SEQUENCE ANALYSIS OF A *COWDRIA RUMINANTIUM* LAMDBA GEM-11 CLONE

By

Trudy Patience

Thesis presented in partial fulfilment of the requirements for the degree of Master of Science in Biochemistry at the University of Stellenbosch

Supervisor: Prof Dirk Uwe Bellstedt

December 2002

DECLARATION

I, the undersigned, hereby declare that the work contained in this thesis is my own original work and that I have not previously in its entirety or in part submitted it at any university for a degree.

OPSOMMING

Hartwater is 'n bedreiging vir vee in Afrika weens die hoë mortaliteitssyfer verbonde aan die siekte. Die intrasellulêre aard van die organisme wat hartwater veroorsaak, *Cowdria ruminantium*, bemoeilik navorsing aangaande die organisme. Dit het tot gevolg dat 'n effektiewe en gebruikersvriendelike entstof moeilik bekombaar is. Daar is onlangs sukses behaal met die konstruksie van twee *C. ruminantium* DNA genoteke, die lambda GEM11 bakteriofaag genoteek en die lambda ZAPII bakteriofaag genoteek. Dit het gelei tot 'n herlewing in die soektog na beskermende gene, wat in 'n entstof teen hartwater gebruik kan word.

In hierdie studie is verskeie molekulêre tegnieke insluitende PKR, klonering en geenopeenvolging bepaling, gebruik om gene te identifiseer in die lambda GEM11 bakteriofaag genoteek wat kodeer vir proteïene wat in entstowwe gebruik kan word as beskerming teen hartwater.

Die *secD* geen is gebruik om die lambda GEM11 bakteriofaag genoteek te sif. Een positiewe plaak is gevind waarna die DNA uit die bakteriofaag plaak geïsoleer en die *C. ruminantium* DNA vanuit die bakteriofaag plaak geamplifiseer is deur gebruik te maak van PKR en spesifieke *C. ruminantium* inleiers. Die *C. ruminantium* DNA is gesif met *Mycoplasma,* bees en *Cowdria* radioaktief gemerkte DNA peilers. Die *C. ruminantium* DNA is vervolgens in twee vektore gekloneer. Die klone is gesif deur middel van restriksie analise. Die DNA volgorde van die klone is bepaal en twee ononderbroke sekwense is geïdentifiseer met 'n gaping in die middel tussen die twee sekwense. Oligonukleotied inleiers is daarna ontwerp om die geenopeenvolging van die gaping tussen die twee sekwense te vul. Hierdeur kon die volledige geenopeenvolging van die genoom van *C. ruminantium* wat in die lambda GEM 11 bakteriofaag plaak voorkom, bepaal word. Hierdie volledige geenopeenvolging is vervolgens geanaliseer en die oop leesrame wat daarin voorkom geïdentifiseer. Vyf leesrame is gevind om homologie met gene wat kodeer vir proteïene wat in bakterieë voorkom, te toon. Twee leesrame het homologie met die gene wat kodeer vir transport proteïene, FtsY en die ABC transporter getoon, en drie leesrame het homologie met gene wat kodeer vir die essensiële ensieme detiobiotin sintetase, prolipoproteïen diasielgliserol transferase en die NADH-ubikinoon oksidoreduktase subeenheid getoon. Dié vyf leesrame het die potensiaal om as entstowwe gebruik te word aangesien al vyf leesrame kodeer vir gene wat 'n belangrike rol speel in die oorlewing van die *C. ruminantium* organisme. Alhoewel die leesrame moontlik

nie so effektief sal wees in 'n DNA entstof nie, toon dit potensiaal om in mutasie-eksperimente gebruik te word. Organismes wat die gemuteerde weergawe van die geen besit sal nie-funksionele proteïene produseer, wat 'n invloed kan hê op die normale fisiologiese funksies van die organisme en dus sal lei tot 'n minder virulente organisme. Die geattenueerde organisme kan moontlik gebruik word om diere te immuniseer en daardeur immuniteit aan diere lewer wat beskerming sal bied teen patogeniese *C. ruminantium* isolate.

ABSTRACT

Heartwater is a major threat to livestock in Africa due to its high mortality rate. The intracellular nature of the causative organism, *Cowdria ruminantium*, makes it difficult to study, hence an effective and user-friendly vaccine has been extremely difficult to obtain. Two *C. ruminantium* DNA libraries have recently been constructed, the lambda GEM11 bacteriophage DNA library and the lambda ZAPII bacteriophage DNA library, and this has lead to a renewed search for protective genes that could be used as a vaccine against heartwater.

In this study, several molecular techniques including PCR, cloning and sequencing were used to identify genes in the lambda GEM11 bacteriophage DNA library that code for proteins, which could be used as vaccines to protect susceptible animals against heartwater.

The lambda GEM11 library was screened with a rickettsial secretory protein gene sequence, known as *secD*. One positive colony was selected from which the bacteriophage DNA was isolated. The *C. ruminantium* DNA was amplified from the bacteriophage DNA by using PCR and *C. ruminantium*-specific primers. The *C. ruminantium* DNA was screened with *Mycoplasma*, bovine and *Cowdria* DNA probes. The amplified DNA was subcloned into two vectors and the clones were screened by restriction analysis to identify clones containing inserts. The appropriate clones were sequenced and overlapping sequences matched, ordered and aligned. Two sequences were continuous with a short sequence of unidentified bases in between. Oligonucleotide primers were designed to amplify the DNA sequence between the two contiguous sequences. This led to the identification of the entire sequence of the *C. ruminantium* genome contained within the bacteriophage plaque. The single contiguous sequence was analysed and the putative protein-coding sequences were obtained and compared to DNA sequences of known organisms using the BLAST program. Five open reading frames were identified with homology to genes encoding specific proteins in bacteria. Two open reading frames showed homology to the genes encoding the transporter proteins, FtsY and the ABC transporter, and three open reading frames were found to be homologous to genes encoding the essential enzymes dethiobiotin synthetase, prolipoprotein diacylglycerol transferase and the putative NADH-ubiquinone oxidoreductase subunit. The five open reading frames encode for genes, which are essential for the normal functioning of the *C. ruminantium* organism. However, these open reading frames might not be effective for use in a DNA vaccine since none of the open reading frames showed homology to obvious

genes that could play a role in immunity and therefore confer protection. The open reading frames can be used in mutagenesis studies to produce attenuated strains of the organism that possess mutated versions of these proteins. These attentuated strains could be used for the vaccination of cattle, and thereby confer protection against viable pathogenic *C. ruminantium* isolates.

Vir Nickey

ACKNOWLEDGEMENTS

I wish to express my sincere appreciation to the following people:

Dr Nicola Collins, my supervisor, for her encouragement, support and guidance.

Prof Dirk Bellstedt, my promoter, for his confidence in me and his support and guidance.

My collegues at the OVI especially Juanita Liebenberg and Erika Faber.

My new collegues at the NIV especially Maria Papathanosopoulos.

My family especially my mom.

My sisters; Merle, Melony, Prescilla, Anita as well as their families.

God, because without faith all would have been impossible.

LIST OF ABBREVIATIONS

AT	adenosine and thymine
BLAST	Basic Local Alignment Search Tool
bp	basepair
cELISA	competitive enzyme-linked immunosorbent assay
CoA	acetyl-coenzyme A
contig	contiguous sequence
cpm	counts per minute
CTL	cytotoxic T lymphocyte
DNA	deoxyribonucleic acid
EC	endothelial cell
ECF	East Coast Fever
ELISA	enzyme-linked immunosorbent assay
E-value	Expect value
gap4	genome assembly program
IFAT	indirect fluorescent-antibody test
IL	interleukin
INF	interferon
IPTG	isopropyl β-D- thiogalactopyranoside
kb	kilobases
kDa	kilodalton
λ	lambda
λ phage	bacteriophage lambda
MAP	major antigenic protein, the protein
map	major antigenic protein, the gene
MCS	multiple cloning site
MHC	major histocompatibility complex
Mϕ	macrophage
MPO	myeloperoxidase
MSP	membrane spanning protein
nip4	nucleotide interpretation program
NK	natural killer

NO	nitric oxide
OD	optical density
OMP	outer membrane protein
ORF	open reading frame
OVI	Onderstepoort Veterinary Institute
PBMC	peripheral blood mononuclear cell
PCR	polymerase chain reaction
pfu/ml	plaque forming units per millilitre
RCF	Relative centrifugal force
RNA	ribonucleic acid
rpm	resolutions per minute
SecD	secretory protein D, the protein
secD	Secretory protein D, the gene
SI	stimulation index
SRP	signal recognition particle-like
T cell	thymus-derived lymphocyte
T_H	thymus-derived lymphocyte helper cell
T_{H1}	thymus-derived lymphocyte helper cell type 1
TCGF	T cell growth factor
TNFα	tumor necrosis factor alpha
X-gal	5-Bromo-4-Chloro-3-Indolyl-β-D-galactoside

LIST OF TABLES

LIST OF FIGURES

TABLE OF CONTENTS

Chapter Three

Heartwater – The Vector

Chapter Four

Heartwater – The Causative Organism

Chapter Five

Heartwater - The Study

Chapter Six

Methodology

Chapter Seven

Results and Discussion

Chapter Eight

Conclusion

Chapter Nine

References

Chapter Ten

Addenda

CHAPTER ONE
INTRODUCTION

Livestock is a major source of nutritional protein in many Third World countries. Since the population of sub-Saharan Africa is increasing at an alarming rate, the quantity of food for human consumption, including meat and milk production, will have to be significantly increased. There are several constraints, which hamper animal production: climate, nutrition, management, genetic productivity limits of local breeds of livestock and disease (Uilenberg *et al.*, 1993).

A considerable increase in animal production is not possible in the arid zones of Africa. It is expected that most of the growth will have to occur in the subhumid areas, where the association of crop and animal production could lead to more intensive and sustainable mixed crop-livestock production systems. The shift from the arid to the subhumid zones will mean that a number of livestock diseases will assume more importance. The tick-borne diseases of cattle and small ruminants, which include theileriosis, cowdriosis, anaplasmosis and babesiosis, are probably the greatest disease constraints to the improvement of livestock productivity in these countries (Uilenberg *et al.*, 1993).

The most significant among these diseases on the African continent, in terms of economic losses and restriction of livestock development are East Coast Fever (ECF), tsetse-transmitted trypanosomiasis and heartwater. These conditions are most severe in animals introduced to the endemic locations from areas that are free of disease. In South Africa, from which ECF has been eradicated and where there is only a limited distribution of tsetse, heartwater is almost certainly the most important livestock disease (Uilenberg *et al.*, 1993; McKeever, 1993; Allsopp, 2000).

Heartwater was first identified in sheep in South Africa in the 1830's. A livestock census performed by the directorate of Animal Health in South Africa indicated that up to 17.5 million head of livestock are at risk within the heartwater endemic area. Approximately half of these animals are in the small-scale farming sector, while the other half are in the commercial sector. Heartwater therefore constitutes a serious threat to food security in rural areas, an economic burden on the production of food for the cities, and an economic constraint on animals and animal products for export (Provost and Bezuidenhout, 1987; Anon, 1996).

CHAPTER TWO
HEARTWATER – THE DISEASE

Heartwater is endemic throughout most of sub-Saharan Africa, except Lesotho. It is also present on islands around Africa such as Zanzibar and Sao Tomè. Heartwater also occurs on islands in the Indian Ocean such as the Comores, Madagascar, La Réunion and Mauritius (Uilenberg, 1983; Camus and Barré, 1988; du Plessis *et al.*, 1989; Flach *et al.*, 1990; Uilenberg, 1990; Uilenberg *et al.*, 1993; Yunker and Allsopp, 1994).

2.1 ECONOMICAL IMPORTANCE

Heartwater is of economical importance as this disease affects sheep, goats and cattle. The heartwater-endemic area of South Africa contains 42% of the country's total cattle population and 52% of the total goat population. Small-scale farmers own at least half of these goats, and it is estimated that 30% of their goats contract heartwater annually. It is estimated that 8.7 million head of livestock in the small-scale farming sector are at risk, which threatens the food security of rural communities. In the commercial sector another 8.8 million is at risk, which threatens food destined for the cities and the income derived from the export of animal products (Provost and Bezuidenhout, 1987; Allsopp, 2000).

The Deputy Director of Veterinary Services in the Eastern Cape estimated that in 1998 an amount of between R8.8 million and R30 million was spent on prophylactics and vaccination. These figures are only estimates as much of the expenditure was on private farms. Despite this there was a loss of 10%, at an estimated cost of R214 million, of total stock in the province due to heartwater (Allsopp, 2000).

In Botswana, the government is currently awarding loans to farmers to purchase flocks of goats in the areas close to the capital and the South African border. After moving heartwater-naive animals from the sandveld to the hardveld many of the flocks suffer 50% losses within months as a result of heartwater (Allsopp, 2000).

It has recently been realised that the loss of wildlife resources, as a result of poaching, is escalating at an alarming rate across Africa. Domestic farmed meat is frequently too expensive for poor communities, and instead they turn to poaching wild animals. This is not confined to conspicuous large mammals, and is therefore more difficult to prevent than the much better known poaching for trophies and traditional medicine. In some areas this is driving small and medium sized mammals to extinction. The issue of food security is therefore crucial to the preservation of the genetic diversity of wildlife. Therefore anything that will improve the health of domestic livestock, will improve food security (BBC News, March 2000).

2.2 VETERINARY IMPORTANCE

Heartwater is a non-contagious, infectious and often fatal disease caused by the rickettsial organism, *Cowdria ruminantium*. The mortality rate due to heartwater is high in imported breeds, which are introduced because of their high potential productivity. Heartwater is one of the main causes for the failure of establishing such breeds and their crosses in endemic areas. Heartwater therefore constitutes a major constraint to increasing the productivity of cattle and also small ruminants by the use of exotic stock in Africa (Spreull, 1922; Alexander, 1931; du Plessis *et al.*, 1983; du Plessis *et al.*, 1984a; Bezuidenhout, 1985a; Uilenberg, 1990; Uilenberg *et al.*, 1993; Uilenberg and Camus, 1993).

Indigenous breeds and populations of cattle and small ruminants in enzootic areas are far less susceptible than animals from other regions. This is due to natural selection over many generations. In areas where adequate tick control and veterinary care do not exist, the only way of living with heartwater at present is to use populations of indigenous breeds originating from enzootic regions (Uilenberg, 1990).

2.3 THREAT TO THE USA

Heartwater also occurs in the New World. Heartwater-infected *Amblyomma variegatum* ticks were introduced from Senegal into the Caribbean in the 19th century and the disease is now established on three islands: Guadeloupe, Marie Galante and Antigua (Uilenberg *et al.*, 1984; Birnie *et al.*, 1985; Burridge, 1985; Uilenberg *et al.*, 1993; Camus and Barré, 1987; Allsopp *et al.*, 1999). The close proximity of these islands to the USA and southern and central America

creates a threat for the spread of the disease to these countries. Potential vectors are present in these countries where the disease is currently absent and the animal population is therefore totally susceptible (Barré *et al.*, 1987; Camus *et al.*, 1996; Musoke *et al.*, 1996; Musoke *et al.*, 1997).

The *A. variegatum* tick has spread at an alarming rate and is fast approaching the American mainland where large areas are climatically suitable for the vector. This rapid spread is due to the introduction of the African cattle egret (*Bubulus ibis*) into America. The cattle egret, a native of Africa, is established in the Western Hemisphere, including in the Caribbean islands. These birds are commonly associated with cattle in pastures and they can serve as hosts for immature *A. variegatum* ticks. Clinical heartwater cases are observed within the distribution area of these ticks (Hancock *et al.*, 1984; Arendt, 1988; Crosby, 1972; Dinsmore, 1972; Heatwole, 1965; Barrè *et al.*, 1988).

Data on migration patterns of these birds suggests that, after domestic animals, cattle egrets are a second potential vehicle for disseminating *A. variegatum* ticks and thus *C. ruminantium* into the Americas. The invasion and multiplication of the cattle egret in the area precede the extension of the tick by a few years. Immature ticks are found on this bird and its migrating route is between the island and the continent. In addition, three American *Amblyomma* species have been shown to be capable of transmitting the disease. The consequence of an invasion would be catastrophic as the tick is not only a vector of heartwater but is also associated with severe bovine dermatophilis, secondary abscesses and screw-worm infestations. This could lead to the spread of heartwater to livestock in North, Central and South America (Corn *et al.*, 1993; Barré *et al.*, 1987; Uilenberg, 1982; Uilenberg *et al.*, 1993; Alderink and McCauley, 1988; Uilenberg *et al.*, 1984).

Previously it has been accepted that heartwater was introduced into the Caribbean through the importation of infested cattle from French West Africa in the early to middle 1800's (Birnie *et al.*, 1985). However it has recently been reported that there are genetic variations in *A. hebraeum* ticks from the Caribbean and Africa. Such observations raise questions about the origins of *A. variegatum* and *C. ruminantium* found in the Caribbean (Estrada-Pena *et al.*, 1994; Reddy *et al.*, 1996). This suggests the possibility of more than one introduction of *C. ruminantium*- infected ticks into the Caribbean from Africa and possibly from different regions within Africa (Estrada-Pena *et al.*, 1994).

2.4 PATHOLOGY

Heartwater is an acute and fatal infectious disease of cattle, sheep, goats and certain wildlife species. The pathology of the disease differs between breeds of ruminants and *Cowdria* isolates (Camus *et al.*, 1996). The name heartwater is derived from the fact that the most striking change in animals that die of the disease is the development of hydropericardium. The pathology of the disease is mainly the result of increased vascular permeability, leading to transudation and oedema of the lungs (Clark, 1962; Owen *et al.*, 1973; du Plessis 1975b; Prozesky and du Plessis, 1984; Prozesky and du Plessis, 1985a/b; Prozesky, 1987a). Heartwater is also characterised by the development of petechia on the conjuctiva of the eye, high fever, diarrhea, respiratory distress, severe nervous disorders such as circling and disorientation and hydrothorax (Clark, 1962; Neitz, 1968; Uilenberg *et al.*, 1983; van de Pypekamp and Prozesky, 1987; Uilenberg *et al.*, 1993).

The average incubation period in naive animals is less than 2 weeks and the mortality rate is between 20-90%. In adult cattle the mortality rate is 82%, in merino sheep it is 95% and in Angora goats as many as 90% of infected goats die. It is estimated that in this region over 175 million cattle are at risk, and perhaps twice as many small stock (Neitz, 1964; Uilenberg, 1983; du Plessis and Malan, 1987d; van de Pypekamp and Prozesky, 1987).

Depending on the age, immune status, individual or breed susceptibility of the animal and virulence of the isolate, the course of the disease may range from peracute to mild. The incubation period of heartwater is influenced by the species of animal affected; the route of infection, virulence of the heartwater isolate and the amount of infective material administered (Alexander, 1931; Neitz, 1968; Uilenberg, 1983; van de Pypekamp and Prozesky, 1987).

2.5 CONTROL

There are currently four basic ways of controlling heartwater: vector control, immunisation, chemotherapy and the use of natural resistance of young animals and indigenous breeds in endemic regions (Uilenberg, 1990; Musoke *et al.* 1996).

