4.1.2 Identification and cloning of the gene encoding p67, the major sporozoite antigen ........................................ 82
4.1.3 Could the p67 gene be used to distinguish between T. p. parva and T. p. lawrencei stocks? .................. 83
4.2 Materials and methods .................................................. 84
  4.2.1 Parasites ................................................................. 84
  4.2.1 PCR amplification of the variable region of the p67 gene .... 84
  4.2.2 Sequencing of the variable region of the p67 gene .......... 84
    4.2.2.1 Direct sequencing of single p67 PCR products .... 87
    4.2.2.2 Cloning and sequencing of multiple p67 PCR products ... 87
    4.2.2.3 p67 Gene sequence analysis ................................. 88
4.3 Results and discussion .................................................. 88
  4.3.1 Amplification of the variable region of T. parva p67 genes . 88
  4.3.2 Variation between T. parva p67 sequences ................... 90
  4.3.3 Sequence analysis of T. p. lawrencei KNP/94_1 p67 genes reveals sequences characteristic of both T. p. parva and T. p. lawrencei .................................................. 97
  4.3.4 p67-Based probes to distinguish between T. p. parva and T. p. lawrencei ........................................... 101
  4.3.5 Summary .................................................................. 102

CHAPTER 5: GENERAL DISCUSSION AND CONCLUSIONS ............ 103
  5.1 Can sporozoite antigen and ribosomal RNA gene sequences be used to distinguish between T. p. parva and T. p. lawrencei? .... 103
  5.2 Are T. p. parva and T. p. lawrencei valid labels? .............. 105
  5.3 Conclusion .................................................................. 107
3.2.2 Cloning and sequencing of the T. p. lawrencei Hluhluwe 3 ITS ............................................. 60
3.2.3 A new strategy for amplification and sequencing of T. parva ITS regions ................................. 60
  3.2.3.1 Amplification of T. parva ITS regions ........... 60
  3.2.3.2 Primers for sequencing T. parva ITS amplification products ............................................. 62
  3.2.3.3 Cloning and sequencing of ITS amplification products ...................................................... 62
3.2.4 Experimental probe design and testing ......................................................................................... 63

3.3 Results and discussion .................................................................................................................... 63
  3.3.1 The search for T. p. lawrencei- and T. p. parva-specific probes in the ITS ................................. 63
    3.3.1.1 Experimental probes based on T. p. lawrencei 7344 and T. p. parva Muguga and Uganda ITS sequences .................................................................................................................. 64
    3.3.1.2 Nucleotide sequence of the T. p. lawrencei Hluhluwe 3 ITS ................................................ 66
    3.3.1.3 An additional putative T. p. lawrencei characterisation oligonucleotide ................................ 67
  3.3.2 Nucleotide sequence of the ITS of a range of T. parva isolates .................................................. 67
  3.3.3 Screening of T. parva isolates with new ITS experimental probes .............................................. 75
  3.3.4 Summary .................................................................................................................................. 80

CHAPTER 4: VARIATION IN THEILERIA PARVA p67 GENES ................................................................. 81
  4.1 Introduction .................................................................................................................................. 81
    4.1.1 Both schizonts and sporozoites stimulate immune responses to T. parva infection ................... 81
2.2.3.3 Primers for amplification of the internal transcribed spacer ........................................ 27
2.2.3.4 Amplification of rRNA genes .................. 28
2.2.4 Cloning of the SSU rRNA amplification product .......... 28
2.2.5 Cloning of the LSU rRNA and ITS amplification products ........................................... 30
2.2.5.1 Preparation of recombinant pUC18 clones ....... 30
2.2.5.2 Sonication of recombinant pUC18 clones ...... 30
2.2.5.3 Preparation of recombinant M13 subclones ..... 32
2.2.6 Sequencing ............................................. 32
2.2.6.1 SSU rRNA templates ............................ 32
2.2.6.2 LSU rRNA and ITS templates .................. 33
2.2.6.3 Sequence compilation and manipulation .......... 33
2.3 Results and discussion .................................. 33
2.3.1 The need to amplify parasite rRNA genes .......... 33
2.3.1.1 Amplification and cloning of the SSU rRNA gene of 
T. p. lawrencei 7344 ........................................ 34
2.3.1.2 Amplification and cloning of the LSU rRNA gene and 
the ITS of T. p. lawrencei 7344 .......................... 34
2.3.2 Sequence of the T. p. lawrencei 7344 transcription unit .... 38
2.3.2.1 The small and large subunits ....................... 41
2.3.2.2 The internal transcribed spacer .................... 43
2.3.3 Summary ............................................... 45

CHAPTER 3: SEQUENCE VARIATION IN THE THEILERIA PAVKA 
INTERNAL TRANSCRIBED SPACER ........................................ 57
3.1 Introduction .............................................. 57
3.2 Materials and methods .................................... 58
3.2.1 Parasites and preparation of DNA ................... 58
CHAPTER 2: SEQUENCE OF THE RIBOSOMAL RNA TRANSCRIPTION UNIT
OF A CLONED T. P. LAWRENCEI PARASITE

2.1 Introduction ........................................... 20
2.1.1 Ribosomes and ribosomal RNA genes .......... 20
2.1.2 Organisation of rRNA genes ................. 20
2.1.3 Using rRNA genes to distinguish between closely related species ....................... 22

2.2 Materials and methods ...................... 24
2.2.1 Parasites ...................................... 24
2.2.2 DNA extraction .............................. 24
2.2.3 Polymerase chain reaction (PCR) amplification of rRNA genes ................. 25
2.2.3.1 Primers for amplification of the small subunit .... 25
2.2.3.2 Primers for amplification of the large subunit .... 27
### TABLE OF CONTENTS

**ABSTRACT** ................................................................. ii

**DECLARATION** ........................................................... iv

**ACKNOWLEDGEMENTS** ..................................................... v

**LIST OF ABBREVIATIONS** ................................................ vi

**TABLE OF CONTENTS** ..................................................... viii

**LIST OF TABLES** ........................................................ xvi

**LIST OF FIGURES** ........................................................ xvii

**CHAPTER 1: INTRODUCTION** ..................................... 1

1.1 History of theileriosis in southern Africa ............... 1
   1.1.1 East Coast fever ........................................... 1
   1.1.2 Corridor disease and January disease ................. 4

1.2 Life cycle ............................................................ 5

1.3 *Theileria* species in southern Africa .................. 5
   1.3.1 *Theileria parva* ........................................ 7
   1.3.2 Other *Theileria* species .......................... 10

1.4 Current epidemiology of theileriosis in South Africa .. 10
   1.4.1 Corridor disease ...................................... 10
   1.4.2 The threat of East Coast fever ....................... 11
   1.4.3 Relocation of buffalo in South Africa ............ 11

1.5 Characterisation of *T. parva* stocks ....................... 13
   1.5.1 Pathogenicity ........................................ 13

viii
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ng</td>
<td>nanogram</td>
</tr>
<tr>
<td>nm</td>
<td>nanometre</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PEG</td>
<td>polyethylene glycol</td>
</tr>
<tr>
<td>pg</td>
<td>picogram</td>
</tr>
<tr>
<td>pmol</td>
<td>picomole</td>
</tr>
<tr>
<td>rDNA</td>
<td>ribosomal DNA</td>
</tr>
<tr>
<td>RFLP</td>
<td>restriction fragment length polymorphism</td>
</tr>
<tr>
<td>rpm</td>
<td>revolutions per minute</td>
</tr>
<tr>
<td>rRNA</td>
<td>ribosomal RNA</td>
</tr>
<tr>
<td>S</td>
<td>Svedburg unit</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SSUrRNA</td>
<td>subunit ribosomal RNA</td>
</tr>
<tr>
<td>TEMED</td>
<td>N,N',N',N'-tetramethyl-ethylenediamine</td>
</tr>
<tr>
<td>Tₚ</td>
<td>dissociation temperature</td>
</tr>
<tr>
<td>Tpr</td>
<td><em>Theileria parva</em> repetitive (gene family)</td>
</tr>
<tr>
<td>v/v</td>
<td>volume/volume</td>
</tr>
<tr>
<td>w/v</td>
<td>weight/volume</td>
</tr>
<tr>
<td>W</td>
<td>watts</td>
</tr>
<tr>
<td>X-gal</td>
<td>5-bromo-4-chloro-3-indolyl-β-D-galactoside</td>
</tr>
</tbody>
</table>
LIST OF ABBREVIATIONS

Aₜₜ absorbance at n nanometres wavelength
anon. anonymous
bp base pairs
BSA bovine serum albumin
°C degrees Celsius
Cl Curie
ddNTPs deoxynucleoside triphosphates (ddATP, ddCTP, ddGTP, ddTTP)
DMSO dimethyl sulfoxide
DNA deoxyribonucleic acid
dNTPs deoxynucleoside triphosphates (dATP, dCTP, dGTP, dTTP)
DTT dithiothreitol
e molar extinction coefficient
ECF East Coast fever
EDTA ethylenediaminetetraacetic acid disodium salt
ETS external transcribed spacer
g force of gravity
g gram
IPTG isopropyl-β-D-thiogalactopyranoside
ITS internal transcribed spacer
kbp kilobase pair
LSUrRNA large subunit ribosomal RNA
M molar
mM millimolar
MAb monoclonal antibody
mg milligram
µg microgram
ml millilitre
µl microlitre
ACKNOWLEDGEMENTS

I would like to express my sincere thanks to the following people:

Dr. Basil Allsopp, Assistant Director, Onderstepoort Veterinary Institute (OVI), supervisor of this project, for all his valuable suggestions, advice and guidance throughout the course of this research, and for his constructive criticism of this thesis.

Prof. Jenny Alexander, Head of the Microbiology Department, University of the Witwatersrand, supervisor of this project, for her help, advice and encouragement.

Dr. Hein Stoltz, Protozoology Section, OVI, for many hours of interesting and informative discussion.

Dr. Hein Stoltz and Ms. Milana Dunsterville, Protozoology Section, OVI, for the supply of *Theileria* cell culture material and blood samples.

Dr. D. W. Verwoerd, Director, OVI for permission to use the results of this research for this thesis.

The Photography Section, OVI, for help with the preparation of some of the figures.

Dr. Richard Bishop and Dr. Vishwanath Nene, International Livestock Research Institute (ILRI), Nairobi, Kenya, for their assistance and advice during my visit to ILRI.

My friends and colleagues in the Molecular Biology Section, OVI, for their suggestions, comments and criticism.

My family and friends for their encouragement and support.

Peter Hawkes for his love, help, and never-ending patience.
DECLARATION

I declare that this thesis is my own, unaided work. It is being submitted for the degree of Doctor of Philosophy in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other University, nor has it been prepared under the aegis or with the assistance of any other body or organisation or person outside the University of the Witwatersrand, Johannesburg, other than as indicated in the acknowledgements which follow.

Nicola E. Collins
15th day of October 1997
probes TPL2 and/or TPL3a, there was no absolute distinction between T. p. parva and T. p. lawrencei isolates using the eleven probes. Nonetheless, this panel of oligonucleotides will be useful in epidemiological surveys to determine the distribution of the different genotypes.

It has been shown that the sequence of the major sporozoite antigen, p67, is conserved amongst T. p. parva isolates but varies in T. p. lawrencei isolates. The variable region of the p67 gene of a number of T. parva isolates was amplified and sequenced to determine whether probes that would distinguish between T. p. parva and T. p. lawrencei could be identified in the p67 gene. The conserved p67 sequence characteristic of T. p. parva was found in two T. parva isolates obtained from cattle in Zambia. Most of the South African and Zimbabwean T. parva isolates from buffalo contained a variable region in their p67 gene sequences, which is characteristic of T. p. lawrencei isolates. However, T. p. lawrencei isolates from the Kruger National Park, which cause Corridor disease in cattle, were found to contain p67 gene sequences characteristic of both T. p. parva and T. p. lawrencei.

The ITSs of both T. p. parva and T. p. lawrencei parasites contained different combinations of identifiable segments of sequence in the variable regions, resulting in a mosaic of ITS sequence segments in any one isolate. A sequence mosaic was also identified in at least one of the T. parva p67 gene sequences. These sequence mosaics indicate that T. parva parasites undergo genetic recombination. It appears, therefore, that the T. p. parva and T. p. lawrencei parasite populations are not yet separate gene pools. Under these circumstances no probe will distinguish between all T. p. parva and all T. p. lawrencei parasites unless it targets a sequence which is directly associated with the mechanisms responsible for their differential behaviour in cattle.
ABSTRACT

The aim of this thesis was to develop DNA probes to distinguish between the protozoan parasites *Theileria parva* *parva* and *T. p. lawrencei* which cause East Coast fever (ECF) and Corridor disease respectively. ECF was eradicated from South Africa in 1954, and today Corridor disease has become the most important form of theileriosis. Although ECF has been eradicated, the vector ticks are still prevalent in South Africa and the cattle population would be highly susceptible to a recurrence of the disease. At present there is no reliable means of distinguishing between *T. p. parva* and *T. p. lawrencei*.

Sequence differences between *T. parva* and other *Theileria* species have previously been found in the small subunit ribosomal RNA (rRNA) gene; probes designed to detect these sequence differences can be used to distinguish between *Theileria* species. We therefore decided to search for differences in the rRNA genes of *T. p. parva* and *T. p. lawrencei*. To this end, the entire rRNA transcription unit was amplified from a cloned *T. p. lawrencei* parasite; the unit comprises the small subunit rRNA (SSUrRNA) gene, the internal transcribed spacer (ITS) and the large subunit rRNA (LSUrRNA) gene. The amplification products were cloned and sequenced, and the *T. p. lawrencei* rRNA sequence was compared to that of *T. p. parva*. While there was little variation in their SSUrRNA and LSUrRNA gene sequences, there was major sequence variation in the ITS. The ITSs from twelve *T. parva* isolates were amplified, cloned and sequenced, and eleven characterisation oligonucleotide probes were identified. The *T. p. parva* isolates screened in this study hybridised with a limited subset of the probes, while the *T. p. lawrencei* isolates hybridised with many more of the probes, indicating that the *T. parva* population in cattle is more homogenous than that in buffalo. There thus appears to have been a selection in cattle of a relatively homogenous subpopulation of *T. parva* from a much larger, more diverse gene pool in buffalo. Although most *T. p. parva* isolates (93.5%) were detected by probe TPP1, and most *T. p. lawrencei* isolates (81.8%) were detected by
The relationship between *Theileria parva parva* and *T. p. lawrencei* as shown by sporozoite antigen and ribosomal RNA gene sequences.

Nicola Elaine Collins

A thesis submitted to the Faculty of Science,
University of the Witwatersrand, Johannesburg,
in fulfilment of the requirements for the degree
of Doctor of Philosophy.

Johannesburg 1997
waterbuck *Kobus defassa* are also reservoirs of the parasite in East Africa (Stagg et al., 1994), but there is as yet no evidence to suggest that common waterbuck (*K. ellipsiprymnus*) are involved in the epidemiology of theileriosis in southern Africa (H. Stoltz, pers. comm.).

1.3.1 *Theileria parva*

It has been recognised that the parasites causing East Coast fever, Corridor disease and January disease are morphologically and serologically indistinguishable, and that they were classified solely on the basis of the clinical and epidemiological features of the infections they produce. Ullenberg (1976) and Lawrence (1979) thus proposed that they be classified as subspecies of *T. parva*. Under this trinomial classification system, *Theileria parva parva* was considered to be the causative agent of East Coast fever, *T. p. lawrencei* was considered to cause Corridor disease, while *T. p. bovis* was considered the causative agent of January disease in Zimbabwe (Lawrence, 1979; Ullenberg, 1981). *T. p. bovis* has also been implicated in Tzaneen disease in South Africa (Stoltz, 1989).

East Coast fever is an acute and usually fatal disease, which can result in mortalities as high as 90%. In its early stages the disease is lympho-proliferative; at the onset of pyrexia, a lympho-destructive phase is initiated. Clinical symptoms include enlarged lymph nodes and pulmonary oedema, and widespread subcutaneous oedema may also be seen. The disease is characterised by high schizont and piroplasm parasitaemias observed in lymphocytes and erythrocytes respectively (Neitz, 1957; Ullenberg, 1981). High piroplasm parasitaemias result in the majority of ticks becoming infected thus enabling effective transmission of the parasite between cattle. The brown ear tick, *Rhipicephalus appendiculatus*, is the main vector, and it has been shown experimentally that *R. zambeziensis* can transmit ECF (Lawrence et al., 1983).
Figure 1.1: Life cycle of *Theileria parva* (from Norval *et al.*, 1992).
placed in a different genus as it was believed that the stage of this parasite in the
eythrocyte did not divide; members of the genus *Gonderia* were believed to multiply
in the erythrocytic stage. Neitz considered *G. bovis* to be different from *G. lawrencei*
based on its ability to persist in carrier cattle; when he demonstrated a carrier state for
*G. lawrencei* (Neitz, 1958a; 1958b) the name *G. bovis* was made synonymous with
*G. lawrencei*. The genus *Gonderia* was later abandoned but the term "bovis" was
later reintroduced to describe the Zimbabwean parasite (see section 1.3).

1.2 Life cycle

A diagrammatic representation of the *Theileria parva* life cycle as it is now known
is shown in fig. 1.1. Sporozoites are introduced into cattle in the saliva of infected
ticks when they take a blood meal. The sporozoites enter peripheral lymphocytes
(Stagg et al., 1981; Fawcett et al., 1982a) which eventually migrate to the lymph
nodes. Here, the sporozoites develop into macroschizonts (Stagg et al., 1981). The
parasites transform the host cells and induce lymphoblastogenesis; parasites and host
cells divide synchronously, and daughter schizonts are distributed into each of the
daughter lymphocytes (Stagg et al., 1980). A proportion of macroschizonts
differentiates to produce merozoites, which are released when the infected
lymphocytes rupture. Merozoites penetrate erythrocytes and develop into plasmodia,
the stage infective for ticks. In the tick, it is thought that a number of sexual stages
(Melhorn and Schein, 1984) leads to infection of the Type III salivary gland acini
with sporozoites (Fawcett et al., 1982b; Binnington et al., 1983).

1.3 *Theileria* species in southern Africa

The African buffalo (*Syncerus caffer*) is the major reservoir for a mixed population
of antigenically diverse theilerioid parasites. There is evidence to suggest that defassa
1.1.2 Corridor disease and January disease

In 1953, there was an outbreak of theileriosis in the corridor between the Umfolozi and Hluhluwe game reserves in Natal (Neitz et al., 1955), although East Coast fever had not been reported from that area for fifteen years. The clinical symptoms and pathology of the disease, named Corridor disease, were distinct from East Coast fever. Schizont and piroplasm parasitoses were low and the parasite could only be transmitted directly from buffalo to cattle; it could not be transmitted between cattle as affected animals usually died before piroplasms appeared (Neitz, 1955; Neitz et al., 1955). Corridor disease was thus considered to be caused by a different theilerial parasite which was named *Theileria lawrencei* (Neitz, 1955).

The symptoms of Corridor disease were similar to those which had been reported by Lawrence (1935; 1936) of a seasonal outbreak of theileriosis which was clinically milder than East Coast fever and associated with the presence of buffalo in Southern Rhodesia (although later cases of this disease (Lawrence, 1937; 1938; 1939) were not always associated with buffalo). Post-mortem results were not typical of those reported for East Coast fever (Lawrence, 1937), schizontes were rarely if ever found and there were also very few piroplasms. These findings, combined with the fact that animals which had recovered from this atypical *Theileria* infection were not immune to an East Coast fever challenge, led Lawrence (1938) to distinguish between East Coast fever and this new form of theileriosis. Mortalities from East Coast fever occurred throughout the year (Edmonds, 1924, cited in Koch, 1990), but in outbreaks of this atypical form of theileriosis mortalities were highest in January and outbreaks stopped before or in March. The disease thus became known as January disease (it has also been called Rhodesian/Zimbabwean malignant theileriosis).

The parasites causing East Coast fever, Corridor disease and January disease were originally attributed to three different pathogenic species: *Theileria parva*, *Gonderia lawrencei* and *G. bovis* (Neitz and Jansen, 1956; Neitz, 1957). *Theileria parva* was
in Lawrence, 1992) had published a description of the piroplasm, the stage of the organism which occurs in erythrocytes and which is infective for ticks. The parasite was thus named *Piroplasma kochi* by Stephens and Christophers in 1903 (cited in Lawrence, 1992). Theliler (1905) renamed the parasite *Piroplasma parvum*; this name was changed to *Theileria parva* by Bettencourt *et al.* in 1907 (cited in Lawrence, 1992). Koch (1903) was also the first to recognize the schizont (or "Koch's bodies"), the stage of the organism which occurs in lymphoid cells.

The effects of the East Coast fever epidemic were ruinous. Commercial transport routes, which relied on ox-wagons, came to a standstill and cattle farmers suffered heavy losses. Mortalities were greater than 90% and thousands of cattle died. A policy of eradication was decided upon in South Africa and the campaign to control and finally to eradicate East Coast fever was to last for more than fifty years. Strict quarantine measures and control of cattle movement effectively stopped the spread of the disease. Infected pastures were cleared of the disease by removing healthy cattle and slaughtering sick animals, then maintaining the area free of cattle for eighteen months. Strict dipping regimes and an intensive surveillance program, involving regular inspection of cattle and examination of blood and spleen smears, were introduced in all potential East Coast fever areas. In later years, outbreaks of East Coast fever were controlled by slaughtering all cattle on infected farms; over 100,000 animals were slaughtered in control operations (Lawrence, 1992). As a result of these rigorous control measures, East Coast fever was eradicated from South Africa between 1948 and 1954 (Anon., 1981), the last case was reported in Southern Rhodesia in 1954 (Adamson, 1954) and in Swaziland in 1960 (Lawrence, 1992). Total mortalities from 1901 to 1960 have been estimated at 1.4 million head (Lawrence, 1992) and the total cost to South Africa at R100 million (Anon., 1981).
subsequent three years this disease spread to Bechuanaland (now Botswana), the Transvaal, Natal, Orange Free State and finally to the Cape of Good Hope, decimating the cattle population as it spread. The Anglo-Boer War (1899-1902) also contributed to the decrease in the cattle population of the region. When peace was re-established in 1902, the numbers of cattle remaining were insufficient to supply the demands of the British army and Boer farmers who were being resettled in the region. To replenish the cattle population in Southern Rhodesia and the Transvaal, large numbers of cattle were imported from countries around the world including Great Britain, the United States of America, Argentina, Australia, Madagascar, and from further north in Africa including Northern Rhodesia (now Zambia), German East Africa (now Tanzania) and Kenya (Lawrence, 1992).

Gray and Robertson (cited in Lawrence, 1992) first reported an outbreak of the disease that would later become known as East Coast fever (ECF) in Southern Rhodesia in 1902, although they mistook it for a virulent form of redwater, caused by *Babesia* (Henning, 1956). It is thought that East Coast fever was introduced by a shipment of cattle from German East Africa (Lawrence, 1992), where the parasite had probably existed for centuries (Norval et al., 1992). East Coast fever also became established in the Lourenço Marques (now Maputo) area and the first outbreak of the disease in South Africa occurred at Komatiport in the eastern Transvaal (Anon, 1981). It spread rapidly along the transport routes further into the Transvaal and into Swaziland. Eventually, Natal, Transkei and finally areas of the eastern Cape Province were affected, despite attempts to prevent the movement of infected cattle into these regions.

Robert Koch, a German microbiologist brought to Rhodesia to consult on East Coast fever, was the first person to recognise that the disease was different from redwater, being caused by a previously unknown parasite (Koch, 1903). In 1898, Koch (cited

---

* These four territories were individual colonies until 1910; at this time they were combined as the Union of South Africa.
CHAPTER 1: INTRODUCTION

Theileriosis is a widespread disease of wild and domestic animals caused by Apicomplexan parasites of the genus *Theileria*. There are two economically important species: *T. annulata* causes tropical theileriosis in areas of North Africa, Mediterranean Europe, the Near and Middle East, India and Central Asia; *T. parva* occurs in parts of Kenya, Uganda, Tanzania, Sudan, Rwanda, Burundi, Zaire, Malawi, Zambia, Zimbabwe, Mozambique and South Africa, where it causes East Coast fever, Corridor disease and January disease. The African buffalo is the reservoir for *T. parva* parasites and transmission to cattle via ticks results in disease (Dolan, 1989).

East Coast fever was eradicated from South Africa in 1954 (Anon., 1981) and today Corridor disease has become the most important form of theileriosis in the country. Although East Coast fever has been eradicated, the vector ticks are still prevalent in South Africa and the cattle population would be highly susceptible to a recurrence of the disease. At present, there is no reliable means of distinguishing between the two forms of the parasite. The aims of this project were therefore to develop DNA probes to distinguish between the causative agents of East Coast fever and Corridor disease.

