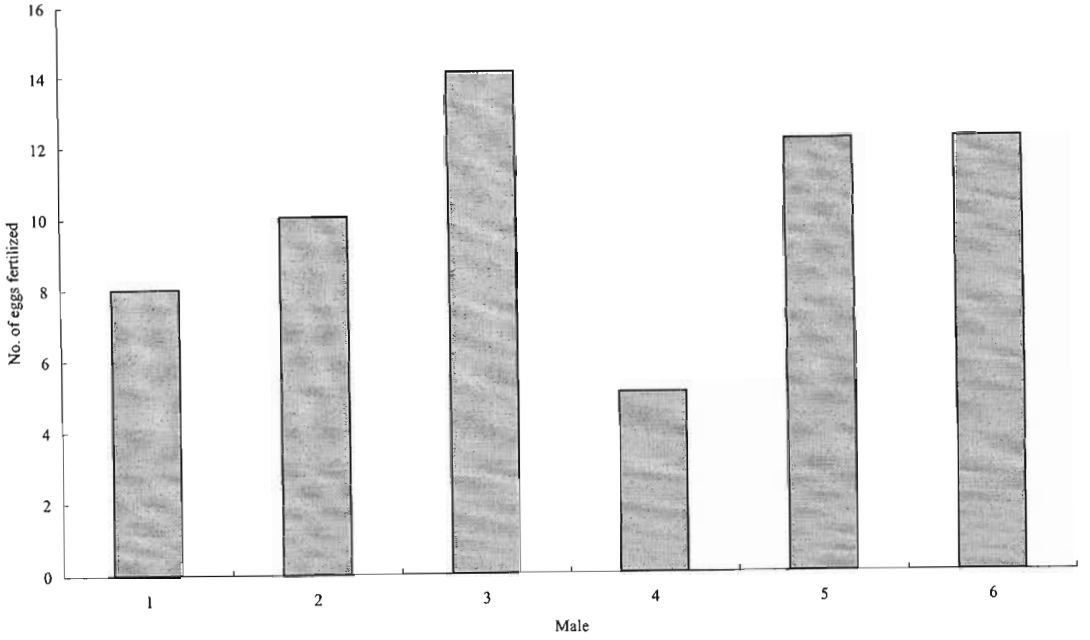


Figure 2 Number of eggs fertilized by individual males during the sampling period, as based on the genotypes of 61 progeny

Discussion

Marshall *et al.* (1998) suggested that the number of loci required to resolve parentage with a given level of confidence depends on factors such as the level of variation at a locus (expected heterozygosity), the number of candidate parents, the proportion of candidate parents sampled and the availability of genetic data from a known parent. According to Dodds (2003), 10-13 markers are required to assign parentage at the 80-95% level without knowledge of sire-dam combinations using a co-dominant marker. This study found that eight markers were sufficient to assign parentage when tested on the sample of progeny with known parentage in the pair-bred sample. This proved to be not the case when applied to the colony. When the relationships between potential parents in the colony were considered, it was found that three of the male parents were full sibs with three of the female parents. This relatedness between parents can account for the low observed heterozygosity at all loci and high null allele frequencies. A null allele occurs because of mutations in one or both primer binding sites, sufficient to prevent effective amplification of the microsatellite allele (Callen *et al.*, 1993). Marshall *et al.* (1998) warn that it is more difficult to identify a null allele with certainty in the absence of known parent – offspring relationships. However, a locus with a large positive estimate of null allele frequency indicates an excess of homozygotes but does not necessarily imply that a null allele is present. Five of the loci used in this study showed a higher heterozygosity when used in a parentage analysis of a wild population (Kimwele & Graves, 2003). The low observed heterozygosity and excess of homozygotes suggest a low amount of genetic variation in this population. Although parentage could be assigned using the loci, the accuracy of these assignments is low. Therefore a greater number of loci are required to assign parentage with a greater accuracy. Progeny that were not assigned to either parent were typed at too few loci to enable assignment of parentage. The high estimated error rate occurs because of the high null allele frequencies observed across the loci. CERVUS suggests excluding loci with high estimated error rates to improve the power of the analysis. However, it was not feasible in this study.

From the frequency distribution in Figure 2, it can be noted that all males contributed to the fertilization of a minimum of 5 eggs during the sampling period. One female was excluded as a parent to all progeny and possibly did not lay any fertile eggs during this period of sampling. A coefficient of variation of 55.8% was calculated for the production of fertile eggs by individual females during the sampling period. This measure of variation is about equivalent to variation in egg production in studies done on pair-breeding females (51.8%: Bunter *et al.*, 2001; 52.9%: Cloete *et al.*, 2004). It has to be stated that these studies included infertile eggs. The genotyping of such eggs should thus receive attention, if a robust system involving all eggs produced is envisaged. In general, it nevertheless seems as if colony females vary just as much as pair-bred females in respect to individual egg production. The main difference is that a low producing female cannot always be identified in a colony, unless resorting to genotyping.

Conclusions

Relatedness between individuals in this colony increased the difficulty in assigning parentage of individual progeny. A greater number of loci are required in addition to the loci that were used, to increase the accuracy with which assignment is done. The high homozygosity and low genetic variation within this colony is indicative of some inbreeding which can potentially be circumvented by the checking of pedigrees of birds for relatedness before accepting them as replacements in colonies. In practice, this may not always be feasible, as a measure of relatedness may be expected in colony birds, unless specific steps (e.g. the mating of male and female birds obtained from unrelated strains) are taken.