2.5.1 Vector control

Vector control is the major prevention method used in Africa and relies on acaricides. The intensive acaricidal tick control is harmful as animals reared under tick-free conditions are totally susceptible to all tick-borne diseases and heavy losses are incurred when tick control breaks down. Tick control can break down for various reasons: ticks become resistant to the acaricides, lack of funds for the purchase of acaricides and maintenance of control facilities, lack of water during drought, political unrest, civil war, etc. Additional concerns relate to environmental contamination such as introducing the pathogenic agent into areas, which are *Cowdria* free (Uilenberg, 1990; Uilenberg *et al.*, 1993; Musoke *et al.* 1996; Musoke *et al.*, 1997).

2.5.2 Immunisation

Current vaccination strategies are based on the initiation of infections with blood- or culture-derived stabilates followed by treatment with tetracycline. Animals are infected with the blood of donor animals (containing viable virulent organisms of the Ball3 isolate) and are subsequently treated with tetracycline on elevation of temperature (van Amstel and Oberem, 1987; Bezuidenhout and Oberem, 1985c; Uilenberg, 1990b; Neitz and Alexander, 1941). The major disadvantage of available methods of immunisation against tick-borne diseases is their dependence on live organisms (Musoke *et al.*, 1997). This method cannot be used in non-endemic areas because it includes viable *Rickettsiae*. The immunisation material must be stored and transported at -40°C, which makes it unsuitable for application in rural areas. It can cause erythrolysis and adverse histamine reactions and the duration of immunity is uncertain. It causes shock in young lambs and kids and the vaccine has to be administered intravenously, which makes it difficult for use by resource poor farmers (Bezuidenhout, 1981; van der Merwe, 1987).

Tick material infected with *C. ruminantium* is currently being used. It is cheaper to use and offers less risk than the use of infected animal blood. The infected tick material can be stocked at -18°C making it easier to use in rural areas. It is currently costing South Africa in the region of R280 000 per annum to produce this vaccine (Martinez *et al*, 1993; Bezuidenhout, 1981; Uilenberg, 1990; Musoke *et al.* 1996; Dr Theo de Waal, personal communication)

The infection and subsequent treatment method has many shortcomings, which includes potential contamination of blood-based vaccines with known pathogens and widespread direct transmission of unidentified or newly emergent pathogens. Vaccination associated deaths do occur and there are limitations on standardisation of this vaccine. Infected tick material is unsuitable for use in countries where potential vectors are present but which do not harbor the disease. Animals that are treated in this manner develop isolate specific immunity and are not protected against strains not included in the vaccine (Rogers *et al.*, 1988; Henry *et al.*, 1983; Potgieter and van Rensburg, 1983; Oberem and Bezuidenhout, 1987b; van der Merwe, 1987).

2.5.3 Chemotherapy

The antibiotic tetracycline has significant effect on heartwater. Long-acting preparations of tetracycline have limited the number of injections needed but may be the cause of local tissue damage. The problem with using tetracycline is that the disease is only obvious when clinical symptoms appear and treatment is often too late to save the animals (Uilenberg, 1990).

2.5.4 Natural resistance

Heartwater is rarely observed in indigenous livestock in endemic areas, and presents as a problem mainly in susceptible animals that have been moved to areas where the agent is present (Provost and Bezuidenhout, 1987; Neitz and Alexander, 1941).

Young cattle possess innate resistance to the causal agent *C. ruminantium*. Calves, up to a few weeks old, often survive infection with no or mild obvious clinical signs. Lambs and kids also exhibit considerable tolerance for a short period after birth. This natural tolerance is independent of the immune status of the parent animal and also occurs in populations that have not been exposed to the disease before. This short period of natural tolerance after birth can be used to immunise the animal with little risk of severe reaction (Neitz and Alexander, 1941; Uilenberg, 1983; Uilenberg 1990).

2.6 VACCINES

Vaccination has proved to be an extremely effective approach to protecting human and animal populations from the ravages of infectious diseases caused by microorganisms. Vaccines promote long-lasting immunity in the absence of debilitating clinical symptoms, and the spread of potentially pathogenic microorganisms in a normally susceptible population can be limited. Vaccination is based on the observation that exposure to a pathogenic entity can

result in the host mounting a long-lasting immunological defence preventing reinfection by the same organism (Winther and Dougan, 1984).

A vaccine is needed to control heartwater, as current methods are expensive, impractical and often hazardous. The ideal vaccine should mimic the immunological stimulus associated with natural infection and should be effective, giving more than 90% protection. The vaccine should have minimal side effects and be absolutely safe to use. Following vaccination against the disease, the individual should have life-long immunity. Immunisation should require only one injection and the vaccine should be compatible with other vaccines. Administration should be easy and by a non-invasive route, and be given as close to birth as possible. The vaccine should have defined and reproducible production conditions. It should be inexpensive, stable and not require a cold chain for transportation, distribution or storage (Helland and Hey, 1990).

The major consideration in designing vaccines for haemoparasites is that many of these organisms have at least two lifecycle stages in the host animal. A successful vaccine may need to include antigens of different lifecycle stages. An effective vaccine need not necessarily mimic immune responses implicated in protection under natural circumstances, as artificial manipulation of the host's immune system using appropriate antigen delivery systems might result in protective responses against antigens that might not be involved in protective immunity in animals immunised with the whole organism (Musoke *et al.*, 1996; Totté *et al.*, 1999).

Vaccines are classified into three groups on the basis of how the antigenic material for vaccination was derived: First-generation, second-generation and third-generation vaccines (Alarcon *et al.*, 1999).

2.6.1 First-generation vaccines

First-generation vaccines are based on whole organisms, in either live attenuated or killed forms. However, vaccines that contain whole organisms have drawbacks, they retain molecules that are not involved in evoking protective immunity and may include contaminants that could trigger allergic or disruptive reactions (Alarcon et al., 1999; Weiner and Kennedy, 1999).

2.6.1.1 Live vaccines

Live vaccines include two types of vaccines: (i) original disease-causing microorganisms are weakened and used as vaccines and (ii) microorganisms that replicate similarly to the natural microorganism *in vivo*, thereby eliciting an immune response similar to that elicited by the natural infection, are used.

The major advantage of using live vaccines is that they are able to elicit life-long immunity with minimal reactigenicity after only two doses. Live vaccines usually elicit humoral immunity (production of antibodies by B lymphocytes that act on pathogens that are outside cells) and cellular immunity (production of cytotoxic T lymphocytes (CTL) that eradicate pathogens that colonise cells) as well as thymus-derived lymphocyte helper cell (T_H) immunity (Ellis, 1999; Weiner and Kennedy, 1999). This dual activity is essential for blocking infection and for ensuring immunity.

The major disadvantage of using live vaccines is that they can cause full-blown illness in subjects that are immune-compromised, as they can replicate in the host. Healthy subjects can contract disease from recently infected subjects and weakened microorganisms might mutate, which could restore virulence (Helland and Hey, 1990; Ellis, 1999; Weiner and Kennedy, 1999; Ogra *et al.*, 1991; Weeks-Levy *et al.*, 1991).

2.6.1.2 Attenuated / Inactivated vaccines

When producing attenuated and inactivated vaccines, the pathogenic capacity is eliminated while the organism's ability to replicate is retained. The original organism is modified by passage, and selection in culture or an appropriate mutant is selected by chemical mutagenesis. The major advantage of using these vaccines are that they are generally inexpensive to produce, give persistent immunity and have a good safety record. However, these vaccines do not enter cells but make antigens that are displayed by the cells, which are subsequently used to inoculate a specific animal. They thus stimulate a humoral response, which leads to antibody production. The major disadvantages with these vaccines are reversion to virulence, possible induction of disease in immunodifficient subjects and the need for a cold chain (Helland and Hey, 1990; Weiner and Kennedy, 1999).

The *C. ruminantium* Senegal isolate became attenuated spontaneously after several passages *in vitro* and was capable of inducing protection in goats and sheep against virulent homologous challenge with a blood stabilate, but not against challenge from *C. ruminantium* isolates of geographically diverse backgrounds. It would therefore appear to be necessary to

attenuate several strains to cover the entire antigenic repertoire. In addition, only certain strains can be attenuated by *in vitro* passage, which limits the use of this method for vaccination (Camus *et al.*, 1996; Jongejan, 1991c; Jongejan *et al.*, 1993; Mahan *et al.*, 1999).

An inactivated vaccine consisting of bovine endothelial cell (EC) culture-derived *C. ruminantium* organisms that are chemically inactivated or lysed, and combined with an adjuvant Montanide ISA 50 is also being used to combat heartwater. This vaccination strategy led to high levels of protection against homologous challenge in goats, sheep and cattle. This vaccine has also been shown to protect sheep against heterologous challenge from isolates from geographically diverse backgrounds. Overall, against a challenge that resulted in a mortality rate of 73% in control animals, a vaccine efficacy of 68% was observed. In goats, sheep and cattle, this vaccine protected 65%, 79% and 100% of the vaccinated animals respectively (Martinez *et al.*, 1994; Mahan *et al.*, 1995b; Mahan *et al.*, 1998a; Totté *et al.*, 1997; Martinez *et al.*, 1996).

The inactivated vaccine does not prevent infection, but prevents death. Vaccinated animals develop a lower rickettsemia, hence a milder disease after challenge. This vaccine has the advantage that it can be modified to include any isolate of *C. ruminantium* and therefore could overcome the recognised lack of cross-protection between some heterologous isolates (Mahan *et al.*, 1995b; Mahan *et al.*, 1998a; Mahan *et al.*, 1999).

2.6.1.3 Killed vaccines

The advantages of using killed vaccines are multiple. Killed vaccines are more stable than attenuated vaccines. They are simple to prepare and contain all the important antigens. Killed vaccines cannot revert back to the pathogenic form, which makes it less hazardous to work with. The disadvantages of using killed vaccines are: the need for careful monitoring to avoid the presence of viable organism makes it difficult to work with. Periodic booster shots are needed as the protection often wears off after time. Killed vaccines cannot make their way into cells and can only stimulate a humoral response against a pathogen and are unable to generate specific CTL responses. A humoral response alone is ineffective against many microorganisms that infiltrate cells (Alarcon *et al.*, 1999; Helland and Hey, 1990; Weiner and Kennedy, 1999).

2.6.2 Second-generation vaccines

Second-generation vaccines consist of defined native or recombinant protein components derived from the organism, rather than the whole organism itself. These components are prepared by biochemical purification or genetic engineering. The advantages of using second-generation vaccines are that they are safe to use as they cannot replicate in the host and cannot multiply or revert to pathogenecity. They are less reactogenic, nontransmissible to another person or animal and in most cases, more technically feasible. The major disadvantage of using second-generation vaccines are that vaccination with defined protein components induces humoral immune responses but generally not CTL responses (Helland and Hey, 1990; Ellis, 1999; Alarcon *et al.*, 1999).

2.6.2.1 Subunit vaccines

Subunit vaccines are the best defined of all vaccines. Subunit vaccines include inactivated toxins from a pathogen, purified surface compounds, or conjugated surface compounds (Helland and Hey, 1990; Ellis, 1999).

Although a number of antigens are recognised by the immune system, not all of the antibodies that bind to the microorganism and its extracellular products will necessarily lead to its destruction by host systems. The viral or bacterial neutralising antibodies that are generated by whole-cell vaccines may be attributable to only one or a few antigens. The basis of developing subunit vaccines is to identify the essential immunogenic components of a virus or bacterium and to devise a way of producing the particular antigen separately from the rest of the cellular components that does not contribute to the protective immune response. The immune system is presented with a simple immunogen, which should produce high levels of specific neutralising antibody, but without the side effects or hazards associated with some whole-cell vaccines (Winther and Dougan, 1984).

The advantages of using subunit vaccines are the fact that they can be produced in large quantities under controlled conditions and are more stable than attenuated vaccines The disadvantage of using subunit vaccines is that they require several immunisations (Helland and Hey, 1990; Ellis, 1999).

2.6.2.2 Synthetic peptide vaccines

Synthetic peptide vaccines are a further refinement to the subunit vaccine concept. This follows on the observation that only small portions of a native antigen are exposed to the immune system. The use of synthetic peptides in vaccine production involves the use of short segments of a specific protein molecule as immunogen, resulting in the production of a highly specific vaccine, free from competing or non-essential antigenic components. One problem is that whereas antibodies directed against native protein usually recognise the intact tertiary protein structure, antibodies raised against a single peptide may only recognise unfolded protein (Helland and Hey, 1990; Winther and Dougan, 1984).

2.6.3 Third-generation vaccines

Third-generation vaccines are nucleic acid vaccines and are based on the genetic material of the infectious organism (Alarcon *et al.*, 1999).

2.6.3.1 Genetic vaccines

A DNA (deoxyribonucleic acid) or RNA (ribonucleic acid) construct encoding the gene of one or more antigenic proteins normally synthesised by the selected pathogen is delivered directly to the cells of the organism to be immunised. The construct is taken up by the host cells and expressed. The endogenously expressed immunogen subsequently induces an immune response in the host. The advantages of genetic vaccines are: (i) Genetic vaccines are able to induce specific CTL as well as humoral responses without the risks associated with live vaccines. (ii) The *in situ* expression of genes by host-cells, leads to the synthesis of proteins that closely resemble the native molecule, which contain the relevant epitopes and are therefore able to induce a more effective immune response and have a standardised method of production. (iii) Genetic vaccines exclude genes that would enable the pathogen to reconstitute itself and cause disease. The major disadvantage of genetic vaccines is that they stimulate synthesis of antigens in cells only (Waine and McManus, 1995; Ellis, 1999; Weiner and Kennedy, 1999; Alarcon *et al.*, 1999).

At present, DNA-based vaccines promise to be the best way to solve the problems of antigen presentation and elicitation of the complex immune mechanisms needed for protection against tick-borne pathogens. The difficulties associated with using genetic vaccines are:

(i) The high level of genetic and antigenic diversity (the existence of complex parasite populations comprising a large number of genotypes, each with the potential to express a

variant form of a particular antigen) of *Cowdria* complicates the development of a genetic vaccine (Preston and Jongejan, 1999a; Totté *et al.*, 1999). (ii) Antigenic variation (the sequential expression of variant forms of large multicopy gene families) (du Plessis *et al.*, 1989; Jongejan *et al.*, 1991b; Preston and Jongejan, 1999a).

Major antigenic protein 1 (MAP1) is the only well characterised antigen shown so far to be partially protective in mice. Intramuscular injection of a DNA vaccine containing the *map1* gene protected 23% to 88% in different groups after challenge with a lethal dose of cell culture-derived *C. ruminantium* organisms. The immunised mice synthesised antibody to MAP1 and mounted a T helper cell type 1 (T_{H1}) response. The analysis of the supernatants of splenocytes cultures of the immunised mice showed the production of interferon gamma (IFN-γ) and interleukin 2 (IL-2). Immunisation with denatured MAP1 did not protect goats and sheep, despite high antibody titres before challenge. The capacity of this vaccine to protect against heterologous challenge remains to be tested, and its protective effect has not been reproduced in ruminants (Nyika *et al.*, 1998; van Kleef *et al.*, 1993).

CHAPTER THREE

HEARTWATER – THE VECTOR

3.1. *AMBLYOMMA HEBRAEUM*

The *C. ruminantium* organism is transmitted exclusively by ticks of the genus *Amblyomma* to wild and domestic ruminants in sub-Saharan Africa (Fig. 1) and the Caribbean (Provost and Bezuidenhout, 1987).

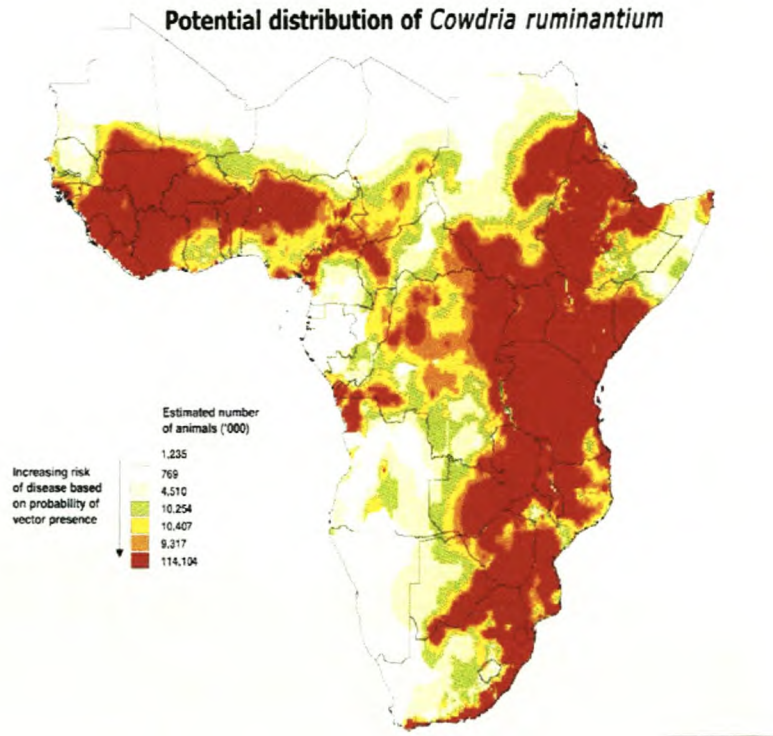


Fig. 1. The potential distribution of *C. ruminantium* (adapted from the poster, Heartwater in sub-Saharan Africa, Minjuaw, 2001)

Heartwater is transmitted by at least ten African tick species of the genus *Amblyomma* and at least three American species of the genus (*A. cajennense, A. maculatum, A. dissimile*) are experimental vectors. In most of Africa, *A. variegatum* is the main vector (Fig. 2) while in Southeastern Africa the most important vector is *A. hebraeum* (Fig. 3) (Lounsbury, 1900; Cowdry, 1925; Uilenberg, 1982; Jongejan, 1992; Uilenberg, 1983; Bezuidenhout, 1987; Prozesky, 1987c; Uilenberg *et al.*, 1993; Uilenberg and Camus, 1993; Yunker, 1995).