1.1 History of theileriosis in southern Africa

1.1.1 East Coast fever

At the turn of the century the cattle population in southern Africa had been dramatically reduced by the rinderpest epidemic, which crossed the Zambezi River and spread into Southern Rhodesia (now Zimbabwe) in 1896 (Henning, 1956). In the
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Life cycle of <em>Theileria parva</em> (from Norval et al., 1992)</td>
<td>6</td>
</tr>
<tr>
<td>2.1</td>
<td>Structure of the <em>T. parva</em> ribosomal RNA transcription unit, showing rRNA coding regions and transcribed spacers</td>
<td>21</td>
</tr>
<tr>
<td>2.2</td>
<td>Strategy used to clone and sequence the <em>T. p. lawrencei</em> SSUrRNA PCR amplification product</td>
<td>29</td>
</tr>
<tr>
<td>2.3</td>
<td>Strategy used to clone and sequence <em>T. p. lawrencei</em> ITS and LSUrRNA PCR amplification products</td>
<td>31</td>
</tr>
<tr>
<td>2.4</td>
<td>PCR products of <em>T. p. lawrencei</em> 7344 rRNA genes</td>
<td>35</td>
</tr>
<tr>
<td>2.5</td>
<td>Two ITS PCR products were generated from <em>T. p. lawrencei</em> 7344 with primers ILO 466 and ILO 464</td>
<td>37</td>
</tr>
<tr>
<td>2.6</td>
<td>Recombinant pUC18 clones containing the LSU-1 PCR product</td>
<td>39</td>
</tr>
<tr>
<td>2.7</td>
<td>Screening of recombinant M13 plaques for M13 clones containing LSU-1 insert</td>
<td>40</td>
</tr>
<tr>
<td>2.8</td>
<td>Alignment of <em>T. p. parva</em> and <em>T. p. lawrencei</em> 7344 rRNA transcription units from the 5' end of the small subunit to the approximate 3' end of the large subunit</td>
<td>46-56</td>
</tr>
<tr>
<td>3.1</td>
<td>Putative ITS characterization oligonucleotides for <em>T. p. parva</em> and <em>T. p. lawrencei</em></td>
<td>65</td>
</tr>
<tr>
<td>3.2</td>
<td>Alignment of ITS sequences obtained in this study with <em>T. p. parva</em> Muguga and Uganda ITS sequences</td>
<td>69-71</td>
</tr>
<tr>
<td>3.3</td>
<td>Candidate region for distinction between <em>T. p. parva</em> and <em>T. p. lawrencei</em> genotypes</td>
<td>76</td>
</tr>
<tr>
<td>4.1</td>
<td>PCR products of the variable region of the <em>p67</em> gene generated with primers 613 and 792</td>
<td>89</td>
</tr>
<tr>
<td>4.2</td>
<td>Nucleotide sequence of the variable region of <em>T. parva</em> <em>p67</em> genes, from primer 613 to the end of the 29 base pair intron</td>
<td>91-92</td>
</tr>
<tr>
<td>4.3</td>
<td>Deduced amino acid sequences of partial <em>T. parva</em> <em>p67</em> gene sequences</td>
<td>93</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 2.1 Oligonucleotides used for amplification of tRNA genes .......................... 26
Table 2.2 Sequence details of the T. p. lawrencei 7344 tRNA transcription unit obtained from cloned PCR products .................................................. 38
Table 2.3 T. p. parva and T. p. lawrencei tRNA sequences aligned in fig. 2.8 .............................................................. 45
Table 3.1 Theileria stocks used as sources of DNA from which ITS regions were sequenced .......................................................... 59
Table 3.2 Oligonucleotides used for amplification, sequencing and probing of the T. parva ITS .......................................................... 61
Table 3.3 Number of ITS sequences cloned and sequenced from each T. parva isolate .......................................................... 68
Table 3.4 Probing of amplified T. parva ITS regions with experimental characterization oligonucleotides .......................................................... 77
Table 4.1 Theileria stocks used as sources of DNA from which the variable region of the p67 gene was amplified and sequenced ........ 85-86
Table 4.2 Oligonucleotides used for amplification and sequencing of the variable region of T. parva p67 genes .................................................. 87
C.8.1.2 pUC18 ........................................ 121
C.8.1.3 Dephosphorylation of vectors .................. 121
C.8.1.4 T-vectors ....................................... 121

C.8.2 Ligations ........................................ 122
C.8.3 Preparation and transformation of competent cells .. 122
C.8.3.1 For transformation of M13, pUC and pGEM-T clones
.......................................................... 122
C.8.3.2 For transformation of pMOSBlue-T clones .... 123
C.8.4 Plate analysis ..................................... 124
C.8.4.1 Controls ........................................ 124

C.9 Preparation of DNA from recombinant plasmid clones .. 125
C.9.1 Small-scale isolation of recombinant plasmid clones .. 125
C.9.1.1 Alkaline lysis ................................ 125
C.9.1.2 Wizard™ plasmid minipreps ..................... 125
C.9.2 Large-scale isolation of recombinant plasmid clones .. 126

C.10 Screening of recombinant plaques .................. 127
C.10.1 Plaque lifts ...................................... 127
C.10.2 Hybridisation with PCR product as probe ........ 127

C.11 Preparation of radiolabelled probes and assay of [α-32P] incorporation ................................................................. 128
C.11.1 Random primed labelling of double stranded DNA .. 128
C.11.2 3' Terminal labelling of oligonucleotides .......... 128
C.11.3 Assay of [α-32P] incorporation of radiolabelled probes .. 129

C.12 Sequencing ........................................... 130
C.12.1 Dideoxy chain termination sequencing ............
C.12.1.1 Preparation of templates for sequencing ....... 129
C.12.1.2 Sequencing reactions .......................... 131
C.12.2 Direct thermocycle sequencing of PCR products .. 132
C.12.2.1 End-labelling of the sequencing primer ......... 133
C.12.2.2 Extension/termination .......................... 133
APPENDIX A: SOURCE OF REAGENTS AND CHEMICALS USED IN THIS STUDY ........................................ 108

APPENDIX B: COMPOSITION OF COMMON SOLUTIONS AND MEDIA USED IN THIS STUDY ........................................ 111

APPENDIX C: DETAILS OF MOLECULAR BIOLOGICAL TECHNIQUES USED IN THIS STUDY ........................................ 114

C.1 Purification of nucleic acids ........................................ 114
C.1.1 Phenol: chloroform extraction ........................................ 114
C.1.2 Ethanol precipitation ........................................ 114
C.1.2.1 Of genomic DNA ........................................ 114
C.1.2.2 Of smaller DNA molecules (≤ 15 kbp) .............. 115

C.2 Quantitation of DNA ........................................ 116
C.2.1 Double-stranded DNA ........................................ 116
C.2.2 Single-stranded oligonucleotides ........................................ 116

C.3 Polymerase chain reaction ........................................ 117
C.3.1 The reaction ........................................ 117
C.3.2 Checking for PCR products ........................................ 117

C.4 Dissociation temperature of oligonucleotide primers and probes ........................................ 118

C.5 Restriction digests ........................................ 118

C.6 Filling in recessed 3' or 5' ends and phosphorylation ........................................ 119

C.7 Purification of PCR products ........................................ 119
C.7.1 "Freeze-squeeze": recovery of DNA from agarose gels using liquid nitrogen ........................................ 119
C.7.2 Wizard™ DNA purification system ........................................ 120

C.8 Cloning of PCR products ........................................ 121
C.8.1 Preparation of vectors ........................................ 121
C.8.1.1 M13mp18 and mp19 ........................................ 121

xiii
communication). This was the first T. p. lawrencei clone to be produced and we were fortunate to have been supplied with this cell line.

1.6 Aims of this project

The aims of this project were to sequence the LSU rRNA genes and the transcribed spacer regions of the T. p. lawrencei clone and the p67 sporozoite antigen genes of some South African T. p. lawrencei isolates, in order to identify variable regions that might be used as targets for probes to distinguish between the causative agents of East Coast fever and Corridor disease.
gene of *T. p. bovis* (Boleni), *T. p. parva* and *T. p. lawrencei* (Hluhluwe 3) were identical, although the full-length SSUrRNA sequence of a cloned *T. p. lawrencei* parasite has never been obtained.

None of the polymorphisms detected by the methods described above correlate with cross-protection, nor do they distinguish between the causative agents of East Coast fever and Corridor disease. Another approach needs to be adopted if we are to distinguish between *T. p. parva* and *T. p. lawrencei*. There is a great deal more variation between species in the expansion segments of their large subunit ribosomal RNA (LSUrRNA) genes. In addition, the non-coding spacer region which separates SSUrRNA and LSUrRNA genes, the internal transcribed spacer (ITS), is under low structural and functional constraint, and is consequently more variable than rRNA genes (Noller, 1984). It has recently been shown that improved discrimination between different *Theileria* species can be obtained using oligonucleotide probes derived from LSUrRNA sequences (Bishop *et al.*, 1995). It therefore seemed likely that unique differences between *T. p. parva* and *T. p. lawrencei* might lie within a variable region of the LSUrRNA gene, or in the ITS. Most of the sequence of the *T. p. parva* rRNA transcription unit (including the SSUrRNA and LSUrRNA genes and the ITS) has recently been obtained (Kibe *et al.*, 1994); this provides a basis for the comparison of new *T. parva* rRNA sequences.

1.5.6 Cloned *T. parva* parasites

The extensive polymorphisms which occur between *T. parva* stocks and the presence of more than one genotype in some stocks have necessitated the generation of parasite clones, a number of which have now been produced (Morzaria *et al.*, 1995). Workers at ILRAD (International Laboratory for Research on Animal Diseases) in Kenya have cloned, in a buffalo lymphoblastoid cell line, a parasite which produces the clinical symptoms of Corridor disease (Dolan and Spooner, personal
piroplasm parasitaemias. In addition, results obtained from stocks containing mixed isolates are difficult to interpret.

1.5.5.3 *Ribosomal RNA genes*

Small subunit ribosomal RNA (SSU rRNA) genes have been used in numerous evolutionary studies of many different organisms, resulting in the availability of a great deal of information on the sequences of ribosomal RNAs. Phylogenetically ordered alignments of both large and small subunit rRNA sequences are available from the Ribosomal Database Project (Maidak et al., 1996). A new rRNA sequence can be compared to existing sequences to identify homologies with other organisms and to locate conserved and variable regions within the molecule.

Portions of the nucleotide sequence of rRNA are highly conserved between all organisms, particularly in regions which determine the secondary structure of the molecule (Noller, 1984). Other regions of rRNA genes which are not under pressure to remain unchanged vary between species. Species-specific variations found in rRNA genes, together with the abundance of ribosomes (and therefore rRNA) in cells, can be used to develop very specific and sensitive probes (Waters and McCutchan, 1989).

Allsopp et al. (1993) turned to the *Theileria* SSU rRNA genes in the hope that these might contain less variability than the *Tpr*, and might define broader groupings of similar *T. parva* parasites. The entire sequence of the *T. parva* Muguga SSU rRNA gene was obtained and a hypervariable region was identified. This region was sequenced from a range of *T. parva* stocks isolated from cattle and buffalo, from *T. annulata, T. mutans, T. buffeli* and *T. tauronagi* samples, and also from an unknown cloned parasite from a buffalo in Kenya which was designated *Theileria* sp. (buffalo). Oligonucleotides were designed for use as probes to distinguish between these six species. However, the sequence of the variable region of the SSU rRNA
1.5.4.5 Two dimensional gel electrophoresis

Stock-specific differences in infection-specific proteins have been detected using two-dimensional gel electrophoresis (Sugimoto et al., 1989) and two-dimensional Western blotting (Kishima et al., 1995). However, these techniques are difficult to perform and are often difficult to interpret.

1.5.5 Molecular genetic characterisation of T. parva

1.5.5.1 Polymorphism detected by DNA probes

Extensive polymorphism has been detected amongst T. parva stocks using repetitive DNA probes (Conrad et al., 1987b; Allsopp and Allsopp, 1988). Allsopp et al. (1989) obtained the first unequivocal discrimination between different T. parva stocks. These workers amplified a variable region of the T. parva repetitive (Tpl') gene family (Baylis et al., 1991) using the Polymerase Chain Reaction (PCR); two oligonucleotide probes were used to distinguish between PCR products. Telomeric and ribosomal DNA probes have been used to characterise T. parva stocks (Bishop et al., 1993a; Kibe et al., 1994). However, to date, oligonucleotide probes based on sequence differences in the Tpr (Bishop et al., 1993a) and rDNA (Kibe et al., 1994) can only distinguish between some T. parva stocks in a positive/negative manner.

1.5.5.2 Genomic polymorphism

Genomic polymorphisms between T. parva stocks and clones have been detected by pulsed field gel electrophoresis (PFGE) of Sfi I-digested DNA (Morzaria et al., 1990) and by using PCR with single arbitrary primers (Bishop et al., 1993b). However, these techniques require purified parasite DNA free from host DNA contamination; this is difficult to obtain, especially from isolates which do not produce high
amongst six enzymes in different *T. parva* stocks, but concluded that the method was not practical for large scale surveys.

### 1.5.4.2 Protein banding patterns using SDS-PAGE

Protein banding patterns of purified theilerial piroplasms and schizonts of different *T. parva* stocks have been studied by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) (Allsopp, 1981; Creemers, 1983). Creemers (1983) was able to detect minor differences in protein bands between *T. p. parva* and *T. p. lawrencei*. Problems experienced in these studies included the difficulty of obtaining parasites free of contaminating bovine proteins.

### 1.5.4.3 Schizont antigens

Differences between schizont antigens have been identified in different *T. parva* stocks by probing Western blots of the relevant proteins with anti-schizont monoclonal antibodies (Minami *et al.*, 1983). Four of these MAbs recognised a polymorphic, immunodominant molecule (Shapiro *et al.*, 1987) which displayed polymorphisms in both size and expression of antibody epitopes among different stocks of *T. parva* (Toye *et al.*, 1991; 1995).

### 1.5.4.4 Sporozoite antigens

The major antigenic determinants of *T. parva* sporozoites have been identified using an antiserum directed against the sporozoite stage of the parasite (Iams *et al.*, 1990). The major target of sporozoite neutralising antibodies, p67, has been cloned and characterised (Nene *et al.*, 1992). The sequence of the gene coding for p67 is conserved amongst *T. parva* parasites isolated from cattle (Nene *et al.*, 1996), but is variable in *T. parva* parasites isolated from buffalo (V. Nene, pers. comm).
1.5.2 Cross immunity trials

Cross-protection between different stocks of *T. parva* can be tested by immunising cattle with one stock and challenging immunised animals with other stocks. Stocks that are able to cause disease in immunised animals are considered immunologically distinct (Norval *et al.*, 1992). However, cross immunity trials are time consuming and, because of the number of animals involved, very expensive. In addition, they are sometimes difficult to interpret (Irvin, 1987), probably because some stocks contain more than one *T. parva* immunotype. There is thus a need for more practical and reliable laboratory methods to characterise immunologically distinct *T. parva* populations.

1.5.3 Screening with monoclonal antibodies

A number of monoclonal antibodies (MAbs) directed against schizont antigens have been used to characterise *T. parva* stocks (Pinder and Hewett, 1980; Minami *et al.*, 1983; Irvin *et al.*, 1983; Conrod *et al.*, 1987a, 1989; Koch *et al.*, 1988; Bishop *et al.*, 1994). *T. parva* schizont-infected cells derived from cell cultures are detected in indirect fluorescent antibody tests using a panel of MAbs. These MAbs detect enormous diversity amongst *T. parva* stocks, but they lack sensitivity and some parasites may be overlooked, as certain subpopulations can be selected during the establishment of cultures (Norval *et al.*, 1992). The results are also rather subjective, difficult to interpret and may not be reproducible (B. A. Allsopp, pers. comm.).

1.5.4 Analysis of proteins

1.5.4.1 Enzyme analysis

Negligible polymorphism has been found between most *T. parva* isoenzymes (Melrose *et al.*, 1980; Musisi *et al.*, 1981). Allsopp *et al.* (1985) found variations...
1.5 Characterisation of *T. parva* stocks

Some cattle recover from infection with *T. parva*; these animals are resistant to challenge with the parasite which caused the infection, but immunity after challenge with an heterologous parasite varies from partial to none (Irvin, 1987). Attempts to immunise cattle against theileriosis by the infection and treatment method have produced similar results (Irvin, 1987; Mutugi et al., 1989). It is increasingly obvious that there is enormous antigenic variation between different stocks of *T. parva* and that mixed infections are common in the field. Therefore, in order to understand the different patterns of theileriosis in southern and eastern Africa and to develop effective vaccines against theileriosis, precise characterisation of different *T. parva* isolates has become necessary. Several different techniques have been used to characterise *T. parva* isolates.

1.5.1 Pathogenicity

The clinical symptoms, parasitological characteristics and pathology of infections with different *T. parva* stocks can be monitored *in vivo*. Differences between *T. p. parva*, *T. p. lawrencei* and *T. p. bewes* are displayed in the length of the prepatent period, time to fever, duration of fever, time to appearance of piroplasms, levels of schizont and piroplasm parasitoses and time to death or recovery. Such pathological features are generally too variable to be of value in parasite strain characterisation, but sufficiently consistent differences have resulted in the trinomial classification system for subspecies of *T. parva* (Irvin, 1987).

However, the severity of disease is dependent on the dose of sporozoites applied (Dolan *et al*., 1984). Thus, a standard dose of sporozoites must be used for infectivity assays and, although it is possible to count total numbers of sporozoites, it is difficult to determine numbers of infective sporozoites (Irvin, 1987).
present, but there is as yet no evidence to suggest that these antelope are involved in the epidemiology of theileriosis in southern Africa (H. Stoltz, pers. comm.).

The movement of buffalo from the large herds in the Kruger National Park is prohibited as both foot and mouth disease and Corridor disease are endemic there. In the Natal parks, buffalo are free from foot and mouth disease but, because Corridor disease is endemic in these areas, the relocation of these animals is also prohibited. The major source of animals free from both diseases has previously been a relatively small herd at the Addo Park in the eastern Cape (Stoltz, 1989). However, these disease-free buffalo are very expensive because of their limited numbers, there are not enough animals at Addo to meet the increased demand and the stock is limited in genetic diversity.

A situation has thus developed where unscrupulous farmers may be tempted to circumvent the regulations and move buffalo from prohibited areas. An outbreak of Corridor disease occurred in the Warmbaths area in early 1994, as a result of such an illegal translocation of buffalo from the foot and mouth control area near the Kruger National Park (Kotze et al., 1994). In addition, it is feared that, with the ever-expanding human population and the consequently increasing needs for grazing for growing herds of cattle, particularly in areas bordering the game reserves, there might be a relaxation of the strict control measures that have previously been applied. The combination of these factors could lead to an increase in the number of Corridor disease outbreaks in the future.

Thus there is an urgent need in South Africa for practical and reliable methods to characterise and distinguish between *T. parva* parasites in the field. It would be very useful to be able to identify *Theileria* parasites in buffalo and to identify *T. p. lawrencei*-free animals for translocation. Also, it is of paramount importance to know whether there is a danger of a recrudescence of East Coast fever from the *T. parva* population that remains in buffalo in South Africa.
confined to well fenced game reserves, and rigorous quarantine and tick control measures are enforced in areas bordering these reserves. These measures have also contributed to preventing the spread of Corridor disease (Stoltz, 1989).

1.4.2 The threat of East Coast fever

Although East Coast fever has been eradicated in South Africa, the vector ticks are still prevalent and the cattle population would be highly susceptible to a recurrence of the disease, which could occur if the parasite was reintroduced into the country from an endemic area. In addition, there is a historical fear that T. p. lawrencei may be able to 'transform' to T. p. parva. Barnett and Brocklesbury (1966) first reported 'transformation' in Kenya during the passage of a T. p. lawrencei isolate through cattle. However, no evidence for 'transformation' has ever been found in South Africa (Bigalke et al., 1976; Potgieter et al., 1988). In Central and East Africa, buffalo carry mixed populations of theilerial parasites (Norval et al., 1992). It seems likely that what has been described as transformation may occur as a result of the reservoir host harbouring a mixed infection, including both T. p. lawrencei and T. p. parva. Passage in cattle may then result in the selection of a subpopulation of parasites characteristic of T. p. parva.

1.4.3 Relocation of buffalo in South Africa

The rising popularity of commercial game farms and private game reserves in South Africa in recent years has led to an increasing demand for buffalo. In order for buffalo to be translocated, the animals must be certified free of both foot and mouth disease and Corridor disease to prevent the spread of these virulent diseases. In addition, Stagg et al. (1994) have shown that, when passaged through defassa waterbuck (Kobus defassa), a subpopulation of parasites that resembles T. p. parva can be selected from the original T. parva population in buffalo. Although defassa waterbuck do not occur in South Africa, common waterbuck (K. ellipsiprymnus) are
1.3.2 Other *Theileria* species

Other *Theileria* species occurring in South Africa include the mildly pathogenic *T. mutans* (Theiler, 1906) and *T. taurotragil* (De Vos and Roos, 1981a), and the non-pathogenic *T. vellifera* (Berger, 1979); these species are of little or no economic importance (Stoltsz, 1989). However, they produce schizonts and piroplasms and in mixed infections they may complicate the diagnosis of *T. p. lawrencei* infections.

*T. mutans* is transmitted by ticks of the genus *Amblyomma*; in South Africa *A. hebraeum* has been shown to be an efficient vector (De Vos and Roos, 1981b). Pathogenic strains of *T. mutans* have been reported in Kenya (Moll *et al.*, 1986) and in Zambia (Musisi *et al.*, 1984). Although *T. mutans* appears to be widely distributed in both cattle and buffalo in South Africa, all South African isolates thus far observed appear to be non-pathogenic in cattle (Stoltsz, 1989).

*T. taurotragil*, transmitted by *R. appendiculatus*, has been implicated in bovine cerebral theileriosis in South Africa (De Vos, 1982). *T. taurotragil* is widespread in Zimbabwe and seriously complicates investigations on *T. p. bovis* (Koch, 1990).

1.4 Current epidemiology of theileriosis in South Africa

1.4.1 Corridor disease

Corridor disease has become the most important form of theileriosis in South Africa, since the eradication of East Coast fever and presumably of its causative agent, *T. p. parva*. Outbreaks of Corridor disease occur periodically on farms bordering the Kruger National Park and the Natal parks, Hluhluwe and Umfolozi; such outbreaks are always associated with the presence of the African buffalo (Bigalke *et al.*, 1976). This animal is also a reservoir of the highly virulent foot and mouth disease virus. In order to prevent the spread of foot and mouth disease in South Africa, buffalo are
of parasites; this would have resulted in subclinical infections and carrier animals which spread the disease.

It has recently been recommended (Anon., cited in Norval et al., 1992) that the use of the trinomial system be discontinued as there are no real biological grounds for distinction between the three subspecies. Different *T. p. parva* stocks have been found to display a wide range of schizont and piroplasm parasites and pathogenicities (Norval et al., 1992) and there are varying degrees of cross-immunity between stocks of *T. p. parva*, *T. p. lawrencei* and *T. p. bovis* (Irvin et al., 1989). It has therefore been suggested that *T. parva* is the only justifiable species and that *T. parva* parasites should rather be classified according to their host of origin, i.e. cattle-derived or buffalo-derived (Anon., cited in Norval et al., 1992). However, this nomenclature raises a serious question. How should one designate a stock isolated from a bovine if the infection had been received from a tick which had become infected from a buffalo? The stock could conceivably be genetically identical to the original stock in the buffalo, although experience suggests that this would not be likely because of genetic recombination in the tick. In any event the different labels would merely reflect the investigator's (unavoidable) ignorance of the provenance of the parasite.

For the purposes of this thesis, which deals with many different *T. parva* stocks from eastern and southern Africa which cause East Coast fever, Corridor disease or January disease, it will be simpler, more convenient and easier for the reader to follow, if the trinomial system is used. The reader should therefore note that, in cases where we had no information about the clinical symptoms caused in cattle by specific *T. parva* stocks we have simply indicated the host of origin but without using the cattle-derived and buffalo-derived terminology, since that has implications of exclusivity to one host. In cases where the clinical symptoms induced in cattle were known we have used *T. p. parva* for East Coast fever-inducing stocks and *T. p. lawrencei* for Corridor disease-inducing stocks.
Corridor disease is also a severe and normally fatal disease in infected cattle. There is a marked lympho-proliferative phase, but the lympho-destructive stage is less severe than in East Coast fever. Clinical symptoms include characteristically swollen lymph nodes and the development of pyrexia. Pulmonary oedema is not necessarily seen, whereas subcutaneous oedema is common. One of the main features distinguishing Corridor disease from East Coast fever is that very few schizonts and piroplasms are produced in cattle. As a result, *T. p. lawrencei* parasites transmitted directly from buffalo, as the low numbers of piroplasms produced in cattle result in a low rate of infection in ticks and therefore poor transmission between cattle. The tick vectors are *R. appendiculatus* and *R. zambeziensis*, but there is evidence to suggest that the latter might be the preferred vector (Blouin and Stoltsz, 1989).

Koch (1990) has described contemporary January disease as follows:

"January disease or theileriosis is a seasonal disease of Zimbabwe, characterized by lymph node swelling and oedema of the lungs, usually with a low morbidity and a high mortality of the affected cattle. Outbreaks are usually associated with high numbers of the vector, adult *Rhipicephalus appendiculatus*. *Rhipicephalus zambeziensis* might sometimes be the vector. Calves in the field have more resistance to the disease than adult cattle. A number of cattle will remain infective to ticks after recovery from infection. The frequency of schizonts in the lymph nodes of affected cattle varies, but the percentage of erythrocytes containing piroplasms is always low."

Koch (1990) postulates that the parasite which causes January disease was a strain of *T. parva* which was selected from the East Coast fever complex introduced into southern Africa. Tick control measures as applied in Rhodesia may have suppressed the disease but allowed infection of ticks and susceptible cattle with small numbers
Fig. 2.3: Strategy used to clone and sequence *T. p. lawrencei* ITS and LSU rRNA PCR amplification products.
2.2.5 Cloning of the LSU rRNA and ITS amplification products

Shotgun cloning (Baird et al., 1987) was used to clone random fragments of the LSU rRNA and ITS PCR products into M13mp18 for sequencing, as illustrated in fig. 2.3.

2.2.5.1 Preparation of recombinant pUC18 clones

After the amplification of each fragment the products from three PCRs were pooled and concentrated by precipitation with sodium acetate and ethanol (Appendix C.1.2.2). The termini of the PCR products were filled in and kinased as described in Appendix C.6. The resulting blunt-ended fragments were purified from residual enzymes, oligonucleotide primers and unincorporated nucleotides by gel-purification (Appendix C.7.1). Two ITS PCR products were obtained and were separated at this stage by fractionation on a 1% agarose/TBE gel and excision of the separated fragments. The shorter of the two ITS PCR products was designated ITS-S and the longer one ITS-L. The PCR products were gel-purified (Appendix C.7.1) and cloned into Sma I-cut and dephosphorylated pUC18 (Appendix C.8). White colonies were picked and minipreparations of recombinant plasmid DNA were made by the alkaline lysis method (Sambrook et al., 1989) (Appendix C.9.1.1). Plasmid DNA was digested (Appendix C.5) with Eco RI and Bam HI and checked by electrophoresis to ensure that the correct inserts had been cloned. A QIAGEN tip 100 (Appendix C.9.2) was used to prepare DNA from three recombinant pUC18 clones of each of the LSU rRNA PCR products and from six clones of each of the ITS PCR products.

2.2.5.2 Sonication of recombinant pUC18 clones

For each cloned ITS and LSU rRNA, the recombinant pUC18 clones were pooled and 10 µg of purified recombinant plasmid DNA was sonicated for 120 seconds at an output of 160 watts in a Branson Cuphorn Sonifier 250. The sonicates were
Fig. 2.2: Strategy used to clone and sequence the *T. p. lawrencei* SSUrRNA PCR amplification product
2.2.3.4 Amplification of rRNA genes

The PCR was performed as described in Appendix C.3. Annealing was performed at 60°C for primers A and B and the ITS primer pair, and at 65°C for both of the large subunit primer pairs. Cycling conditions were: 72°C for 1.5 minutes, 95°C for 1 minute and 60°C (small subunit and ITS) or 65°C (large subunit) for 1 minute, for 30 cycles. A positive *T. p. parva* Muguga LN/MI/1 and a negative calf thymus DNA (Sigma) control were included with each PCR performed.

2.2.4 Cloning of the SSUrRNA amplification product

A scheme showing the strategy used to clone the amplified SSUrRNA gene into the sequencing vectors M13mp18 and mp19 (Messing, 1983) is shown in fig. 2.2.

5 μl aliquots of the SSUrRNA product were initially digested with the enzymes included in the primers to determine which enzymes could be used for directional cloning (i.e., those that did not cut the PCR product internally). Restriction digests were performed as described in Appendix C.5. *Bam* HI and *Eco* RI were the enzymes of choice for cloning the SSUrRNA gene.

The products from three PCRs were pooled, concentrated by precipitation with ammonium acetate and propanol (Appendix C.1.2.2), and double-digested with *Bam* HI and *Eco* RI (Appendix C.5). Digested products were separated from residual enzymes, oligonucleotide primers and unincorporated nucleotides by gel-purification using a modification of the "freeze-squeeze" technique (Koenen, 1989) (Appendix C.7.1), they were ligated into *Bam* HI- and *Eco* RI-cut, dephosphorylated M13mp18 or mp19 and transformed into competent *E. coli* TG1 cells (Appendix C.8). Colourless, recombinant phage plaques were picked and single-stranded templates were prepared for sequencing as described in Appendix C.12.1.1.
2.2.3.2 Primers for amplification of the large subunit

Six eukaryotic LSU rRNA sequences, available at the time, were extracted from the international databases EMBL and Genbank. These included two protozoan LSU rRNA sequences, *Tetrahymena pyriformis* (Engberg et al., 1990) and *Procratium nicans* (Leeners et al., 1989), *Caenorhabditis elegans* 28S rRNA (Ellis et al., 1986) and three mammalian 28S rRNA sequences, *Mus musculus* (Hussouna et al., 1984), *Rattus norvegicus* (Chan et al., 1983; Hadjiolov et al., 1984) and *Homo sapiens* (Gonzalez et al., 1985). CLUSTALV (Higgins et al., 1992) was used to align these 28S sequences with the *T. p. parva* LSU rRNA sequence (Kibe and Bishop, pers. comm.); two Theileria-specific primer pairs, ILO 465 and ILO 467, and NC1 and NC2, were identified for amplification of the large subunit. ILO 465 and ILO 467 were designed to amplify the 5' half of the large subunit (designated LSU-I); NC1 and NC2 the 3' half (designated LSU-II). Primer sites are illustrated in fig. 2.1. ILO 465 anneals to the coding strand approximately 140 bp downstream of the 5' end of the large subunit. ILO 467 is complementary to the non-coding strand and is located in a *Theileria*-conserved region approximately 1.5 kbp downstream from ILO 465. NC1, complementary to the coding strand, lies approximately 90 bp upstream from ILO 467. NC2 anneals to the non-coding strand at the 3' end of the large subunit gene. Details of PCR primers are in table 2.1.