The second main practical implication of this study was to assess the fertilization potential of individual males and, more importantly, the production potential of females within the colony mating system. It appeared as if the production of fertile eggs by individual females considered were as variable as egg production in pair-bred females. In order to determine overall egg production, the parentage analysis will have to be carried out with a higher accuracy, incorporating all eggs laid. The further development of a robust DNA fingerprinting protocol should thus receive serious attention, if extrapolation to the broader industry is considered.

Acknowledgements

We express our sincerest gratitude to all those responsible for the maintenance and recording of the experimental birds at the Oudtshoorn Experimental farm. We are grateful to the Klein Karoo Cooperation for financial support and the usage of the resource flock. The study was partially funded by the THRIP programme of the South African Department of Trade and Industry.

References

BARKER, G.C., 2002. Microsatellite DNA: a tool for population genetics. Trans. R. Soc. Trop. Med. Hyg. 96, S1/21-S1/24.

BELLO, N., & SANCHEZ, A., 1999. The identification of a sex-specific DNA marker in the ostrich using a random amplified polymorphic DNA (RAPD) assay. Mol. Ecol. 8, 667-669.

BUNTER, K.L., 2002. The Genetic Analysis of Reproduction and Production Traits Recorded for Farmed Ostriches (*Struthio camelus*). Ph.D. dissertation, University of New England, Armidale, Australia.

BUNTER, K.L., CLOETE, S.W.P., VAN SCHALKWYK, S.J. & GRASER, H.-U., 2001. Factors affecting reproduction in farmed ostriches. Proc. Assoc. Advmnt Anim. Breed. Genet. 14, 43-46.

CALLEN, D.F., THOMPSON, A.D., SHEN, Y., PHILLIPS, H.A., RICHARDS, R.I., MULLEY, J.C. & SUTHERLAND, G.R., 1993. Incidence and origin of "null" alleles in the (AC)n microsatellite markers. Am. J. Hum. Genet. 52, 922-927.

CLOETE, S.W.P., BUNTER, K.L., BRAND, Z. & LAMBRECHTS, H., 2004. (Co)variances for reproduction, egg weight and chick weight in ostriches. S. Afr. J. Anim. Sci. 34(Supplement 2) 6, 17-19.

DODDS, K.G., 2003. The number of markers required for parentage assignment. Proc. Assoc. Advmnt Anim. Breed. Genet. 15, 39-42.

HUYNEN, L., MILLAR, C.D. & LAMBERT, D.M., 2002. A DNA test to sex ratite birds. Mol. Ecol. 11, 851-856.

KIMWELE, C.N. & GRAVES, J.A., 2003. A molecular genetic analysis of the communal nesting of the ostrich (*Struthio camelus*). Mol. Ecol. 12, 229-236.

KIMWELE, C.N., GRAVES, J.A., BURKE, T. & HANOTTE, O., 1998. Development of microsatellite markers for parentage typing of chicks in the ostrich *Struthio camelus*. Mol. Ecol. 7, 247-255.

KUMARI, P., & KEMP, S.J., 1998. Polymorphic microsatellite markers in the ostrich (*Struthio camelus*). Mol. Ecol. 7, 133-140.

MALAGO, W., FRANCO, H.M., MATHEUCCI, E., MEDAGLIA, A. & HENRIQUE-SILVA, F., 2002. Large scale sex typing of ostriches using DNA extracted from feathers. BMC Biotechnology 2, 19.

MARSHALL, T.C., SLATE, J., KRUUK, L.E.B. & PEMBERTON, J.M., 1998. Statistical confidence for likelihood-based paternity inference in natural populations. Mol. Ecol. 7, 639-655.

MINE, O.M., MOCHAKANA, M.E., MPAPHO, T., MOTLHANKA, D.T.M. & KGWATALALA, P., 2002. Application of a sex identification technique in juvenile ostriches and its potential application in Botswana. S. Afr. J. Anim. Sci. 32, 160-163.

TANG, B., HUANG, Y.H., LIN, L., HU, X.X., FENG, J.D., YAO, P., ZHANG, L. & LI, N., 2003. Isolation and characterization of 70 novel microsatellite markers from ostrich (Struthio camelus) genome. Genome 46, 833-840.

VAN MARLE-KOSTER, E. & NEL, L.H., 2003. Genetic markers and their application in livestock breeding in South Africa: A review. S. Afr. J. Anim. Sci. 33, 1-10.

VAN SCHALKWYK, S.J., CLOETE, S.W.P. & DE KOCK, J.A., 1996. Repeatability and phenotypic correlations for live weight and reproduction in commercial ostrich breeding pairs. Br. Poult. Sci. 37, 953-962.

WARD, W.K., MATTHEWS, M.E., MURRAY, N.D. & ROBINSON, N.A., 1994. An ostrich dinucleotide repeat polymorphism at the VIAS-OS2 locus. Anim. Genet. 25, 291.

WARD, W.K., MCPARTLAN, H.C., MATTHEWS,M.E. & ROBINSON, N.A., 1998. Ostrich microsatellite polymorphisms at the VIAS-OS4, VIAS-OS8, VIAS-OS14, VIAS-OS22, and VIAS-OS29 loci. Anim. Genet. 29, 331.