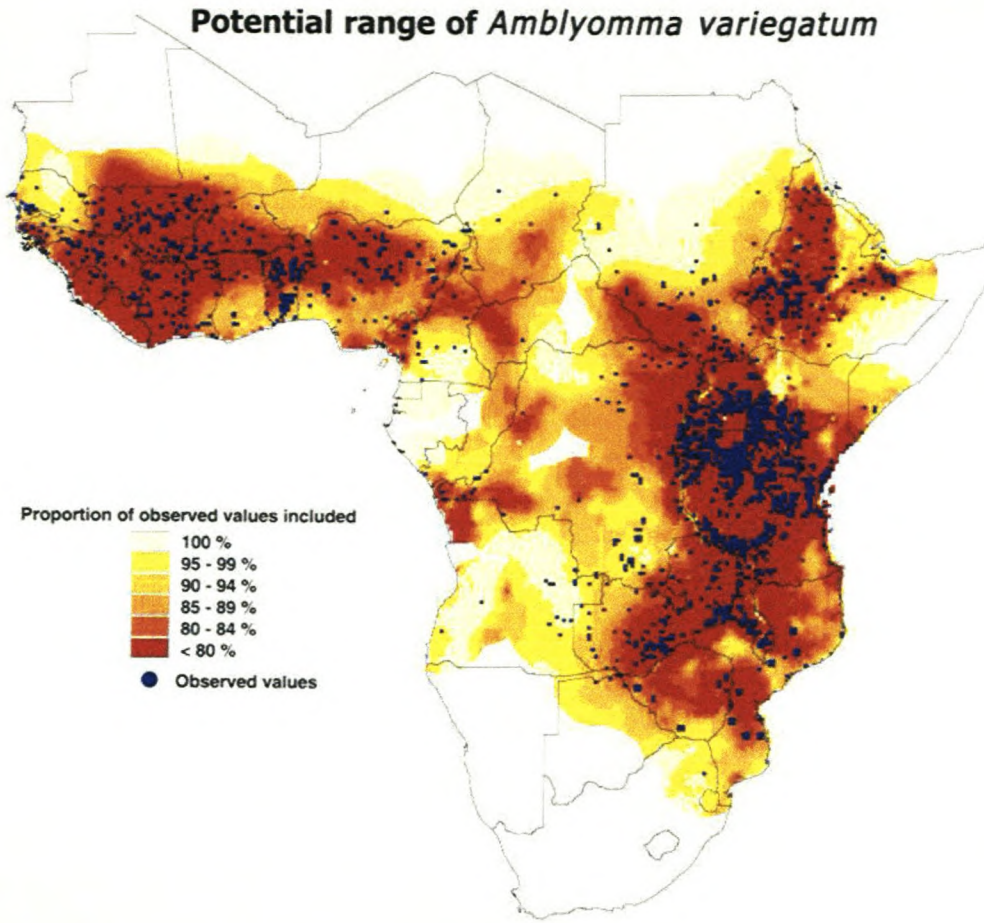


Fig. 2. The potential distribution of *A. variegatum* (adapted from the poster, Heartwater in sub-Saharan Africa, Minjuaw, 2001).

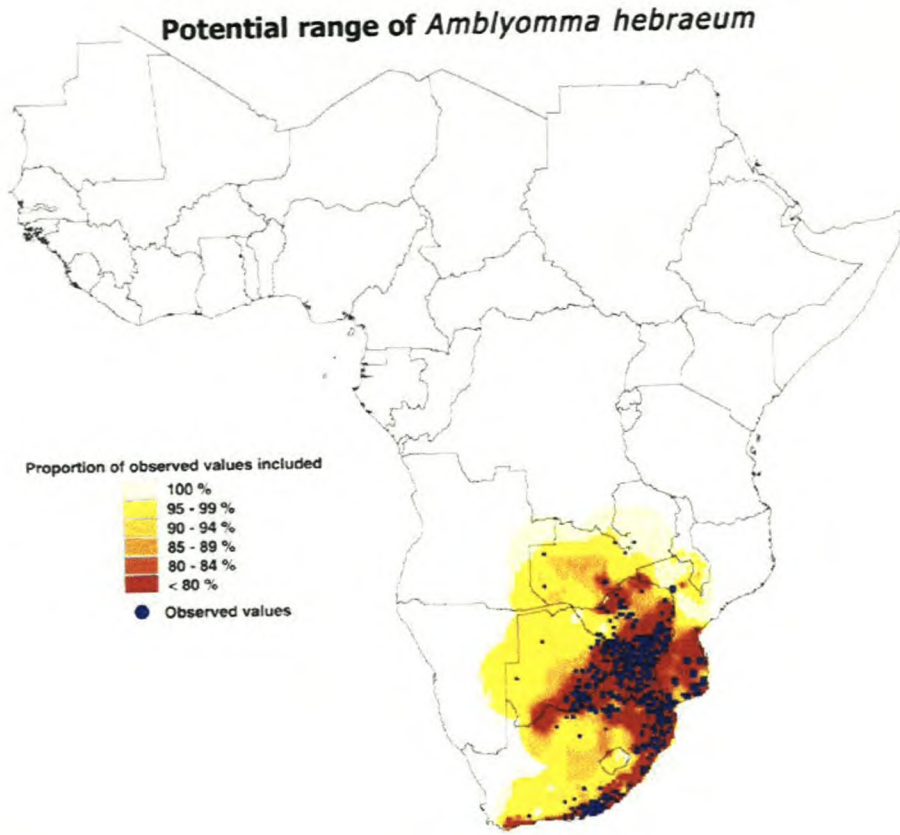


Fig. 3. The potential distribution of *A. hebraeum* (adapted from the poster, Heartwater in sub-Saharan Africa, Minjuaw, 2001)

Wildlife species are suspected to be important reservoirs for *C. ruminantium* by being alternative hosts for *Amblyomma* ticks in areas where tick control methods on domestic ruminants are exercised, thus contributing to the spread of heartwater. Natural or experimental infection with heartwater has been reported in *Antidorcas marsupialis* (springbok), *Taurotragus oryx* (eland), *Tragelaphus spekei* (sitatunga) and *Kobus leche* (lechwe), whereas subclinical infection has been demonstrated in *Syncerus caffer* (African buffalo) and *Giraffa camelopardalis* (giraffe). The crowned guinea fowl (*Numida meleagis*) and leopard tortoise (*Geochelone pardalis*) can serve as subclinical carriers of *C. ruminantium* and act as a source of the organism for ticks (Andrew and Norval, 1989; Oberem and Bezuidenhout, 1987a; Uilenberg and Camus, 1993; Norval *et al.*, 1994; Kock *et al.*, 1995).

3.2 PHYLOGENY

Amblyomma hebraeum (Fig. 4) is classified in the family of hardback ticks Ixodidae. The lineage of *A. hebraeum* is as follows (National Library of Medicine website[1]):

Superkingdom: Eukaryota

Kingdom: Metazoa

Phylum: Arthropoda

Subphylum: Chelicerata

Class: Arachnida

Subclass: Acari

Superorder: Parasitiformes

Order: Ixodida

Family: Ixodidae

Genus: Amblyomma

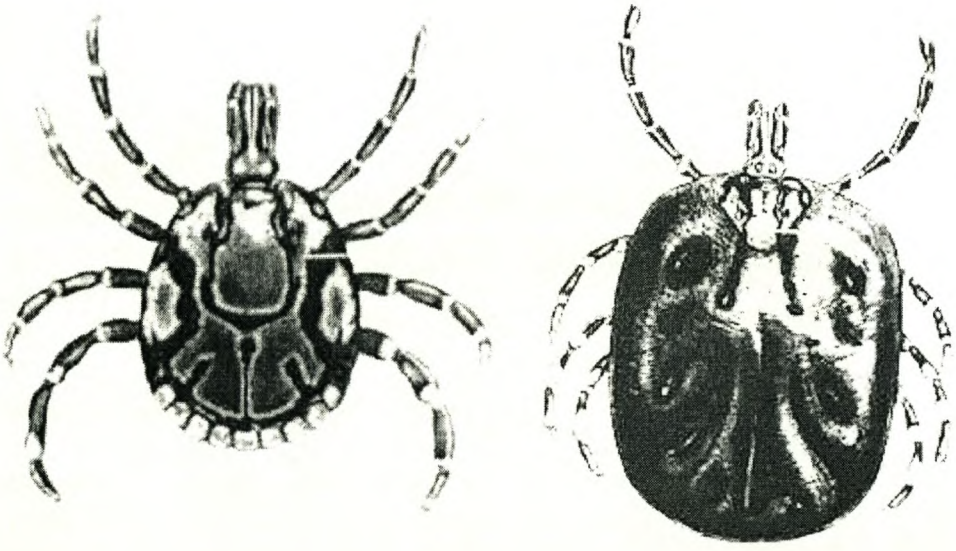


Fig. 4. *A. hebraeum* male tick (left) and female tick (right) (Proagri website[2]).

[1] http://www.ncbi.nlm.nih.gov
[2] http://www.proagri.co.za

3.3 LIFE CYCLE

All hard ticks share a similar basic life cycle, with three mobile stages; the small six-legged larvae, the slightly larger eight-legged nymph and the easily recognisable eight-legged adult. The vectors of heartwater are 3-host ticks (Fig. 5), where the larvae, nymph and adults each feed on separate hosts, after which the ticks detach and spend long periods in the vegetation. The larvae feed on its first ruminant host after which it drops off its host and moults into a nymph. The nymph feeds on the second ruminant host, subsequently drops off and moults into an adult. The adult feeds on the third ruminant host. The engorged adult drops off its host and the female adult lays her eggs on the ground (Preston and Jongejan, 1999b.) Transmission is transstadial and *C. ruminantium* infection in *A. hebraeum* is located in the midgut (Bezuidenhout, 1987; Yunker *et al.*, 1993; Jongejan and Uilenberg, 1994).

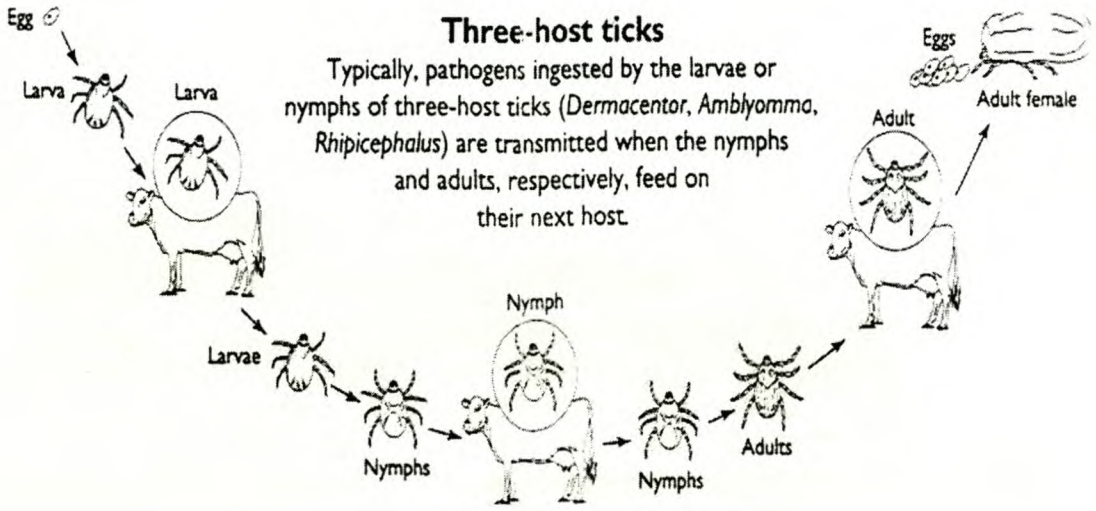


Fig. 5. The three-host tick transmission of *C. ruminantium* (Preston and Jongejan, 1999b).

There are two moulting phases between the three mobile stages, during which the outer skin or cuticle is shed and replaced, together with internal reorganisation taking place. Eggs are laid on the ground in a cool, moist place under a tuft of grass or fallen leaves. Larvae hatch from the eggs and climb up and cling to vegetation while waiting for passing hosts. The larvae pierce the host's skin with their mouthparts to feed on a blood meal. Larvae subsequently

moult and change to nymphs. The nymphs feed on a blood meal, engorge, moult and change into adults. The adult stage is the only stage that is sexually differentiated. The adults mate on the host. The female consumes a large meal of blood, becomes engorged and drops of the host to lay 2 000 – 20 000 eggs, after which she dies (Proagri website[2]).

Ticks acquire the infection via the blood meal while feeding on a reacting animal or a subclinically infected host during the period in which a heartwater reaction or circulation of the organism would be expected. The organisms initially develop in the midgut epithelial cells and subsequent stages invade and develop in the haemolymph and salivary glands. The vertebrate host is infected transsalivaril. Re-attachment of an infected tick on a susceptible vertebrate host results in transfer of the organism into the bloodstream of the host, where a new development cycle occurs. Transmission to the final host seems to be co-ordinated with the feeding cycle of the ticks (Lounsbury, 1900; Neitz, 1968; Bezuidenhout, 1987; Kocan and Bezuidenhout, 1987; Kocan *et al.*, 1987a).

[2] http://www.proagri.co.za

CHAPTER FOUR
HEARTWATER – THE CAUSATIVE ORGANISM

4.1 *COWDRIA RUMINANTIUM*

The causative organism of heartwater is *Cowdria ruminantium*; hence the disease is also known as cowdriosis. *C. ruminantium* lives in blood-sucking arthropods (ticks of the *Amblyomma* species), which serve as vectors or primary hosts. Although *C. ruminantium* mostly infects domestic ruminants it can also infect birds, reptiles and rodents. Several species of game animals are also susceptible to heartwater (Cowdry, 1925; Walker and Olwage, 1987; Young and Basson, 1973; Uilenberg, 1983; Uilenberg *et al.*, 1993).

There are important differences between different isolates (Table 1) of *C. ruminantium* in virulence and infectivity for various hosts, as well as in antigenic composition shown in *in vivo* cross-protection studies. These differences hamper the development of a vaccine that would be protective against all heartwater outbreaks (Uilenberg *et al.*, 1993).

Table 1. *C. ruminantium* isolates (adapted from PhD Thesis of de Villiers 2001).

Isolate	Geographic area	Reference
Antigua	Antigua	Birnie *et al.*, 1985
Mochudi	Botswana	Mahan *et al.*, 1998d
Sunnyside	Botswana	Mahan *et al.*, 1998d
Burkina Faso	Burkina Faso	Jongejan and Bekker, 1999
Cameroun	Cameroun	Jongejan *et al.*, 1988
Comoro	Comoro Islands	du Plessis *et al.*, 1989
Pokoase 417	Ghana	Bell-Sakyi *et al.*, 1996
Pokoase 418	Ghana	Bell-Sakyi *et al.*, 1996
Sankat 429	Ghana	Bell-Sakyi *et al.*, 1996
Sankat 430	Ghana	Bell-Sakyi *et al.*, 1996
Gardel	Gaudeluope	Uilenberg *et al.*, 1985
Isiolo	Kenya	Mahan *et al.*, 1998a
Kiswani	Kenya	Kocan *et al.*, 1987a
Mali	Mali	Logan *et al.*, 1985

Table 1. *C. ruminantium* isolates continued (adapted from PhD Thesis of de Villiers 2001)

Isolate	Geographic area	Reference
Bela Vista	Mozambique	Bekker *et al.*, 2000
Bom Pastor	Mozambique	Bekker *et al.*, 2000
Porto Henrique	Mozambique	Bekker *et al.*, 2000
Umpala	Mozambique	Jongejan *et al.*, 1988
Nigeria D225	Nigeria	Ilemobade and Leeflang, 1977
Sao Tomé	Sao Tomé	Uilenberg *et al.*, 1982
Senegal	Senegal	Jongejan *et al.*, 1988
Ball3	South Africa	Haig, 1952
Blaauwkrantz	South Africa	Zweygarth, personal communication
Breed	South Africa	du Plessis *et al.*, 1983
Germishuys	South Africa	du Plessis *et al.*, 1989
Kümm	South Africa	du Plessis and Kümm, 1971
Kwanyanga	South Africa	Mackenzie and van Rooyen, 1981
Mara 87/7	South Africa	du Plessis *et al.*, 1989
Nonile	South Africa	Mackenzie and McHardy, 1984
Omatjenne	South Africa	du Plessis, 1990
Pretoria North	South Africa	Allsop and Allsopp, 2001
Rietgat	South Africa	Mahan *et al.*, 1998d
Skukuza	South Africa	Peter *et al.*, 1999
Vosloo	South Africa	du Plessis, 1993b
Warmbaths	South Africa	Mahan *et al.*, 1998d
Welgevonden	South Africa	du Plessis, 1985
Um Banein	Sudan	Jongejan *et al.*, 1984
Mpisi	Swaziland	Mahan *et al.*, 1998d
Big Bend	Swaziland	Mahan *et al.*, 1998d
Tanga	Tanzania	Smith *et al.*, 1998
Gamela	Zambia	Mahan *et al.*, 1998d
Lutale	Zambia	Jongejan *et al.*, 1988
Beatrice	Zimbabwe	Mahan *et al.*, 1998
Crystal Springs	Zimbabwe	Byrom *et al.*, 1991
Finale	Zimbabwe	Smith *et al.*, 1998

Table 1. *C. ruminantium* isolates continued (adapted from PhD Thesis of de Villiers 2001)

Isolate	Geographic area	Reference
Highway	Zimbabwe	Byrom *et al.*, 1991
Hunyani	Zimbabwe	Smith *et al.*, 1998
Kwekwe	Zimbabwe	Smith *et al.*, 1998
Lemco T3	Zimbabwe	Byrom *et al.*, 1991
Mbizi	Zimbabwe	Byrom *et al.*, 1991
Mubayira	Zimbabwe	Smith *et al.*, 1998
Nyatsanga	Zimbabwe	Byrom *et al.*, 1991
Palm River	Zimbabwe	Byrom *et al.*, 1991
Plumtree	Zimbabwe	Mahan *et al.*, 1994b
Rusape	Zimbabwe	Smith *et al.*, 1998
Zvimba	Zimbabwe	Norval *et al.*, 1994

4.2 PHYLOGENY

The lineage of *Cowdria ruminantium* is as follows (National Library of Medicine website[1]):

Superkingdom: Bacteria

Phylum: Proteobacteria

Class: Alpha subdivision

Order: Rickettsiales

Family: Rickettsiaceae

Tribe: Ehrlichieae

Genus: Cowdria

Cowdria is grouped as a separate genus in the order *Rickettsiales* (Fig. 6).

Cowdria is closely related to and almost inseparable from *Rickettsia, Ehrlichia* and *Anaplasm,* seriously complicating the diagnosis and serology of heartwater. *C. ruminantium* is the only described species in the genus *Cowdria* and is most closely related to *Ehrlichia ewingi* (Byrom and Yunker, 1990; Dame *et al.*, 1992; van Vliet *et al.*, 1992; Ristic and Huxoll, 1984; Allsopp *et al.*, 1997).