2.2.3.3 Primers for amplification of the internal transcribed spacer

Primers for amplification of the ITS were designed in regions which were specific for *Theileria* at the 3' end of the small subunit and at the 5' end of the large subunit (Kibe et al., 1994). The 3' end of ILO 466 anneals to the coding strand 90 bp upstream of the 3' end of the SSUrRNA gene. ILO 464 is complementary to the non-coding strand approximately 140 bp downstream of the 5' end of the large subunit (fig. 2.1) (ILO 464 is the reverse complement of primer ILO 465, which was used to amplify LSU-I).
<table>
<thead>
<tr>
<th>Primer</th>
<th>Length (bp)</th>
<th>GC content</th>
<th>Strand</th>
<th>T_m (°C)</th>
<th>Sequence 5' - 3'</th>
<th>Reference/Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>35</td>
<td>19</td>
<td>+</td>
<td>69.0</td>
<td>CCB AAT TCG TCG ACA ACC TGG TTG ATC CTG   CCA GT</td>
<td>For use with primer B to amplify SSU rRNA gene (Sogin, 1990)</td>
</tr>
<tr>
<td>B</td>
<td>39</td>
<td>22</td>
<td>-</td>
<td>71.8</td>
<td>CCG GGG ATC CAA GCT TGA TCC TGC TGC AGB   TTC ACC TAC</td>
<td>For use with primer A to amplify SSU rRNA gene (Sogin, 1990)</td>
</tr>
<tr>
<td>ILQ 464</td>
<td>22</td>
<td>11</td>
<td>+</td>
<td>71.0</td>
<td>GGS AAC GTC TAG GSA AAG TTT G</td>
<td>For use with ILQ 464 to amplify internal transcribed spacer (Kibb et al., 1994)</td>
</tr>
<tr>
<td>ILQ 466</td>
<td>25</td>
<td>11</td>
<td>-</td>
<td>70.7</td>
<td>CCG TTG GCA AAT TCA ACG CAT TTT C</td>
<td>For use with ILQ 466 to amplify internal transcribed spacer (Kibb et al., 1994)</td>
</tr>
<tr>
<td>ILQ 463</td>
<td>25</td>
<td>11</td>
<td>+</td>
<td>70.7</td>
<td>SAA AAT CCG TTS AAT TGG CCA ACG G</td>
<td>For use with ILQ 467 to amplify 3' half of LSU rRNA gene. Reverse complement of ILQ 464</td>
</tr>
<tr>
<td>ILQ 467</td>
<td>25</td>
<td>12</td>
<td>-</td>
<td>72.4</td>
<td>TGC GTC GAS AAG TTC ACT TCT GTT G</td>
<td>For use with ILQ 468 to amplify 3' half of LSU rRNA gene</td>
</tr>
<tr>
<td>NCI</td>
<td>23</td>
<td>14</td>
<td>+</td>
<td>75.3</td>
<td>CSG CAC GTC AAC TGC GTS TGA G6</td>
<td>For use with NCI2 to amplify 3' half of LSU rRNA gene</td>
</tr>
<tr>
<td>NCI2</td>
<td>25</td>
<td>15</td>
<td>-</td>
<td>77.3</td>
<td>GAS CAC CTC GGG TAG AAT CTC ACG G</td>
<td>For use with NCI to amplify 3' half of LSU rRNA gene</td>
</tr>
</tbody>
</table>

*Temperature at which 50% of the probe / target DNA duplexes are dissociated. See Appendix C.4 for calculation of dissociation temperatures.
with 40 ml of 0.1% (v/v) nonidet P40 in a 50 ml polypropylene tube, and the cells were pelleted by centrifugation at 1600g for 20 minutes. The pelleted cells were lysed in 5 ml of lysis buffer [10 mM Tris-HCl pH 8.0, 10 mM NaCl, 10 mM EDTA, 1% w/v SDS] at 70°C for five minutes. Most of the protein was degraded initially by incubation at 37°C overnight, with proteinase K (Boehringer Mannheim) added to the mixture to a final concentration of 0.1 mg/ml. After the addition of 1 ml of ANB [10 mM sodium acetate, 0.1 M sodium chloride, 1 mM EDTA, 0.5% SDS], residual proteins were removed by extraction with phenol and chloroform (Sambrook et al., 1989) (Appendix C.1.1). DNA was recovered by precipitation with sodium acetate and ethanol (Sambrook et al., 1989) (Appendix C.1.2.1) and dissolved in 300-500 μl of TE (Appendix B). The concentration of the purified DNA was determined spectrophotometrically (Appendix C.2.1).

2.2.3 Polymerase chain reaction (PCR) amplification of rRNA genes

The polymerase chain reaction (PCR) (Saiki et al., 1988) was used to amplify the T. p. lawrencei 7344 rRNA transcription unit from the mixture of bovine and parasite DNA extracted from cell culture. Oligonucleotide primers were synthesized by the Biochemistry Department, University of Cape Town.

2.2.3.1 Primers for amplification of the small subunit

Primers A and B, designed to amplify and clone eukaryotic 16S-like rRNA genes (Medlin et al., 1988; Sogin, 1990), were used to amplify the T. p. lawrencei 7344 SSUrRNA gene. Primer A anneals to the coding strand at the 5' end of the small subunit, while primer B is complementary to the non-coding strand at the 3' terminus (fig. 2.1). Primer A contains Eco RI and Sal I restriction sites; primer B includes sites for Smal, Bam HI, Hind III and Pst I. Details of PCR primers are in table 2.1.
genetic studies (reviewed in Hillis and Dixon, 1991). It therefore seemed likely that differences between *T. p. parva* and *T. p. lawrencei* might lie within a variable region of the LSUrRNA gene, or in the ITS. Most of the sequence of the *T. p. parva* rRNA transcription unit (including the SSUrRNA and LSUrRNA genes and the ITS) has recently been obtained (Kibe *et al.*, 1994); this provides a basis for the comparison of new *T. parva* rRNA sequences.

In this chapter, I describe the sequencing of the transcription unit of the cloned *T. p. lawrencei* parasite.

### 2.2 Materials and methods

#### 2.2.1 Parasites


*T. p. parva* Muguga LN/MI/I was used as a positive control in PCRs. The *T. p. parva* Muguga stock was isolated from Kilambu district, Kenya (Brocklesby *et al.*, 1961), and causes East Coast fever. Muguga LN/MI/I was obtained from schizont-infected lymph node material from a bovine infected at ILRAD with *T. p. parva* Muguga (T. Dolan, pers. comm.).

#### 2.2.2 DNA extraction

Genomic DNA, which includes both bovine and *T. p. lawrencei* DNA, was extracted from schizont-infected lymphoblastoid cell cultures. Ten ml of culture was washed...
Species-specific variations found in rRNA genes, together with the abundance of ribosomes (and therefore rRNA) in cells, can be used to develop very specific and sensitive probes. This technique has been successfully applied in distinguishing between species and strains of various Apicomplexan parasites, e.g. *Plasmodium* (Waters and McCutchan, 1989), *Babesia* (Dalrymple, 1990; Persing et al., 1992; Allsopp, 1994) and *Theileria* (Allsopp et al., 1993; Allsopp, 1994).

Allsopp et al. (1993) sequenced a hypervariable region of the small subunit rRNA (SSUrRNA) gene from various *T. parva* stocks isolated from cattle and buffalo, from *T. annulata*, *T. mutans*, *T. buffeli* and *T. tauronigrum* samples, and also from an unknown cloned parasite from buffalo in Kenya which was designated *Theileria* sp. (buffalo). Oligonucleotides were designed for use as probes to distinguish between these six species. The *T. parva* probe did not distinguish between *T. p. bovis* (Boleni), *T. p. parva* and *T. p. lawrencei* (Hluhluwe 3) - the variable region of their SSUrRNA gene sequences was identical. The full length SSUrRNA gene sequence of a cloned *T. p. lawrencei* parasite had not been obtained when this work began, although it seemed likely that it would be identical. It was necessary, however, to confirm this experimentally.

Assuming that the SSUrRNA sequences of the two parasites were found to be identical, we would need to look for a region of greater variability in order to distinguish *T. p. parva* and *T. p. lawrencei* parasites. There is a great deal more variation in the expansion segments or divergent (D) domains of large subunit rRNA (LSUrRNA) genes (Noller, 1984; Hillis and Dixon, 1991) and it has recently been shown that improved discrimination between different *Theileria* species can be obtained using oligonucleotide probes derived from LSUrRNA sequences (Bishop et al., 1995). In addition, ITS regions, which are non-coding spacer regions under low structural and functional constraint, accumulate mutations at a higher rate and are consequently more variable than rRNA gene coding regions. Variations in ITS regions have been used to identify species or strains, and as markers in population
by the ITS1 and the ITS2; finally there is the large subunit rRNA gene. The transcription unit is transcribed as a single precursor RNA molecule (45S); before leaving the nucleus the primary transcript is cleaved to release the individual rRNA molecules. The leader and most of the ITS regions are degraded to nucleotides in the nucleus.

2.1.3 Using rRNA genes to distinguish between closely related species

Nucleotide sequence analyses of both the small and large subunit rRNAs have shown that parts of the sequences are highly conserved between all organisms, particularly in the regions which determine the secondary structural core of the molecule (Gerbi, 1986). The features of rRNA structure that have been conserved in evolution are presumably needed for rRNA function in the ribosome; any changes which deleteriously affect this function would be harmful and selected against. Evolutionarily conserved segments within the rRNAs of many different organisms are correlated with important functions of ribosomes, including sequences important for interactions with tRNA and mRNA, and ribosomal subunit association. Other regions of the rRNA molecule which are not under such constraints are variable. It is precisely these characteristics that make rRNAs good candidates for a number of applications, including phylogenetic studies and the development of species-specific probes. Small subunit ribosomal RNA (SSUrRNA) genes have been used in numerous evolutionary studies of many different organisms, resulting in a great deal of information on the sequences of ribosomal RNAs. Phylogenetically ordered alignments of both large and small subunit rRNA sequences are available from the Ribosomal Database Project (Maidak et al., 1996). A new rRNA sequence can be compared to existing sequences to identify homologies with other organisms and to locate conserved and variable regions within the molecule for the development of amplification oligonucleotides and characterisation probes.
Fig. 2.1: Structure of the *T. parva* ribosomal RNA transcription unit, showing rRNA coding regions (open bars) and transcribed spacers (solid bars). ETS: external transcribed spacer. ITS: internal transcribed spacer. Arrows indicate approximate positions of primers used to amplify rRNA genes.
CHAPTER 2: SEQUENCE OF THE RIBOSOMAL RNA TRANSCRIPTION UNIT OF A CLONED T. P. LAWRENCEI PARASITE

2.1 Introduction

2.1.1 Ribosomes and ribosomal RNA genes

Ribosomes are complex ribonucleoprotein organelles, responsible for translating the information encoded by messenger RNA and synthesising proteins (Alberts et al., 1989). The structure of ribosomes is similar in prokaryotes and eukaryotes, and comprises a small and a large subunit which consist of ribosomal RNA (rRNA) and protein molecules. Because the different rRNA species were initially separated by density gradient centrifugation, they are referred to by their sedimentation coefficients measured in Svedberg units (S). In eukaryotes, the small subunit of the ribosome is made up of 17S or 18S rRNA and a number of proteins. The large subunit consists of three rRNA molecules, the 5S, 5.8S and 26S or 28S, and a number of proteins.

2.1.2 Organisation of rRNA genes

The general pattern of rRNA gene organisation is identical in most eukaryotes. Multiple copies of the genes coding for rRNA are organised in a tandemly arranged series (Long and Dawid, 1980). The repeating units which make up the tandem clusters can be divided into the transcription unit, which is transcribed into rRNA, and the non-transcribed spacer, which separates transcription units. The transcription unit (fig. 2.1) is made up as follows: firstly there is a 5' leader sequence, the external transcribed spacer (ETS); this is followed by the small subunit rRNA gene; thirdly we find the internal transcribed spacer (ITS), which contains the 5.8S gene flanked
SSUrRNA sequence analysis can be used as a very effective tool to elucidate phylogenetic relations between organisms, but this technique cannot always be used to distinguish between species. Therefore, although the SSUrRNA genes of *T. p. parva* Muguga and the cloned *T. p. lawrencei 7344* are identical, this does not necessarily mean that they are the same species. The sequence similarities between the SSUrRNA and LSUrRNA genes of *T. p. parva* Muguga and *T. p. lawrencei 7344* do, however, indicate that these two organisms are very closely related.

2.3.2.2 *The internal transcribed spacer*

The sizes of the two ITS sequences from the 3' end of the SSU to the putative 5' end of the LSU were 1120 bp and 1070 bp. These sequences have been submitted to Genbank with accession numbers U03601 (ITS-L) and U03602 (ITS-S). There were three regions of variability between the two sequences, located just downstream of the SSUrRNA gene and on either side of the 5.8S gene. Deletions, insertions and substitutions were all found, and both sequences differed from the *T. p. parva* Muguga and Uganda ITS sequences (Kibe et al., 1994), although ITS-S tended to be more similar to those than ITS-L. The ITS was the first region in the *T. p. lawrencei* and *T. p. parva* rRNA transcription units in which major sequence variations were identified.

The 5.8S gene sequence was highly conserved between the two *T. p. lawrencei 7344* rRNA units and those of *T. p. parva* Muguga and *T. p. parva* Uganda. This is different from the situation in *Plasmodium falciparum* in which four different 5.8S genes have been found (Snippen-Lentz et al., 1990). Only one of these is transcribed in the erythrocytic stage of the parasite.

The presence of two polymorphic ITS sequences in the cloned *T. p. Lawrencei* parasite suggests the existence of at least two rRNA coding units in *T. parva*. Kibe et al. (1994) showed that there were two rRNA transcription units in *T. p. parva*.
Muguga LSUrRNA gene sequence have insertions of 3 and 6 bp when compared to that of T. p. lawrencei 7344 (fig. 2.8).

The fact that the SSUrRNA genes of T. p. parva Muguga and a cloned T. p. lawrencei are identical and their LSUrRNA genes are very similar would seem to indicate that these two organisms belong to the same species. However, although SSUrRNA sequence analysis has proven very useful for establishing phylogenetic relationships across a broad range of taxonomic levels, there are no universally applicable criteria for deciding how many sequence differences in SSUrRNA sequences are sufficient to justify establishment of a new species (Stackebrandt and Goebel, 1994). In fact, studies to date indicate that "the resolving power of this method dissipates somewhere near the species level" (Wilson, 1995). SSUrRNA sequence analyses of certain organisms fail to distinguish between recognized species; examples include Mycobacterium avium and M. paratuberculosis which have identical SSUrRNA sequences (Rogall et al., 1990; Stahl and Urbance, 1990) and Brithia phagocytophila and E. equi which have a single base pair difference in their SSUrRNA genes (Maidak et al., 1996). On the other hand, unique SSUrRNA sequences have been found amongst serovars of the same species, e.g., Franciseilla tularensis (Forsman et al., 1994), and in Escherichia coli there are five different SSUrRNA sequences with sequence differences ranging up to 1.2% (Maidak et al., 1996). To further complicate the situation, more than one copy of a SSUrRNA gene can be present within the same organism, e.g., in Halobacterium marismortui the two rRNA copies are different for 5.1% of their sequence (Maidak et al., 1996). In a recent large-scale taxonomic study of the Mycobacteria in which DNA-DNA hybridisation, SSUrRNA sequences, antibody binding, and chemotaxonomic and phenotypic characteristics were compared, Wayne et al. (1996) could identify "no single 16S rRNA interstrain nucleotide sequence difference value that unequivocally defined species boundaries".

42
CLUSTALV (Higgins et al., 1992) was used to align sequences obtained from all T. p. lawrencei 7344 PCR products with the sequences of the T. p. parva Muguga rRNA transcription unit (Allsopp et al., 1993; Kibe et al., 1994; Kibe and Bishop, pers. comm.). The sequences aligned are listed in table 2.3 (p. 45). The alignment is shown in fig. 2.8 (pp. 46-55). The SSUrRNA gene is located between columns 1 and 1742. The positions of the 5.8S and LSUrRNA genes were estimated by comparison of our sequence with other 5.8S and LSUrRNA sequences, so start and end positions are approximate. The approximate location of the T. parva 5.8S gene (between columns 2367 and 2546 in fig. 2.8) was determined by comparison with two other protozoan 5.8S sequences, Prorocentrum micans (Maroteaux et al., 1985) and Giardia ardeae (van Kaftlen et al., 1991a; 1991b). The estimated start of the 5.8S gene was 14 nucleotides upstream of the highly conserved sequence, GGAT. All known 5.8S rRNA sequences begin 14 to 18 bp upstream of this sequence (Naasr, 1984). Sequence homology with two other protozoan LSUrRNA sequences, Tetrahymena pyriformis (Engberg et al., 1990) and Prorocentrum micans (Lenzes et al., 1989), allowed a reasonable estimate of the 3' end of the T. parva LSUrRNA gene (column 2929 in fig. 2.8), but the position of the 3' terminus was more difficult to determine due to decreased sequence homology in this region.

2.3.2.1 The small and large subunits

The T. p. lawrencei 7344 SSUrRNA gene sequence was 1741 bp in length and was identical to that of T. p. parva Muguga (fig 2.8). Although the T. p. parva Muguga SSUrRNA gene has previously been sequenced (Allsopp et al., 1993; Allsopp, 1994), this represents the first T. p. lawrencei SSUrRNA gene sequence, as a cloned T. p. lawrencei parasite had not previously been available.

There were only minor sequence differences between the T. p. lawrencei 7344 and T. p. parva Muguga LSUrRNA gene sequences. Two regions in the T. p. parva
Screening of recombinant M13 plaques for M13 clones containing LSU-1 insert. Plaques were lifted in duplicate onto Hybond N+ filters. Duplicate plaque lifts (1 and 2) were from two plates (A and B) which contained 42 and 48 clear plaques respectively. LSU-1 PCR product was labelled using the Multiprime DNA labelling kit (Amersham) and used as a probe to identify 16 (plate A) and 21 (plate B) M13 plaques containing LSU-1 inserts. Plaques which were detected by the probe on both filters were picked for sequencing.
Recombinant pUC18 clones containing the LSU-1 PCR product. PCR products were end-filled and kinased and ligated into Smal I-cut pUC18. Recombinant clones were digested with Bam HI and Eco RI to ensure that they contained the correct insert (lanes 2-5). The LSU-1 PCR product (lane 6) contained no Bam HI sites (lane 8) but two Eco RI sites (lane 9). λ Hind III molecular weight markers (lane M) were used to estimate the size of the PCR products.
Recombinant pUC18 clones of the LSU-1 amplification product are shown in fig. 2.6 (lanes 2-5).

Random DNA fragments generated by sonication were subcloned into M13mp18 for sequencing. The sonicate comprises rRNA insert and pUC18 fragments in a ratio of approximately 1:2 (the inserts ranged in size from 1.2 - 1.8 kbp, pUC18 is 2.69 kbp). Therefore, approximately two thirds of recombinant M13 plaques obtained would contain pUC18 fragments. It was thus necessary to screen recombinant plaques to select only those that had rRNA inserts. PCR products were radio-labelled and used to probe plaque lifts of recombinant M13 clones. Positive recombinant plaques detected by radiolabelled LSU-1 amplification product are illustrated in fig. 2.7. These plaques were picked for sequencing.

2.3.2 Sequence of the T. p. lawrencei 7344 transcription unit

The sequence of each PCR product was obtained at least once on each strand. The lengths of the consensus sequences obtained for each portion of the rDNA unit and the average number of gel characters per consensus character are shown in table 2.2.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Amplification primers</th>
<th>Total length (bp)</th>
<th>Average number of gel characters per consensus character</th>
</tr>
</thead>
<tbody>
<tr>
<td>SS1</td>
<td>A &amp; B</td>
<td>1741</td>
<td>2.86</td>
</tr>
<tr>
<td>LSU-1</td>
<td>ILO 465 &amp; ILO 467</td>
<td>1564</td>
<td>6.07</td>
</tr>
<tr>
<td>LSU-2</td>
<td>NC1 &amp; NC2</td>
<td>1761</td>
<td>5.30</td>
</tr>
<tr>
<td>ITS-S</td>
<td>ILO 466 &amp; ILO 464</td>
<td>1262</td>
<td>7.28</td>
</tr>
<tr>
<td>ITS-L</td>
<td>ILO 466 &amp; ILO 464</td>
<td>1315</td>
<td>7.06</td>
</tr>
</tbody>
</table>
Two ITS PCR products were generated from *T. p. lawrencei* 7344 (Lanes 2 and 3) with primers ILO 466 and ILO 464. Only one ITS PCR product was obtained from *T. p. parva* Muguga (lanes 4 and 5). Lane 1: calf thymus negative control reaction. *λ* Hind III molecular weight markers (lane M) were used to estimate the size of the PCR products.
NC1 and NC2 amplified approximately 1.8 kbp at the 3' terminus (fig. 2.4, Panel C, lane 1). A single *T. p. lawrencei* 7344 16S rRNA amplification product was obtained with each set of primers. The *T. p. parva* Muguga positive control sample yielded very similar products (fig. 2.4, Panels B and C, lane 2). No amplification product was obtained in either of the bovine negative control reactions (fig. 2.4, Panels B and C, lane 3).

Two ITS amplification products differing by approximately 50 bp were obtained from *T. p. lawrencei* 7344 (Fig. 2.5, lanes 2 and 3). The longer product was approximately 1.3 kbp and was designated ITS-L; the shorter one was approximately 1.25 kbp in length and was designated ITS-S. *T. p. parva* positive control samples yielded a single PCR product (lanes 4 and 5). No amplification product was obtained in the bovine negative control reaction (lane 1).

Shotgun cloning and sequencing of randomly-generated DNA fragments is a quick and efficient method of gathering sequence information (Bankier et al., 1987). A linear DNA fragment is not randomly sheared by sonication as the centre of the molecule experiences the maximum shear energy; linear fragments tend to be halved and subfragments halved again. To increase the random nature of the process, DNA must be in a circular form for sonication. Thus, as sonication was used to shear DNA in this study, the *T. p. lawrencei* 7344 ITS and LSU rRNA PCR products were initially cloned into pUC18 and the circular recombinant plasmids were sonicated.

*Taq* DNA polymerase preferentially adds a single adenosine nucleotide to the 3' ends of the double stranded DNA molecules that it produces during a PCR (Clark, 1988). These 3' overhangs had to be filled in and the ends of the molecules kinased before they could be cloned into *Sma* I-cut pUC18 (the 3' A overhang produced by *Taq* polymerase was later utilised for direct cloning of PCR products into a new generation of T-vectors which became available, as described in Chapters 3 and 4).
Figure 2.4: PCR products of *T. p. lawrencei* 7344 rRNA genes (lane 1, panels A, B and C). Panel A: SSU rRNA gene amplified with primers A and B. Panel B: 5' half of the LSU rRNA gene (LSU-1) amplified with primers IL0 465 and IL0 467. Panel C: 3' half of the LSU rRNA gene (LSU-2) amplified with primers NC1 and NC2, respectively. Positive *T. p. parva* Muguga (lane 2, panels A, B and C) and negative calf thymus (lane 3, panels A, B and C) controls were included with each PCR performed. *ϕX174* lac II (lanes M1) and λ Hind III (lanes M2) molecular weight markers were used to estimate the size of the PCR products.

35
in order to eliminate the possibility of obtaining sequence errors originating early in any one of the reactions. 3-6 clones in each orientation were sequenced, and a consensus sequence determined. This eliminated minor sequence heterogeneities resulting from Taq polymerase errors as well as from individual variation between parasites. The sequencing of more than one clone also allowed for detection of heterogeneous copies of the gene.

2.3.1.1 Amplification and cloning of the SSUrRNA gene of T. p. lawrencei 7344

Primers A and B specifically amplified the SSUrRNA genes of T. p. lawrencei 7344 and T. p. parva Muguga (fig. 2.4, Panel A, lanes 1 and 2, respectively) from the mixture of bovine and parasite DNA extracted from cell culture. Bovine SSUrRNA genes did not amplify (fig. 2.4, Panel A, lane 3), indicating that there are sufficient mismatches between the sequences of 16S-like genes and mammalian 18S rRNA genes at the primer sites to prevent priming on the bovine rRNA transcription unit. The T. p. lawrencei 7344 sample yielded a single amplification product of approximately 1.8 kbp (fig. 2.4, Panel A, lane 1); this size corresponds closely with that obtained from the T. p. parva Muguga positive control (fig. 2.4, Panel A, lane 2).

The T. p. lawrencei 7344 SSUrRNA gene contained Hind III and Sma I restriction enzyme sites; enzymes which did not cut the PCR product internally, Eco RI and Bam HI, were therefore chosen for cloning. Eight SSUrRNA clones were pooled and sequenced in both orientations.

2.3.1.2 Amplification and cloning of the LSUrRNA gene and the ITS of T. p. lawrencei 7344

Primers ILO 465 and ILO 467 specifically amplified approximately 1.6 kbp of the 5' end of the T. p. lawrencei 7344 LSUrRNA gene (fig. 2.4, Panel B, lane 1); primers
2.2.6.2 LSUrRNA and ITS templates

The M13-20 forward sequencing primer was used to sequence 48 random clones for LSU-1, 61 clones for LSU-2, 37 for ITS-S and 38 for ITS-L.

2.2.6.3 Sequence compilation and manipulation

Sequences were read using a Graf/Bar MarkII sonic digitiser and the READGEL program (D. Judge, Department of Biochemistry, University of Cambridge, pers. comm.). Sequences were assembled and edited using the Staden-Plus Program Package (Amersham), and aligned using CLUSTALV (Higgins et al., 1992).

2.3 Results and discussion

2.3.1 The need to amplify parasite rRNA genes

It is difficult to obtain *T. p. lawrencei* DNA free from host DNA. With other theilerias and *T. p. parva* it is possible to extract substantially pure parasite DNA from preparations of piroplasms. However, *T. p. lawrencei* produces very low piroplasm parasitemias and it is therefore not possible to purify enough piroplasms to extract DNA. The polymerase chain reaction was therefore utilised to amplify *T. p. lawrencei* SSUrRNA and LSUrRNA genes specifically from the mixture of bovine and parasite DNA extracted from cell culture.

*Taq* polymerase has an error rate of approximately $1.0 \times 10^{-4}$ errors per base pair per cycle, including both base substitutions and frameshift mutations (Tindall and Kunkel, 1988; Keohavong and Thilly, 1989; Cariello et al., 1991; Ling et al., 1991; Lundberg et al., 1991; Barnes, 1992). Therefore, when sequencing DNA obtained by means of PCR, it is necessary to sequence a number of clones to detect nucleotide misincorporations by *Taq* polymerase. Products from a number of PCRs were pooled.
separated out on a 0.8% agarose/TBE gel with \( \phi X174 \) Hae III size markers (New England Biolabs). The smear between 300 and 800 bp was excised and gel-purified as described in Appendix C.7.1.

2.2.5.3 Preparation of recombinant M13 subclones

The ends of the DNA fragments obtained by sonication were repaired and kinased (Appendix C.6), ligated into Sma I-cut and dephosphorylated M13mp18 and transformed into competent E. coli TG1 cells (Appendix C.8). The recombinant plaques were screened to select only those that had rRNA inserts. Plaques were lifted onto Hybond N+ (Amersham) filters (Appendix C.10.1). The Multiprime DNA labelling kit (Amersham) was used to label the LSU-1, LSU-2, ITS-S and ITS-L PCR products (Appendix C.11.1) and recombinant plaques were identified by hybridising these probes with DNA immobilized on the filters (Appendix C.10.2). Recombinant phage plaques were picked and single-stranded templates were prepared for sequencing as described in Appendix C.12.1.1.

2.2.6 Sequencing

The dideoxy chain termination method (Sanger et al., 1977) was used to sequence single stranded templates as described in Appendix C.12.1. A number of clones in each orientation were sequenced to obtain a consensus.

2.2.6.1 SSUrRNA templates

Aliquots of eight single-stranded templates were pooled and 5 µl aliquots of this pool were used in annealing reactions. The M13-20 forward primer (Messing, 1983) and a series of 5 primers for each strand (Elwood et al., 1985) were used to sequence the full length SSUrRNA gene in both directions.
Figure 2.8 continued, page 10 of 11.
<table>
<thead>
<tr>
<th>Ensembl</th>
<th>Loci</th>
<th>Description</th>
<th>Chromosome</th>
<th>Start</th>
<th>Stop</th>
<th>Length</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ensembl</td>
<td>Loci</td>
<td>Description</td>
<td>Chromosome</td>
<td>Start</td>
<td>Stop</td>
<td>Length</td>
<td>Gene</td>
</tr>
<tr>
<td>Ensembl</td>
<td>Loci</td>
<td>Description</td>
<td>Chromosome</td>
<td>Start</td>
<td>Stop</td>
<td>Length</td>
<td>Gene</td>
</tr>
<tr>
<td>Ensembl</td>
<td>Loci</td>
<td>Description</td>
<td>Chromosome</td>
<td>Start</td>
<td>Stop</td>
<td>Length</td>
<td>Gene</td>
</tr>
<tr>
<td>Ensembl</td>
<td>Loci</td>
<td>Description</td>
<td>Chromosome</td>
<td>Start</td>
<td>Stop</td>
<td>Length</td>
<td>Gene</td>
</tr>
<tr>
<td>Ensembl</td>
<td>Loci</td>
<td>Description</td>
<td>Chromosome</td>
<td>Start</td>
<td>Stop</td>
<td>Length</td>
<td>Gene</td>
</tr>
</tbody>
</table>

**Figure 2.8 continued, page 6 of 11.**
Figure 2.8 continued, page 4 of 11.
<table>
<thead>
<tr>
<th>Column 1</th>
<th>Column 2</th>
<th>Column 3</th>
<th>Column 4</th>
<th>Column 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data A</td>
<td>Data B</td>
<td>Data C</td>
<td>Data D</td>
<td>Data E</td>
</tr>
<tr>
<td>123</td>
<td>456</td>
<td>789</td>
<td>012</td>
<td>345</td>
</tr>
<tr>
<td>678</td>
<td>901</td>
<td>234</td>
<td>567</td>
<td>890</td>
</tr>
</tbody>
</table>

Failure 2.8 continued,
Figure 2.8: Alignment of T. p. parva (Tpp) and T. p. lawrencei 7344 (Tpl) rRNA transcription units from the 5' end of the small subunit to the approximate 3' end of the large subunit. Details of sequences are in table 2.3. Genbank accession numbers of published sequences are given. SSU: small subunit; ITS: internal transcribed spacer; ITS-L: longer ITS; ITS-S: shorter ITS; LSU: large subunit. Positions of amplification primers are highlighted.
2.3.3 Summary

There was little variation between the SSUrRNA and LSUrRNA gene sequences of *T. p. lawrencei* 7344 and *T. p. parva* Muguga. However, major sequence variation was identified in the ITS regions. We hoped to identify regions within the ITS which could be used as target sequences for distinguishing between these two parasites; this work will be discussed in the following chapter.

<table>
<thead>
<tr>
<th>Sequence code</th>
<th>Stock</th>
<th>Genbank accession no.</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tpp_SSU_L02366</td>
<td><em>T. p. parva</em> Muguga</td>
<td>L02366</td>
<td>Allopp et al. (1993)</td>
</tr>
<tr>
<td>Tpp_SSU_L28999</td>
<td><em>T. p. parva</em> Muguga</td>
<td>L28999</td>
<td>Kibe et al. (1994)</td>
</tr>
<tr>
<td>Tpp_Ug_ITS_L26332</td>
<td><em>T. p. parva</em> Uganda</td>
<td>L26332</td>
<td>Kibe et al. (1994)</td>
</tr>
<tr>
<td>Tpp_LSU_S_L28036</td>
<td><em>T. p. parva</em> Muguga</td>
<td>L28036</td>
<td>Kibe et al. (1994)</td>
</tr>
<tr>
<td>Tpp_LSU_link</td>
<td><em>T. p. parva</em> Muguga</td>
<td></td>
<td>Kibe and Bishop (pers. comm.)</td>
</tr>
<tr>
<td>Tpp_LSU_S_L28998</td>
<td><em>T. p. parva</em> Muguga</td>
<td>L28998</td>
<td>Kibe et al. (1994)</td>
</tr>
<tr>
<td>Tpl_SSU</td>
<td><em>T. p. lawrencei</em> 7344</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tpl_ITS_L_U03601</td>
<td><em>T. p. lawrencei</em> 7344</td>
<td>U03601</td>
<td>Kibe et al. (1994)</td>
</tr>
<tr>
<td>Tpl_ITS_S_U03602</td>
<td><em>T. p. lawrencei</em> 7344</td>
<td>U03602</td>
<td>Kibe et al. (1994)</td>
</tr>
<tr>
<td>Tpl_LSU_S'</td>
<td><em>T. p. lawrencei</em> 7344</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tpl_LSU_S''</td>
<td><em>T. p. lawrencei</em> 7344</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Muguga but that their sequences were identical. They were mapped to two different chromosomes and were not tandemly repeated. This unusual arrangement of rRNA genes is similar in some other Apicomplexa, including *Plasmodium falciparum* (Langsley *et al.*, 1983; McCutchan *et al.*, 1988), *Plasmodium berghei* (Dame and McCutchan, 1983), *Babesia bovis* (Dalrymple, 1990) and *Babesia bigemina* (Reddy *et al.*, 1991).

In *Plasmodium*, the sequences of the different rRNA transcription units are distinct and different ones are transcribed at different life cycle stages (Langsley *et al.*, 1983; Gunderson *et al.*, 1987; Waters *et al.*, 1989). The polymorphism observed between the two rDNA units in the cloned *T. parva* parasite can be used in studies to determine whether stage-specific transcription of the rRNA coding units occurs in *T. parva*. The sequence differences between the two ITS sequences of *T. p. lawrencei* 7344 have been used to design primers which will differentially amplify the ITS and LSU rRNA gene of the two different rRNA transcription units from this parasite. Primers used for amplifications were the reverse complements of *T. parva* characterisation oligonucleotides 2_7344/TSS (specific for the short ITS of *T. p. lawrencei* 7344) or 3_7344/TSL (specific for the long ITS of *T. p. lawrencei* 7344) in conjunction with NC2 (Details of primers can be found in Chapter 3, table 3.2). Sequence analysis of the two amplicons revealed that there were differences between the LSU rRNA coding regions of the two *T. p. lawrencei* 7344 transcription units (R. Bishop and E. Gobright, pers. comm.). Work is currently in progress to determine whether the transcription units are stage-specifically transcribed. Oligonucleotide probes that will detect one or the other transcription unit have been designed and probing conditions to detect the transcription units at different life cycle stages are being determined. Preliminary results suggest that both transcription units were transcribed into rRNA in *T. p. lawrencei* 7344 schizont RNA (R. Bishop, pers. comm.).
3.1.3 Additional putative *T. p. lawrencei* characterisation oligonucleotide

A region conserved between *T. p. lawrencei* 7344 ITS-L and both *T. p. lawrencei* Hluhluwe 3 ITS sequences was identified and a new oligonucleotide, 5_TPL/TS, was synthesised (Table 3.2). This oligonucleotide was used to screen ITS amplification products of the same panel of *T. p. parva* and *T. p. lawrencei* isolates previously used (fig 3.1, panel 5). Probe 5_TPL/TS detected all of the *T. p. lawrencei* isolates except for KNP 2 and Warmbaths 1. It did not detect *T. p. parva* Muguga, but it did hybridise strongly to the South African isolate, *T. p. parva* Schoonspuit.

This probe was closer to the *T. p. lawrencei*-specific probe that we were seeking, but it still did not detect all of the South African *T. p. lawrencei* isolates. We therefore decided to sequence the ITS regions of a number of other *T. parva* isolates to determine whether there exist any regions in the ITS which could be used to distinguish absolutely between the causative agents of East Coast fever and all the causative agents of Corridor disease.

3.2.2 Nucleotide sequence of the ITS of a range of *T. parva* isolates

At this stage sufficient ITS sequence data had been accumulated to plan a new rapid strategy for cloning and sequencing *T. parva* ITS amplification products. Restriction enzyme sites were added to the ITS amplification oligonucleotides, such that the PCR products could be digested and cloned directionally into the M13 sequencing vectors. Four internal sequencing primers were identified which, along with the M13-20 forward sequencing primer, could be used to obtain the entire sequence of the non-coding strand of each cloned ITS relatively quickly.

ITS sequences were amplified, cloned and sequenced from the *T. parva* isolates listed in table 3.1. Many of the isolates contained more than one ITS sequence (Table 3.3). This might indicate that these isolates contain mixed populations of *T. parva*.
probes may offer information on the relatedness between isolates but did not distinguish between the causative agents of East Coast fever and Corridor disease. As a further step towards refining probes that might define a broader grouping of *T. p. lawrencei* isolates, we sequenced the ITS of *T. p. lawrencei* Hluhluwe 3.

### 3.3.1.2 Nucleotide sequence of the *T. p. lawrencei* Hluhluwe 3 ITS

Two *T. p. lawrencei* Hluhluwe 3 ITS amplification products were generated and they were cloned and sequenced separately. The two ITS sequences from the 3' end of the SSU rRNA gene to the putative 5' end of the LSU rRNA gene were 1210 bp (ITS-L) and 1144 bp (ITS-S) in length. The *T. p. lawrencei* Hluhluwe 3 ITS sequences comprised an arrangement of conserved and variable regions similar to that previously observed in the ITSs of *T. p. lawrencei* 7344 and *T. p. parva* Muguga and Uganda. *T. p. lawrencei* Hluhluwe 3 ITS-L contained a unique insert, approximately 70 bp in length, located in the conserved region between V1 and V2 (see the region between columns 206 and 287 in fig. 3.2, no. 8L), which accounted for the large size difference between the two *T. p. lawrencei* Hluhluwe 3 ITS sequences. There were also other minor deletions and insertions as well as nucleotide substitutions (fig. 3.2).

Similarities to each of the *T. p. lawrencei* 7344, *T. p. parva* Muguga and *T. p. parva* Uganda sequences were observed in different regions of the *T. p. lawrencei* Hluhluwe 3 ITS. For example, in the first variable region (V1), *T. p. lawrencei* Hluhluwe 3 ITS-S (no. 8S in fig. 3.2) was similar to the *T. p. parva* Muguga ITS, whereas in the second variable region (V2) it resembled *T. p. lawrencei* 7344 ITS-L (no. 7L in fig. 3.2). As a result of further variation downstream of the 5.8S gene it was similar to *T. p. lawrencei* 7344 ITS-S (no. 7S in fig. 3.2).
Figure 3.1: Putative ITS characterisation oligonucleotides for *T. p. parva* (1_TPP/TSS) and *T. p. lawrencei* (2_7344/TSS, 3_7344/TSL, 4_7344/TS, 5_TPL/TS). ITS PCR products were slot-blotted onto Hybond N+ and probed with 3' end-labelled oligonucleotides. Details of probes are in table 3.2.
3.3.1.1 Experimental probes based on T. p. lawrencei 7344 and T. p. parva Muguga and Uganda ITS sequences

One putative T. p. parva and three putative T. p. lawrencei characterisation oligonucleotides were designed in the ITS region. Details of the probes are shown in Table 3.2. Probe 1_TPP/TS was designed in a region of the ITS that was specific to T. p. parva Muguga and Uganda. Probe 2_7344/TSS was unique to the shorter ITS of T. p. lawrencei 7344 while probe 3_7344/TSL was unique to the longer one. Probe 4_7344/TS was designed in a region that was conserved between the two T. p. lawrencei 7344 ITS sequences. ITS amplification products of a number of T. p. parva and T. p. lawrencei isolates were slot-blotted and probed with the four radiolabelled characterisation oligonucleotides (fig. 3.1).

Probe 1_TPP/TS hybridised strongly to T. p. parva Muguga and Schoonspruit samples (fig. 3.1, panel 1) and weakly to only one of the South African T. p. lawrencei field isolates (Warmbaths 3). This oligonucleotide appeared to be a good candidate for a T. p. parva-specific probe.

Probe 2_7344/TSS (fig. 3.1, panel 2) detected only T. p. lawrencei 7344; it did not hybridise to any of the other isolates tested. Probe 3_7344/TSL (fig. 3.1, panel 3) detected T. p. lawrencei samples 7344 and KNP 2 and hybridised weakly to Warmbaths 1, but it also detected T. p. parva Schoonspruit. Probe 4_7344/TS (fig. 3.1, panel 4) detected T. p. lawrencei 7344 and hybridised weakly to T. p. lawrencei KNP 2, but did not detect any of the other T. p. lawrencei samples tested. None of these probes detected the South African T. p. lawrencei isolate, Hluhluwe 3, indicating that this isolate differs from T. p. lawrencei 7344 in the variable region of the ITS where the probes are located.

The sequence differences detected by probes 2_7344/TSS, 3_7344/TSL and 4_7344/TS appeared to be isolate-specific rather than subspecies-specific. These
3.2.4 Experimental probe design and testing

The ITS sequences were aligned using CLUSTALW (Thompson et al., 1994) and the Genetics Data Environment (GDE) alignment editor (Smith, 1992), and the alignment is shown in figure 3.2. Eleven negative strand characterization probes were designed in unique regions (Table 3.2).

A modification of the slot blot hybridisation of PCR products described by Saiki et al., (1986) was used to screen ITS amplification products with radiolabelled oligonucleotide probes. PCR amplification of the ITS region of a number of T. parva isolates was performed as described previously, using unmodified primers ILO 466 and ILO 464. A 10 μl aliquot of each PCR product was denatured and slot-blotted onto replicate Hybond N+ filters (Amersham) as described in Appendix C.13. One filter was prepared for each probe to be used. The filters were pre-hybridised in 0.5 M sodium phosphate buffer (pH 7.5) and 7% SDS (modified from Church and Gilbert, 1984) for at least one hour at 65°C in a Hybaid hybridisation oven. Oligonucleotide probes were 3' end-labelled with [α-32P] using terminal transferase (Appendix C.11.2) and were added directly to the hybridisation solution. Hybridisation was carried out at Tm-20°C overnight. The filters were washed in 4xSSC (Appendix B), three times at room temperature for 10 minutes, and once at Tm-10°C for 1 minute. Filters were autoradiographed with Cronex 4 X-ray film (Du Pont) at -80°C with an intensifying screen for 2.5 hours and overnight.

3.3 Results and discussion

3.3.1 The search for T. p. lawrencei- and T. p. parva-specific probes in the ITS

This section describes how the search for characterisation probes developed. At each stage the probes designed showed where more data were required, modifications were then made and further probes designed as the extra data became available.
3.2.3.2 Primers for sequencing T. parva ITS amplification products

The T. p. lawrencei 7344 and Hluhluwe 3 ITS sequences were aligned with those of T. p. parva Muguga and Uganda (Kibe et al., 1994). Four internal primers (Table 3.2) were designed for obtaining the sequence of the non-coding strand of the entire ITS, along with the M13 -20 forward sequencing primer. The sequencing primers were complementary to the coding strand and were located approximately 250 bp apart in conserved regions of the T. parva ITS.

3.2.3.3 Cloning and sequencing of ITS amplification products

Allquotes (5 μl) of PCR products were digested with Eco RI, Sac I, Xba I and Hind III (Appendix C.5), to select enzymes for directional cloning which did not cut the PCR product internally. None of the ITS amplification products contained internal Eco RI or Hind III sites.

For each isolate, amplification products from three PCR reactions were pooled, digested with Eco RI and Hind III and gel-purified using the Wizard™ DNA purification system (Appendix C.7.2). The resultant sticky-ended PCR products were ligated into appropriately digested M13mp19 (Appendix C.8). Single-stranded templates of six clones from each isolate were sequenced by the dideoxy chain termination method (Sanger et al., 1977) using the M13 -20 forward primer and the four internal sequencing primers (Appendix C.12.1). Autoradiographs of sequencing gels were scanned and sequences determined using an automatic base scanner and semi-automatic base-calling software (Amersham); gel readings were assembled and edited using XBAP (Staden, 1994).
Table 3.2: Oligonucleotides used for amplification, sequencing and probing of the *T. parva* ITS.

<table>
<thead>
<tr>
<th>Identity</th>
<th>Strand</th>
<th><strong>T_m</strong> (°C)</th>
<th>Sequence 5'→3'</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Amplification oligonucleotides</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mlLO 466</td>
<td>-</td>
<td>82.6</td>
<td>Cgg Aat Tgc Agc Tgg Gga Aag Tct Agg Gaa Gtt Ttg</td>
<td>Modified from Kiba et al. (1994), Include restriction enzyme sites for cloning</td>
</tr>
<tr>
<td>mlLO 464</td>
<td>+</td>
<td>80.3</td>
<td>Gct Cta Gaa Gct Tcc Gtt Ggc Aaa Ttc Aag Gga Ttt Tc</td>
<td>Modified from Kiba et al. (1994), Include restriction enzyme sites for cloning</td>
</tr>
<tr>
<td><strong>Sequencing oligonucleotides</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TPTS1</td>
<td>+</td>
<td>59.3</td>
<td>Tag Ttt Tag Tgg Gaa Gat G</td>
<td>Sequencing primer for <em>T. parva</em> ITS</td>
</tr>
<tr>
<td>TPTS2</td>
<td>+</td>
<td>63.1</td>
<td>Bag Tao Cog Ggt Aag G</td>
<td>Sequencing primer for <em>T. parva</em> ITS</td>
</tr>
<tr>
<td>TPTS3</td>
<td>+</td>
<td>64.1</td>
<td>Ggt Gga Tgt Gtt Ggc Tc</td>
<td>Sequencing primer for <em>T. parva</em> ITS</td>
</tr>
<tr>
<td>TPTS4</td>
<td>+</td>
<td>59.3</td>
<td><strong>3</strong> Atg Agt Gaa Aot Aag G</td>
<td>Sequencing primer for <em>T. parva</em> ITS</td>
</tr>
<tr>
<td><strong>Experimental probes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1_TPLVS</td>
<td>-</td>
<td>59.3</td>
<td>Taa Tag Gtt Aat Aaa Tta Aco Ga</td>
<td>Specific for <em>T. parva</em> Uganda and Muguga ITS</td>
</tr>
<tr>
<td>2_7344TS</td>
<td>-</td>
<td>62.1</td>
<td>Taa Tag Taa Tgg Tga Taa Tag Tcc G</td>
<td>Specific for <em>T. p. lawrencei</em> 7344 ITS-S</td>
</tr>
<tr>
<td>3_7344TS</td>
<td>-</td>
<td>72.4</td>
<td>Aca Agt Atg Tgg Gta G</td>
<td>Specific for <em>T. p. lawrencei</em> 7344 ITS-L</td>
</tr>
<tr>
<td>4_7344TS</td>
<td>-</td>
<td>62.1</td>
<td>Aatt Aac Cct Aaa Ggg Ttw Cta A</td>
<td>Specific for <em>T. p. lawrencei</em> 7344 ITS-S and ITS-L</td>
</tr>
<tr>
<td>5_TPLVS</td>
<td>-</td>
<td>63.2</td>
<td>Gtt Att Aca Atg Saa Taa Caa Taa Tt</td>
<td>Specific for <em>T. p. lawrencei</em> 7344 and Hluhluwe 3 ITS</td>
</tr>
<tr>
<td>TPP1</td>
<td>-</td>
<td>55.0</td>
<td>Agg Tta Ata Aat Taa Ooc A</td>
<td>Specific for <em>T. p. parva</em> isolates</td>
</tr>
<tr>
<td>TPP2</td>
<td>-</td>
<td>50.7</td>
<td>Tgg Taw Taa Aat Taa Coa A</td>
<td>Specific for Zimbabwean <em>T. p. parva</em> isolates</td>
</tr>
<tr>
<td>TPL1</td>
<td>-</td>
<td>59.3</td>
<td>Cta Aac Cgt Aaa Aco Ga</td>
<td>Specific for group 1 <em>T. p. lawrencei</em> isolates</td>
</tr>
<tr>
<td>TPL2</td>
<td>-</td>
<td>63.6</td>
<td>Tar Gtc Gga Aqa Cyc Waa A</td>
<td>Specific for group 2 <em>T. p. lawrencei</em> isolates</td>
</tr>
<tr>
<td>TPL3a</td>
<td>-</td>
<td>lowest 62.4</td>
<td>Aac Och Aaw Ggg Ttw Cta Aa</td>
<td>Specific for group 3 <em>T. p. lawrencei</em> isolates</td>
</tr>
<tr>
<td>TPL3b</td>
<td>-</td>
<td>57.1</td>
<td>Gct Aaa Aac Ggt Ttc Taa A</td>
<td>Specific for group 3 <em>T. p. lawrencei</em> isolates</td>
</tr>
</tbody>
</table>

* **T_m**: the temperature at which 50% of the probe/target DNA duplexes are dissociated. See Appendix C4 for calculation of dissociation temperatures.
3.2.2 Cloning and sequencing of the *T. p. lawrencei* Hluhluwe 3 ITS

*T. p. lawrencei* Hluhluwe 3 is a South African *T. p. lawrencei* isolate, derived from buffalo, which has remained unchanged after successive passages through cattle (Potgieter *et al.*, 1988). The ITS of this isolate was amplified with primers ILO 466 and ILO 464 as described in section 2.2.3 and two amplification products were obtained. They were isolated by fractionation on a 1% agaroseTBE gel and the two fragments were excised separately as previously described (section 2.2.5.1). Each fragment was cloned directly into pGEM-T (Promega) as described in Appendix C.8. Shotgun cloning and sequencing (Bankier *et al.*, 1987) was used to obtain the sequences of the PCR products as described in section 2.2.5.

3.2.3 A new strategy for amplification and sequencing of *T. parva* ITS regions

After obtaining the full sequence of the *T. p. lawrencei* Hluhluwe 3 ITS we altered our methods in order to gather information from the other isolates more rapidly. The amplification primers were modified to facilitate more rapid cloning, internal sequencing primers were designed to avoid the need for shotgun sequencing, and a decision was taken to sequence only one strand of each amplicon.

3.2.3.1 Amplification of *T. parva* ITS regions

*T. parva*-specific oligonucleotide primers, ILO 466 and ILO 464 (Kibe *et al.*, 1994) were modified such that the resulting ITS amplification products could be digested with a selection of restriction enzymes. *Eco* RI and *Sac* I sites were included at the 5' end of ILO 466; *Xba* I and *Hind* III were added to the 5' end of ILO 464 (Table 3.2). PCR amplifications were performed and the amplicons analysed as described in Appendix C.3. Cycling conditions were: 72°C for 1.5 minutes, 95°C for 1 minute and 60°C for 1 minute, for 30 cycles.
Table 3.1: *Theileria* stocks used as sources of DNA from which ITS regions were sequenced

<table>
<thead>
<tr>
<th>Species origin</th>
<th>Stock / Isolate code</th>
<th>Remarks</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>T. parva</em> isolated from cattle</td>
<td>Muguga</td>
<td>Field isolate from Kiambu District, Kenya, East Coast fever reference stock.</td>
<td>Brocklesby et al. (1961)</td>
</tr>
<tr>
<td></td>
<td>LN/M/IA</td>
<td>Schizont-infected lymph node material from a bovine infected at ILRAD, Nairobi, with <em>T. parva</em> Muguga.</td>
<td>T. T. Doolan (pers. comm.)</td>
</tr>
<tr>
<td></td>
<td>Muguga SM</td>
<td>Schizont-infected lymph node material from a bovine infected at the CTVM, Edinburgh, with <em>T. parva</em> Muguga.</td>
<td>C. G. D. Brown (pers. comm.)</td>
</tr>
<tr>
<td></td>
<td>Schoonspruit</td>
<td>Isolated in 1937 from natural bovine ECF infection in Transvaal, South Africa.</td>
<td>Nalix (1948)</td>
</tr>
<tr>
<td></td>
<td>T.p./CHI/16</td>
<td>Isolated from naturally infected bovine on Chikaya farm, Zimbabwe. Typed as ‘cattle-derived’ using MAb profile and Tpr &amp; rRNA probes.</td>
<td>Bishop et al. (1994)</td>
</tr>
<tr>
<td></td>
<td>T.p./CHI/23</td>
<td>Animal experimentally infected by ticks collected from reacting animal at Chikaya farm. Severe reaction, died. Types as ‘cattle-derived’ using Tpr &amp; rRNA probes, but ‘buffalo-derived’ using MAb profile.</td>
<td>Bishop et al. (1994)</td>
</tr>
<tr>
<td></td>
<td>T.p./AYR/12</td>
<td>Isolated from naturally infected bovine on Ayrshire farm, Zimbabwe. Typed as ‘cattle-derived’ using MAb profile and Tpr &amp; rRNA probes.</td>
<td>Bishop et al. (1994)</td>
</tr>
<tr>
<td><em>T. parva</em> isolated from buffalo</td>
<td>7344/G5/F5/ES8</td>
<td>Clone generated from an isolate made in Laitlapa District, Kenya, from naturally infected buffalo 7014. Classical Corridor disease in cattle.</td>
<td>Kautki et al. (1990); Morar et al. (1995)</td>
</tr>
<tr>
<td></td>
<td>Hluhluwe 3</td>
<td>Isolated from naturally infected buffalo in Natal, South Africa. Extensive passage through cattle, stable Corridor disease clinical picture.</td>
<td>Polglaser et al. (1988)</td>
</tr>
<tr>
<td></td>
<td>KNP 2</td>
<td>Isolated from naturally infected buffalo in the Kruger National Park, eastern Transvaal, South Africa. Classical Corridor disease in cattle.</td>
<td>H. Stoltz (pers. comm.)</td>
</tr>
<tr>
<td></td>
<td>T.p./BAL/93</td>
<td>Produced severe reaction in cattle, typed as ‘buffalo-derived’ using MAb profile.</td>
<td>G. K. Kasiri and P. Spooner (pers. comm.)</td>
</tr>
<tr>
<td></td>
<td>T.p./MAS/38</td>
<td>Natural infection in bovine kept on buffalo grazed paddock, Masari farm, Zimbabwe. Typed as ‘buffalo-derived’ using MAb profile and Tpr &amp; rRNA probes.</td>
<td>Bishop et al. (1994)</td>
</tr>
<tr>
<td></td>
<td>T.p./MAS/40</td>
<td>Natural infection in bovine kept on buffalo grazed paddock, Masari farm, Zimbabwe. Typed as ‘buffalo-derived’ using MAb profile and Tpr &amp; rRNA probes.</td>
<td>Bishop et al. (1994)</td>
</tr>
</tbody>
</table>
the general degree of conservation in the ITS of *Drosophila* and suggested that approximately 40% of spacer sequences are not free to diverge. The non-constrained parts of the ITS evolve with a rate close to the neutral rate of evolution in *Drosophila*.

The variability displayed in the ITS can thus be used in phylogenetic studies and the development of species- and strain-specific probes. ITS regions have been used to distinguish between species of yeasts (Botelho and Planta, 1994; Mitchell *et al*., 1992), *Anopheles* mosquitoes (Porter and Collins, 1991), trichostrongyle nematodes (Gasser *et al*., 1994), trematodes (Adlar et al., 1993; Anderson and Barker, 1993) and a number of protists (Cai *et al*., 1992; Cevallos *et al*., 1993; Chambers *et al*., 1986; Goggin, 1994; Lee and Taylor, 1992).

In an attempt to find oligonucleotide probes to distinguish between *T. p. lawrencei* and *T. p. parva* parasites, we sequenced the ITS of a number of *T. parva* parasites from South Africa, Zimbabwe and Kenya.

### 3.2 Materials and methods

#### 3.2.1 Parasites and preparation of DNA

DNA extractions were performed as described in section 2.2.2 and Appendix C.I. DNA from the Zimbabwean *T. parva* isolates was provided by R. Bishop (International Livestock Research Institute, Nairobi, Kenya). ITS sequences were amplified from DNA extracted from the eleven *T. parva* isolates from South Africa, Kenya and Zimbabwe detailed in Table 3.1. *T. parva* DNA samples used in characterization studies were extracted either from schizont-infected lymphoblastoid cell cultures or from whole blood.
CHAPTER 3: SEQUENCE VARIATION IN THE THEILERIA PARVA INTERNAL TRANSCRIBED SPACER

3.1 Introduction

In the previous chapter I described the finding that the SSUrRNA genes of *T. p. parva* and a *T. p. lawrencei* clone are identical and the L`UrRNA genes are very similar, but that there is considerable variation in the internal transcribed spacer. We therefore decided to search for *T. p. lawrencei* and *T. p. parva* specific probes within the ITS.