[1] http://www.ncbi.nlm.nih.gov

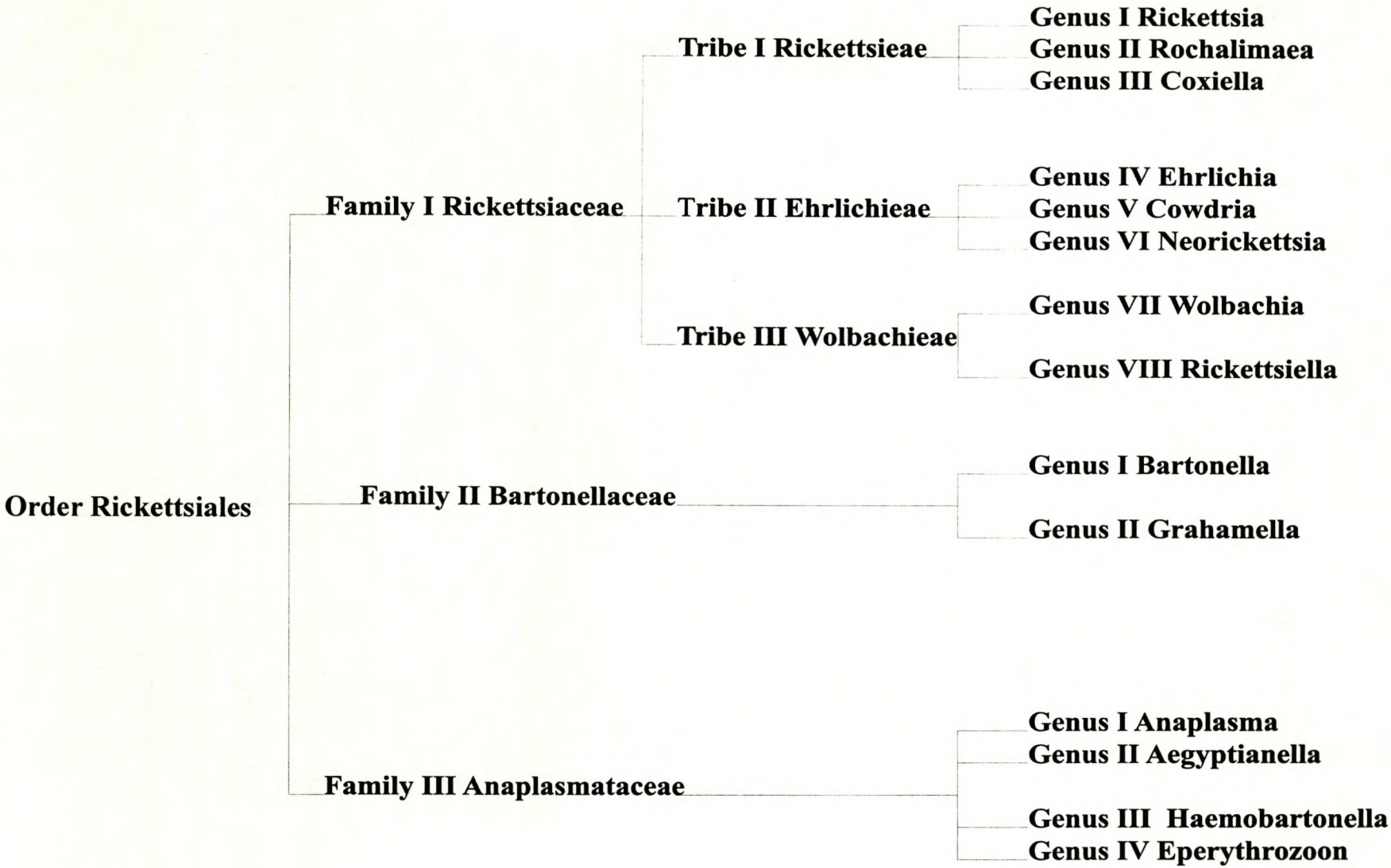


Fig. 6. The Phylogenetic tree of order *Rickettsiales* showing the phylogenetic grouping of the genus *Cowdria* (Holt and Krieg, 1979).

4.3 MORPHOLOGY

C. ruminantium is a Gram-negative obligate intracellular parasite, which only grows and reproduces within an eukaryotic host cell. It is maintained in the environment through a cycle involving mammalian reservoirs and insect vectors. It is an important pathogen commonly found in domestic animals such as sheep and cattle. The organisms only grow in nucleated cells such as vascular ECs, monocytes and neutrophils and resemble viruses in their intracellular existence. The organism consists of a plasma membrane containing DNA, RNA, functioning ribosomes and enzymes participating in metabolic pathways, and reproduce by binary fission. Carrier-mediated transport systems in the plasma membrane enable the organism to absorb and directly utilise host cell nutrients and coenzymes. The organism lacks the glycolytic pathway and therefore oxidises glutamate and tricarboxylic acid cycle intermediates such as succinate as an energy source (Prescott *et al.*, 1993).

Different morphological forms of *C. ruminantium* are found. Small, medium-sized and large forms of *C. ruminantium* were identified in the vertebrate host and electron-dense and reticulated forms in ticks and cultured ECs of vertebrate host. The reticulate body (reticulate form) appears to be the predominant vegetative form and is the intracellular stage capable of multiplication within the host cell. The developmental form of *C. ruminantium*, the elementary body (electron dense form or initial body) is the extracellular stage capable of infecting host cells. After further subdivision and organisation it gives rise to colonies of organisms (Pienaar, 1970; Prozesky and du Plessis, 1985a; Prozesky *et al.*, 1986; Kocan *et al.*, 1987a; Prozesky and du Plessis, 1987c).

4.4 LIFE CYCLE

The life cycle of *C. ruminantium* has only recently become apparent and follows a *Chlamydia*-like developmental cycle in ECs (Fig. 7). The cycle starts with the entry of an electron-dense elementary body, the infectious stage of the organism, into the intracytoplasmic vacuole of an EC. The elementary body divides by binary fission to produce a large colony containing metabolically active organisms (reticulate bodies) of similar size and shape. After 5 to 6 days the cells disrupt and numerous elementary bodies are released to initiate a new infectious cycle. It is during this period that the organisms can be taken up by feeding ticks (Jongejan *et al.*, 1991d).

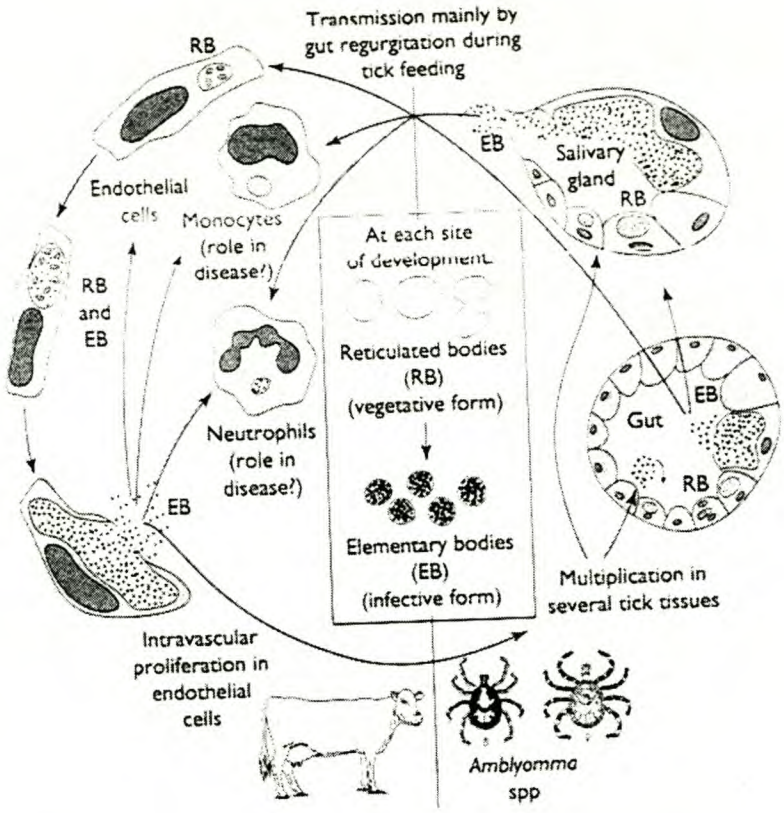


Fig. 7. The life cycle of *C. ruminantium* (Preston and Jongejan, 1999b).

C. ruminantium enters the host cell by inducing phagocytosis and escape the phagosome by an as yet unknown mechanism. *C. ruminantium* grows in membrane bound vacuoles within the host cell cytoplasm (Kocan and Bezuidenhout, 1987; Pienaar, 1970; Kocan *et al.*, 1987a; Uilenberg, 1983). Following infection, *C. ruminantium* initially replicates in the vacuoles of the cytoplasm of reticulo-endothelial cells and macrophages (Mϕ) in regional lymh nodes. The infective organisms are drained from the site of infection by the afferent lymphatic or phagosytosed by leukocytes. After multiplication in the leukocytes, the cells burst and the organisms are released into the general circulation to invade ECs of the blood vessels of various organs. This results in increased vascular permeability, which in turn results in the common postmortem finding of excessive fluid in the sac surrounding the heart, from which heartwater derives its name. The fluid also accumulates within the lungs, which appear "wet" and heavy, or within the chest cavity itself, outside the lungs. *C. ruminantium* multiplies in vacuoles in the cytoplasm of ECs by binary fission. Aside from incurring damage from cell lysis, the host cells are harmed by cell wall toxicity which appears to be related to the mechanism of penetration by the organism into host cells (Cowdry, 1926; Jackson and Neitz,

1932; du Plessis, 1970a; Pienaar, 1970; du Plessis, 1975; Uilenberg *et al.*, 1993; Prozesky and du Plessis, 1985a; Kocan *et al.*, 1987a; Prozesky *et al.*, 1986; Prozesky and du Plessis, 1987; Prescott *et al.*, 1993).

After recovery from heartwater, animals develop a long-lasting immunity to homologous challenge (partial or total lack of cross-protection between isolates has been demonstrated) and remain asymptomatic carriers for many months. Animals surviving the disease may have persistent infections; thus, they are capable of serving as reservoirs for transmission of *C. ruminantium* to ticks (du Plessis *et al.*, 1984a; Neitz and Alexander, 1941; Uilenberg, 1983; Andrew and Norval, 1989; Jongejan *et al.*, 1991b).

4.5 MOLECULAR BIOLOGY

The disease is poorly understood at the molecular level due to the fact that the causative agent is a fragile intracellular rickettsia. The molecular genetic analysis of *C. ruminatium* has been severely limited by three important factors. Firstly, by the difficulty of obtaining *C. ruminantium* DNA, free of host DNA contamination. Secondly, by the absence of a reliable cDNA library which could be used to identify and isolate genes of interest. Thirdly, by the lack of probes needed to isolate any such genes. However, significant advances have been made at molecular level with the *in vitro* propagation of the organism. The subsequent isolation of *Cowdria* DNA, free of contaminating host DNA, has led to the discovery of several immunogenic proteins. It has allowed the construction of several *Cowdria* genomic libraries, including the lambda ZAPII (λZAPII) library, and a map of the *Cowdria* genome (Fig. 8) (Bezuidenhout *et al.*, 1985b; Byrom and Yunker, 1990; Brayton *et al.*, 1997; Mahan, 1995a; de Villiers *et al.*, 2000).

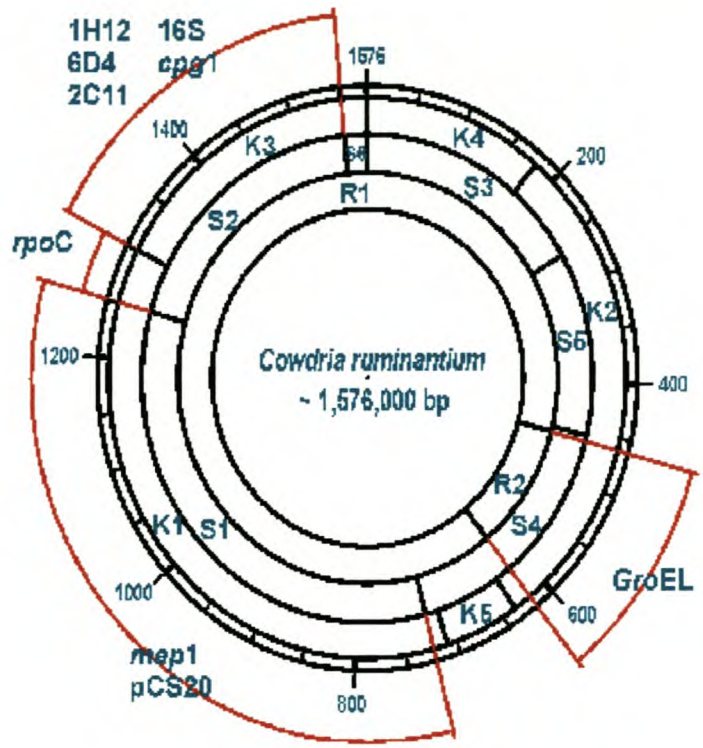


Fig. 8. The physical and genetic map of the genome of *C. ruminantium,* showing the positions of the genes *GroEL*, *map1*, *pCS20*, *rpoC*, *cpg1* and *16S* as well as the restriction sites of Ksp I and Sma I (de Villiers *et al.*, 2000).

It was found that *C. ruminantium* has an A-T (adenosine and thymine) rich genome, greater than 70 mol%, showing similarity to *Rickettsia prowazekii*. The small size of the genome 1.576 x 10^6 bp ± 143 bp (basepairs) is consistent with the organism's obligate intracellular lifestyles. Most of the immunogenic proteins, which have been identified are surface-exposed, making them important in the investigation of vaccine and diagnostic test development (Mahan *et al.*, 1992; van Vliet *et al.*, 1994; Waghela *et al.*, 1991; Anderson *et al.*, 1998; Brayton *et al.*, 1997; de Villiers *et al.*, 2000; Mahan, 1995a).

4.5.1 Major Antigenic Protein 1 / MAP1

The first protein examined was the 32 kDa (kilodalton) protein. This protein was shown to be the most immunodominant protein and subsequently named MAP1 (major antigenic protein 1). MAP1 is a major structural surface protein. The gene encoding the protein MAP1 is 854 bp long and has a 70 mol% A-T content. The gene shows homology with the gene encoding the MSP-4 protein (membrane spanning protein) of *Anaplasma marginale*, with the *Ehrlichia*

canis p28 gene and the *E. chaffeensis* omp1 gene family encoding the outer membrane protein-1. The *map1* gene is a member of a multigene superfamily (McBride *et al.*, 1999; Sulsona *et al.*, 1999).

The protein shows structural and size variation (31-33 kDa) when subjected to different electrophoretic conditions. This protein amino acid sequence is conserved within all tested *C. ruminantium* isolates originating from different geographic regions. MAP1 is antigenically conserved in nine strains of *C. ruminantium* from Africa and the Caribbean (Barbet *et al.*, 1994; van Vliet *et al.*, 1995; Jongejan *et al.*, 1991a; Jongejan and Thielemans, 1989; Rossouw *et al.*, 1990; van Vliet *et al.*, 1994).

The antigenic conservation within widely different geographical isolates has generated considerable interest in this gene as a candidate for development of serodiagnostic tools and vaccines although its size variation could affect the value of this protein as an immunogen (Barbet *et al.*, 1994; Reddy *et al.*, 1996; Jongejan *et al.*, 1993; Muller Kobold *et al.*, 1993; de Vries *et al.*, 1993; Mahan *et al.*, 1993; Mahan 1995a; Jongejan and Thielemans, 1989a).
Homologues of the *map1* gene are either dispersed throughout the genome or arranged tandemly. MAP1 is polymorphic and contains three highly immunogenic hypervariable regions. The variability in nucleotide sequences of MAP1 amongst four African and two Caribbean strains ranged from 0.6-14.0%. These nucleic acid differences translated to amino acid substitutions, deletions, or insertions at three hypervariable regions of the gene. (Reddy *et al.*, 1996; Coutinho *et al.*, 1991; Vachiéry *et al.*, 1998; Mahan *et al.*, 1994a)

4.5.2 Major Antigenic Protein 2 / MAP2

A second immunodominant protein, designated MAP2 (major antigenic protein 2) was identified and the gene encoding the protein was cloned from a pUC_{13} library. The gene is 672 bp long and has a 74 mol% A-T content. The 21 kDa protein is 55.5% identical to the MSP-5 protein of *A. marginale*. MAP2 contains a N-terminal signal peptide sequence, suggesting it could either be a secreted or a surface protein. Examination of the MAP2 of five geographically diverse strains of *C. ruminantium* revealed a highly conserved protein with only three amino acid substitutions (amino acid variability between 0-1.44%). This conservation makes it suitable for vaccines testing or diagnostic test development. MAP2 is an immunoreactive protein consistently recognised by antibodies present in sera obtained from sheep, goats and cattle (Mahan *et al.*, 1994a; Mahan, 1995a; Bowie *et al.*, 1999).

4.5.3 pCS20

pCS20 was one of the first genes of *C. ruminantium* to be cloned. It was cloned from a pUC_{19} library of the Crystal Springs (CS) strain of Zimbabwe. The pCS20 DNA probe is 1306 bp long. It has an A-T content of 70 mol% and contains two open reading frames (ORFs). This probe cross-reacts with different *C. ruminantium* strains and can detect *C. ruminantium* infection in infected *A. hebraeum* and *A. variegatum* ticks that have been fed on clinically ill animals. It is conserved among *C. ruminantium* isolates from Zimbabwe, South Africa, Nigeria, the Carribbean and Kenya. pCS20 is the most sensitive indicator of *Cowdria* genotypes and has superior specificity. It is currently being used for diagnostic purposes (Waghela *et al.*, 1991; Mahan *et al.*, 1992; Yunker, 1993; Peter *et al.*, 1995).

4.5.4 16S rRNA

The 16S rRNA coding region has been cloned from several different isolates of *Cowdria,* by polymerase chain reaction (PCR) amplifications. This gene has been used to elucidate the phylogenetic relationship of *Cowdria* to other *Rickettsiales*. Cross-reactions provided further evidence for the proposed close relationship between *Cowdria* and *Ehrlichia* species suggested by analysis of 16S ribosomal DNA sequences (Dame *et al.*, 1992; van Vliet *et al.*, 1992; Allsopp *et al.*, 1996).

4.5.5 Other genes

The *gro*E operon was cloned from a *C. ruminantium* (Welgevonden) λZAPII library. The *ftsZ* gene, a homolog of the *E. coli* gene encoding a cell division protein has been cloned and sequenced from the Senegal isolate. The *pCR9* gene was cloned from the *Cowdria ruminantium* pUC_{19} library (Lally *et al.*, 1995; Jongejan and Bekker, 1999; Waghela *et al.*, 1991).

4.6 DIAGNOSIS

Different *Cowdria* genotypes of differing immunogenecities exist in the field and detailed information on the distribution of these geno- and immunotypes is essential if any large-scale vaccination programme is to be carried out (Allsopp *et al.*, 1997; du Plessis *et al.*, 1989; Allsopp *et al.*, 1999).

Detection of *C. ruminantium* in *Amblyomma* ticks is essential for developing an understanding of the epidemiology of heartwater and in developing effective control measures. In addition, the detection and characterisation of heartwater-associated pathogens in local tick populations is an important guide to the threat posed to livestock in the areas concerned (Peter *et al.*, 1995; Allsopp *et al.*, 1999).

C. ruminantium is closely related to *Rickettsia*, *Anaplasma* and *Ehrlichia* species. This close phylogenetic relationship causes false positives in results, which are often seen with sera from areas that are totally heartwater-free, and where no *Amblyomma* ticks are found. This lack of specificity is caused by cross-reacting antigens, which results in false positive and negative reactions (Dame *et al.*, 1992; Byrom and Yunker, 1990; Jongejan *et al.*, 1988).

Cross-reactions have been demonstrated with anti-sera to various known *Ehrlichia* species. The identities of the agents responsible for the false positives are not always certain. There is a serious need for improved diagnostic tests in order to illuminate the complex epidemiological situation. Improved diagnostics using superior disease control strategies would allow detection of the disease in infected and carrier animals, which would in turn help reduce animal losses (du Plessis, 1981; Byrom and Yunker, 1990; Jongejan *et al.*, 1991b; Dame *et al.*, 1992; Martinez *et al.*, 1993; Mahan, 1995a).