Various studies indicate that there are different rates of evolution in different regions of rDNA (reviewed in Hillis and Dixon, 1991). In subgroups of *Drosophila melanogaster* variation has been shown to be greatest in the non-transcribed spacer, the internal transcribed spacer exhibits less variation, while the rRNA coding regions are almost identical (Coen et al., 1982a; 1982b; Strachan et al., 1982). A similar situation obtains in the rDNA within and between other species. Sequence divergence is found in the transcribed spacers of *Xenopus laevis* and *X. mulleri* (Brown et al., 1972) and *X. laevis* and *X. borealis* (Furlong and Maden, 1983; Furlong et al., 1983), whereas the rRNA genes of these species have only rare microheterogeneities. Amongst primates, there is greater sequence divergence in the transcribed spacers than in the 28S genes, where substitutions are clustered in the variable regions (Gonzalez et al., 1990).

Although ITS regions display greater variability than the rRNA coding regions, functional analyses of the ITS in yeasts have shown that some regions in the ITS are conserved and are essential for correct and efficient processing of rRNA genes (Musters et al., 1990; van der Sands et al., 1992). Swidtscher et al. (1994) assessed
Figure 2.8 continued, page 11 of 11.
All of the South African *T. p. lawrencei* isolates were detected by one or more of the *T. p. lawrencei* probes, but none of them was detected by TPP1, which detected almost all *T. p. parva* and *T. p. bovis* isolates. If TPP1 is truly detecting only *T. p. parva* and *T. p. bovis* isolates, this might indicate that there are no such parasites in South Africa, at least not amongst the isolates that we have screened in this study. All the isolates of the South African *T. p. lawrencei* Hluhluwe 3 stock, made after different numbers of passages, were consistently detected by the same probes (TPP2, TPL1 and TPL3a). These results suggest that the genotype of this isolate has changed little after extensive passage through cattle which supports reports that it has never "transformed" to a form that could cause ECF (Potgieter et al., 1988; H. Stoltsz, pers. comm.).

All of the isolates obtained from cattle in Zimbabwe were detected by TPP1, except for *T.p./CHI/23*. This correlates with a previous study (Bishop et al., 1994) in which all of the same isolates, except for *T.p./CHI/23*, were typed as 'cattle-derived' using anti-schizont MAbs and RFLP analysis and probing with *T. parva* Tpr repetitive and ribosomal DNA probes. *T.p./CHI/23* was obtained from Chikaya farm, where the cattle were not exposed to *T. parva* infection from buffalo, yet screening with a panel of anti-schizont MAbs indicated that this isolate was more similar to isolates from buffalo from elsewhere in Zimbabwe (Bishop et al., 1994). However, the rRNA gene pattern was similar to that of isolates from cattle and although the Tpr RFLP pattern differed from the predominant genotype seen in most isolates from cattle, it was not similar to that of the other isolates putatively from buffalo. ITS sequence analysis indicated that this isolate contained three distinct ITS sequences (section 3.3.2).

*T. parva* Boleni (Lawrence and Mackenzie, 1980) was the only isolate made from cattle in Zimbabwe which was detected by one of the *T. p. lawrencei* ITS probes, TPL1, as well as by the *T. p. parva* probe, TPP1. Bishop et al. (1994) were also able to distinguish between this stock and other Zimbabwean isolates from cattle by using a telomeric probe. *T. parva* Boleni is being used for immunization in Zimbabwe as
Table 3.4: Probing of amplified *T. parva* ITS regions with experimental characterization oligonucleotides

<table>
<thead>
<tr>
<th>T. PARVA ISOLATES</th>
<th>PROBES</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Origin</strong></td>
<td><strong>Name / isolate code</strong></td>
</tr>
<tr>
<td><strong>Kenya</strong></td>
<td>Mugaai LNMM/1</td>
</tr>
<tr>
<td></td>
<td>Mugaai SM</td>
</tr>
<tr>
<td></td>
<td>Mugaai BH/400</td>
</tr>
<tr>
<td></td>
<td>Mugaai BH/559/FR/57</td>
</tr>
<tr>
<td></td>
<td>Mugaai BH/571</td>
</tr>
<tr>
<td></td>
<td>Mugaai BH/471</td>
</tr>
<tr>
<td><strong>Uganda</strong></td>
<td>Uganda BR3372</td>
</tr>
<tr>
<td></td>
<td>Uganda BH2527</td>
</tr>
<tr>
<td><strong>Zambia</strong></td>
<td>Zambwa V1</td>
</tr>
<tr>
<td></td>
<td>Zambwa B105</td>
</tr>
<tr>
<td></td>
<td>Zambwa L1</td>
</tr>
<tr>
<td><strong>South Africa</strong></td>
<td>Schoomerezi</td>
</tr>
<tr>
<td><strong>Zimbabwe</strong></td>
<td><em>Vaccino strain</em></td>
</tr>
<tr>
<td></td>
<td>Botale BH/9339</td>
</tr>
<tr>
<td></td>
<td>Tp/CHU/16</td>
</tr>
<tr>
<td></td>
<td>Tp/CHU/17</td>
</tr>
<tr>
<td></td>
<td>Tp/CHU/18</td>
</tr>
<tr>
<td></td>
<td>Tp/CHU/20</td>
</tr>
<tr>
<td></td>
<td>Tp/CHU/21</td>
</tr>
<tr>
<td></td>
<td>Tp/CHU/22</td>
</tr>
<tr>
<td></td>
<td>Tp/CHU/23</td>
</tr>
<tr>
<td></td>
<td>Tp/AYR/11</td>
</tr>
<tr>
<td></td>
<td>Tp/AYR/12</td>
</tr>
<tr>
<td></td>
<td>Tp/AYR/13</td>
</tr>
<tr>
<td></td>
<td>Tp/AYR/14</td>
</tr>
<tr>
<td></td>
<td>Tp/KAL/20</td>
</tr>
<tr>
<td></td>
<td>Tp/AVB/9</td>
</tr>
<tr>
<td></td>
<td>Tp/RU/1994</td>
</tr>
<tr>
<td><strong>Kenya</strong></td>
<td><em>Vaccino strain</em></td>
</tr>
<tr>
<td><strong>Buffalo</strong></td>
<td>Hihiliwe 3</td>
</tr>
<tr>
<td></td>
<td>Hihiliwe #1</td>
</tr>
<tr>
<td></td>
<td>Hihiliwe #1</td>
</tr>
<tr>
<td></td>
<td>Hihiliwe #7</td>
</tr>
<tr>
<td></td>
<td>KNP 2</td>
</tr>
<tr>
<td></td>
<td>KNP/46</td>
</tr>
<tr>
<td></td>
<td>KNP/23</td>
</tr>
<tr>
<td></td>
<td>Waramba I</td>
</tr>
<tr>
<td></td>
<td>Waramba 2</td>
</tr>
<tr>
<td></td>
<td>Waramba 3</td>
</tr>
<tr>
<td><strong>South Africa</strong></td>
<td><em>Vaccino strain</em></td>
</tr>
<tr>
<td></td>
<td>Tp/KAL/21</td>
</tr>
<tr>
<td></td>
<td>Tp/KAL/27</td>
</tr>
<tr>
<td></td>
<td>Tp/KAL/33</td>
</tr>
<tr>
<td></td>
<td>Tp/IXMAS/07</td>
</tr>
<tr>
<td></td>
<td>Tp/IXMAS/38</td>
</tr>
<tr>
<td></td>
<td>Tp/IXMAS/39</td>
</tr>
</tbody>
</table>

77
Table 3.4: Probing of amplified *T. parva* ITS regions with experimental characterization oligonucleotides

<table>
<thead>
<tr>
<th>T. PARVA ISOLATES</th>
<th>PROBES</th>
<th>T. P. Improved</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Origin</strong></td>
<td><strong>Name / Isolate code</strong></td>
<td><strong>PROBE</strong></td>
</tr>
<tr>
<td>Kenya</td>
<td>Muguga INMA21</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Muguga BM1</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Muguga IDQ33</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Muguga SI500</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Makhekele G23/G56/G67</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Makhekele G371</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Makhekele P771</td>
<td>-</td>
</tr>
<tr>
<td>Uganda</td>
<td>Uganda BH151</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Uganda BH157</td>
<td>-</td>
</tr>
<tr>
<td>Malawi</td>
<td>Malawi</td>
<td>-</td>
</tr>
<tr>
<td>Zambia</td>
<td>Zambia V1</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Zambia S105</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Zambia L1</td>
<td>-</td>
</tr>
<tr>
<td>South Africa</td>
<td>Zimbabwe</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(vaccine strain)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>KNP2</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>KNP/94_1</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>KNP/94_2</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>KNP/94_3</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>KNP/94_4</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Warmbaths 1</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Warmbaths 2</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Warmbaths 3</td>
<td>-</td>
</tr>
<tr>
<td>Zimbabwe</td>
<td>Bolcal D770</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Bolcal RH2 3369</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Tp.CH115/16</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Tp.CH115/17</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Tp.CH115/18</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Tp.CH128</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Tp.CH121</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Tp.CH122</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Tp.CH123</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Tp.AYR111</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Tp.AYR112</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Tp.AYR113</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Tp.AYR114</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Tp.GAL50</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Tp.AV109</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Tp.HH94/6</td>
<td>-</td>
</tr>
<tr>
<td>Kenya</td>
<td>Tj44/GG5/95/28</td>
<td>-</td>
</tr>
<tr>
<td>Zambia</td>
<td>Hlahlahwe 3</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Hlahlahwe 63</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Hlahlahwe 5</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Hlahlahwe 65</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Hlahlahwe 67</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>KNP 2</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>KN994_1</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>KN994_2</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>KN994_3</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>KN994_4</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Warmbaths 1</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Warmbaths 2</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Warmbaths 3</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(990) (1993)</td>
<td>-</td>
</tr>
<tr>
<td>Zimbabwe</td>
<td>TP.BAL2/112</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>TP.BAL2/17</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>TP.BAL2/19</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>TP.BAL2/101</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>TP.BAL2/103</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>TP.AM8017</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>TP.AM8029</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>TP.AM8039</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>TP.AM8040</td>
<td>-</td>
</tr>
</tbody>
</table>

77
Fig. 3.3: Candidate region for distinction between T. p. parva and T. p. lawrencei genotypes. Boxes are drawn around the sequences of the experimental probes. From a total of nine ITS sequences from five isolates obtained from cattle we found three different sequences in this region, whereas of fourteen sequences from seven isolates obtained from buffalo there were thirteen different sequences. This is representative of the rest of the ITS; the isolates from buffalo were more variable.
3.3.3 Screening of *T. parva* isolates with new ITS experimental probes

One variable region in the ITS1 appeared to be a likely target for distinguishing between *T. p. parva* and *T. p. lawrencei* isolates (Fig. 3.3). The *T. p. lawrencei* isolates could be divided into three groups on the basis of sequence differences in this region. Six new experimental characterisation oligonucleotides were designed (Table 3.2). The results of screening *T. parva* ITS amplification products with the six new ITS probes are shown in Table 3.4.

The putative *T. p. parva* probe, TPP1, hybridised to all of the isolates originating from cattle, with the exception of *T. parva* Mu4 and *T. p. CHI/23*. TPP1 did not detect any of the South African *T. p. lawrencei* isolates, but it did hybridise with two of the stocks which had putatively been isolated from buffalo in Zimbabwe. That TPP1 detects both *T. p. parva* and *T. p. bovis* isolates but not *T. p. lawrencei* isolates suggests that the former are more closely related to each other than they are to *T. p. lawrencei*. There was no pattern which distinguished *T. p. parva* isolates from all *T. p. bovis* isolates, although TPP2 detected nine out of the thirteen Zimbabwean isolates from cattle. TPP2 was designed specifically to detect isolates, obtained from cattle in Zimbabwe, whose sequences were slightly different from those of other *T. p. parva* isolates (Fig. 3.3). However, this probe also hybridised to *T. p. lawrencei* Hluhluwe 3.

The oligonucleotide designed to detect group 1 *T. p. lawrencei* isolates, TPL1, detected seven *T. p. lawrencei* isolates but also a number of *T. p. parva* isolates. Between them, probes TPL2, TPL3a and TPL3b detected 19 of the 22 *T. p. lawrencei* isolates tested. Probes TPL2 and TPL3a did not hybridise to any of the isolates originating from cattle, except for *T. p. CHI/23* from Zimbabwe. Probe TPL3b detected seven *T. p. parva* isolates.
We have demonstrated that the ITSs of both *T. p. parva* and *T. p. lawrencei* parasites contain different combinations of identifiable segments of sequence in the variable regions, resulting in a mosaic of ITS sequence segments. This suggests that the *T. parva* parasites have undergone genetic recombination. Recombination events between different *T. p. parva* genotypes have been unequivocally demonstrated. Unique genetic markers were used to study isolates made after joint passage of the original genotypes through ticks (Bishop, pers. comm.). This finding adds support to our conclusion.

Some studies of ITS sequences indicate that, in many species, intraspecific variation in ITS sequences is undetected or low (Porter and Collins, 1991; Lee and Taylor, 1992; Botelho and Planta, 1994; Gasser et al., 1994; Zhuo et al., 1994). This is probably due to homogenisation which is thought to occur in tandemly repeated genes and results in new mutations spreading through the rRNA genes in a tandem array. The mechanisms involved include unequal crossing over, gene conversions and transpositions (Dover, 1982). In species where new mutations have arisen but have not yet undergone homogenisation, there may be extensive intraspecific ITS variation. This variation can be used to distinguish between isolates and strains. ITS sequence differences have been used to distinguish between closely related bacteria; notably species and strains of mycobacteria can be distinguished in this way (De Smet et al., 1995; Frothingham and Wilson, 1993; 1994). Two or more distinct ITS1 types were shown to exist in isolates of *Bchiococcus granulosus* (Bowles and McMannus, 1993), which is very similar to the situation that we have found in the *T. parva* ITS. Sequence divergence between a pathogenic and a non-pathogenic strain of *Entamoeba histolytica* has been found in the ITS2 (Cevallos et al., 1993). Major differences in ITS1 and ITS2 sequences correlate with pathogenicity grouping of isolates of *Leptosphaeria maculans* (Morales et al., 1993). Three distinct ITS types were found in *Trichoderma harzianum* isolates; group 2 isolates, which were aggressive colonizers of mushroom compost, could be distinguished from the other two groups using the ITS sequences (Muthumeenkshi et al., 1994).
homogeneous (Bishop et al., 1994). In the same study T. p. /CHV/23, another isolate originating from Chikeya farm, was found to be atypical, in that its Mab profile was similar to Zimbabwean stocks which were putatively derived from buffalo, but the rRNA gene pattern was similar to that of isolates from cattle. The \textit{T}pr RFLP pattern differed from the predominant genotypes seen in most isolates from cattle, but it was not similar to that of the other isolates from buffalo. We have found three distinct \textit{ITS}s in T. p. /CHV/23 (table 3.3), which is thus likely to be from a mixed infection. The three \textit{ITS}s were all different from the two \textit{ITS}s of T. p. /AYR/12 and T. p. /CHV/16 (fig 3.2).

T. p. /MAS/38 and T. p. /MAS/40, two isolates obtained from buffalo on Masuri farm, also each contained two \textit{ITS}s. The \textit{ITS}-L sequences shared 99.3\% identity but the \textit{ITS}-S sequences were identical (nos. 12L, 12S, 13L and 13S in fig. 3.2). Differences in the \textit{ITS}-L sequences of these isolates occurred as a few point mutations which suggests that these sequences have only very recently diverged from a common genotype.

By contrast two isolates obtained from buffalo on a second farm, Bally Vaughin, T. p. /BAL/25 and T. p. /BAL/93 each contained two \textit{ITS}s and all four of these \textit{ITS}s were distinct (nos. 10L, 10S, 11L and 11S in fig. 3.2). It is interesting to note that the longer \textit{ITS} sequence of each of these isolates contained the large insert of approximately 70 bp which was previously observed in the conserved region between \textit{V1} and \textit{V2} in \textit{T. p. lawrencei} Hluhluwe \textit{ITS}-L (see the region between columns 206 and 287 in fig. 3.2). This insert was also found in \textit{T. p. lawrencei} KNP 2 \textit{ITS}-L (no. 9L in fig. 3.2). The sequence of this insert was virtually identical in these four isolates, differing by only two or three point mutations per sequence, so it is reasonable to assume that these isolates shared a common ancestor in the relatively recent past.
The full length of the *T. parva* ITS sequences from the 3' end of the SSUrRNA gene to the putative 5' end of the LSUrRNA gene varied between 1058 and 1210 bp. The ITS1 varied between 500 and 600 bp in length and contained two variable regions as described in Chapter 2. The *Theileria* 5.8S gene was approximately 180 bp in length and was highly conserved between all of the *T. parva* isolates studied. The ITS2 was approximately 300 bp long and could be further divided into two variable regions (V3 and V4) separated by a conserved sequence (fig. 3.2).

Few of the ITS sequences were identical to any other over their full length. However, each ITS had sections homologous to those of a number of other ITS sequences within the variable regions. For example, *T. p. parva* Schoonspruit ITS (No. 3 in fig. 3.2) was initially similar to *T. p. parva* Muguga and Uganda at the beginning of V1, but in the middle of V1 (see the region between columns 61 and 111 in fig. 3.2) it changed to resemble *T. p. lawrencei* 7344 ITS-L (no. 7L in fig. 3.2). At the start of V2, *T. p. parva* Schoonspruit was once again identical to *T. p. parva* Muguga and Uganda, while in the middle of V2 its sequence became more similar to those of *T. p. lawrencei* 7344 ITS-L, *T. p. lawrencei* Hluhluwe 3 ITS-L and ITS-S (see column 619, fig. 3.2). In V3 and V4 the *T. p. parva* Schoonspruit ITS again resembled that of *T. p. parva* Muguga and Uganda.

The ITS sequences obtained from the Zimbabwean isolates, *T.p./AYR/12* and *T.p./CHU/16* were almost identical. The isolates were obtained from cattle on two different farms, Ayrshire and Chikaya, where there were no buffalo, nor ticks which may have fed on buffalo. Each of these isolates contained two ITS sequences, a longer one (L) and a shorter one (S) (table 3.3 and nos. 4L, 4S, 5L and 5S fig. 3.2). T.p./CHU/16 ITS-L was 98.3% identical to *T.p./AYR/12* ITS-L, while T.p./CHU/16 ITS-S and T.p./AYR/12 ITS-S were identical. This finding correlates with a previous study in which RFLP analysis and probing with *T. parva* Tpr repetitive and ribosomal DNA probes and screening with a panel of anti-schizont MAbs indicated that *T. parva* parasites isolated from cattle on nine farms in Zimbabwe were relatively
Figure 3.2 continued, page 5 of 7.
Figure 3.2: Alignment of ITS sequences obtained in this study with *T. p. parva* Muguga (Genbank accession number L26331) and Uganda (Genbank accession number L26332) ITS sequences (Kibe *et al.*, 1994). ITS sequences are numbered and correlate to ITS sequences listed in table 3.3. Variable regions in the *T. parva* ITS1 can be found between columns 10 and 567 (V1) and between columns 619 and 752 (V2). The 5.8S gene is located between columns 756 and 935. The ITS2 of *T. parva* can also be divided into two variable regions between columns 936 and 1079 (V3) and between columns 1174 and 1270 (V4). - ; sequence not done.
parasites, as most of them were field isolates and not cloned parasites. However, we had already demonstrated sequence polymorphism between the two ITS sequences within a cloned parasite (Chapter 2; Kibe et al., 1994), so the presence of more than one ITS sequence was not necessarily an indication of the presence of a mixed isolate. An alignment of these ITS sequences with *T. p. parva* Muguga and Uganda ITS sequences (Kibe et al., 1994) is shown in fig. 3.2.

Table 3.3: Number of ITS sequences cloned and sequenced from each *T. parva* isolate

<table>
<thead>
<tr>
<th>Species origin</th>
<th>Isolate</th>
<th>No. of ITS sequences</th>
<th>Length (bp)</th>
<th>Designation</th>
<th>No. in fig. 3.2</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>T. parva</em> isolated from cattle</td>
<td>Muguga LNMU1</td>
<td>1</td>
<td>1286</td>
<td>ITS-L</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Muguga SM</td>
<td>1</td>
<td>1286</td>
<td>ITS-L</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Schoonspruit</td>
<td>1</td>
<td>1310</td>
<td>ITS-L</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>T.p./AYR/12</td>
<td>2</td>
<td>1329</td>
<td>ITS-L</td>
<td>4-S</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1318</td>
<td>ITS-S</td>
<td>4-S</td>
</tr>
<tr>
<td></td>
<td>T.p./CHI/16</td>
<td>2</td>
<td>1328</td>
<td>ITS-L</td>
<td>5-L</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1317</td>
<td>ITS-S</td>
<td>5-S</td>
</tr>
<tr>
<td></td>
<td>T.p./CHI/23</td>
<td>3</td>
<td>1342</td>
<td>ITS-L</td>
<td>6-L</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1304</td>
<td>ITS-S1</td>
<td>6-S1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1305</td>
<td>ITS-S2</td>
<td>6-S2</td>
</tr>
</tbody>
</table>

| *T. parva* isolated from buffalo | 7544/05/55/ES | 2 | 1316 | ITS-L | 7-L |
| | | | 1262 | ITS-S | 7-S |
| | Hluhluwe 3 | 2 | 1405 | ITS-L | 8-L |
| | | | 1339 | ITS-S | 8-S |
| | KNP 2 | 2 | 1354 | ITS-L | 9-L |
| | | | 1267 | ITS-S | 9-S |
| | T.p./BAL/25 | 2 | 1396 | ITS-L | 10-L |
| | | | 1368 | ITS-S | 10-S |
| | T.p./BAL/93 | 2 | 1357 | ITS-L | 11-L |
| | | | 1294 | ITS-S | 11-S |
| | T.p./MAS/38 | 2 | 1333 | ITS-L | 12-L |
| | | | 1253 | ITS-S | 12-S |
| | T.p./MAS/40 | 2 | 1334 | ITS-L | 13-L |
| | | | 1253 | ITS-S | 13-S |
PCR amplification with primers 613 and 792 generated products approximately 1000 bp in length from the *T. p. lawrencei* Hluhluwe series (fig. 4.1, lanes 5-8) and the Zimbabwean isolates from buffalo (data not shown). However, four p67 gene amplification products were generated from each of the four new *T. p. lewisi* KNP/94 stocks (fig. 4.1, lanes 1-4). The shortest of these was 800 bp in length (this PCR product was designated S), there was a medium-length product of approximately 900 bp (designated M) and a doublet at approximately 1000 bp (designated L1 and L2). There are therefore at least four different parasite genotypes in each of the four *T. p. lawrencei* KNP/94 stocks, as the p67 gene is a single copy gene (Nene et al., 1992). It is possible that these were not all derived from *T. parva* (see section 4.3.3). That an identically-sized set of four different PCR products were amplified from each of these stocks, might indicate that there was a uniform mix of parasite genotypes in the buffalo herd from which these stocks were isolated. ITS probing data (Chapter 3) also indicated that these four buffalo were carrying a uniform population of parasites: all four were detected by the *T. p. lawrencei* ITS probe TPL2, but by no other probes.

4.3.2 Variation between *T. parva* p67 sequences

The *T. parva* p67 nucleotide sequences obtained in this study were aligned with p67 nucleotide sequences previously obtained by V. Nene and E. Gobright (pers. comm.). An alignment of the variable region of the p67 gene sequences from primer 613 to the end of the 29 bp intron (Nene et al., 1992) is shown in fig 4.2. The sequence of the 29 bp intron is shown in lower case letters. An alignment of the amino acid sequences deduced from these nucleotide sequences is shown in fig. 4.3. Amino acid residues that differ from those of the *T. p. parva* Muguga 1'67 amino acid sequence (Nene et al., 1992) are highlighted. Differences between *T. parva* p67 sequences are more easily visualised in the deduced amino acid sequence shown in fig. 4.3, and I will therefore refer to this figure in the following discussion. (Note: numbers given in brackets after isolate names refer to rows in fig. 4.3).
Fig. 4.1: PCR products of the variable region of the p67 gene generated with primers 613 and 792. Four products were generated from the four *T. parva* KNP/94 isolates (lanes 1-4, note: running of this gel for an extended period revealed that the upper band was a doublet). A single product of approximately 1000 bp was generated from the *T. p. lawrencei* Hluliwe series (lanes 5-8). The amplification products from the two Zambian *T. parva* stocks isolated from cattle were approximately 900 bp (lanes 11 and 12). A negative calf thymus DNA control (lane 9) and a positive *T. p. parva* Muguga control (lane 10) were included. φX174 *Hae* III markers (lane M) were included for estimation of PCR product size.
4.2.2.3 p67 gene sequence analysis

Auto-radiographs of sequencing gels were scanned and sequences determined using an automatic base scanner and semi-automatic base-calling software (Amersham); gel readings were assembled and edited using XBAP (Staden, 1994). The p67 gene sequences were aligned using CLUSTALW (Thompson et al., 1994) and GDE (Smith, 1992) and amino acid sequences were deduced in GDE.

4.3 Results and discussion

4.3.1 Amplification of the variable region of T. parva p67 genes

Amplification products of the variable region of the p67 gene generated with primers 613 and 792 are shown in fig. 4.1. The amplification products from the two Zambian T. parva isolates isolated from cattle were approximately 900 bp in length (fig. 4.1, lanes 11 and 12); this product size is characteristic of other T. parva isolates from cattle (Nene et al., 1996). An amplification product of the variable region was not obtained from the Zimbabwean isolates T.p./CHI/16 and T.p./AYR/12 (data not shown). This suggests that the sequences of their p67 genes may vary at the primer sites. An amplification product of approximately 1000 bp was obtained from the Zimbabwean isolate from cattle, T.p./CHI/23 (data not shown). This product size is characteristic of T. parva isolates from buffalo. Screening with a panel of anti-schizont MAb (Bishop et al., 1994), as well as screening with ITS probes (Chapter 3) have indicated that this isolate is more similar to T. parva isolates from buffalo than to those isolated from cattle. However, the rRNA gene pattern was similar to that of isolates from cattle and although the Tpr RFLP pattern differed from the predominant genotype seen in most isolates from cattle, it was not similar to that of the other isolates putatively from buffalo (Bishop et al., 1994).
Table 4.2: Oligonucleotides used for amplification and sequencing of the variable region of *T. parva* p67 genes.

<table>
<thead>
<tr>
<th>Identity</th>
<th>Strand</th>
<th>Tm (°C)</th>
<th>Sequence 5' → 3'</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>613</td>
<td>+</td>
<td>62.4</td>
<td>AGA AAC ACA ATO CGA AGT TC</td>
<td></td>
</tr>
<tr>
<td>792</td>
<td>-</td>
<td>61.7</td>
<td>CTT TTA GTA COT TGG CG</td>
<td>None et al., 1996</td>
</tr>
<tr>
<td>145</td>
<td>-</td>
<td>69.5</td>
<td>GTT CTT TCG COT TCA TAT GCC C</td>
<td></td>
</tr>
</tbody>
</table>

4.2.2.1 Direct sequencing of single p67 PCR products

Oligonucleotide primers 613 and 145 (table 4.2) were end-labelled and used for direct thermocycle sequencing of single PCR products according to the fluo™ DNA Sequencing System (Promega) (Appendix C.12.2).

4.2.2.2 Cloning and sequencing of multiple p67 PCR products

The four PCR products of different sizes which were generated from *T. p. lawrencei* KNP/94_1 were cloned directly into pMOSBlue-T (Amersham). This vector contains 3' thymidine overhangs which hybridise to the 3' adenosine nucleotides added to the ends of PCR products by *Taq* polymerase during the PCR (Clark, 1988). White colonies were picked and minipreparations of recombinant plasmid DNA were made using the Wizard™ Minipreps DNA Purification System (Promega) (Appendix C.9.1.2). Plasmid minipreparations were digested (Appendix C.4) with *Bam* HI and *Hind* III and examined by electrophoresis to ensure that the correct inserts had been cloned; clones containing inserts of all four sizes were obtained. The nucleotide sequences of the four *T. p. lawrencei* KNP/94_1 partial p67 gene clones were obtained by double stranded sequencing of plasmid minipreparations using the dideoxy chain termination method (Sanger et al., 1977) (Appendix C.12.1) with the T7 promoter primer and the U-19mer primer (Amersham).
<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>T.p./BAL/25**</td>
<td>Animal experimentally infected by ticks collected from buffalo grazed paddock, Bally Vaughan Creme Park, Zimbabwe.</td>
<td>Bishop et al. (1994)</td>
</tr>
<tr>
<td>T.p./BAL/93**</td>
<td>Produced severe reaction in cattle, typed as buffalo-derived using monoclonal antibodies</td>
<td>O. K. Kanhai and P. Spooner, pers. comm.</td>
</tr>
<tr>
<td>T.p./MAS/38**</td>
<td>Natural infection in bovines kept on buffalo grazed paddock, Mauret farm, Zimbabwe.</td>
<td>Bishop et al. (1994)</td>
</tr>
<tr>
<td>T.p./MAS/40**</td>
<td>Natural infection in bovines kept on buffalo grazed paddock, Mauret farm, Zimbabwe.</td>
<td>Bishop et al. (1994)</td>
</tr>
<tr>
<td>Hluhluwe 37</td>
<td>Isolated from naturally infected buffalo in Natal, South Africa. Noa.</td>
<td>Potgieter et al. (1988)</td>
</tr>
<tr>
<td>Hluhluwe 3 53**</td>
<td>3, 5, 6 &amp; 7 seroaggregates through cattle. Stable Corridor disease</td>
<td></td>
</tr>
<tr>
<td>Hluhluwe 3 65**</td>
<td>clinical picture.</td>
<td></td>
</tr>
<tr>
<td>KNP 25</td>
<td>Isolated from naturally infected buffalo in the Kruger National Park, eastern Transvaal, South Africa. Classical Corridor disease in cattle.</td>
<td>H. Stoltez (pers. comm.)</td>
</tr>
<tr>
<td>KNP/94_1**</td>
<td>Isolated from naturally infected cattle buffalo in the Kruger National Park, eastern Transvaal, South Africa. Corridor disease in cattle. Also positive for Theileria sp. (buffalo) and T. mutans.</td>
<td>H. Stoltez (pers. comm.)</td>
</tr>
<tr>
<td>KNP/94_2**</td>
<td>Isolated from naturally infected cattle buffalo in the Kruger National Park, eastern Transvaal, South Africa. Corridor disease in cattle. Also positive for Theileria sp. (buffalo) and T. mutans.</td>
<td>H. Stoltez (pers. comm.)</td>
</tr>
<tr>
<td>KNP/94_3**</td>
<td>Isolated from naturally infected cattle buffalo in the Kruger National Park, eastern Transvaal, South Africa. Corridor disease in cattle. Also positive for Theileria sp. (buffalo) and T. mutans.</td>
<td>H. Stoltez (pers. comm.)</td>
</tr>
<tr>
<td>KNP/94_4**</td>
<td>Isolated from naturally infected cattle buffalo in the Kruger National Park, eastern Transvaal, South Africa. Corridor disease in cattle. Also positive for Theileria sp. (buffalo) and T. mutans.</td>
<td>H. Stoltez (pers. comm.)</td>
</tr>
</tbody>
</table>
Table 4.1: *Theileria* stocks used as sources of DNA from which the variable region of the p67 gene was amplified and sequenced. The variable region of the p67 gene was amplified from stocks marked with an asterisk (*) in this study. p67 sequence data was obtained from stocks marked with a double cross (**) in this study. p67 gene sequences have already been obtained from stocks marked with a cross (†) (V. Nene and E. Gobright, pers. comm.).

<table>
<thead>
<tr>
<th>Species origin</th>
<th>Stock</th>
<th>No. in Figs. 4.2 and 4.3</th>
<th>Remarks</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Uganda†</td>
<td>2</td>
<td>Isolated from a beef ranch in northwestern Uganda.</td>
<td>Minami et al. (1983); Morzarte et al. (1995)</td>
</tr>
<tr>
<td></td>
<td>Madiabum†</td>
<td>3</td>
<td>Isolated from Kirinyaga District, Kenya.</td>
<td>Irvin et al. (1983)</td>
</tr>
<tr>
<td></td>
<td>Waterbuck-passaged**</td>
<td>4</td>
<td>Waterbuck experimentally infected with <em>T. parva</em> derived from buffalo 7014 transmitted a subpopulation of parasites which caused BCF in cattle.</td>
<td>Stagg et al. (1994)</td>
</tr>
<tr>
<td></td>
<td>Selousplains†</td>
<td>5</td>
<td>Isolated in 1937 from natural bovine BCF infection in Transvaal, South Africa.</td>
<td>Nelte (1948)</td>
</tr>
<tr>
<td></td>
<td>Zambia L1**</td>
<td>6</td>
<td>Isolated from a farm near Chipata, East Zambia. BCF clinical picture.</td>
<td>D. Clayson, pers. comm.</td>
</tr>
<tr>
<td></td>
<td>Zambia S105**</td>
<td>7</td>
<td>Isolated from a farm at Lumuli, East Zambia. BCF clinical picture.</td>
<td>D. Clayson, pers. comm.</td>
</tr>
<tr>
<td></td>
<td>T.p./CHI/16**</td>
<td>8</td>
<td>Isolated from naturally infected bovine on Chikuyu farm, Zimbabwe.</td>
<td>Bishop et al. (1994)</td>
</tr>
<tr>
<td></td>
<td>T.p./CHI/23**</td>
<td>8</td>
<td>Animal experimentally infected by ticks collected from vouching animal at Chikuyu farm. Sero reaction, died.</td>
<td>Bishop et al. (1994)</td>
</tr>
<tr>
<td></td>
<td>T.p./AYR/12**</td>
<td>Isolated from naturally infected bovine on Ayrshire farm, Zimbabwe.</td>
<td>Bishop et al. (1994)</td>
<td></td>
</tr>
<tr>
<td><em>T. parva</em> isolated from buffalo</td>
<td>7013**</td>
<td>10</td>
<td>Isolated from naturally infected cattle 7014 captured in the Nanyuki area, Kenya.</td>
<td>Kariuki, et al. (1990); Mozanbo et al. (1990)</td>
</tr>
<tr>
<td></td>
<td>7014**</td>
<td>11</td>
<td>Isolated from naturally infected carrier buffalo 7014 captured in the Nanyuki area, Kenya.</td>
<td>Kariuki, et al. (1990); Mozanbo et al. (1990)</td>
</tr>
</tbody>
</table>
isolated from cattle and buffalo are consistent and could be used for the development
of probes to distinguish between the two forms.

4.2 Materials and methods

4.2.1 Parasites

*T. parva* stocks used as sources of DNA from which the variable region of the p67
gene was amplified in this study are marked with an asterisk (*) in table 4.1. Those
from which sequences of the variable region of the p67 gene were subsequently
obtained are marked with a double cross (**). Stocks from which p67 gene sequences
have previously been obtained (Nene et al., 1996; V. Nene and E. Gobright, pers.
comm.) are marked with a cross (†).

4.2.1 PCR amplification of the variable region of the p67 gene

Primers 613 and 792 (table 4.2, Nene et al., 1996) were used to amplify the variable
region of the p67 gene from the eighteen *T. parva* stocks marked with an asterisk in
table 4.1. A positive *T. parva* Muguga and a negative calf thymus DNA control were
included. The PCR was performed as described in Appendix C.3. Annealing was
performed at 55°C. Cycling conditions were: 72°C for 1.5 min, 93°C for 1 min and
55°C for 1 min, for 30 cycles. PCR products were purified using Wizard™ PCR
Minipreps (Promega) as described in Appendix C.9.1.2.

4.2.2 Sequencing of the variable region of the p67 gene

Most *T. parva* stocks yielded a single p67 PCR product (fig. 4.1), which was
sequenced directly. Four differently sized products were generated from each of the
*T. p. lawrencei* KNP/94 isolates (fig. 4.1), these fragments therefore had to be cloned
individually prior to sequencing.
mass estimated from SDS-polyacrylamide gels is larger than the calculated mass. Nene et al. (1995) attempted to express the protein in the correct conformation using the baculovirus expression system, but this system resulted in only a very small amount of the protein in the correct form in insect cells and large amounts of a partially modified form.

4.1.3 Could the p67 gene be used to distinguish between T. p. parva and T. p. lawrencei stocks?

In order to be effective, a candidate vaccine antigen must provide protection against a broad range of T. parva stocks. To determine whether there was variation in p67 between different T. parva stocks, the gene coding for p67 was amplified from a number of different T. parva stocks and sequenced (Nene et al., 1996). The sequence of the gene coding for p67 was found to be conserved amongst T. parva parasites isolated from cattle studied to date, but there was a variable region in the p67 gene of T. parva parasites isolated from buffalo (Musoke et al., 1993; Nene et al., 1996). This evidence suggests that the variable region of the p67 gene could be a possible target region for the development of DNA probes to distinguish between T. p. parva and T. p. lawrencei stocks. In most stocks isolated from buffalo there were a number of nucleotide differences which translated to several amino acid substitutions, as well as an additional segment of 129 bp which had no homologue in stocks isolated from cattle. The only exception was the p67 gene of the T. p. lawrencei KNP 2 stock, isolated from buffalo, which had a number of nucleotide differences but which did not contain the additional segment of 129 bp (V. Nene, pers. comm.).

Four new T. parva isolates were recently obtained from buffaloes in the southern part of the Kruger National Park. We sequenced the variable region of the p67 gene from these isolates as well as from a number of Zimbabwean T. parva isolates, to determine whether the differences previously observed between T. parva stocks
sporozoites from the same stock as well as from a number of other stocks. Immunity in these animals correlated with the level of anti-sporozoite antibodies produced: those with high antibody titres were completely protected, those with low titres developed a parasitosis. The broad range of protection induced by sporozoite antigens suggests that there is a protective antigenic determinant that is conserved between the sporozoites of different *T. parva* stocks. This would be an excellent candidate for the development of a recombinant vaccine.

4.1.2 Identification and cloning of the gene encoding p67, the major sporozoite antigen

The major antigenic determinants of *T. parva* sporozoites were identified using C16, an antiserum directed against sporozoite antigens (Lam et al., 1990; Musoke and Nene, 1990). The major target antigen of sporozoite neutralising antibodies, p67, was identified and the gene which codes for p67 was cloned and characterised (Nene et al., 1992). The deduced amino acid sequence of the p67 gene contains 709 residues and it is predicted to be a membrane protein. The gene contains a short (29 bp), AT-rich intron. It is a single copy gene which was mapped to a 780 kbp SfiI fragment, located on *T. parva* chromosome 3 (Morzaria and Young, 1992). The p67 gene is expressed only in the sporozoite stage.

The *T. parva* Muguga p67 gene was cloned into pGEX and pMGI and expressed in *E. coli* (Nene et al., 1992; Musoke et al., 1992). Recombinant p67 expressed from pMGI was produced as a fusion protein with the first 85 amino acid residues of a nonstructural gene (NS1) of influenza virus A. This fusion protein, NS1-p67, was shown to induce protection against homologous challenge in six out of nine cattle (Musoke et al., 1992). Six out of eleven cattle were protected against heterologous challenge with the *T. parva* Muguge stock (Musoke et al., 1993). However, NS1-p67 does not have the same characteristics as the native form of p67: Tp M12, a MAb which binds to native p67, does not react with NS1-p67 and the relative molecular
CHAPTER 4: VARIATIONS IN THEILERIA PARVA p67 GENES

4.1 Introduction

4.1.1 Both schizonts and sporozoites stimulate immune responses to T. parva infection

It has been shown that both cell-mediated responses directed against schizonts (Eugui and Emery, 1981) and humoral responses directed against sporozoites (Musoke et al., 1982) are responsible for immunity in cattle which have recovered from T. parva infection. It is believed that the cell-mediated response is most important for control of the parasite, as the appearance of cytotoxic T lymphocytes coincides with the clearance of parasitised cells (Eugui and Emery, 1981; Morrison et al., 1987). The effector cells responsible for clearance of parasitised cells are MHC-I-restricted, CD8+ cytotoxic T lymphocytes (Morrison and Goddeeris, 1990), which are often parasite-stock specific. A limitation of the current infection and treatment method of immunisation is that it induces cell-mediated immunity and is therefore usually parasite stock specific (Radley, 1981).

Musoke et al. (1984) showed that sera from cattle immunised with T. parva Muguga were able to neutralise the infectivity of sporozoites from both homologous and heterologous stocks in vitro. Monoclonal antibodies (MAbs) raised against sporozoite antigens display a similar ability to neutralise sporozoites from both homologous and heterologous stocks in vitro (Musoke et al., 1984; Dobbelare et al., 1984; 1985). Musoke et al. (1984) obtained comparable results in vivo: cattle were immunized by infection and treatment, using a method of immunization that favoured the production of antibodies. Those animals were protected from challenge with
The mosaic of sequences present in the ITS suggests that *T. parva* parasites undergo genetic recombination. Therefore, the finding that the probes for *T. p. lawrencei* and *T. p. parva* isolates are not conserved between all *T. parva* isolates that we have screened is probably due to recombination. We cannot identify probes in the ITS which give an unequivocal differentiation between the two forms, but we have a panel of oligonucleotides which can assign a probability to the origin of a *T. parva* parasite. Extrapolating from the results we have obtained (29 out of 31 *T. parva* isolates from cattle were detected by TPP1; 19 out of 23 isolates from buffalo were detected by TPL1 and/or TPL3a) we can say that if probe TPP1 hybridises to an isolate, there is a 93.5% probability that it is a *T. p. parva* isolate. Conversely, if an isolate is detected by probes TPL2 and/or TPL3a, we have found an 81.8% probability that it is a *T. p. lawrencei* isolate. However, there will be small numbers of isolates that react with both types of probes and it would not be possible to assign an origin to isolates such as this. Despite these reservations, our panel of 11 oligonucleotides will be invaluable for epidemiological surveys to determine the distribution of most of the different genotypes.

### 3.3.4 Summary

We had hoped that variability within the ITS would be such that a clear-cut distinction between *T. p. lawrencei* and *T. p. parva* isolates could be made, as has been done previously to distinguish between isolates of other organisms (Cevallos *et al.*, 1993; Morales *et al.*, 1993; Muthumeenakshi *et al.*, 1994). It appears, however, that *T. p. lawrencei* and *T. p. parva* parasites still undergo genetic recombination, the two populations are, therefore, not yet separate gene pools and absolute discrimination is not possible. Under these circumstances no probe will distinguish between all *T. p. lawrencei* and *T. p. parva* parasites unless it targets a sequence which is directly associated with the mechanisms responsible for their differential behaviour in cattle.
It has been shown to provide protection against a broad range of Zimbabwean T. parva isolates (Koch et al., 1990; Hove et al., 1995). It is interesting that T. parva Boteni has the same ITS profile as isolates obtained from cattle in Malawi and Zambia and the South African T. p. parva Schoongrultz.

All of the Zimbabwean isolates that have previously been typed as 'buffalo-derived' (Bishop et al., 1994), were detected by one or more of the T. p. lawrencei ITS probes, with the exception of T.p/BAL/27. This isolate was only detected by TPP1, one of the T. p. parva ITS probes. T.p/MAS/37 was also detected by TPP1, although it also hybridised to TPL3a and TPL3b. If the sequence detected by TPP1 is truly a marker for T. p. parva and T. p. bovis isolates, these results suggest that parasites with a similar genotype to that of T. p. bovis are also present in buffalo in Zimbabwe. It is possible that these parasites were introduced via ticks into the buffalo through contact with infected cattle, resulting in a mixed population of T. p. lawrencei and T. p. bovis parasites in the buffalo. Alternatively, these parasites could be representative of an ancestral form of T. parva occurring in buffalo which, when introduced into cattle, has the potential to adapt to cattle and cause January disease. However, this seems unlikely as January disease and East Coast fever were never recognised in southern Africa prior to the introduction of T. p. parva from East Africa.

In general, the T. p. parva isolates hybridised with a limited subset of the probes, while the T. p. lawrencei isolates hybridised with many more of the probes, resulting in a more variable pattern of reactivity. These results support previous suggestions (Koch et al., 1990; Bishop et al., 1994) that there appears to have been a selection in cattle of a relatively homogeneous subpopulation of T. parva from a much larger and more diverse gene pool in buffalo. Unfortunately, however, there is no absolute distinction between T. p. parva and T. p. lawrencei isolates using these probes. The possibility that some of the isolates could be derived from mixed stocks cannot be ruled out, although some isolates which gave mixed results were cloned parasites (e.g. T. parva Uganda).
Research Institute, Kenya, where workers have already produced a number of *T. parva* parasite clones (Morzaria *et al.*, 1995).

### 4.3.5 Summary

The p67 gene of *T. parva* stocks isolated from cattle is, with some exceptions, highly conserved, while it is very much more variable in *T. parva* stocks isolated from buffalo. This is consistent with the ITS probing results which also indicate that the *T. parva* population in buffalo is more diverse. In other parts of Africa, e.g. in Zimbabwe (Bishop *et al.*, 1994) and possibly in East Africa, there appears to have been a selection of a homogenous subpopulation of *T. parva*, able to replicate and maintain itself in cattle populations, from a much larger and more diverse gene pool in buffalo. It seems likely that the two forms of the parasite are in the process of speciation, with *T. p. parva* types adapting to cattle as their host. The *T. p. parva*-like p67 genes that we have found in *T. parva* isolates from buffalo in the Kruger Park may represent a subpopulation of *T. parva* parasites which gave rise to ECP-causing *T. parva* parasites in other parts of Africa, i.e. it may be representative of an ancestral *T. p. parva* genotype.
parasites could then be used to infect cattle, to determine if one is more likely to become established in cattle. The p67 gene could be used as a marker for distinguishing between these cloned parasites, using PCR to obtain amplification products of the different p67 genes.

It is also of interest to note that T. p. parva-like p67 sequences were found in isolates from buffalo. If we did isolate a T. parva parasite which caused ECF from the mixture of T. parva genotypes in the buffalo, should we call this isolate buffalo-derived or cattle-derived? This is an obvious problem with the cattle-derived / buffalo-derived terminology. T. p. parva isolates presumably originated in buffalo and there is no reason to suppose that T. parva isolates which could cause ECF may not still be found in buffalo.

4.3.4 p67-Based probes to distinguish between T. p. parva and T. p. lawrencei

The presence of an additional sequence in the p67 gene in T. p. lawrencei stocks suggests a possible region for the development of DNA probes to distinguish between T. p. parva and T. p. lawrencei stocks. For T. p. parva stocks a probe could be designed which spans the region of the additional sequence in T. p. lawrencei isolates, while for T. p. lawrencei stocks, a conserved region in this insert could be used. However, an obvious problem with this strategy is that the probe for T. p. parva would detect all of the KNP isolates, which contain a T. p. parva-like p67 gene. Conversely the probe for T. p. lawrencei isolates would detect at least two Zimbabwean isolates, T.p./CHI/23 and T.p./HUN/96, both isolated from cattle.

These isolates could (and probably do) contain a mixture of T. parva genotypes. This highlights the necessity of obtaining cloned parasites, with well-defined disease syndromes, from field isolates. It is essential that we know exactly what sort of a parasite we are working with before we can begin to attempt to develop probes to distinguish between them. Work to this end is well under way at the International
destroyed). Thus, *T. parva* parasites have not had the opportunity to become established in cattle in South Africa since the eradication of ECF. However, this does not address the question of why ECF did not arise in South Africa prior to its introduction in the early part of this century. Perhaps the *T. parva* parasites present in buffalo had not accumulated the necessary mutations which made them fit to survive in cattle in other parts of Africa. Norval *et al.* (1991) have proposed that 'transformation' has never occurred in southern Africa because of the limited contact between cattle and buffalo and the seasonal occurrence of the tick vectors.

Despite numerous passages through cattle, 'transformation' has never been reported in the South African isolate *T. p. lawrencei* Hluhluwe 3 (Potgieter *et al.*, 1988). ITS and p67 data reported here support previous indications that this isolate has remained genetically unchanged after successive passages in susceptible splenectomised cattle. Both ITS and p67 sequence data suggest that *T. p. parva*-like sequences are not present in this isolate. However, p67 data indicate that the *T. p. lawrencei* isolates from the Kruger National Park do contain *T. p. parva*-like p67 sequences. It would be interesting to infect cattle with the *T. p. lawrencei* KNP/94 isolates and to determine if a *T. parva* parasite with a *T. p. parva*-like p67 type is able to become established in cattle, and whether that subpopulation would cause typical ECF.

Unfortunately, there would be practical difficulties in passaging these parasites through cattle, as the *T. p. lawrencei* parasites which cause Corridor disease are likely to kill the cattle before any other subpopulation of (e.g. *T. p. parva*-like) parasites could become established. In addition, the outcome of such an experiment could depend on the initial dose of sporozoites of each type (Dolan *et al.*, 1984). It would therefore be necessary to make parasite clones, each one containing one of the four different p67 types. Cattle could then be infected separately with equivalent doses of each cloned parasite and the disease caused by each could be determined. This would also confirm the hypothesis that L2 and S are clones of the p67 genes of *Theileria* sp. (buffalo). Mixtures of equal numbers of sporozoites from the cloned
the PCR products (M) was characteristic of *T. parva* stocks isolated from cattle in that it lacked the 129 bp region specific for *T. p. lawrencei*, although there were a number of amino acid substitutions.

Although all of the *T. p. lawrencei* isolates from the Kruger Park contained a *T. p. parva*-like p67 gene, none of them was detected by *T. p. parva* ITS probes (Chapter 3). *T. p. lawrencei* KNP 2 was detected by TPL3a and TPL3b, while the four *T. p. lawrencei* KNP/94 isolates were detected by TPL2 (Chapter 3). In addition, these isolates cause classical Corridor disease when transmitted to cattle by infected ticks (H. Stoltz, pers. comm.), even though *T. p. parva*-like forms of the p67 gene were present in the buffalo. It is tempting to speculate that the *T. p. parva*-like p67 gene may represent a subpopulation of *T. p. parva*-like parasites. The differences found in the deduced amino acid sequence could indicate that it is an ancestral *T. p. parva*-like genotype which may have given rise to ECF-causing *T. parva* parasites in other parts of Africa. Alternatively, the *T. p. parva* parasite which was introduced into South Africa in the early part of this century may have been transmitted into the buffalo population; mutations could have resulted in the slight sequence differences we have observed in the p67 gene. Either way, this raises the issue of whether *T. parva* parasites which could cause ECF exist in South African buffalo.

It is thus of importance to establish what the *T. p. parva*-like p67 sequences in the *T. p. lawrencei* KNP/94 isolates actually represent. Are they in fact indicative of a subpopulation of *T. parva* parasites that could cause ECF? This could well be the case because the strict control measures implemented in the areas surrounding the Kruger Park have prevented *T. parva* parasites from spreading into cattle herds. That there have not been any outbreaks of ECF in South Africa since 1954 can be attributed to the regular dipping regimes, which destroy potentially infective ticks before sporozoites can be transmitted, and to the strict control measures associated with outbreaks of Corridor disease (herds are isolated and infected animals
completely specific for *T. parva* p67 genes (V. Nene, pers. comm.). It has already been shown that the SSUrRNA genes of *T. parva* and *Theileria* sp. (buffalo) are very similar (Allsopp *et al.*, 1993); it is thus not unlikely that the p67 gene sequences of these two species could also be very similar. It thus seems possible that L2 and S represent the p67 genes of *Theileria* sp. (buffalo).

It is possible that the cell cultured *T. p. lawrencei* KNP 2 is representative of only a portion of the *T. parva* parasites that were present in the original buffalo from which the *T. p. lawrencei* KNP 2 isolate was made. *T. p. lawrencei* KNP 2 was isolated from an African buffalo cow captured in the southern part of the Kruger park. *Rhipicephalus zambeziensis* nymphs were fed on this buffalo; adult ticks reared from these nymphs transmitted Corridor disease to an adult cow, B9678-2. Schizont-infected lymphoblastoid cells were established *in vitro* from lymph node aspirate obtained from B9678-2 (H. Stoltsz, pers. comm.). The variable region of the p67 gene amplified from DNA extracted from these cell cultures was found to be similar to the p67 gene of *T. p. parva*; it did not contain the region specific for *T. p. lawrencei*, although a number of amino acid substitutions distinguished it from other *T. p. parva* p67 sequences (V. Nene and E. Gobright, pers. comm.). Selection of a subpopulation of parasites able to survive in cattle may have occurred in the *T. p. lawrencei* KNP 2 isolate, either *in vivo* or in cell culture (Note: no attempts have been made to infect cattle with cell-cultured *T. p. lawrencei* KNP 2, so we do not know what clinical symptoms it would produce). If the p67 gene is a determinant of the ability of *T. parva* parasites to survive in different hosts, such selection could favour those with sequences characteristic of *T. p. parva* isolates.

The p67 genes of the new *T. p. lawrencei* KNP/94 isolates were amplified from DNA extracted directly from blood samples taken from four buffalo which were also captured in the southern part of the Kruger park. This amplification yielded four PCR products. One of these (L1) was characteristic of other *T. parva* stocks isolated from buffalo; it contained the 129 bp region specific for *T. p. lawrencei*. Another of
4.3.3 Sequence analysis of *T. p. lawrencei* KNP/94_1 p67 genes reveals sequences characteristic of both *T. parva* and *T. p. lawrencei*

Four different p67 amplification products (fig. 4.1) were generated from the four *T. p. lawrencei* KNP/94 stocks. The four products of different sizes (called L1, L2, M and S) from *T. p. lawrencei* KNP/94_1 were cloned into pMOSBlue-T and sequenced separately. The deduced amino acid sequences are shown in fig. 4.3 (Note: numbers given in brackets after isolate names in the following discussion refer to rows in this figure). Sequence analysis of the cloned *T. p. lawrencei* KNP/94 (no. 19) p67 amplification products showed that the longest, L1, was similar to those of the Zimbabwean stocks isolated from buffalo (nos. 13-16); it contained the additional 44 amino acid *T. p. lawrencei*-specific region and the amino acid substitutions (between the 7th and 17th amino acid residues of this region) which were characteristic of the Zimbabwean samples. The deduced amino acid sequence of clone M was similar to those of the stocks isolated from cattle in that it did not contain the *T. p. lawrencei*-specific region, but like that of *T. p. lawrencei* KNP 2 (no. 18), there were a number of amino acid substitutions when compared to *T. parva* stocks isolated from cattle. The deduced amino acid sequences of clones L1 and M contained a number of identical amino acid substitutions when compared to the *T. parva* stocks isolated from cattle.

The sequences of *T. p. lawrencei* KNP 94_1 p67 clones L2 and S were completely different from any of the other *T. parva* p67 gene sequences. The four *T. p. lawrencei* KNP/94 buffalo samples were diagnosed as positive for *Theileria* sp. (buffalo) and *T. mutans* as well as *T. parva* (A. K. Mynhardt, pers. comm.) by PCR amplification of a variable region of the SSU rRNA gene and probing with species-specific probes (Allsopp et al., 1993). Nene et al. (1992) found that a p67 gene probe did not hybridise to DNA isolated from *T. mutans* or *T. tauronagel*, although at low hybridisation stringency, a 3.4 kbp Eco RI DNA fragment of *T. annulata* was detected. No studies have been done to determine whether primers 613 and 792 are
11), 7344 (no. 12) and the Zimbabwean isolates (nos. 13-16), were more similar to those of the *T. parva* stocks isolated from cattle in this region.