Traditionally diagnosis has relied on microscopy of Giemsa-stained brain smears or histopathological sections, as *C. ruminantium* is not detectable by light microscopic examination of stained blood smears. This method is impractical for routine use and does not have the sensitivity required to detect the organisms in animals that have recovered but may be carriers of the disease (Cowdry, 1925; Jackson and Neitz, 1932; Purchase, 1945; Uilenberg, 1983; Prozesky, 1987b; Andrew and Norval, 1989). The use of light-, electron- or fluorescence microscopy of tick tissues, or enzyme-linked immunosorbent assay (ELISA) techniques to detect infections are cumbersome, time consuming, have unknown sensitivities and specificities and are therefore unsuitable for routine screening of tick infections (Andrew and Norval, 1989; Kocan and Bezuidenhout, 1987; Kocan *et al.*, 1987a; Kocan *et al.*, 1987b; Yunker *et al.*, 1987; Neitz *et al.*, 1986). Detection of *C. ruminantium* has been achieved by inoculation of tick extract into, or by feeding ticks on susceptible small ruminants or mice. However, this method is laborious and expensive (Andrew and Norval, 1989; Birnie *et al.*, 1985; Camus and Barré, 1987; Camus and Barré, 1992; Norval *et al.*, 1990; du Plessis, 1985).

4.6.1 Serological Tests

Serological tests that have been developed for heartwater are based on detection of antibodies by immunofluorescence or by ELISAs with *in vitro*-cultured *C. ruminantium* organisms. Two types of serological tests have been developed: (i) Indirect fluorescent-antibody test (IFAT) which detect antibodies against the whole organism and (ii) immunoblotting and ELISAs which detect antibodies produced against defined proteins (du Plessis and Malan, 1987d; Jongejan *et al.*, 1989b; Logan *et al.*, 1986; Semu *et al.*, 1992; Jongejan and Thielemans, 1989; Rossouw *et al.*, 1990; Jongejan *et al.*, 1991a; Martinez *et al.*, 1993).

IFATs have poor specificity as the antibodies cross-reacts with *Ehrlichia phagocytophila* (Jongejan *et al.*, 1989b; Logan *et al.*, 1987; Logan *et al.*, 1986; du Plessis *et al.*, 1987c).

Various serological ELISAs have been developed for *C. ruminantium* diagnosis but these were hampered by the close relationship between this organism and other *Ehrlichia* species, resulting in cross-reactivity of antibodies (Jongejan *et al.*, 1991a; du Plessis *et al.*, 1993a; Jongejan *et al.*, 1993; du Plessis *et al.*, 1987c; du Plessis and Malan, 1987b; Logan *et al.*, 1986; Mahan *et al.*, 1993).

IFATs, ELISAs and immunoblotting have neither the sensitivity nor the specificity for use as diagnostic tools. Immunoblotting seems to be the most sensitive of the three techniques (Mahan *et al.*, 1993; Jongejan *et al.,* 1991a). Current serological tests are therefore of limited reliability and cannot be used to detect the prevalence of heartwater-causative organisms in ticks (Allsopp *et al.*, 1999).

4.6.1.1 Major Antigenic Protein 1 / MAP1

A competitive ELISA (cELISA) as well as an immunoblotting technique has been developed. These techniques are based on the *C. ruminantium* antigen derived from EC cultures and mediated by monoclonal antibodies against MAP1 (Jongejan and Thielemans, 1989; Jongejan *et al.*, 1991a; Mahan *et al.*, 1993).

It was proposed that MAP1 may be a good diagnostic antigen because it is immunodominant and antigenically conserved in *C. ruminantium* isolates. This protein was used in a specific cELISA and immunoblotting assays because of its immunodominance. However, the MAP1 protein targets a variable *Cowdria* antigen. As a result, serological tests based on the detection of antibodies directed against the MAP1 protein often have low specificities. Additionally, the MAP1 protein-based immunoblotting assay detects false positives due to serological cross-

reactions of *C. ruminantium* with antigenically related organisms such as *Ehrlichia* spp. (Mahan *et al.*, 1993; Jongejan *et al.*, 1993b; de Vries *et al.*, 1993; du Plessis *et al.*, 1993a; Jongejan and Thielemans, 1989; Rossouw *et al.*, 1990). Demonstration of homology in the MAP1 coding sequence between *C. ruminantium* and *Ehrlichia* spp. and divergence in hypervariable regions between strains suggest that MAP1 may not be a suitable candidate for specific serodiagnosis of heartwater. Thus current serological tests for cowdriosis using MAP1 are unsatisfactory as the MAP1 antibody lacks specificity, i.e. false positive sera react with the whole *C. ruminantium* organism and the reported immunoreactive MAP1 (du Plessis *et al.*, 1987c).

4.6.1.2 Major Antigenic Protein 1-B / MAP1-B

Complete, truncated or recombinant MAP1 reacts with sera from ruminants from heartwater-free areas and with several members of the genus *Ehrlichia* (Katz *et al.*, 1996; Katz *et al.*, 1997; Mahan *et al.*, 1993; Jongejan *et al.*, 1993). The gene encoding MAP1 has been cloned and expressed in *Escherichia coli*, which allowed the identification of the immunogenic regions on MAP1 and their evaluation of their possible use in the construction of a specific serological test for cowdriosis. Two immunogenic regions designated MAP1-A and MAP1-B were identified. Epitopes conserved between *C. ruminantium* and *Ehrlichia* spp. were found on MAP1-A. MAP1-B contained one or more *C. ruminantium* specific epitopes. MAP1-B is conserved between isolates and invoked strong antibody responses, making it ideal for use in serological tests. Recombinant MAP1-B antigen was recognised by antibodies from nine isolates of *C. ruminantium* originating from different geographic regions. The recombinant MAP1-B based ELISA has significantly improved the efficiency of the test in detecting *C. ruminantium*-specific antibodies compared with previously described serological tests. Although MAP1-B appeared to be specific to *C. ruminantium*, the MAP1-B based assay failed to eliminate all false-positive reactions (van Vliet *et al.*, 1994; van Vliet *et al.*, 1995; Reddy *et al.*, 1996; Katz *et al.*, 1996; Katz *et al.*, 1997). It nevertheless has potential as a tool in the study of the epidemiology of heartwater in Africa and the Caribbean. An indirect ELISA based on MAP1-B has been developed with significantly reduced cross-reactivity (van Vliet *et al.*, 1995; Reddy *et al.*, 1996).

4.6.1.3 Major Antigenic Protein 2 / MAP2

Although MAP2 meets the criteria i.e. immunoreactivity and conservation, for a suitable diagnostic antigen, research has shown MAP2 to be unsuitable for diagnostic tests due to a lack of specificity. Data showed reactivity between recombinant MAP2 and sera from known heartwater-free animals (Mahan *et al.*, 1994a; Mahan, 1995a; Bowie *et al.*, 1999).

4.6.2 Nucleic based assays

The homology at gene level and translation to antigenically conserved proteins between the genus *Cowdria*, *Anaplasma* and *Ehrlichia*, influences the design of specific diagnostic tests for heartwater, anaplasmosis and ehrlichiosis. Nucleic acid-based assays would therefore have an application in the epidemiological studies of heartwater (Mahan *et al.*, 1995a; Mahan *et al.*, 1992). The isolation of *C. ruminantium* genes encoding immunoreactive proteins may enable the development of diagnostic tests with improved sensitivity and specificity. Such an immunoreactive protein must, however be conserved among *C. ruminantium* strains and sufficiently different from the homologues present in related species (Bowie *et al.*, 1999).

Nucleic based assays using DNA and RNA probes are currently being used for the detection of *C. ruminantium*. DNA probes detect *C. ruminantium*-infected ticks irrespective of the stage at which infection occurred. The sensitivity of these assays are maximised by using PCR to amplify DNA sequences, which are specific to an organism. PCR also permits the diagnosis of persisting infection after recovery (Mullis and Faloona, 1987; Saiki *et al.*, 1988; Waghela *et al.*, 1991; Mahan *et al.*, 1992; Mahan, 1995a).

4.6.2.1 DNA probes

The most widely used DNA probes are based on three gene families, namely the *pCS20, 16S*, and the *map1* gene (Waghela *et al.*, 1991; Brayton *et al.*, 1997).

4.6.2.1.1 pCS20

The pCS20 probe is sensitive and specific for *Cowdria* DNA but does not distinguish between pathogenic and non-pathogenic genotypes. It cross-reacts with *C. ruminantium* strains from Zimbabwe, South Africa, Nigeria, the Caribbean and Kenya. It shows no cross-hybridisation with DNA from *Ehrlichia* spp. This probe is an ideal diagnostic tool of superior specificity and can be randomly labelled to yield a probe with high specific activity. It is therefore the

probe of choice for the initial screening of *Cowdria* of large numbers of samples (Allsopp *et al.*, 1999).

The pCS20 probe can detect low levels of *C. ruminantium* in infected *A. hebraeum* and *A. variegatum* ticks that have been fed on clinically ill animals and is therefore suitable for both prospective and retrospective studies which require detection of *C. ruminantium* in ticks. This is especially important for epidemiological studies to determine field tick infection rates. The pCS20 probe detects *C. ruminantium* during the febrile reaction, which is the first clinical sign of disease. This was the first method to diagnose heartwater at this early phase. However, this probe gives no phylogenetic information. The sensitivity of this test has been improved by amplification of *C. ruminantium*-specific pCS20 sequences by PCR (Yunker *et al.*, 1993; Waghela *et al.*, 1991; Uilenberg *et al.*, 1983; Mahan *et al.*, 1992; Peter *et al.*, 1995; Peter *et al.*, 2000; Allsopp *et al.*, 1999; Mahan *et al.*, 1993).

4.6.2.1.2 16S

The *16S rRNA* gene has been used to develop probes for the study of the epidemiology of heartwater. These probes detect the different *C. ruminantium* genotypes or isolates, and any Group II and Group III *Ehrlichia* species other than *C. ruminantium* (Allsopp *et al.*, 1996; Allsopp *et al.*, 1997; Dame *et al.*, 1992; van Vliet *et al.*, 1992; Allsopp *et al.*, 1999).

The 16S oligonucleotide probes are technically demanding to use because, among the different *Cowdria* genotypes, there are few nucleotide sequence differences in the most variable region (V1 loop) of the rRNA gene. The necessity for end labelling means that the probes carry less label than a random-labelled probe. The hybridisation signal strength may therefore be low. The 16S probe is the only probe, which gives any phylogenetic information (Allsopp *et al.*, 1999).

4.6.2.1.3 map1

The *map1* gene exhibits extensive sequence polymorphism between isolates and it may therefore be possible to design map1 probes which will provide immunotypic information about the distribution of *Cowdria* genotypes in any given area (Kock *et al.*, 1995; Reddy *et al.*, 1996; Allsopp *et al.*, 1999).

The *map1* gene can be labelled to high specific activity. However, its detection of false positives in sera from animals in heartwater-free areas seriously limits its use in serodiagnosis. The map1 probe is less sensitive than the pCS20 probe and does not provide any genotype information (Mahan *et al.*, 1993).

4.6.2.2 PCR based assays

PCR based assays allow the detection of early stage infections that are below the detection limit of the pCS20 DNA probe. PCR assays are highly specific and sensitive. These assays are specific for *C. ruminantium* and do not detect other hemoparasitic tick transmitted parasites of livestock, or DNA from phylogenetically closely related serologically cross-reactive *Ehrlichia canis* organisms. PCR assays detect organisms in infected sheep before the febrile reaction (Uilenberg, 1983; Mahan *et al.*, 1992; van Vliet *et al.*, 1992; Kelly *et al.*, 1994; Peter *et al.*, 1995).

4.6.2.2.1 pCS20

The pCS20 PCR assay is currently the most reliable and best-characterised test for *C. ruminantium* infection in ticks. The absence of cross-reaction with closely related organisms such as *Ehrlichia chaffeensis* and *E. canis* increases the value of this test, particularly as *C. ruminantium* serological assays developed to date are limited by either poor specificity or low sensitivity. The pCS20 PCR assay is more sensitive than *C. ruminantium* PCR assays based on the 16S rRNA and the *map1* gene (Mahan *et al.*, 1998b; Peter *et al.*, 2000; Peter *et al.*, 1995; Bowie *et al.*, 1999; du Plessis *et al.*, 1993; Jongejan *et al.*, 1993, Katz *et al.*, 1996; Katz *et al.*, 1997; Mahan *et al.*, 1998c; Mahan *et al.*, 1993; van Vliet *et al.*, 1995; Allsopp *et al.*, 1999).

The ability of the pCS20 PCR assay to detect DNA of *C. ruminantium* originating from locales throughout the distribution of heartwater and in field ticks collected from diverse sites in southern Africa demonstrates conservation of the primer sequences and the wide applicability of the assay (Peter *et al.*, 2000).

PCR amplifications together with pCS20 DNA probe hybridisation can be used to diagnose heartwater in animals and ticks as well as in facilitating the understanding of the epidemiology of heartwater, which could result in improved future control strategies (Mahan *et al.*, 1992).

4.6.2.2.2 16S

The *16S* ribosomal RNA gene was amplified by PCR and cloned from several different isolates of *C. ruminantium* and subsequently used to elucidate the phylogenetic relationship of *C. ruminantium* to other *Rickettsiales* (Dame *et al.*, 1992; van Vliet *et al.*, 1992; Allsopp *et al.*, 1996; Allsopp *et al.*, 1997; Allsopp *et al.*, 1999)

4.6.2.2.3 map1

The *map1* gene has been used to design primers complementary to the gene sequence, for PCR amplifications in the detection of low level *C. ruminantium* infection in carrier animals and ticks. This gene has been used successfully to differentiate between different isolates to improve knowledge of the epidemiology of the disease (Kock *et al.*, 1995; Reddy *et al.*, 1996).

4.7 IMMUNITY

Immunity against *C. ruminantium* is poorly understood because of its obligate intracellular nature, making it difficult to culture the organism *in vitro.* Since it was discovered that some strains are pathogenic to mice, a laboratory model has been developed. Understanding the basis of protective immunity to *C. ruminantium* will facilitate the development of an effective subunit vaccine against heartwater and will contribute to a better definition of protective immune mechanisms to obligate intracellular pathogens in general (du Plessis and Kümm, 1971; Brayton *et al.*, 1998; Totté *et al.*, 1999).

4.7.1 Humoral vs. Cellular Immunity

C. ruminantium survives within vascular ECs of infected animals and the agent is thus hidden from the humoral immune response when it is most pathogenic (Clark, 1962; du Plessis, 1975b; Prozesky, 1987b).

After inoculation with *C. ruminantium,* antibody responses are detected in cattle at the height of the febrile reaction. These responses are induced by the rickettsaemia that follows rupture of infected cells. Infection with *Cowdria* elicits antibody responses in ruminants that recover from the disease and are also detected in cattle after inoculation with *C. ruminantium*. There is no apparent correlation between antibody titre as measured by IFATs and the immune status of the animal and no evidence that antibodies influence the course of infection. Therefore

antibody titres are unlikely to be a measure of protective immunity, even though murine and bovine antisera are capable of neutralising *C. ruminantium* infection of bovine ECs *in vitro*. In some instances, hyperimmune sera were shown to neutralise the infection *in vitro*, whereas in others no significant effect was observed. Transfer of immune serum or gamma globulin failed to protect animals or even modify the course of the disease. Although these results are not sufficient to exclude the existence of protective antibodies, they underline the limitations of using immune sera to study protective mechanisms (Martinez *et al.*, 1994; Byrom *et al.*, 1993; du Plessis, 1970b).

Cross-protection studies have shown the existence of fully, non- or partially cross-protective isolates, indicating that antigens responsible for protection are polymorphic. The inaccessibility of the organism to serum antibody during much of the infection period suggests that humoral responses play a small role in protection (Totté *et al.*, 1999). Together with the intracellular location of the organism, these observations have led to the belief that a cell-mediated immune response are expected to play a pivotal role in protection (Semu *et al.*, 1992; Alexander, 1931; du Plessis, 1970b; du Plessis *et al.*, 1984a; du Plessis *et al.*, 1984b; Musoke *et al.*, 1996; Musoke *et al.*, 1997; McKeever, 1993).

Protective immunity against *C. ruminantium* is therefore mainly cell-mediated and requires the activation of thymus-derived lymphocyte cell (T cell) responses. Humoral immunity on its own does not appear to play a major role although it might not be completely irrelevant. Antibodies may have a potential role in opsonisation, complement-mediated killing and antibody-dependent cell-mediated cytotoxicity (Byrom *et al.*, 1993; du Plessis *et al.*, 1991; Uilenberg *et al.*, 1993; Totté *et al.*, 1999; du Plessis, 1970b; Totté *et al.*, 1997; Mahan *et al.*, 1996; Mahan *et al.*, 1994b).

4.7.2 The role of Cytokines

Protection against *C. ruminantium* involves immune T cells, mainly T helper (T_H) cells that release cytokines. The mechanism by which these cytokines cause inhibition of *C. ruminantium* growth is twofold. Early multiplication of the organism is inhibited and lysis of intracellular *Cowdria* organisms or specific lysis of infected cells is induced. T cell cytokines inhibit the progression of *C. ruminantium* infections in bovine ECs *in vitro*. Thus cytokines, if induced during heartwater infections, would be expected to have an antagonistic effect on *C.*

ruminantium multiplication and hence affect the severity of the disease (du Plessis *et al.*, 1987a).

Cytokines are important mediators of protective and pathological immune responses. Studies suggest a highly complex response with several factors exerting a variety of effects that can result in resistance but also exacerbate disease (Preston and Jongejan, 1999a). An individual cytokine is able to stimulate the production of many others, generating a network that orchestrates the host's immune response to infection. Paradoxically, the uncontrolled or excessive production of cytokines could contribute to the pathophysiology of heartwater thereby exacerbating the disease (Preston and Jongejan, 1999a, Totté *et al.*, 1999).

4.7.3 Interferon – The major role player

The most important cytokine involved is interferon. Interferons play a major role in protection against intracellular pathogens, including *Cowdria*-related organisms such as *Rickettsia* and *Chlamydia*. The interferon system also plays a key role in natural resistance to *C. ruminantium* infections (Totté *et al.*, 1999; Mahan *et al.*, 1999).

IFN-γ is an important mediator of immunity against other rickettsial agents, intracellular bacteria and protozoa and is central to the induction of cell-mediated immune response pathways (Byrne and Turco, 1988; Sher and Coffman, 1992; Totté *et al.*, 1993a; Mahan *et al.*, 1996). IFN-γ production by $CD4^+$ cells, γα T cells and possibly $CD8^+$ cells, natural killer (NK) cells and Mϕ may play a role as there are at least four pathways (Fig. 9) through which it could contribute towards an immune response: (i) IFN-γ leads to the upregulation of the major histocompatibility complex (MHC) class I and class II. The MHCs subsequently presents the *Cowdria* antigens and this presentation leads to the generation of memory T cells. (ii) IFN-γ production activates phagocytes, which leads to increased phagocytosis and the release of nitric oxide (NO) and myeloperoxidase (MPO). (iii) IFN-γ directly upregulates the production of NO, which subsequently causes the direct killing of extracellular *Cowdria*. (iv) IFN-γ leads to the killing of *Cowdria* ECs through an as yet unknown mechanism.