The deduced amino acid sequence of *T. p./HUN/96* p67 (no. 9) (E. Gobright and V. Nene, pers, comm.) contained regions homologous to the p67 sequences of a number of the other isolates, resulting in a sequence mosaic much like that found in the ITS sequences (Chapter 3). The *T. p./HUN/96* p67 amino acid sequence had a large unique insert of 23 amino acid residues. The amino acid sequence approximately 30 to 50 residues prior to this insert contained seven substitutions which were characteristic of the other Zimbabwean isolates (nos. 13-16). Immediately after the unique insert, the *T. p./HUN/96* amino acid sequence contained ten substitutions, eight of which were also found in the *T. p. lawrencei* KNP 2 sequence (no. 18). The 81 residue variable region immediately upstream of the sequence specific to *T. p. lawrencei* stocks was identical in *T. p./HUN/96* and the *T. p. lawrencei* Hluhluwe series (nos. 17, 17#3, 17#5, 17#6 and 17#7). In the region specific to *T. p. lawrencei* stocks, the *T. p./HUN/96* amino acid sequence initially resembled that of the *T. p. lawrencei* Hluhluwe series, although there were three amino acid differences, but at the 3' end of this region *T. p./HUN/96* contained a number of amino acid substitutions in common with 7013 (no. 10). After the region specific to *T. p. lawrencei* stocks, there were three amino acid substitutions in the *T. p./HUN/96* sequence which were also found in 7013. Thus, the deduced amino acid sequence of the *T. p./HUN/96* p67 gene resembled first that of the Zimbabwean stocks isolated from buffalo, then *T. p. lawrencei* KNP 2, then *T. p. lawrencei* Hluhluwe 3, then again the Zimbabwean isolates and finally *T. p. lawrencei* 7013. This sequence mosaic, like that found in the ITS sequences, suggests that there has been genetic recombination amongst *T. parva* parasites.
which is attributable to a single base pair difference (fig. 4.2). One would expect to find at least one identical p67 gene in T. p. lawrencei 7344 and 7014, as 7344 was cloned from the 7014 isolate. The single base pair difference may therefore be attributable to a Taq polymerase copying error. When compared to 7014 and 7344, 7013 contained a deletion of three amino acid residues in this region; other nucleotide differences in the 7013 p67 gene sequence in this region resulted in seven substitutions in its deduced amino acid sequence.

The p67 amino acid sequences of the Zimbabwean stocks isolated from ticks or cattle in areas where buffalo were present, were identical to each other in the T. p. lawrencei-specific region, but were different from the other T. p. lawrencei isolates (fig. 4.3). The first six amino acid residues were conserved between all of the T. parva stocks isolated from buffalo (except for the T. p. lawrencei KNP isolates, see discussion in section 4.3.3). The subsequent amino acid sequence was variable until the final 24 residues where each sequence differed by only a few amino acid substitutions, except for T. p. lawrencei 7013 which had seven amino acid substitutions.

The p67 deduced amino acid sequences were identical in the four cattle passages, #3, #5, #6 and #7, of the T. p. lawrencei Hluhluwe 3 isolate (17#3, 17#5, 17#6 and 17#7, fig. 4.3). The same ITS probes (namely, TPP2, TPL1 and TPL3a) detected DNA isolated from these four cattle passages of the T. p. lawrencei Hluhluwe 3 isolate (see Chapter 3). These results suggest that the genotype of this isolate has remained unchanged after extensive passage through cattle. There was a variable region in the T. p. lawrencei Hluhluwe 3 p67 sequence, which began approximately 81 residues (column 136, fig. 4.3) upstream of the additional sequence which is specific to most T. p. lawrencei stocks. The sequence of this portion of the p67 variable region in T. p. lawrencei Hluhluwe 3 was very similar to that of T. p. lawrencei 7013 (no. 10). The p67 gene sequences of the other T. parva stocks isolated from buffalo, 7014 (no.
The p67 amino acid sequences of the two Zambian T. parva isolates from cattle (nos. 6 and 7) were identical to those of other T. parva isolates from cattle; they did not contain the additional sequence that is characteristic of T. parva stocks isolated from buffalo and there were no amino acid substitutions. However, the p67 sequences of two of the isolates from Zimbabwe were aberrant. The p67 amino acid sequence of T.p./CHI/23 (no. 8) isolated from cattle was identical to those obtained from the Zimbabwean T. parva isolates from buffalo. It contained the additional sequence that is characteristic of T. p. lawrencei isolates and there were a number of amino acid substitutions. This provides more evidence to suggest that this isolate is more like other Zimbabwean T. parva isolates from buffalo, even though it was obtained from Chikeya farm where there were no buffalo, nor ticks which may have fed on buffalo (Bishop et al., 1994). T.p./HUN/96 (no. 9), another Zimbabwean isolate from cattle, also contained the additional sequence that is characteristic of T. p. lawrencei isolates. However, ITS probing results are more consistent with the provenance of this isolate: it was detected by TPP1, a T. p. parva-specific probe (Chapter 3).

Sequence analysis of the p67 amplification products generated from the T. p. lawrencei Hluhluwe series (nos. 17, 17#3, 17#5, 17#6 and 17#7) and the Zimbabwean T. parva stocks putatively isolated from buffalo (nos. 13-16) revealed that they all contained a number of amino acid substitutions when compared to isolates from cattle (fig. 4.3), as well as the additional sequence which is typical of most other T. parva stocks isolated from buffalo (Nene et al., 1996). This T. p. lawrencei-specific sequence, located between columns 227 and 270 in fig. 4.3, is also present in T. p. lawrencei 7013 (no. 10), 7014 (no. 11) and 7344 (no. 12), but is absent from T. p. lawrencei KNP 2 (no. 18). The length of this sequence ranges from 40 amino acid residues in 7013, to 43 amino acids in 7014, 7344 and the Hluhluwe series, or 44 residues in the Zimbabwean isolates.

Within the additional sequence specific to buffalo-derived stocks, T. p. lawrencei 7014 and T. p. lawrencei 7344 differed by a single amino acid residue (fig. 4.3),
Fig 4.3: Deduced amino acid sequences of partial T. parva p67 gene sequences shown in fig. 4.2. Numbering (1-19) corresponds to T. parva stocks listed in table 4.1. Amino acid residues which differ from those of the T. p. parva Muguga (no. 1) p67 amino acid sequence are highlighted. - : sequence not obtained.
Fig. 4.2: Nucleotide sequence of the variable region of *T. parva* p67 genes, from primer 613 to the end of the 29 base pair intron (shown in lower case letters). Numbering (1-19) of *T. parva* stocks corresponds to numbers shown in table 4.1. ~: sequence not obtained.
APPENDIX C: DETAILS OF MOLECULAR BIOLOGICAL TECHNIQUES USED IN THIS STUDY

C.1 Purification of nucleic acids

C.1.1 Phenol: chloroform extraction

Nucleic acids were purified from proteins by phenol and chloroform extractions as described by Sambrook et al. (1989). The total volume of the solution was determined and an equal volume of water-saturated phenol was added. The solution was mixed very gently to form an emulsion when extracting genomic DNA, or vortexed to form an emulsion when isolating small DNA molecules (<10 kb). The phases were separated by centrifugation at 1600g for 10 minutes in a Beckman JA-14 rotor for large volumes, or at 12000g for 3-5 minutes in a microcentrifuge for smaller volumes. The upper, aqueous phase was transferred to a fresh tube. Deproteinization is more efficient when two organic solvents are used, so if large amounts of protein were present (e.g. in extractions of DNA from cell culture or blood samples) a further extraction with an equal volume of phenol: chloroform: isoamyl alcohol (25:24:1 (v/v)) was performed. Traces of phenol were removed by extraction with an equal volume of chloroform: isoamyl alcohol (24:1 (v/v)).

C.1.2 Ethanol precipitation

C.1.2.1 Of genomic DNA

Genomic DNA was precipitated by adding 0.1 volumes of 3 M sodium acetate (pH 5.2) and 2.5 volumes of absolute ethanol at room temperature (Sambrook et al., 1989). The solution was mixed to consolidate the strands of DNA, which were then
2xTY

bacto-tryptone  16 g
yeast extract   10 g
sodium chloride  5 g
Make to one litre with distilled water

10xM9 salts

sodium hydrogen phosphate  60 g
potassium dihydrogen phosphate  30 g
ammonium chloride  10 g
sodium chloride  5 g

Glucose / minimal medium (GMM) plates

Autoclave the following reagents separately and cool before mixing aseptically:
agar  15 g in 900 ml water
10xM9 salts  100 ml
1 M magnesium sulfate  1 ml
0.1 M calcium chloride  1 ml
1 M thiamine HCl  1 ml
20% glucose  10 ml
**Luria-Bertani (LB) medium**

- bacto-tryptone 10 g
- yeast extract 5 g
- sodium chloride 10 g

Make to one litre with distilled water

**LB plates**

- bacto-tryptone 10 g
- yeast extract 5 g
- sodium chloride 10 g
- agar 15 g

Make to one litre with distilled water

**H agar plates**

- bacto-tryptone 10 g
- sodium chloride 8 g
- agar 12 g

Make to one litre with distilled water

**H top agar**

- bacto-tryptone 10 g
- sodium chloride 8 g
- agar 8 g

Make to one litre with distilled water
APPENDIX B: COMPOSITION OF COMMON SOLUTIONS AND MEDIA USED IN THIS STUDY

TE

10 mM Tris-HCl, 1 mM EDTA, adjust to pH 8.0

10xTBE

Tris-HCl 108 g
boric acid 55 g
EDTA 9.3 g
Make to one litre with distilled water

20xSSC

sodium chloride 175.3 g
sodium citrate 88.2 g
Adjust to pH 7.0
Make to one litre with distilled water

Hybridisation buffer

1 M sodium phosphate buffer (pH 7.5) is made by mixing 840 ml of 1 M di-sodium hydrogen phosphate and 160 ml of 1 M sodium dihydrogen orthophosphate. Hybridisation buffer [0.5 M sodium phosphate buffer (pH 7.5) and 7% SDS] is then made by mixing equal quantities of 1 M sodium phosphate buffer and 14% SDS.
tris-HCl  
triton X-100  
urea  
X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside)  
xylene cyanol FF  
yeast extract

Boehringer Mannheim  
BDH  
USB  
Promega  
Sigma  
Difco

Bacterial strain

Escherichia coli TGI [supE hsdS6 thiA(lac-proA65) traD36 proAB lacI lacZΔM15] maintain on GMM plates
ethidium bromide
ethylenediaminetetra-acetic acid disodium salt (EDTA)
ficoll
formamide
glucose
herring sperm DNA
isoamyl alcohol
isopropyl-β-D-thiogalactopyranoside (IPTG)
lithium chloride
magnesium chloride
magnesium sulfate
mineral oil
nonidet P40
phenol
polyethylene glycol (PEG) 8000
potassium acetate
potassium chloride
potassium dihydrogen orthophosphate
propan-2-ol
sodium acetate
sodium chloride
tri-sodium citrate
sodium dodecyl sulfate (SDS)
sodium dihydrogen orthophosphate
di-sodium hydrogen phosphate
sodium hydroxide
tetracycline
N,N,N',N'-tetramethyl-ethylenediamine (TEMED)
thiamine HCl
trichloroacetic acid

Sigma
BDH
Sigma
Saarchem
Merck
Boehringer Mannheim
Merck
Boehringer Mannheim
Sigma
BDH
Merck
Saarchem
Merck
BDH
Merck
Saarchem
BDH
Sigma
Merck
BDH, Saarchem
Boehringer Mannheim
Sigma
Saarchem
APPENDIX A: SOURCE OF REAGENTS AND CHEMICALS USED IN THIS STUDY

acetic acid (glacial)  Saarchem
acrylamide  Merck
N,N'-methylene-bis-acrylamide  Merck
agar  Difco
agarose type I-A: low BEO  Sigma
ammonium acetate  BDH
ammonium chloride  BDH
ammonium persulfate  Sigma
ampicillin  Boehringer Mannheim
bacto-tryptone  Difco
boric acid  BDH, Merck
bovine serum albumin (BSA), acetylated  Promega
bromophenol blue  Sigma
calcium chloride  Merck
chloroform  Merck
deoxynucleoside triphosphates (dNTPs)  Boehringer Mannheim
dideoxynucleoside triphosphates (ddNTPs)  Boehringer Mannheim
7-deaza-2'-dGTP  Boehringer Mannheim
[α-32P] dCTP (10 mCi/ml, 3000 Ci/mmol)  Amersham
[α-35S] dATP (10 mCi/ml, 1000 Ci/mmol)  Amersham
[y-33P] dATP (10 mCi/ml, 5000 Ci/mmol)  Amersham
dimethyl sulfoxide (DMSO)  Merck
dimethydichlorosilane  BDH
dL-dithiothreitol (DTT)  Sigma
ethanol  BDH
5.3 Conclusion

By using both ITS probes and p67 gene sequences we were able to distinguish between most T. p. lawrencei and T. p. parva genotypes, but there are exceptions which do not conform to the patterns we have observed. These exceptions could be the result of genetic recombination between T. p. lawrencei and T. p. parva parasites. The mosaic of ITS sequences in T. parva, with different combinations of similar sequences present in the ITS of both T. p. lawrencei and T. p. parva, indicates that there has been genetic recombination amongst T. parva parasites. A similar mosaic effect was observed in the p67 gene of at least one of the Zimbabwean T. parva isolates, T.p/HUN/96. What appears to be happening is that a relatively homogeneous subpopulation of T. parva parasites is becoming adapted to cattle and is beginning to be distinguishable from the larger and more diverse gene pool in buffalo. However, genetic recombination with the parent gene pool still takes place, so the gene pools of T. parva parasites in cattle and in buffalo are not yet completely separate. Under these circumstances, no DNA probe will distinguish between all T. p. parva and T. p. lawrencei genotypes unless it targets a sequence which is directly associated with the mechanisms responsible for their differential behaviour in cattle.

Such a difference might be found in the gene (or genes) responsible for the increased number of schizonts and piroplasms seen in T. p. parva infections, or in a promoter which increases the expression of such a gene. However, this gene could be involved at any number of stages in the life cycle of the parasite; e.g. in determining the ability of schizonts to replicate, in determining the proportion of macroschizonts which differentiate to produce merozoites, or in the ability of merozoites to penetrate erythrocytes. It will undoubtedly be very difficult to identify such a gene (or genes).
black-backed gull to the east, a complete series of interbreeding intermediates between the two species can be traced in a ring around the world (Ridley, 1993).

To a much lesser extent the situation in *T. parva* may be similar, but with the continuum of types spread across cattle and buffalo hosts rather than geographically. *T. p. parva* appears to be adapting to cattle as host, and in the process it may be diverging from the *T. p. lawrencei* population in buffalo. However, the degree to which *T. p. parva* and *T. p. lawrencei* can diverge may be limited by the overlap between cattle and buffalo and subsequent interbreeding in the tick host which they still share. Nonetheless, there are distinguishing features between most *T. p. parva* and most *T. p. lawrencei* parasites: there are recognizable clinical differences between the diseases they cause, and ITS characterisation probes and p67 gene sequence differences distinguish them to an extent.

Since there are no formal criteria for defining subspecies of protozoa, they are commonly described on the basis of differences in host range, geographical distribution, pathological effects and epidemiological features. For example, *Trypanosoma brucei* is divided into three subspecies, namely *T. b. brucei*, a parasite of nonhuman animals, *T. b. gambiense* which causes trypanosomiasis in humans in West and central Africa and *T. b. rhodesiense* which causes trypanosomiasis in humans in East Africa. Most *Trypanosoma brucei* isolates can be grouped into one of these subspecies using isoenzyme analysis (Godfrey et al., 1990) but a small number cannot be grouped by these means.

Thus, *T. p. lawrencei* and *T. p. parva* may be valid sub-species labels which really do describe the majority of *T. parva* parasites. Although this terminology may result in misclassification of some isolates, it conveys more information about the parasite than the "cattle-derived/buffalo-derived" terminology.
in most T. p. parva stocks and also a number of nucleotide differences which translate to several amino acid substitutions (Nene et al., 1996). Most T. parva isolates can be classified using these p67 data, which correlates in general with the grouping assigned to them using ITS probes. However, we found exceptions which cannot be classified using p67 genes (e.g. T. parva isolates from the Kruger National Park which contain both types of p67 genes).

5.2 Are T. p. parva and T. p. lawrencei valid labels?

The Penguin Dictionary of Biology (Abercrombie et al., 1983) defines a subspecies as follows:

"Subdivision of a species forming a group whose members resemble each other in certain characteristics, and differ from other members of the species, though there may be no sharp dividing line. Polymorphism is excluded. While breeding is possible and in many cases occurs between members of different subspecies of the same species, it does not occur as freely as within the confines of the subspecies. Because reproductive isolation is incomplete, subspecies nearly always grade into each other. The partial reproductive isolation is commonly due to the occupation of different geographical areas. Some subspecies are probably new species in the making."

This definition is directly applicable to the different forms of T. parva.

In addition, there is extensive variation within many species. In the case of "ring species" this variation may result in populations which are classified as separate species (Ridley, 1993). For example, in Europe the lesser black-backed gull and the herring gull are two distinct species. These birds do not interbreed, they choose different nesting sites and there are a number of phenotypic differences between them. If one follows the distribution of the herring gull to the west and the lesser
Bishop and E. Gobright, pers. comm.). Work is currently in progress to determine whether the transcription units are stage-specifically transcribed. Preliminary results suggest that both transcription units are transcribed into rRNA in *T. p. lawrencei* 7344 schizont RNA (R. Bishop, pers. comm.).

Although we could not identify probes in the ITS which give an unequivocal differentiation between the two forms, we can use the patterns of reactivity obtained upon screening with ITS characterisation probes to assign probabilities to the types of *T. parva* parasites present in an isolate. TPP1 detected most *T. p. parva* genotypes, although some *T. parva* stocks isolated from cattle were also detected by TPL1 and TPL3b. Most *T. parva* stocks isolated from buffalo were detected by one or more of the probes TPL1, TPL2, TPL3a and TPL3b, but not by TPP1. Thus, extrapolating from the results we have obtained, we can say that if probe TPP1 hybridises to an isolate, there is a 93.5% probability that it is a *T. p. parva* isolate. Conversely, if an isolate is detected by probes TPL2 and/or TPL3a, we have found an 81.8% probability that it is a *T. p. lawrencei* isolate. *T. parva* stocks isolated from buffalo that are detected by both types of probes could contain a mixture of *T. p. parva* and *T. p. lawrencei* genotypes. Alternatively this pattern of reactivity would result from a single genotype containing segments of sequence found in both *T. p. parva* and *T. p. lawrencei*, but it would not be possible to classify such isolates. It is important to note that there are exceptional isolates which do not react with ITS probes as expected (e.g. *T. parva* Mariakani G29/E3/F8/E7, isolated from cattle, reacts with only TPL3b). Despite these reservations, the panel of *T. parva* ITS characterisation probes developed in this study represents a practical laboratory method for characterisation of *T. parva* parasites and should be useful for epidemiological surveys of the prevalence of the different genotypes.

The p67 gene in most *T. parva* stocks isolated from cattle is conserved, while it is very much more variable in *T. parva* stocks isolated from buffalo. The p67 gene of most *T. p. lawrencei* stocks contains an additional sequence which has no homologue.
CHAPTER 5: GENERAL DISCUSSION AND CONCLUSIONS

5.1 Can sporozoite antigen and ribosomal RNA gene sequences be used to distinguish between *T. p. parva* and *T. p. lawrencei*?

The SSUrRNA and 5.8S genes of the cloned *T. p. lawrencei* parasite, 7344, were identical to those of *T. p. parva* Muguga and there were only minor sequence heterogeneities between their LSUrRNA genes. Although SSUrRNA sequence analysis can be used as a very effective tool to elucidate phylogenetic relationships between organisms, it cannot always be used to distinguish between species (see discussion in section 2.3.2.1). Therefore, although the SSUrRNA genes of *T. p. parva* Muguga and the cloned *T. p. lawrencei* 7344 are identical, this does not necessarily mean that they are the same species, even though there are sufficient sequence differences in the SSUrRNA gene to distinguish between *T. parva* and other *Theileria* species (Allsopp, et al., 1993; Allsopp, 1994; Allsopp, et al., 1994). The sequence similarities between the SSUrRNA and LSUrRNA genes of *T. p. parva* Muguga and *T. p. lawrencei* 7344 do, however, indicate that these two organisms are very closely related.

Two polymorphic ITS sequences were identified in *T. p. lawrencei* 7344. Kibo et al. (1994) have shown that there are two identical rRNA transcription units in *T. parva* Muguga, which are located on different chromosomes and are not tandemly repeated. The sequence differences between the two ITS sequences of *T. p. lawrencei* 7344 have been used to design primers to differentially amplify the ITS and LSUrRNA gene of the two different rRNA transcription units from this parasite. Sequence analysis of the two amplicons revealed that there were differences between the LSUrRNA coding regions of the two *T. p. lawrencei* 7344 transcription units (R.
Purification Resin was added to the cleared cell lysate and the mixture was mixed by inverting the tube. The resin/NA complex was loaded onto a Wizard™ minicolumn using a 5 ml syringe and washed with 2 ml of Column Wash Solution. The minicolumn was supported in a 1.5 ml microcentrifuge tube and centrifuged for two minutes to dry the resin. The minicolumn was transferred to a fresh microcentrifuge tube and 50 μl of TE pH 8.0 was added. After 1 minute the plasmid DNA was eluted by microcentrifugation for 30 seconds at 12000g. Plasmid minipreps were digested with appropriate restriction enzymes to ensure that the correct insert had been cloned.

C.9.2 Large-scale isolation of recombinant plasmid clones

DNA was prepared from recombinant plasmid clones using the QIAGEN Midi protocol, according to the methods stipulated in The QIagenologist: application protocols. QIAGEN columns contain an anion-exchange resin consisting of a silica gel with a hydrophobic coating which selectively binds to nucleic acids. 0.5 ml of the overnight culture of the chosen plasmids was inoculated into 50 ml of LB containing 50 μg/ml ampicillin and grown overnight at 37°C with vigorous shaking. The cells were pelleted by centrifugation at 3000g for 10 minutes at 4°C in a Beckman JA-14 rotor. Alkaline lysis was used to prepare a cell lysate which was loaded onto an equilibrated QIAGEN tip-100. The plasmid DNA was adsorbed to the resin, and residual proteins and RNA were removed by washing with buffer QC. Double stranded DNA was eluted at a higher salt concentration using buffer QF. The DNA was precipitated with 0.5 volumes of propan-2-ol, and collected by centrifugation at 15000g for 30 minutes at 4°C in a Beckman JA-20 rotor. The DNA pellet was resuspended in 0.5 ml of TE pH 8.0.
C.9 Preparation of DNA from recombinant plasmid clones

C.9.1 Small-scale isolation of recombinant plasmid clones

White colonies were picked, inoculated into 3 ml of LB containing the appropriate selection antibiotic at 50 µg/ml, and grown overnight at 37°C with vigorous shaking. The cells were pelleted by centrifugation at 12000g in a microcentrifuge and the supernatant was removed by aspiration. Minipreparations of plasmid DNA were prepared initially by the alkaline lysis method according to Sambrook et al. (1989) and later using Wizard™ Plasmid minipreps.

C.9.1.1 Alkaline lysis

The cell pellet was resuspended in 100 µl of ice-cold Solution I [50 mM glucose, 25 mM Tris-HCl (pH 8.0), 10 mM EDTA (pH 8.0)]. The cells were lysed by addition of 200 µl of Solution II [0.2 M sodium hydroxide, 1% (w/v) SDS] and the solution was neutralised by addition of 150 µl of ice-cold Solution III [Solution III is made by mixing 60 ml 5 M potassium acetate with 11.5 ml of glacial acetic acid and 28.5 ml distilled water; the resulting solution is 3 M with respect to potassium and 5 M with respect to acetate]. After incubation on ice for 5 minutes, cell debris was pelleted by centrifugation at 12000g for 5 minutes at 4°C. The cleared lysate was transferred to a fresh tube, plasmid DNA was extracted with phenol and chloroform/IAA and recovered by ethanol precipitation. Plasmid minipreps were digested with appropriate restriction enzymes to ensure that the correct insert had been cloned.

C.9.1.2 Wizard™ plasmid minipreps

Cells were resuspended, lysed and the solution neutralised using the buffers supplied with the Wizard™ Minipreps DNA Purification system. 1 ml of Wizard™ PCR
15 μg/ml tetracycline and spread with X-Gal and IPTG as described above. Plates were incubated inverted overnight at 37°C.

C.8.4 Plate analysis

Numbers of blue and clear plaques (for M13 cloning) or blue and white colonies (for plasmid cloning) on each plate were counted and recorded.

C.8.4.1 Controls

(i) Transformation control

Numbers of blue plaques or colonies were counted on the transformation control plate and the transformation efficiency per μg of vector DNA was calculated.

(ii) Ligation controls

A very small number of blue plaques or colonies (or none at all) on ligation control plates would indicate that restriction enzyme digestion and dephosphorylation were satisfactory.

a) Cut vector control: a large number of blue plaques or colonies would indicate that the vector had not been efficiently linearised.

b) Cut and religated vector control: the presence of clear plaques or white colonies would indicate exonuclease damage to the cut ends of the vector during restriction enzyme digestion or dephosphorylation.
DNA mixed with 200 μl of competent cells) was included to determine the transformation efficiency of the competent cells.

(i) **M13 clones**

The cells/ligation mix were warmed to 37°C before adding to 3 ml of top H agar at 50°C containing 0.1 mg/ml IPTG and 0.25 mg/ml X-gal. The top agar was poured onto prewarmed H plates and incubated overnight at 37°C. Recombinant plaques were colourless.

(ii) **pUC18 and pGEM-T clones**

1 ml of 2xTY was added, and the cells were grown at 37°C for 1 hour with gentle shaking to allow expression of the antibiotic resistance genes. LB plates containing 50 μg/ml ampicillin were prepared. 100 μl of 100 mM IPTG and 20 μl of 50 mg/ml X-gal were spread over the surface of the plates and allowed to absorb for 30 minutes at 37°C prior to use. 50 μl and 300 μl aliquots were plated out and plates were incubated inverted overnight at 37°C.

### C.8.3.2 For transformation of pMOSBlue-T clones

Cloning of PCR products into pMOSBlue-T was performed as described in the pMOSBlue T-vector kit instruction manual (Amersham). MOSBlue competent cells supplied with the kit were transformed as follows. 1 μl of each ligation mix was added to 20 μl of high efficiency MOSBlue competent cells. After a 30 minute incubation on ice, the cells were heat shocked at 42°C for 40 seconds and placed on ice for 2 minutes. 80 μl of SOC medium at room temperature was added to each tube and transformed cells were grown at 37°C for 1 hour with gentle shaking. 50 μl of transformed cells were plated onto LB plates containing 50 μg/ml ampicillin,
These 3' A overhangs were utilised for direct cloning of PCR products into the thymidine tailed vectors, pGEM-T (Promega) or pMOSBlue-T (Amersham).

C.8.2 Ligation

Ligation reactions were set up with molar ratios of vector to insert of 1:1 and 1:3. Ligations were performed in 10 μl volumes using the following conditions: 30 mM Tris-HCl (pH 7.8); 10 mM magnesium chloride; 10 mM DTT; 0.5 mM ATP and 3 units of T4 DNA ligase (Promega). Ligation reactions were incubated at 4°C overnight.

Two controls were included for M13 and pUC vectors: one with 10 ng of cut vector to indicate restriction enzyme cutting efficiency, and one with 10 ng of cut and religated vector to indicate efficiency of dephosphorylation.

C.8.3 Preparation and Transformation of Competent Cells

C.8.3.1 For transformation of M13, pUC and pGEM-T clones

Competent bacterial cells were prepared according to the method of Chung et al. (1989) which utilizes PEG and divalent cations to induce competence. Escherichia coli T1 cells were inoculated into Luria-Bertani (LB) broth and grown at 37°C with shaking to the early exponential growth phase (A600 = 0.3-0.4). Cells were pelleted by centrifugation at 1000g for 10 minutes and resuspended in 0.1 volumes of ice-cold transformation and storage solution, TSS [LB broth containing 10% (w/v) PEG 8000, 5% (v/v) DMSO, 50 mM magnesium chloride, pH 6.5].

5 μl of each ligation and control reaction was mixed with 200 μl of competent cells, and stored on ice for 30 minutes to allow transformation to take place (no heat-shock stage was necessary). A transformation control (50 pg of the relevant, intact vector
C.8 Cloning of PCR products

C.8.1 Preparation of vectors

C.8.1.1 M13mp18 and mp19

2 μg of RF (replicative form) DNA of each M13 strain mp18 and mp19 (Boehringer Mannheim) was double digested with appropriate enzymes for directional cloning of PCR products. 2 μg of RF DNA of M13mp18 was digested with Sma I for shotgun cloning of sonicated PCR products.

C.8.1.2 pUC18

5 μg of pUC18 DNA (Boehringer Mannheim) was digested with Sma I for cloning of blunt-ended PCR products.

C.8.1.3 Dephosphorylation of vectors

Religation of M13 and pUC18 vectors is greatly reduced by dephosphorylation. 0.5 units of alkaline phosphatase (Boehringer Mannheim) was added to the digestion and the reaction was incubated for a further 30 minutes at 37°C. The cleaved, dephosphorylated vector was purified by phenol and chloroform extraction, and precipitated with sodium acetate and ethanol. The pellet was reconstituted in 20 μl of TE and the concentration determined by spectrophotometry.

C.8.1.4 T-vectors

Tag DNA polymerase preferentially adds a single adenosine nucleotide to the 3' ends of the double stranded DNA molecules that it produces during a PCR (Clark, 1988).
slice was placed in a 0.75 ml microcentrifuge tube prepared as described. The agarose matrix was disrupted by incubation in liquid nitrogen for 5 minutes. The tube was then placed in a 1.5 ml microcentrifuge tube and centrifuged at 12000g for 5 minutes to collect the aqueous solution containing the DNA. 0.1 volumes of 4 M lithium chloride were added and the DNA was extracted with phenol and chloroform/IAA. The purified PCR product was precipitated with three volumes of ethanol at -70°C for 15 minutes and washed with 70% ethanol. The pellet was dried and resuspended in 10-50 μl of T1 pH 8.0, depending on the starting concentration of PCR product.

C.7.2 Wizard™ DNA purification system

The Wizard™ purification system (previously Magic™) utilises silicates, which in the presence of the chaotrope, guanidine hydrochloride, will bind to double stranded DNA. DNA fragments smaller than 175 bp do not bind efficiently to the resin; this method can therefore be used to remove primers and unincorporated nucleotides directly. RNA and proteins are also washed from the resin/DNA complex.

300 μl (the products of three PCRs) were added to 100 μl of Direct Purification Buffer and vortexed briefly to mix. 1 ml of Wizard™ PCR Purification Resin was added and the mixture was vortexed for 30 seconds to mix. The resin/DNA complex was loaded onto a Wizard™ minicolumn using a 5 ml syringe and washed with 2 ml of 80% propan-2-ol. The minicolumn was supported in a 1.5 ml microcentrifuge tube and centrifuged for 30 seconds to dry the resin. The minicolumn was transferred to a fresh microcentrifuge tube and 50 μl of T1 pH 8.0 was added. After 1 minute the DNA was eluted by microcentrifugation for 30 seconds at 12000g.
C.6 Filling in recessed 3' or 5' ends and phosphorylation

Both reactions were carried out in polynucleotide kinase (PNK) buffer [10x stock: 0.5 M Tris, 0.1 M MgCl₂, 50 mM DTT, 1 mM spermidine, 1 mM EDTA (Sambrook et al., 1989)]. The end-filling reaction was performed with both Klenow and T4 DNA polymerase. 1 µg of DNA was mixed with 2 µL of 10x PNK buffer. 2.5 mM dNTPs were added to a final concentration of 0.1 mM. 10 units of both Klenow (Promega) and T4 DNA polymerase (Promega) were added and the volume made up to 20 µL with distilled water. The reaction was incubated at 11°C for 2 hours and then heat inactivated at 75°C for 10 minutes (Ausubel et al., 1996).

5' phosphate groups were then attached to the repaired termini using T4 polynucleotide kinase (Promega). 100 mM ATP was added to a final concentration of 1.0 mM, 1 µL of 10x PNK buffer and 10 units of enzyme were added and the volume was made up to 30 µL with distilled water. The reaction was incubated at 37°C for one hour and stopped by heat inactivation at 75°C for 10 minutes.

C.7 Purification of PCR products

Modified PCR products were separated from residual enzymes, oligonucleotide primers, and unincorporated nucleotides by gel-purification using a modification of the "freeze-squeeze" technique using liquid nitrogen (Koenen, 1989), or, later, by using the Magic or Wizard DNA purification system (Promega).

C.7.1 "Freeze-squeeze": recovery of DNA from agarose gels using liquid nitrogen

A hole made in the base of a 0.75 ml microcentrifuge tube with a hot needle, was plugged with plastic wool. PCR products were fractionated on a 1% agarose/TBE gel, the fragment of interest was excised in as small a volume as possible and the gel
All gels were photographed using a Polaroid camera or a video documentation system (OmnisScience).

C.4 Dissociation temperature of oligonucleotide primers and probes

The dissociation temperature (Tₘ) of an oligonucleotide is the temperature at which 50% of the oligonucleotide/target DNA duplexes are dissociated. The Tₘ of oligonucleotide primers for PCR and *T. parva* ITS characterisation probes were calculated using the following formula (Davis *et al.*, 1986):

\[
Tₘ = 16.6 \log_{10} [\text{Na}^{+}] + 0.41 Pₚ + 81.5 - 675/L
\]

*Pₚ* is the percentage of G and C bases in the oligonucleotide

*L* is the probe length in base pairs

C.5 Restriction digests

Enzymes used in this study were *Bam HI*, *Eco RI*, *Hind III*, *Pst I*, *Sal I* and *Sma I* (Promega or Boehringer Mannheim). Reactions were performed in 20 µl volumes: 2 µl of the appropriate 10x restriction buffer was mixed with the DNA to be digested and 8-12 units (usually 1-1 µl) of enzyme was added. The digests were incubated for at least 1 hour at 37°C, except for *Sma I* which was incubated at 30°C. Reactions were usually heat inactivated at 75°C for 10 minutes. 2 µl of loading buffer was added to each digest and the reactions were analysed by agarose gel electrophoresis.

Double digestions with *Eco RI* and *Bam HI* were performed in Boehringer Mannheim buffer B using the conditions described above.
C.3 Polymerase chain reaction

C.3.1 The reaction (Saiki et al., 1988)

The magnesium chloride concentration which yielded the best product for each set of primers was determined by performing the reactions with magnesium chloride concentrations increasing in 0.5 mM steps from 1.5 mM to 3.0 mM. PCR's were performed in 100 μl volumes using the following conditions: 50 mM potassium chloride; 10 mM Tris-HCl (pH 9); 1.5-3.0 mM magnesium chloride; 0.1% Triton X-100; 200 μM of each dNTP; 1 μM of each oligonucleotide primer and 50-200 ng of purified DNA. The reactions were overlaid with 200 μl of mineral oil and denatured at 95°C for 5 minutes on a Hybaid HB-TR1 programmable heating block.

All primer pairs were designed to have balanced dissociation temperatures (Tm) (see section C.4 for calculation of Tm value). Annealing was performed at (Tm-10) to (Tm-5)°C. The heating block was programmed to hold at the annealing temperature (AT) during the initial annealing stage while 2.5 units of Promega Taq polymerase were added to each reaction. Cycling conditions were: 72°C for 1.5 minutes, 95°C for 1 minute and AT for 1 minute, for 30 cycles, unless otherwise stipulated. Positive and negative controls were included with every PCR performed.

C.3.2 Checking for PCR products

PCR products were removed from beneath the mineral oil layer and transferred to a fresh microcentrifuge tube. A 10 μl aliquot was mixed with 2 μl of 6x gel-loading buffer [0.25% bromophenol blue, 0.25% xylene cyanol, 15% Ficoll (Sambrook et al., 1989)] and loaded on a 1% agarose gel containing 0.5 μg/ml of ethidium bromide (Boehringer Mannheim). 0.2 μg of λHind III or φX174Hae III size markers were also loaded as appropriate to determine the size of the products. Electrophoresis was performed in 1xTBE buffer and DNA was visualised using a UV transilluminator.
C.2 Quantitation of DNA

C.2.1 Double-stranded DNA

The absorbance readings at 260 nm ($A_{260}$) and 280 nm ($A_{280}$) of an appropriate dilution of each DNA sample were measured using a Beckman DU®640 spectrophotometer, which determines the concentrations of nucleic acids and proteins automatically using coefficients determined by Warburg and Christian (cited in Dawson et al., 1969). The Warburg-Christian assay calculates protein and nucleic acid concentrations using the following equations, which are based upon the absorptivities of yeast enolase and RNA:

$$[\text{Protein}] = (1552 \times A_{260}) - (757.3 \times A_{280}) \text{ \( \mu g/ml \)}$$

$$[\text{Nucleic acid}] = (-36.0 \times A_{260}) + (62.9 \times A_{280}) \text{ \( \mu g/ml \)}$$

C.2.2 Single-stranded oligonucleotides

The $A_{260}$ of a 1M solution using a 1 cm light path is equal to its molar extinction coefficient (molar e) (Dawson et al., 1969). Molar e for an oligonucleotide is the sum of the molar extinction coefficients of the individual nucleotides which make up the oligonucleotide, and was calculated for each oligonucleotide from the following nucleotide contributions: T = 8400, C = 7050, G = 12010, A = 15200 (Sober, 1970).

The absorbance reading at 260 nm ($A_{260}$) of an appropriate dilution (usually 1:100) of each oligonucleotide was measured using a Beckman DU®640 spectrophotometer. The concentration of each oligonucleotide was inferred from its $A_{260}$ measurement multiplied by its dilution factor.
removed and transferred to a microcentrifuge tube by gentle suction with a Gilson blue tip. The DNA was washed twice with 80% ethanol, excess ethanol was poured off gently and the microcentrifuge tube was incubated open in a 37°C oven for 10-30 minutes to evaporate remaining ethanol. The purified nucleic acid was dissolved in an appropriate volume of TE.

C.1.2.2 Of smaller DNA molecules (≤15 kbp)

DNA was usually precipitated by adding 0.1 volumes of 3 M sodium acetate (pH 5.2) and 2.5 volumes of absolute ethanol and incubated at -70°C for at least 30 minutes. DNA was pelleted by centrifugation at 12000g for 20 minutes at 4°C in a microcentrifuge. A final wash was performed in 0.5 ml of 70% ethanol, and the DNA recovered by centrifugation as above for 10 minutes. The microcentrifuge tube was incubated open in a 37°C oven for 10 minutes to evaporate remaining ethanol. The purified nucleic acid was dissolved in an appropriate volume of TE.

Ammonium acetate reduces the co-precipitation of dNTPs, but cannot be used to precipitate DNA that is to be phosphorylated, as ammonium ions inhibit the activity of T4 polynucleotide kinase (Sambrook et al., 1989). To minimise the total volume, PCR products that were not going to be phosphorylated were precipitated by adding 0.4 volumes of 10 M ammonium acetate and an equal volume of propan-2-ol (Brown, 1990). After a 10 minute incubation at room temperature, DNA was recovered by centrifugation at 12000g for 10 minutes in a microcentrifuge. The DNA pellet was washed, dried and resuspended as described above.
REFERENCES


ALLSOPP, B. A., CARRINGTON, M., BAYLIS, H., SOHAL, S., DOLAN, T. AND JAMS, K., 1989. Improved characterization of Theileria parva isolates using the
To whom it may concern

ANIMAL ETHICS CLEARANCE

Project OV9/3: "The relationship between Theileria parva parva and T. p. lawrencei as shown by sporozoite antigen and ribosomal RNA gene sequences"

It is hereby confirmed that Ms Nicola Collins did not work directly with animals at any stage during the course of the above project and therefore did not require animal ethics clearance. DNA was extracted from blood samples collected by Dr Hein Stoltz for the purpose of establishing Theileria parasites in cell culture under the auspices of project OV8/7. All experimental animal work under this project was authorised by the Onderstepoort Animal Ethics Committee.

Yours sincerely

[Signature]

DR J.D. BEZUIDENHOUT
CHAIRMAN: ANIMAL ETHICS COMMITTEE
C.13 Preparation of slot-blot

To 10 μl aliquots of ITS PCR products, 90 μl of TE pH8 and 6 μl of 0.5 M EDTA were added. DNA was denatured by addition of 8 μl of 6 M NaOH. 110 μl of 2 M ammonium acetate was added and the sample was blotted onto Hybond N+ (Amersham) filters using a Schleicher and Schuell Minifold II slot-blotting apparatus. After a brief wash in 6xSSC for 15 seconds, DNA was fixed to the filters by transferring them to Whatman 3MM paper soaked with 0.4 M NaOH. Filters were washed briefly in 6xSSC as before, prehybridised and probed with 3' labelled oligonucleotides as described in the text (Chapter 3).
C.12.3.2 Loading and running of gels

Sequencing reactions were incubated at 80°C for 5 minutes to denature the DNA strands immediately prior to loading onto the gel. The sequencing reactions were held on ice during loading; 1 μl of each sequencing reaction was loaded.

Gradient gels were run at 33 W in 0.5x TBE until the bromophenol blue just reached the bottom of the gel. Linear gels were electrophoresed at 38 W in 0.5x TBE for five and a half hours; the upper reservoir buffer was changed after two and a quarter hours to replenish depleted ions.

Gels were fixed in 10% (v/v) acetic acid, 10% (v/v) propan-2-ol, transferred to Whatmann 3MM paper and dried under vacuum at 80°C on a DrygelSrb. SE1160 slab gel drier. Dried gels were placed in direct contact with X-ray film and autoradiographed for 24-48 hours at room temperature.

C.12.4 Compilation and manipulation of sequences

Sequences were read manually using a Graf/Bar MarkII sonic digitiser and the READGEL program (D. Judge, Department of Biochemistry, University of Cambridge, pers. comm.), or autoradiographs of sequencing gels were scanned and sequences determined using an automatic base scanner and semi-automatic base-calling software (Amersham). Sequences were assembled and edited using the Staden-Plus Program Package (Amersham) or, later XBAP (Staden, 1994), CLUSTALV (Higgins et al., 1992), CLUSTALW (Thompson et al., 1994) and GDE (Smith, 1992) were used to align sequences.
Where it was necessary to extend the readings, the reactions were run on linear gels (6% acrylamide, 7 M urea, containing 0.5x TBE) which usually allowed the reading of a further 100-200 bp. The dimensions of the gels were: 480 x 200 x 0.4 mm.

C.12.3.1 Pouring of gels

Plates were cleaned with "Handy Andy", washed with distilled water and finally cleaned with ethanol. The cured plate was siliconized with dimethyldichlorosilane in a fume cupboard, excess dimethyldichlorosilane was washed off with ethanol. Plates were taped together with autoclave tape using 0.4 mm-thick plastic strips as spacers.

Gradient gels were poured as follows. In one beaker, 8 ml of lower gel mix (6% acrylamide, 7 M urea, 5x TBE, 0.05% (w/v) bromophenol blue) was combined with 250 μl of 1.6% (w/v) ammonium persulfate (APS) and 6 μl of TEMED. In a second beaker, 50 ml of upper gel mix (6% acrylamide, 7 M urea, 0.5x TBE) was combined with 1.56 ml of 1.6% APS and 37.5 μl of TEMED. 42 ml of the upper gel mix was taken up in a 50 ml glass syringe; the lower gel mix was then carefully drawn into the same syringe so that it formed a layer under the first gel mix. The plates were held at 45°C and the gel solution was introduced. The plates were squeezed and tilted from side to side to even out the gradient. The gel was topped up with upper gel mix if necessary. A 48 well shark's tooth comb (Bio-Rad) was reversed, inserted into the top of the gel and clamped in position to form the well.

Linear gels were poured by polymerizing 50 ml of upper gel mix alone.
C.12.2.1 End-Labelling of the sequencing primer

To prepare enough primer for sequencing six templates, primer labelling reactions were performed in 10 µl volumes using 10 pmol of primer, 10 pmol [γ-32P]dATP (10 mCi/µl, 5000 Ci/mmol), 1 µl T4 polynucleotide kinase 10x buffer and 5 units of T4 polynucleotide kinase. The reaction was incubated at 37°C for 30 minutes, and the kinase was inactivated at 90°C for 2 minutes. This reaction was scaled up or down according to the number of sequencing reactions to be performed; sufficient primer was labelled so that there would be 1.5 pmol of labelled primer per template to be sequenced.

C.12.2.2 Extension/termination

For each template to be sequenced, four columns of a 96-well plate were labelled A, C, G and T. 2.0 µl of the appropriate d/ddNTP Mix were added to each well.

To 4-40 fmol (2-4 µl of a PCR product purified using the Wizard™ DNA Purification System) of each template, were added: 5 µl fmol™ sequencing 5x buffer, 1.5 µl labelled primer and sterile water to 16 µl. 5 units of sequencing grade Taq DNA polymerase were added, and 4 µl of this enzyme/primer/template mix were added to each of the four wells of the microtitre plate containing d/ddNTP mixes. A drop of mineral oil was added to each well and the microtitre plate was placed in a thermal cycler preheated to 95°C for two minutes. Cycling conditions were: 95°C for 30 seconds, 42°C for 30 seconds, 70°C for 1 minute for 30 cycles. Reactions were stopped by addition of 3 µl of fmol™ sequencing stop solution.

C.12.3 Sequencing gels

Sequencing reactions were run on gradient gels (6% acrylamide, 7 M urea with a gradient of 0.5-5x TBE) which allowed 300-350 bp from the primer to be read.
water to 5 µl. 5 µl of this master mix was added to each annealed primer/template and these labelling reactions were incubated at room temperature for 5 minutes.

(iii) Termination reaction

The labelling reaction is terminated by the addition of one of the four 2',3'-dideoxyribonucleotide triphosphates (ddNTPs), which lack the 3'-OH group needed for DNA synthesis.

Columns of a disposable 96-well MicroTest III flexible assay plate (Falcon) were labelled A, C, G and T. 2.5 µl of the appropriate termination mixes were added to the side of the wells in each column. The four termination mixes contained an excess of each dNTP (80 µM) and the appropriate ddNTP at a concentration of 8 µM. 3.5 µl of each labelling reaction was added to each well along a row of the plate. The reactions were mixed by centrifuging the plate for 10 seconds at 1000 rpm in a Hermle Z220 benchtop centrifuge with a plate-spinning attachment, then covered with cling-film and incubated at 37°C for 10 minutes. The reactions were stopped by addition of 4 µl of stop solution [95% formamide; 20 mM EDTA; 0.05% bromophenol blue; 0.05% xylene cyanol]. Plates were stored at -20°C.

C.12.2 Direct thermocycle sequencing of PCR products

Direct thermocycle sequencing of PCR products using end-labelled primers was performed using the *final* DNA Sequencing System (Promega). Sequencing primers are end-labelled and the DNA template and labelled primer are repeatedly annealed and enzymatically extended/terminated.
C.12.1.2 Sequencing reactions

The dideoxy chain termination method (Sanger et al., 1977) was used to sequence single-stranded and double-stranded templates. Buffers and reactions were as described in the *Sequenase: Step-by-step Protocols for DNA sequencing with Sequenase® Version 2 T7 DNA Polymerase* manual (United States Biochemical).

(i) Annealing reaction

During the annealing reaction, the oligonucleotide primer anneals to a specific site, which may be a conserved region in the vector or a conserved region in the sequence of the insert itself. Synthesis of the complementary strand is then initiated at this site.

2 µl of 5x annealing buffer and 1 µl of the appropriate primer (0.5 pmol/µl) was added to 5 µl of each single-stranded M13 preparation (approximately 1 µg) or to each denatured plasmid template. Sterile distilled water was added to bring the total volume to 10 µl. Primer was annealed to template at 65°C for 10 minutes, then allowed to cool to room temperature slowly over 20 minutes.

(ii) Labelling reaction

During the labelling reaction, complementary DNA strands are synthesized by Sequenase and [α-35S]dATP is incorporated, which allows for visualisation of the sequence by autoradiography.

A master mix, sufficient for 5 µl for each annealing reaction, was prepared with the following components in each 5 µl aliquot: 1 µl of 0.1 M DTT, 0.4 µl of 5x labelling mix [7.5 µM dCTP; 7.5 µM dTTP; 7.5 µM 7-deaza-2'-dGTP], 0.5 µl of [α-35S]dATP (10 mCi/ml, 1000 Ci/µmol), 2 units of Sequenase Version 2, and sterile distilled
The presence of insert was confirmed for each phage before continuing with single-stranded DNA preparation. 15 μl of the phage supernatant was mixed with 5 μl of loading buffer containing 5% SDS and run on a 1% agarose/TBE gel, using the supernatant from a blue non-recombinant plaque as a control. The SDS in the loading buffer lyses the coat proteins, allowing the DNA to be visualized; DNA from recombinant phages has a reduced mobility compared to the non-recombinant control.

Phages were precipitated by addition of 200 μl of PEG/NaCl [20% (w/v) PEG 8000; 2.5 M NaCl] and pelleted by centrifugation. The supernatant was removed by aspiration and the phage pellet was resuspended in 100 μl of TE pH 8.0. Phenol and chloroform extractions were performed as described previously (Appendix C.1.1) to remove coat proteins. Single-stranded DNA was precipitated by addition of sodium acetate and ethanol (Appendix C.1.2.2), and resuspended in 30 μl of TE pH 8.0.

(ii) Denaturing double-stranded templates

Alkaline-denaturation was used to separate the strands of double-stranded plasmid templates prior to sequencing. 3-5 μg of plasmid DNA (usually 12-16 μl of a plasmid miniprep.) was denatured by the addition of 0.1 volumes of 2 M sodium hydroxide and 2 mM EDTA, the mixture was incubated at 65°C for two minutes, neutralised by the addition of 0.1 volumes of 5 M sodium acetate (pH 5.0) and the DNA precipitated by the addition of 2.5 volumes of absolute ethanol. The denatured DNA strands were recovered by centrifugation at 12000g for 20 minutes; the pelleted DNA was washed once with 70% ethanol and redissolved in 7 μl sterile distilled water.
The reaction was terminated by heating to 70°C for 10 minutes. The probe reaction was diluted with 80 μl TE and 1 μl was removed for Cerenkov counting (Appendix C.10.3) before adding the probe directly to the hybridisation buffer.

C.11.3 Assay of [α-32P] incorporation of radiolabelled probes

Probe activity was determined by Cerenkov counting as follows: 1 μl of the diluted probe reaction was mixed with 0.5 ml carrier DNA [200 μg/ml sheared herring sperm DNA in 0.1% aqueous SDS]. 0.5 ml of 20% trichloroacetic acid was added and the mixture was allowed to stand at room temperature for 10 minutes to precipitate the DNA. The mixture was filtered through a glass fibre filter (Whatmann GF/C) to collect precipitated DNA. The filter was washed three times with 10% trichloroacetic acid to remove unincorporated label; it was air-dried and the activity measured with a Packard Tri-Carb™ 1600CA liquid scintillation analyser.

C.12 Sequencing

C.12.1 Dideoxy chain termination sequencing

C.12.1.1 Preparation of templates for sequencing

(i) Preparation of single-stranded template

Single-stranded template for the sequencing reaction was prepared as described in the M13 Cloning and Sequencing Handbook (Amersham). Recombinant plaques were picked with a sterile wooden toothpick and inoculated into 2 ml of 2xTY containing a 1:100 dilution of an overnight E. coli TG1 culture. The cultures were grown at 37°C for 5 hours with vigorous shaking (300 rpm). Bacterial cells were pelleted by centrifugation at 12000g in a microcentrifuge and the supernatant containing single-stranded phage was poured into a fresh microcentrifuge tube.
C.11 Preparation of radiolabelled probes and assay of \([\alpha-^{32}P]\) incorporation

C.11.1 Random primed labelling of double stranded DNA

The Multiprime DNA labelling kit (Amersham) was used to label PCR products. Random sequence hexanucleotide primers are used to prime DNA synthesis at random points along the length of the denatured PCR product. Labelled nucleotides are incorporated into the DNA that is subsequently synthesised by Klenow.

25 ng of each PCR product was denatured in a heating block at 100°C for 2 minutes and chilled for 2 minutes on ice. The labelling reaction was performed in a 50 μl volume. 10 μl of Solution 1 (a buffer containing unlabelled dNTPs) and 5 μl of Solution 2 (containing hexanucleotide primers) were added. 5 μl of \([\alpha-^{32}P]\) dCTP (10 mCi/ml, 3000 Ci/mmole) was added followed by 2 units of Klenow. The reaction was incubated at 37°C for 1 hour. The probe reaction was stopped by addition of 5 μl of 0.2 M EDTA, and 1 μl was removed for Cerenkov counting (Appendix C.11.3). Labelled probes were denatured by heating to 100°C in a heating block for 2 minutes, before adding directly to the hybridisation buffer.

C.11.2 3' Terminal labelling of oligonucleotides

Terminal deoxynucleotidyl transferase catalyses the addition of mononucleotides from a dNTP to the 3'-hydroxyl termini of a DNA molecule, accompanied by the release of inorganic phosphate (Ausubel et al., 1994). The enzyme can therefore be used to label the 3' termini of oligonucleotides with \([\alpha-^{32}P]\) "tail". Labelling reactions were performed in 20 μl volumes as follows: 4 μl of 5x terminal transferase buffer (Promega); 100 μg/ml BSA; 2 pmoles primer and 1.6 μl \([\alpha-^{32}P]\) dCTP (10 mCi/ml, 3000 Ci/mmole). 1 μl of terminal deoxynucleotidyl transferase (10-20 units/μl) (Promega) was added and the reaction was incubated at 37°C for 1 hour.
C.10  Screening of recombinant plaques

C.10.1  Plaque lifts

Plaques were lifted onto Hybond N+ (Amersham) filters using the method described by Sambrook et al. (1989). Duplicate Hybond N+ filters, marked asymmetrically, were placed onto the surface of the agar plates for 1 minute to allow transfer of DNA. The asymmetric markings were duplicated on the base of the petri dish. The filters were transferred to 3MM paper (Whatmann) soaked in denaturing solution [0.5 M sodium hydroxide; 1.5 M sodium chloride] for 5 minutes. Filters were then transferred to 3MM paper soaked in neutralizing solution [1.5 M sodium chloride; 0.5 M Tris-HCl (pH7.4)] for 5 minutes. Filters were transferred to 3MM paper soaked in 2x SSC before fixing the DNA to the membrane by transferring to 3MM paper soaked in 0.4 M sodium hydroxide for 10 minutes.

C.10.2  Hybridisation with PCR product as probe

Filters were prehybridised in 0.5 M sodium phosphate buffer (pH 7.5) and 7% SDS (modified from Church and Gilbert, 1984) in bottles in a Hybaid hybridisation oven at 65°C with rotation for at least 1 hour. Positive plaques were selected by hybridising the DNA immobilized on the filters with randomly labelled PCR product (Appendix C.11.1) at 65°C overnight. Filters were washed twice in 10 ml 2xSSC at 65°C for 10 minutes followed by one wash in 10 ml 2xSSC containing 0.1% SDS at 65°C for 10 minutes. Filters were exposed to Cronex 4 X-ray film (Du Pont) overnight at -70°C with an intensifying screen. Positive plaques were identified by aligning the autoradiographs with the markings on the petri dishes. Only those plaques which had hybridised to both of the duplicate filters were picked.


two single copy units in *Theileria parva* ribosomal RNA genes. *Molecular and Biochemical Parasitology*, 66: 249-259.


CAL, J., COLLINS, M. D., MCDONALD, V. AND THOMPSON, D. E., 1992. PCR cloning and nucleotide sequence determination of the 18S rRNA genes and internal transcribed spacer 1 of the protozoan parasites *Cryptosporidium parvum* and *Cryptosporidium muris.* *Biochimica et Biophysica Acta, 1131: 317-320.*


polymerase chain reaction and oligonucleotide probes. Molecular and Biochemical Parasitology, 35: 137-148.


Para tology, 102: 347-356.


ROGALL, T., WOLTERS, J., FLOHR, T. AND BOTTGER, E. C., 1990. Towards a phylogeny and definition of species at the molecular level within the genus