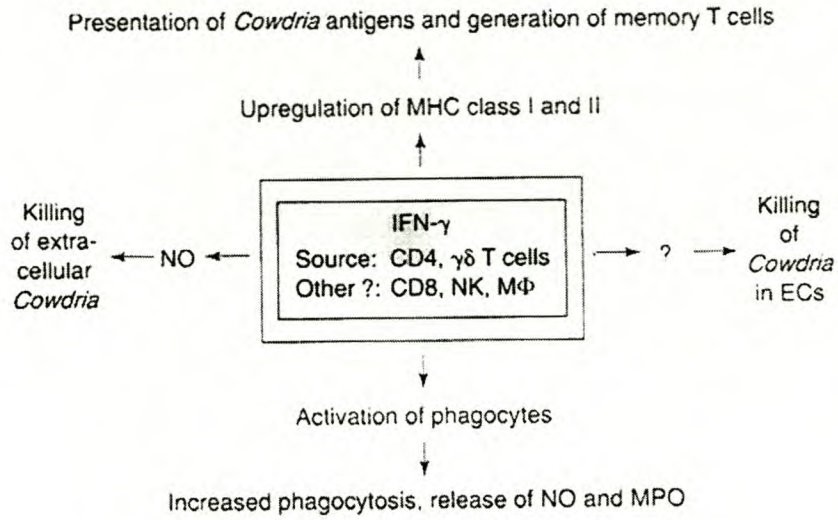


Fig. 9. Putative roles of IFN-γ in protection against *C. ruminantium* (Totté *et al.*, 1999).

A direct effect on bovine and caprine ECs that renders these cells unsuitable for *C. ruminantium* growth was demonstrated for recombinant and endogenous IFN-γ produced by concanavalin A-stimulated T cells. Research showed IFN-γ completely abrogated the growth of *C. ruminantium* regardless of the strains and EC lines used (Totté *et al.*, 1993; Totté *et al.*, 1996; Mahan *et al.*, 1994b; Mahan *et al.*, 1996;).

Although IFN-γ is a potent inhibitor of *C. ruminantium in vitro*, it can be highly cytotoxic for infected and uninfected ECs. A net positive or negative effect of this cytokine most likely depends on the level and the timing of its production during infection. Early induction of IFN-γ can be expected to control the infection, whereas in late-stage clinical disease, IFN-γ could have detrimental consequences (Totté *et al.*, 1999). Antigens such as MAP1 and MAP2, capable of inducing the production of T helper cell type 1 (T_{H1}) cytokines, including IFN-α (interferon alpha) and IFN-γ, by immunocompetent cells, have clear potential in the development of a recombinant vaccine against heartwater (Mahan *et al.*, 1999). Cattle that resisted a lethal dose challenge of *C. ruminantium* were shown to produce IFN-α whereas animals that died did not. However, even at high concentrations recombinant IFN-α could not prevent the growth of *C. ruminantium in vitro* in bovine ECs suggesting that other factors were involved in protection (Totté *et al.*, 1994).

It is obvious that IFN-γ and IFN-α can prevent the development of *C. ruminantium* in susceptible animals, as both IFN-α and IFN-γ has an inhibitory effect on *C. ruminantium*. The cytokine response is highly complex with several factors exerting a variety of effects that can result in resistance to *C. ruminantium* infection (Totté *et al.*, 1993b) as IFN-γ lyse rickettsial agents or cells infected with them. This mechanism may also be responsible for the inhibition of *C. ruminantium* growth by IFN-γ *in vitro* (Hanson, 1991; Mahan *et al.*, 1994b).

It can be concluded that protection against heartwater after immunisation by infection and subsequent treatment or after spontaneous recovery, is mediated at least in part, by generation of *C. ruminantium*-specific T cell responses resulting in the secretion of IFN-γ (Mwangi *et al.*, 1998a). IFN-γ has been shown to inhibit *C. ruminantium* growth *in vitro* and its administration in mice can prevent death against challenge that is lethal for untreated control mice. Hence, IFN-γ is likely to be an important mediator of immunity for controlling infection in immunised animals (Mahan *et al.*, 1994b; Mahan *et al.*, 1996; Totté *et al.*, 1996; Totté *et al.*, 1994b; Totté *et al.*, 1997). New vaccine strategies for heartwater should therefore focus on immune responses that enhance production of *Cowdria*-inhibitory cytokines such as IFN-γ (Mahan *et al.*, 1996; Totté *et al.*, 1996).

4.7.3.1 IFN and Endothelial cells

Very little direct lysis or cytopathic changes are observed in infected ECs of moribund animals and growing evidence suggests that immune effectors, such as cytokines, play a role in the pathology (Totté *et al.*, 1999). ECs have been shown to be capable of presenting antigens to T cells in other systems (Vora *et al.*, 1994). *C. ruminantium* infection of bovine ECs induces cytokine production. *C. ruminantium* infection in bovine brain microvessel ECs (Fig. 10) leads to the *de novo* synthesis of IL mRNA especially IL-1β, IL-6 and IL-8, which is potentiated by IFN-γ (Bourdoulous *et al.*, 1995; Totté *et al.*, 1999). IL-1 and IL-6 can act as co-stimulatory signals for T and B cell activation. Therefore brain ECs, which constitute one of the main targets of *C. ruminantium* in ruminants, could contribute to the development of a protective immune response against the pathogen (Weaver and Unanue, 1990; Totté *et al.*, 1999). This infection strongly inhibits IFN-γ induced MHC class II and constitutive MHC class I surface expression. This effect can be partly attributed to a disruption in the intracellular trafficking of MHC molecules as shown by permeabilisation studies. Finally, under certain conditions, *C. ruminantium* seems to be able to induce the production of nitric oxide (NO) by infected ECs (Totté *et al.*, 1999).

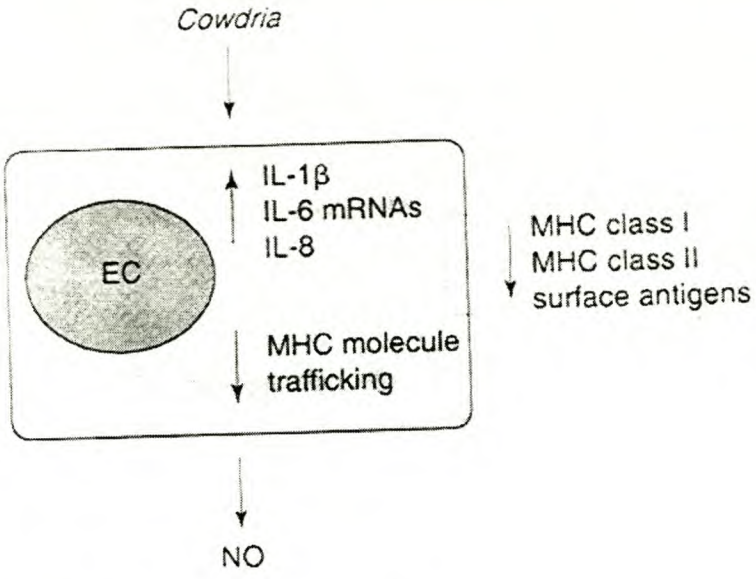


Fig. 10. Pleiotropic effects of *C. ruminantium* on ECs

Immunisation of cattle against heartwater by infection and subsequent treatment results in the generation of T lymphocytes that recognise infected ECs and monocytes. Recognition of infected ECs is dependent on induction of MHC class II expression by treatment with T cell growth factors (TCGF) before infection and fixation with glutaraldehyde before inclusion in proliferative assays (Mwangi *et al.*, 1998a).

4.7.3.2 IFN and Nitric oxide

NO plays a role in resistance to *C. ruminantium* infections and in the pathology of the disease. NO, a powerful vasodilator synthesised in the endothelium, is intimately related to mechanisms of oedema formation and hyperaemia (Furchgott and Zawadski, 1980; Moncada *et al.*, 1988; Moncada *et al.*, 1990). *C. ruminantium* organisms are killed by events involving IFN-γ and NO *in vitro,* as NO is both microbicidal and cytotoxic. IFNs may play an important role in resistance to intracellular viruses and rickettsias such as *C. ruminantium* possibly by elevating NO production as NO produced endogenously or from exogenous sources, kill *C. ruminantium* in a dose-dependent fashion (Adams *et al.*, 1990; Adams *et al.*, 1991; Totté *et al.*, 1993a; Booccvar *et al.*, 1994; Radi *et al.*, 1994; Mutunga *et al.*, 1998).

NO released by ECs or by other cells such as monocytes may play a role in reducing the infectivity of extracellular elementary bodies and thereby reducing the spread of infection. NO is capable of reducing both the viability and infectivity of *C. ruminantium* in bovine pulmonary ECs *in vitro* (Mutunga *et al.*, 1998).

Innate immune mechanisms, involving Mϕ and their products (in particular NO) and in some cases NK cells, play a significant role in recovery from infection and protection against challenge with *C. ruminantium.* NK cells and Mϕ produce cytokines and thus play a crucial role in innate immunity to intracellular pathogens (Preston and Jongejan, 1999a; Totté *et al.*, 1999).

NO-dependent non-specific immunity exists involving the reticulo-endothelial system. This immunity is a primary defence mechanism against tumor cells and several microbial pathogens and a potent agent in non-specific defence mechanisms by up-regulating the release of inflammatory mediators (Liew *et al.*, 1991; Woodman *et al.*, 1991; Moncada and Palmer, 1993). However, NO precipitates the loss of intracellular iron resulting in the inhibition of certain crucial enzymes involved in mitochondrial respiration. This suggests that NO may contribute to the mitochondrial changes observed in the alveolar ECs infected with *C. ruminantium* (Prozesky and du Plessis, 1985b; Busse *et al.*, 1995).

The induction of NO synthesis by IFN-γ may be an important determinant of the outcome of *C. ruminantium* infection. This is supported by the following sequence of events: rickettsemia occurs 3-6 days before the onset of fever and pathology, but terminal collapse, reduced blood pressure and peak pulmonary oedema occur a few days after the first immune response are detectable and when rickettsemia has declined but EC infection is highest (Mutunga *et al.*, 1998).

IFN-γ has a profound effect on monocytes and neutrophils, including increased phagocytosis and release of reactive oxygen intermediates (NO) and lysosomal enzymes (MPO). The release of MPO is caused by IFN-γ in combination with other cytokines such as tumour necrosis factor alpha (TNF-α). MPO might be involved in protection through a direct toxic effect on the extracellular form of *Cowdria.* MPO might also stimulate monocytes to produce IFN-α (Lefkowitz *et al.*, 1993; Lieser *et al.*, 1995; Goff *et al.*, 1996).

4.7.3.3 IFN and Major Histocompatibility Complex

Another important role for IFN-γ is the up-regulation of MHC class I, and MHC class II expression on a wide variety of cells including ECs (Coutinho *et al.*, 1991). Both IFN-γ induced MHC class II and constitutive MHC class I surface expression is strongly inhibited upon infection of ECs with *Cowdria* in a dose-dependent manner (see 4.7.3.1 Fig. 10). In

addition inhibition of synthesis appears to occur at high infection rates (Byrne and Turco, 1988; Coutinho *et al.*, 1991; Totté *et al.*, 1993b; Totté *et al.*, 1994; Totté *et al.*, 1996; Totté *et al.*, 1997; Mwangi *et al.*, 1998a; Vachiéry *et al.*, 1998).

Cowdria-infected ECs induce MHC class II-restricted proliferation of peripheral blood mononuclear cells (PBMCs) from cattle immunised by infection and subsequent treatment, provided they were treated with TCGF to up-regulate MHC class II expression. Monocytes have also been shown to present *Cowdria* antigens in association with MHC class II molecules (Totté *et al.*, 1997; Mwangi *et al.*, 1998a), as extracellular *C. ruminantium* are found in the blood and are available for phagocytosis by monocytes. Thus, up-regulation of MHC class II expression by IFN-γ on monocytes is likely to favour presentation of *Cowdria* antigens to the immune system (Totté *et al.*, 1999).

4.7.3.4 IFN and $CD4^+$ and γα T cells

Antigen processing and presentation by infected cells may be essential for the induction of T cell responses to the agent during infection. These responses are partly mediated by class II MHC-restricted $CD4^+$ T cells. $CD4^+$ T cells are considered to be prominent in the regulation of cellular immune responses to other intracellular pathogens (Kaufmann, 1993). $CD4^+$ T cells stimulate antibody production, proliferation of CTLs, activation of Mϕ and the production of cytokines, which are capable of directly inhibiting the growth of the pathogen (Preston and Jongejan, 1999a).

The major cellular immune responses of cattle to *C. ruminantium* infection are in the $CD4^+$ and γδ T cell compartments. Both αβ (mainly $CD4^+$) and γδ T cells are induced in cattle undergoing *C. ruminantium* infection. γδ T cells represent a more primitive T cell population that acts as a first line of defence against certain pathogens (Mwangi *et al.*, 1998a; Janeway *et al.*, 1993). $CD4^+$ T cells and γδ cells from immune animals proliferate *in vitro* in response to *C. ruminantium* -infected ECs and monocytes, producing IFN-α and IFN-γ. A marked rise in $CD8^+$ T cells that control the infection occurs *in vivo*. It is possible that these cells play a role through the killing of infected cells (Preston and Jongejan, 1999a).

Immunisation with different types of *C. ruminantium* material leads to the development of different immune responses. (i) PBMCs obtained from cattle immunised with live *C. ruminantium* showed a MHC class II-restricted proliferation. This proliferation is specific and

in response to stimulation by autologous infected ECs or monocytes and leads to the secretion of IFN-γ. Thus during infection with live virulent *C. ruminantium*, T cell responses might be preferentially directed at certain epitopes expressed by infected cells but absent from the extracellular form of the organism. γα T cells are induced *in vitro* in cattle undergoing immunisation with live *C. ruminantium*. These cells proliferate in response to autologous and heterologous infected ECs and monocytes (thus in a MHC- unrestricted manner) and produce IFN-γ (Totté *et al.*, 1997; Mwangi *et al.*, 1998a; Totté *et al.*, 1999). MAP1 induces the proliferation of $CD4^+$ T cells. $CD4^+$ and γα T cells expressed high levels of IFN-γ. These responses are fundamental to the control of intracellular pathogens (Mwangi *et al.*, 1998a; Mahan *et al.*, 1999). (ii) Immunisation with killed *C. ruminantium* leads to the production of antibodies as well as $CD4^+$ T cells, which are preferentially directed against other immunogenic proteins such as MAP1 and MAP1-B. $CD4^+$ T cells proliferate and produce IFN-γ in response to *Cowdria* lysates and to *C. ruminantium*-primed autologous monocytes. MAP1 and MAP1-B repeatedly induced γα cells to proliferate (Mwangi *et al.*, 1998b; Totté *et al.*, 1998; Totté *et al.*, 1999). (iii) T_H cells are capable of producing IFN-γ in response to Cowdria antigens derived from both killed and live *C. ruminantium*. In addition, γα T cells specific for Cowdria- infected ECs and for MAP1 and MAP2 are also induced during immunisation with both live and killed *C. ruminantium* and have been shown to produce T_{H1} cytokines (Totté *et al.*, 1999). MAP2 induces proliferation of γα T cells *in vitro*. Upon stimulation, MAP2-specific γα T cell lines strongly express IFN-γ and IFN-α mRNA, but express little IL-2 and no IL-4 or IL-10. This is consistent with a T_{H1} response. (iv) A $CD4^+$ T cell response was induced in cattle vaccinated with inactivated *C. ruminantium* in Freund's complete adjuvant and these cells proliferate when stimulated with lysates of *C. ruminantium* and produced IFN-γ, which has been shown to be inhibitory for growth of *C. ruminantium* (Mahan *et al.*, 1994b; Mahan *et al.*, 1996; Totté *et al.*, 1996).

CHAPTER FIVE
HEARTWATER – THE STUDY

5.1 AIM OF THE STUDY

Current methods employed in the control of heartwater are inadequate and there is an urgent need for a cheaper, safer and easier way to control this disease. The development of recombinant vaccines against and new and better diagnostic tools for *C. ruminantium,* will require a detailed knowledge of the genes responsible for protective immunogenicity and of genes suitable for characterising the organism. To achieve these goals, a thorough understanding of the organism at the molecular level is needed. This led to the formation of a consortium, which is engaged in a project to sequence the entire genome of the Welgevonden isolate of *C. ruminantium*[3]. The project is coordinated from the Onderstepoort Veterinary Institute (ARC-OVI). The consortium includes the University of Stellenbosch, the Centre de Coopération Internationale en Recherche Agronomique pour le Développement – Département Elevage et Médecine Vétérinaire des pays tropicaux (CIRAD-EMVT) and Utrecht University.

This study was focussed on the sequencing of a lambda GEM-11 clone from the lambda GEM-11 library of the Welgevonden isolate of *C. ruminantium* which could have lead to the identification of genes that encode for proteins for use in effective vaccines.

5.2 STUDY HISTORY

Onderstepoort Veterinary Institute (OVI) has been at the centre of heartwater research since its discovery in the 1800's. Sir Arnold Theiler, founder of the OVI, was one of the first veterinarians to diagnose this disease. In 1985 scientists at the OVI reported a culture system for *C. ruminantium*. *C. ruminantium* organisms were cultivated *in vitro* in a calf EC line, which made it possible to study the organism intensively. In recent years the Molecular Biology division at the OVI has been at the forefront of developing an effective vaccine against heartwater (Bezuidenhout *et al.*, 1985).

[3] http://www.arc.agric.za/cowdriagenome.htm

In order to identify potential genes for a recombinant *C. ruminantium* vaccine, Brayton *et al.*, (1998), reported that 14% of outbred mice were protected against a lethal *C. ruminantium* challenge after immunisation with a *Salmonella* vaccine delivery system containing cloned mini-libraries of *C. ruminantium*. Due to the shortcomings of the mouse model system the results were inconclusive.

5.3 DNA LIBRARIES

The intracellular location of *C. ruminantium* makes it difficult to separate the organism from host cell components. It is therefore difficult to obtain *C. ruminantium* DNA uncontaminated by host cell DNA. Few *C. ruminantium* libraries are available and consequently few genes have been cloned. The value of a genomic library is that it collectively represents the whole genome from which any gene of interest can be identified and isolated for further characterisation. Genome sequencing of an organism also requires a good representative genomic library.

AT rich genomes are difficult to clone into cosmids and are more stable in phage systems or in a system which maintains the cloned DNA at a low copy number. An additional problem when cloning DNA of bacterial origin is that the cloned promoters may be active in the host cell leading to expression of the foreign protein, which may be toxic to the host cell (Reddy, 1995; Brayton *et al.*, 1997). *Cowdria* DNA is unstable in plasmid, cosmid and BAC vectors due to its high AT content (70%) as inserts could form AT stem-loops, which are excised or rearranged. In addition, *Cowdria* promoters are active in *Escherichia coli* and some inserts may be expressed and kill the host cell (Brayton *et al.*, 1999, Dr. Nicola Collins, personal communication).

Two *C. ruminantium* DNA libraries have been constructed at the OVI, namely the lambda ZAPII (λZAPII) library and the lambda GEM11 (λGEM11) library. The DNA libraries were derived from the Welgevonden isolate of *C. ruminantium*. The Welgevonden isolate is the most virulent of all the *C. ruminantium* isolates and induces cross-protective immunity against more isolates than any other isolate (du Plessis *et al.*, 1989; Brayton *et al.*, 1997).

The λZAP II library was constructed in a bacteriophage vector, the Lambda ZAP® II vector (Stratagene, Fig. 11). It contains 10^4 genome equivalents. The library is free from *Mycoplasma* clones with a 3% proportion of bovine clones. It is estimated that 98% of the clones contain inserts and the average insert size is ~ 3 kb.

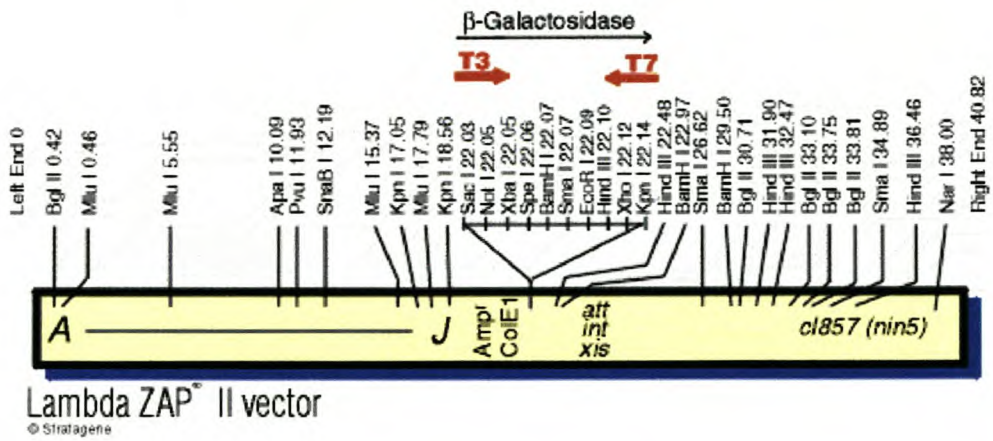


Fig. 11. Restriction map of the Lambda ZAP® II insertion vector (Stratagene).

A second DNA library the λGEM11 library was constructed in a bacteriophage vector, the Lambda GEM® 11 vector (Promega, Fig. 12). It harbours inserts of *C. ruminantium* ranging from 15-23 kb.

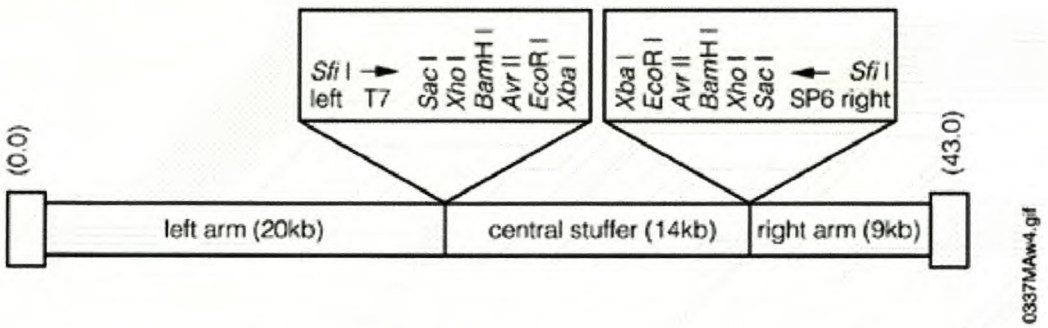


Fig. 12. Restriction map of the Lambda GEM® 11 insertion vector (Promega).

5.4 THE λGEM11 CLONE ANALYSED IN THIS STUDY

During routine sequencing of the λZAPII library, a gene was identified which had homology (score 46, E-value (expect value) 1e-04) to the *secD* gene (secretory protein D gene), of *Rickettsia prowazekii*. This gene encodes the protein-export membrane protein SecD, which is one of the components of the secretory system within the cytoplasmic membrane of bacteria

(Matsuyama *et al.*, 1992; Matsuyama *et al.*, 1993; Pogliano and Beckwith, 1994). As the SecD protein is anchored in the cytoplasmic membrane of bacteria and therefore on the surface, it was thought that it could be exploited as a tool for use in producing a vaccine against *C. ruminantium*.

This thesis describes a study of *C. ruminantium* DNA isolated from the λGEM11 library. The λZAPII clone containing the *secD* gene was used to identify the λGEM11 clone isolated in this study. To this end, the λGEM11 library was screened with a *secD* probe. The *secD* probe identified a bacteriophage plaque from which phage DNA was isolated. The *C. ruminant* insert DNA was amplified from the phage DNA and screened with *Mycoplasma,* Bovine and *Cowdria* probes. The amplified DNA was fragmented by nebulisation and the ends repaired. The resulting blunt-ended fragments were subcloned into pUC_{18} and pMOS*Blue* vectors. The clones were screened by restriction analysis to identify clones containing inserts, which were subsequently sequenced in both directions using the Sanger dideoxy chain termination method (Sanger *et al.*, 1977). Sequences were read using an Automated Sequencer, the ABI™ 377 (Applied Biosystems). Gel readings were assembled and edited using the Staden Genome Assembly Program (Staden, 1998). Two continuous sequences, contiguous sequence were identified with a short sequence of unidentified bases in between. Oligonucleotide primers were designed to amplify the DNA sequences between the two continuous sequences. A single contig was obtained. The entire sequence was analysed and open reading frames (ORFs) were identified using the ORF Finder program[4]. BLAST[5] (Basic Local Alignment Search Tool) was used to screen public databanks for homologues of the sequences in order to identify putative *C. ruminantium* genes (Altschul *et al.*, 1990). Figure 13 shows a flow diagram depicting the overall research strategy of the project.

[4] http://www.ncbi.nlm.nih.gov/gorf/
[5] http://www.ncbi.nlm.nih.gov/BLAST/

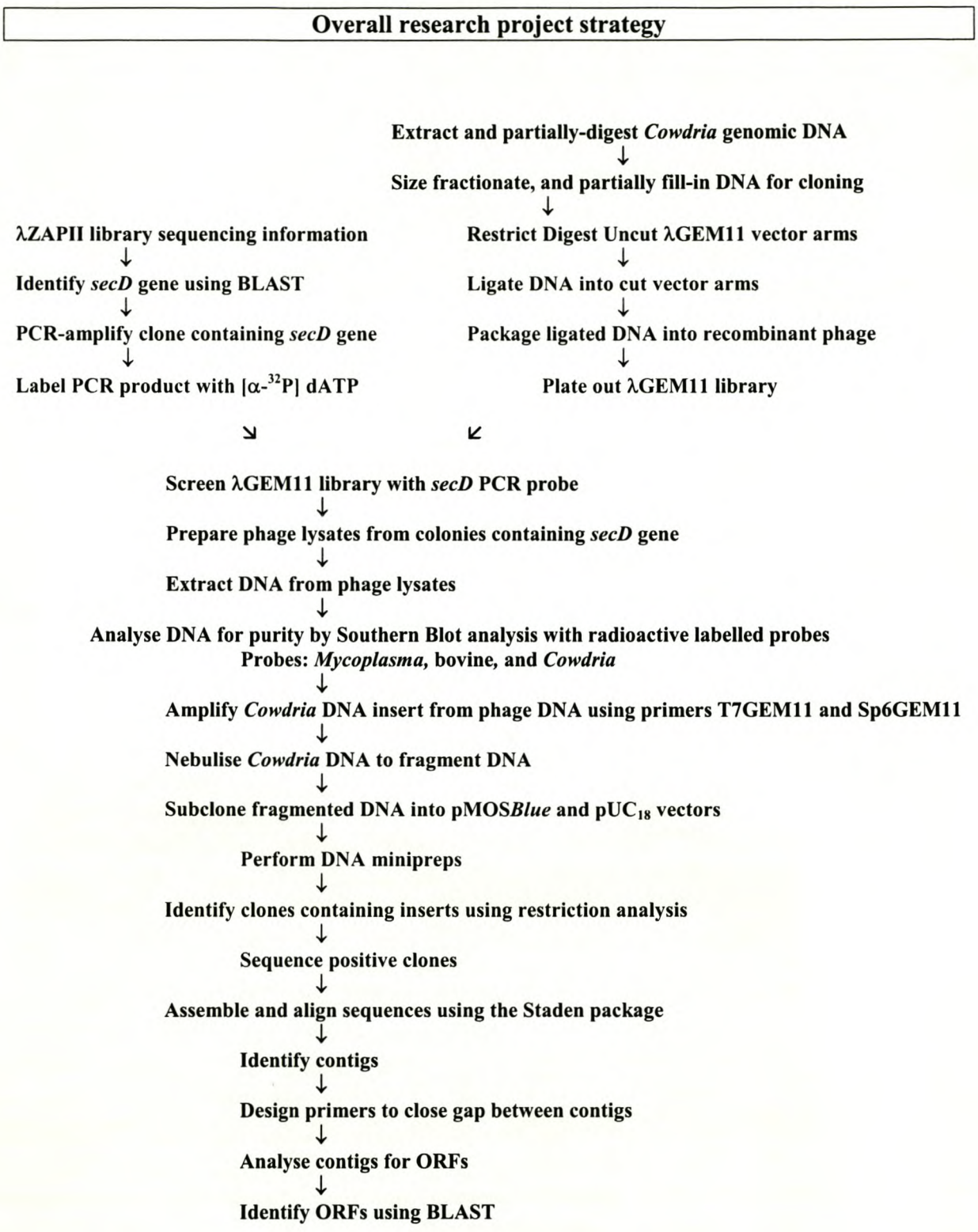


Fig. 13. Flow diagram illustrating the overall reseach strategy of the project.

CHAPTER SIX
METHODOLOGY

6.1 λGEM11 LIBRARY PREPARATION

The *E. coli* KW251 stock was streaked out on LB^{tet} (tetracycline at 15 µl/ml) plates and grown overnight at 37°C. A total volume of 6 ml LB medium supplemented with 0.2% maltose and 10 mM $MgSO_4$ was inoculated with a single *E. coli* KW251 colony. The inoculate was incubated at 37°C at a shaking speed of 225 resolutions per minute (rpm) until the OD_{600} (optical density at 600 nm) reached 0.6-0.8. The λGEM11 library was titred by adding 1 µl of the λGEM11 library stock solution to 300 µl bacterial cells. This was adsorbed at 37°C for 30 minutes without shaking. A total volume of 3 ml LB top agar supplemented with 10 mM $MgSO_4$ was added and the mixture was poured onto LB^{tet} (tetracycline at 15 µl/ml) plates. The plates were incubated overnight at 37°C and the titre of the phage was calculated.

6.1.1 λ phage titre calculation

$$\text{Number of plaques x dilution factor} \; / \; \text{volume of extract plated} = \text{pfu / ml (plaque forming units / ml)}$$

6.2 WL1M636 PCR AMPLIFICATION

Clone WL1M636, the λZAPII phage clone containing the gene encoding *secD*, was amplified by PCR using AmpliTaq Gold™ (Applied Biosystems). Primers WL1F and T7 were used, WL1F primer sequence: 5'- GCT CTA GAA CTA GTG GAT CCC -3'. T7 primer sequence: 5'- GTA ATA CGA CTC ACT ATA GGG C -3' (Integrated DNA Technologies). Originally primer pairs T7 and T3 were used to amplify the *Cowdria* DNA insert contained in the λZAPII vector. However amplification using these primers was suboptimal. This led to the design of primer WL1F which was closer to the multiple cloning site (MCS) than the T3 primer (see Fig. 15, page 66). The DNA clone WL1M636 was used as template of which 1 µg was added to a PCR mixture (Table 2) in a final volume of 100 µl. Amplification (Table 3) was performed on a GeneAmp® PCR System 9700 thermocycler (Applied Biosystems). The PCR product, referred to as WL1M636, was purified using the Concert PCR purification kit

(GibcoBRL) and the DNA concentration was determined on a Beckman spectrophotometer (Model Du-600). Negative PCR controls were included in all PCR amplifications.

Table 2. AmpliTaq Gold™ PCR mixture for amplifying WL1M636

PCR Component	**Concentration**
10 x PCR Buffer II	1x
$MgCl_2$	1.5 mM
dNTPs	200 µM of each dNTP
Primers	0.5 µM of each Primer
Taq	2.5 U

Table 3. AmpliTaq Gold™ cycling conditions for amplifying WL1M636

Temperature	**Time**	**Cycles**
94°C	12 min	1
94°C 58°C 72°C	20 sec 30 sec 1 min 30 sec	10
94°C 58°C 72°C	20 sec 30 sec 2 min 30 sec	6
94°C 58°C 72°C	20 sec 30 sec 5 min	6
94°C 58°C 72°C	20 sec 30 sec 7 min 30 sec	6
68°C	7 min	1
4°C	hold	

6.3 WL1M636 PCR PRODUCT SEQUENCING PROCEDURE

Knowledge of the nucleotide sequence allows one to predict the encoding gene sequence and hence the amino acid sequence of the encoded protein (Winther and Dougan, 1984).

The WL1M636 PCR product was sequenced in both forward and reverse directions. Forward primer M13-47 5'– CAG GAA ACA GCT ATG AC –3' and reverse primer M13-48 5'– GTA AAA CGA CGG CCA GT –3'. A total concentration of 90 ng of the PCR product was added to 8 µl terminator premix (A-dye terminator, C-dye terminator, G-dye terminator, T-dye terminator, dITP, dATP, dCTP, dTTP, Tris-HCl, pH 9.0, $MgCl_2$, thermal stable pyrophosphate, and AmpliTaq DNA polymerase FS) (ABI Prism™ BigDye Terminator Cycle Sequencing Ready Reaction kit, Applied Biosystems). A total primer concentration of 3.2 pmol was added to a final volume of 20 µl. The DNA sequencing reaction (Table 4) was performed in an automated GeneAmp® PCR System 9700 thermocycler (Applied Biosystems). The reaction mixture was ethanol precipitated to remove unincorporated deoxynucleoside triphosphate and oligonucleotides, and vacuum dried. The dry pellet was dissolved in a mixture of 25 mM EDTA pH 8 containing 50 mg/ml Blue Dextran and formamide in a ratio of 5:1. The sequencing reactions were electrophoresed on an ABI™ 377 Automated Sequencer (Applied Biosystems).

Table 4. ABI Prism™ BigDye Terminator cycling conditions for sequencing the WL1M636 PCR product

Temperature	Time	Cycles
94°C	1 min	1
94°C 55°C 72°C	1 min 1.5 min 3 min	30
4°C	hold	

6.4 λGEM11 LIBRARY SCREENING

The λGEM11 library was plated on a series of agar culture plates. The colonies were transferred onto a moist membrane and then exposed to the probe of interest, in this instance the amplified WL1M636 PCR product. The ability of a single stranded DNA probe to bind to similar strands of DNA was exploited to find the clone of interest (Prichard, 1997).

A volume of 6 ml LB medium supplemented with 0.2% maltose and 10 mM $MgSO_4$ was inoculated with a single *E. coli* KW251 colony. The culture was incubated at 37°C at 225 rpm until an OD_{600} of 0.6-0.8 was reached. A total volume of 1 µl of the S_2 λGEM11 library stock was added to 300 µl bacterial cells and adsorbed at 37°C for 30 minutes without shaking, after which 3 ml LB top agar supplemented with 10 mM $MgSO_4$ was added and the mixture poured onto five LB^{tet} (tetracycline at 15 µl/ml) plates. The plates were incubated overnight at 37°C, after which they were incubated at 4°C for 1 hour. Plaque lifts were performed in duplicate, using Magna lift cellulose acetate filters (Separations Scientific). The filters were baked at 80°C for 2 hours. Plaque lift filters were prehybridised in 15 ml Church and Gilbert HYB prehybridisation solution at 65°C for at least 60 min. The radioactively labelled PCR-probe was added to the pre-hybridisation solution and hybridised overnight at 65°C. The filters were washed, sealed in plastic tubing and exposed at -70°C as required. The phage plaque of interest was isolated and resuspended in 100 µl SM buffer. A few drops of chloroform were added and the phage lysate was gently mixed and stored at 4°C.

E. coli KW251 stock was streaked out on LB^{tet} (tetracycline at 15 µl/ml) plates and grown overnight at 37°C. LB medium (6 ml), supplemented with 0.2% maltose and 10 mM $MgSO_4$, was inoculated with a single *E. coli* KW251 colony and incubated at 37°C at 225 rpm until the OD_{600} reached 0.6-0.8. The phage lysate (100 µl) was added to 100 µl bacterial cells, and adsorbed at 37°C for 30 minutes without shaking. After incubation 3 ml LB top agar, supplemented with 10 mM $MgSO_4$, was added and the mixture poured onto LB^{tet} (tetracycline at 15 µl/ml) plates. The plates were incubated overnight at 37°C, after which they were cooled at 4°C for 1 hour. Plaque lifts were performed using Magna lift cellulose acetate filters (Seperations Scientific). The filters were baked at 80°C for 2 hours. A secondary screening was done as described above. A total volume of 5 ml SM buffer was added to the plate and left overnight at 4°C. The LB top agarose was washed gently by pipetting SM buffer over the

agarose with a Pasteur pipette. The SM buffer was removed from the plate and 100 µl chloroform was added. The phage lysate was gently mixed and stored at 4°C.

6.4.1 Preparation of plaque lifts

Three trays lined with benchcoat were prepared, each with it's own piece of Whatmann paper (Seperations Scientific) labelled denature, neutralise and SSC respectively. The circular Magna lift cellulose acetate filters (Seperations Scientific) were marked with pencil. The filters were placed writing side down on plate, making sure there were no bubbles. The filters were marked with needle prick orientation marks right through the agar and membrane. The bottom of the plate was also marked with a pen. The membranes were left on the plate for ten minutes at room temperature. The Whatmann paper (Seperations Scientific) was saturated with the respective solutions (denaturing solution: 0.5 M NaOH and 1.5 M NaCl; neutralising solution: 0.5 M Tris-HCl pH 7.6 and 1.5 M NaCl; 2XSSC). The filters were placed DNA side upwards on the Whatmann paper saturated with denaturing solution and incubated. After 15 minutes the filters were placed on the Whatmann paper saturated with neutralising solution and incubated for 15 minutes. The filters were subsequently moved the Whatmann paper saturated with 2XSSC and incubated for ten minutes. The filters were left to dry on clean Whatmann paper.

6.4.2 Preparation of the [α-^{32}P] dATP probe

The Megaprime™ DNA labelling kit (AEC Amersham) was used to radioactively label the PCR product with [α-^{32}P] dATP. The probe was prepared by mixing together 25 ng of the WL1M 636 PCR product with 17 pmol of primers. This mix was denatured at 100 °C for 5 minutes. Each of the dNTPs (10mM) (dCTP, dTTP, dGTP) was added as well as 10x reaction buffer, 1 U Klenow and 3000Ci/mmol [α-^{32}P] dATP. The reaction was incubated at 37°C for one hour. The reaction was stopped by the addition of 5 µl of 0.2 M EDTA. A Cerenkov count was performed and the probe was heated at 100°C for 5 minutes after which it was immediately placed on ice for 2 minutes. The probe was added to the prehybridised filters and hybridisation was carried out overnight at 65°C.

6.4.3 Cerenkov count

The probe (1 µl) was diluted 1/10 of which 4.5 µl was spotted onto a GF-C filter. The filter was left to dry. This represented the total amount of radionucleotide in the reaction i.e total

count. To determine the amount of incorporated radionucleotide in the reaction, 4.5 µl of the diluted probe was added to 0.5 ml carrier DNA (200 µg/ml sheared calf thymus DNA in 0.1% aqueous SDS). A total volume of 0.5 ml 20% trichloroacetic acid (TCA) was added to precipitate the DNA. The reaction was mixed by inverting the tube and left at room temperature for 10 minutes. A second GF-C filter was placed into a slot of the vacuum filtration apparatus and assembled. The filter was prewet with 10% TCA. The precipitated DNA was subsequently filtered and the unincorporated radionucleotides were removed by washing the filter three times with 1 ml 10% TCA. The filter was washed with 1 ml ethanol and left to dry. Both the filters (total count and incorporated count) were placed into scintillation vials and a Cerenkov count was performed.

$$\text{Percentage incorporation} = {}^{\text{incorporated count}} / {}_{\text{total count}} \times 100$$

6.5 PROCEDURE FOR THE INCREASE OF λGEM11 CLONE TITRE

E. coli KW251 stock was streaked out on LB^{tet} (tetracycline at 15 µl/ml) plates and grown overnight at 37°C. LB medium (6 ml) supplemented with 0.2% maltose and 10 mM $MgSO_4$ was inoculated with a single *E. coli* KW251 colony (300 µl cells / lysate / no of plates). The culture was incubated at 37°C at 225 rpm until an OD_{600} of 0.6-0.8 was reached. Different volumes of phage lysate, 200 µl, 300 µl and 500 µl respectively was added to 100 µl bacterial cells and adsorbed at 37°C for 30 minutes without shaking. To each of these 3 ml LB top agar supplemented with 10 mM $MgSO_4$ was added and poured onto separate LB^{tet} (tetracycline at 15 µl/ml) plates. Plates were incubated overnight at 37°C, after which 5 ml SM buffer was added to confluent plates and incubated at 4°C overnight. The LB top agarose was gently washed by pipetting SM buffer over the agarose with a Pasteur pipette. The SM buffer was collected from the plates and 100 µl chloroform was added. The phage lysate was gently mixed and stored at 4°C. The phage titre was determined as before. A titre of 2×10^7 is required to yield sufficient DNA.

6.6 PHAGE DNA ISOLATION

E. coli KW251 stock was streaked on LB^{tet} (tetracycline at 15 µl/ml) plates and grown overnight at 37°C. LB medium (6 ml) supplemented with 0.2% maltose and 10 mM $MgSO_4$ was inoculated with a single *E. coli* KW251 colony and incubated at 37°C at 225 rpm until the OD_{600} reached 0.6-0.8. A total of 2 x 10^7 phage was adsorbed to 10^{10} bacterial cells at 37°C for 30 minutes without shaking. A 10 µl aliquote of this mixture was added to 2.5 ml LB medium supplemented with 10 mM $MgSO_4$ and incubated overnight at 37°C shaking at 225 rpm. The phage DNA was isolated according to a λ phage DNA Quickprep (University of Zurich) method and the DNA concentration was determined. The DNA (1 µg) was digested with Eco RI and Xho I (Roche) at 37°C overnight. The digested DNA was heated at 75°C for 15 min and electrophoresed on a 0.7% agarose/TBE gel.

6.6.1 λ phage DNA Quickprep, University of Zurich method

The overnight culture (15 ml) was cooled to room temperature and two drops of chloroform was added. A final concentration of 1 U of DNAse (10 U/µl) was added and the culture was incubated at room temperature at 150 rpm for 25 minutes. The culture was divided into Eppendorf tubes and centrifuged on a benchtop micro centrifuge (Labotec, Mikro 20 Hettich) at 13 000 rpm for 5 minutes. A new Eppendorf tube was pre-filled with 200 µl STE buffer to which 600 µl of each of the culture supernatant was added, mixed and left at 70°C for 15 minutes. The reaction was cooled to room temperature and 150 µl of 8 M KAc was added. The reaction was mixed and incubated on ice for 15 minutes after which it was centrifuged for 15 minutes at 13 000 rpm. The clear supernatant was extracted with equal amounts of P: C: I (phenol: chloroform: isoamylalcohol / 25: 24: 1), vortexed for 5 seconds and centrifuged at 13 000 rpm for 10 minutes. The top phase was added to a new tube to which 420 µl isopropanol was added. The reaction was mixed and left to incubate at room temperature for 10 minutes after which it was centrifuged at 13 000 rpm for 8 minutes. The supernatant was carefully removed and the pellet was washed with ice-cold (-20°C) 500 µl 70% ethanol and dried at 37°C for 10 minutes. The pellet was resuspended in 50 µl TE buffer, pH8, containing 0.3 mg/ml RNAse A and digested at 37°C for 30 minutes. The DNA was pooled and aliquoted in 100µl aliquots. To each aliquote 40 µl of 5 M NH_4Ac and 200 µl of isopropanol was added. The sample was incubated at room temperature for 10 minutes then centrifuged at 13 000 rpm

for 10 minutes. The supernatant was removed and the pellet was washed with 500 µl 70% ethanol. The pellet was left to dry and subsequently dissolved in 20 µl water.

6.7 PCR AMPLIFICATION AND ISOLATION OF *C. RUMINANTIUM* DNA

The isolated phage DNA was amplified using *TaKaRa LA Taq* ™ (TaKaRa Biomedicals). λGEM11 primers, SP6GEM11 primer sequence: 5'-CCA TTT AGG TGA CAC TAT AG-3' and T7GEM11 primer sequence: 5'-CTA ATA CGA CTC ACT ATA GG-3' (Integrated DNA technologies) was used to amplify the *C. ruminantium* insert from the phage miniprep DNA. Purified DNA at a final concentration of 1 µg was added to a PCR mixture in a final volume of 50 µl (Table 5). PCR amplification (Table 6) was performed using a GeneAmp® PCR System 9700 thermocycler (Applied Biosystems). The PCR product was purified using a Concert PCR purification kit (GibcoBRL) and the DNA concentration was determined using a Beckman spectrophotometer (Model Du-600). The PCR product (1 µg) was digested with Eco RI and Xho I (Roche) at 37°C overnight. The digested PCR products were electrophoresed on a 0.7% agarose/TBE gel.

Table 5. *TaKaRa LA Taq*™ PCR mixture for amplifying *C. ruminantium* DNA insert from phage miniprep DNA

PCR Component	**Concentration**
10 x LA PCR BufferII	1x
$MgCl_2$	2.5 mM
dNTPs	400 µM of each dNTP
Primers	0.4 µM of each Primer
Taq	2.5 U

Table 6. *TaKaRa LA Taq*™ cycling conditions for amplifying *C. ruminantium* DNA insert from phage miniprep DNA

Temperature	Time	Cycles
94°C	2 min	1
94°C 47°C 68°C	10 sec 30 sec 15 min	10
94°C 47°C 68°C	10 sec 30 sec 15 min + 20 sec/cycle	15
68°C	7 min	1
4°C	hold	

A series of PCR amplifications was performed to obtain sufficient DNA i.e. 1 µg. The individual PCR products were digested with Hind III to verify that the same PCR product was obtained for each individual PCR.

6.8 *C. RUMINANTIUM* PCR PRODUCT SCREENING PROCEDURE

The PCR product obtained above (1 µg) was digested with 10 U/µl of Hind III and Bgl II restriction enzymes (Roche), and was electrophoresed on a 0.7% agarose/TBE gel. Southern blots were prepared using Hybond N nylon membranes (AEC Amersham). Four membranes were prepared, one for each radioactive probe. The membranes were prehybridised in 15 ml Church and Gilbert HYB prehybridisation solution at 65°C for at least 60 min. The Megaprime™ DNA labelling kit (AEC Amersham) was used according to the manufacturers instructions to radioactively label *Mycoplasma* DNA, bovine DNA and *Cowdria* DNA (50 ng/µl in 100 µl) with [α-^{32}P] dATP (see 6.4.2). Each probe was added to an individual filter in pre-hybridisation solution and hybridised overnight at 65°C. The filters were washed and sealed in plastic tubing. The filters were exposed at -70°C as required.

6.9 PREPARATION OF *C. RUMINANTIUM* DNA FOR CLONING

The PCR product was diluted to 10 µg/ml in TE and 2 ml was nebulised in a Medel jet nebuliser reservoir (Medel, Italy) for 5 minutes at 100 kPa. Nebulised DNA was precipitated and resuspended in 50 µl TE buffer. The recessed 3'-termini of the nebulised DNA was filled in using the Klenow fragment of *E. coli* DNA polymerase (AEC Amersham) (nebulised DNA 5 µg, 10xDNA polymerase buffer, 10 mM dNTPs, 20 U Klenow) in a final volume of 200 µl for 1 hour at room temperature. The reaction was stopped by the addition of 50 mM EDTA to the reaction and heat inactivation at 72°C for 15 minutes. The DNA was electrophoresed on a 0.7% agarose/TAE gel. The desired DNA fragment was cut out of the gel and extracted using the BioRad prep-a-gene kit (Promega). The DNA concentration was determined using a Beckman spectrophotometer (Model Du-600).

6.9.1 Precipitation of nebulised DNA

A total of 0,1 volumes of 3 M NaOAc pH 5.2 was added to the nebulised DNA and mixed. To this mixture 2.5 volumes of ice-cold absolute ethanol (calculated after the salt addition) was added. The reaction was mixed and incubated at -70°C for one hour after which the reaction was centrifuged at maximum speed on a benchtop micro centrifuge (Labotec, Mikro 20 Hettich) for 20 minutes at 4°C and the supernatant was discarded. The pellet was washed with 500 µl of ice-cold 70% ethanol and centrifuged at maximum speed on a benchtop micro centrifuge for ten minutes at 4°C. The supernatant was discarded and the pellet was left to dry at 37°C. The pellet was resuspended in water. The ends of the precipitated DNA were filled in with Klenow (AEC Amersham) as described above.

6.10 PREPARATION OF pUC_{18} VECTOR

The pUC_{18} vector (5 µg) (Roche) was cut with Bam HI (15 U) (Roche) in a final volume of 50 µl at 37°C for 1.5 hours. The Bam HI-cut vector was dephosphorylated for 30 minutes at 37°C using alkaline phosphatase (Promega) (5 µg DNA, 10 X buffer, 1 U alkaline phosphatase) in a final volume of 100 µl. The reaction was stopped by the addition of 50 mM EDTA and heat inactivation at 72°C for 15 minutes. The ends were filled in with Klenow (AEC Amersham) (5 µg Bam HI-cut dephosphorylated DNA, 10xDNA polymerase buffer, 10 mM dNTPs, 20 U Klenow) in a final volume of 100 µl. The vector was electrophoresed on a

0.7% agarose/TAE gel. The appropriate bands were gel purified using the BioRad prep-a-gene kit (Promega). The DNA was quantitated on a Beckman spectrophotometer (Model Du-600).

6.11 *C. RUMINANTIUM* DNA CLONING PROCEDURE

Molecular cloning techniques are based on introducing DNA from microorganisms or higher eukaryotes into a range of bacteria, yeast or mammalian cells (Winther and Dougan, 1984).

Both the pMos*Blue* vector as well as the pUC_{18} vector were used as cloning vectors. For cloning into the pMosBlue vector, a molar ratio of 1 mol of vector to 60 mol of insert is recommended for the cloning reaction. Thus 50 ng of pMosBlue vector (2 887 bp) was added to 1.61 µg of insert DNA (average size of 1 550 bp) for cloning purposes. For cloning into the pUC_{18} vector, a molar ratio of 1 mol of vector to 4 mol of insert is recommended for the cloning reaction. Thus 50 ng of pUC_{18} vector (2 686 bp) was again added to 115 ng of insert DNA (average size of 1 550 bp) for cloning purposes.

The pMOS*Blue* blunt ended cloning kit (AEC Amersham) was used according to the manufacturer's instructions. A ligation reaction was set up at 15°C for the pUC_{18} vector using T4 ligase. DNA was added in the above mentioned ratios for each cloning vector in a final volume of 10 µl to which 1 U of ligase and 10x buffer was added. The electrocompetent *E. coli* cells were transformed with these ligation reactions. The pMOS*Blue* transformation reactions were plated on $LB^{amp/tet}$ (50 µg/ml ampicillin and 15 µg/ml tetracycline) plates and the pUC_{18} transformation reactions on LB^{amp} plates. The LB plates contained 40 µl X-gal (50 mg/ml) and 40 µl IPTG (100 mM). The plates were incubated overnight at 37°C. The white colonies were selected and grown overnight at 37°C with vigorous shaking in 5 ml LB containing the respective antibiotics for the respective vectors. The DNA was isolated from the overnight cultures using the QIAprep miniprep kit (QIAGEN). The isolated DNA was digested overnight at 37°C with Eco RI and Xba I (Roche). The digestion products were electrophoresed on a 0.7% agarose/TBE gel. The clones containing inserts were sequenced.

6.11.1 Preparation of electrocompetent cells

E. coli KW251 was freshly streaked out and a single colony was inoculated into 15 ml LB medium. The culture was grown overnight at 37°C with moderate shaking at 225 rpm. Two tubes of 5 ml each of the overnight culture was inoculated into 2 x 500 ml prewarmed LB^{tet} (tetracycline at 15 µl/ml) broth in sterile 2 l flasks and incubated at 37°C, 300 rpm until the OD_{600} reached 0.5 – 0.6. The culture was poured into four prechilled 250 ml centrifuge bottles and incubated on ice for 20 minutes. The cells were pelleted at 5 520 x g (relative centrifugal fields) for 10 minutes at 4°C. The supernatants were discarded and the pellets were pooled and resuspended in 10 ml ice-cold dH_2O. The cells were washed by gently mixing it with 240 ml ice-cold dH_2O. The cells were pelleted at 5 520 x g for 10 minutes at 4°C. The supernatant was discarded and the pellet was resuspended by swirling in the remaining liquid. The cells were washed with 240 ml ice-cold dH_2O. The cells were pelleted at 5 520 x g for 10 minutes at 4°C. The supernatant was discarded immediately and the pellet was resuspended in 10 ml SOC medium containing 10% glycerol and transferred to sterile 50 ml centrifuge tubes. The cells were incubated for one hour. The cells were pelleted at 4 350 x g for 10 minutes at 4°C and the supernatant was discarded. The pellets were pooled and resuspended in 800 µl SOC medium containing 10% glycerol and aliquoted into sterile microcentrifuge tubes and snap freezed.

6.12 PROCEDURE FOR SEQUENCING pUC18 and pMOS*Blue* clones

The positive clones were sequenced in both forward and reverse directions with forward primer M13-47 5'– CAG GAA ACA GCT ATG AC –3' and reverse primer M13-48 5'– GTA AAA CGA CGG CCA GT –3' as described in 6.3. Sequences were read using an automatic base scanner and automatic base calling software (Amersham). Gel readings were assembled and edited using the Staden Genome Assembly Program (Staden, 1998) including gap4 (genone assembly program) and nip4 (nucleotide interpretation program).

6.13 CLOSING THE GAP

The sequence assembly led to the identification of two continuous sequences with a single gap between them. In order to close the gap primers were designed using the Intergrated DNA Technologies Inc. Oligoanalyser 2.5 program[5]. Primer WL2TP1_gapF was designed complementary to the sequence at the 3' end of contiguous sequence (contig) 1 and primer WL2TP1_gapR was designed complementary to the sequence at the 5' end of contig 2. A PCR was performed using both *Cowdria* genomic DNA and the phage miniprep DNA. The phage miniprep DNA and the *Cowdria* genomic DNA was amplified by PCR using the Expand High Fidelity PCR System™ (Roche) and the specially designed primers WL2TP1_gapF: 5'- CAC TTA CAC CAA TGC CAC AC -3' and WL2TP1_gapR: 5'- TTA CCG CCA CCC TAA CAT ATA G -3' (Integrated DNA technologies). The DNA (1 µg) was added at in a final volume of 100 µl (Table 7). PCR amplification (Table 8) was performed on a GeneAmp® PCR System 9700 thermocycler (Applied Biosystems). The PCR product was purified using a Concert PCR purification kit (GibcoBRL) and the DNA concentration was determined using a Beckman spectrophotometer (Model Du-600).

Table 7. Expand High Fidelity PCR System™ (Roche) for amplifying phage miniprep DNA and *Cowdria* genomic DNA

PCR Component	Concentration
10 x LA PCR BufferII	1x
$MgCl_2$	2.5 mM
dNTPs	400 µM of each dNTP
Primers	0.4 µM of each Primer
Taq	2.5 U

[5] http://www.idtdna.com/program/oligocalc/oligocalc.asp

Table 8. Expand High Fidelity PCR System™ (Roche) cycling conditions for amplifying phage miniprep DNA and *Cowdria* genomic DNA

Temperature	Time	Cycles
94°C	2 min	1
94°C 50°C 72°C	10 sec 30 sec 15 min	10
94°C 50°C 72°C	10 sec 30 sec 15 min + 20 sec/cycle	15
72°C	7 min	1
4°C	hold	

The PCR products obtained were sequenced in both forward and reverse directions with forward primer M13-47 5'– CAG GAA ACA GCT ATG AC –3' and reverse primer M13-48 5'– GTA AAA CGA CGG CCA GT –3'. Two of the clones that were previously obtained by the cloning of the *C. ruminantium* DNA into the pUC_{18} and pMOS*Blue* vectors and which were subsequently sequenced (6.11 and 6.12) were also sequenced with the specially designed primers, as they were found to be present in both contigs and would therefore breach the gap. The sequencing was done as described earlier (6.3). Sequences were read using an automatic base scanner and automatic base calling software (Amersham). Gel readings were assembled and edited using the Staden Genome Assembly Program (Staden, 1998) including gap4 and nip4.

CHAPTER SEVEN
RESULTS AND DISCUSSION

7.1 λGEM11 LIBRARY TITRATION

The four λGEM11 library stocks designated S_1-S_4 was titred (Table 9). The λGEM11 library S_2 was used for making plaque lifts. Five plates were made designated T_1-T_5. The number of plaques obtained can be seen in Table 10. Plaque lifts were made (in duplicate) of these five plates and were screened with the WL1M636 PCR product (see 7.4).

Table 9. Titres of the λGEM11 library stocks

λGEM11 library stock	**Phage titre in pfu/ml**
S_1	8 000
S_2	150 000
S_3	105 000
S_4	90 000

Table 10. Plaques obtained from λGEM11 library stock two, S_2

S_2 Plates	**Number of plaques**
T_1	130
T_2	127
T_3	129
T_4	144
T_5	150

7.2 ANALYSIS OF WL1M636 AMPLIFICATION

PCR amplification of WL1M636 yielded a PCR product of approximately 900 bp (Fig.14). A few PCR amplifications were performed to obtain sufficient DNA to use in further experimentations. After purification the total yield of the PCR product was 9.2 mg; $OD_{260nm/280nm} = 1.58$.

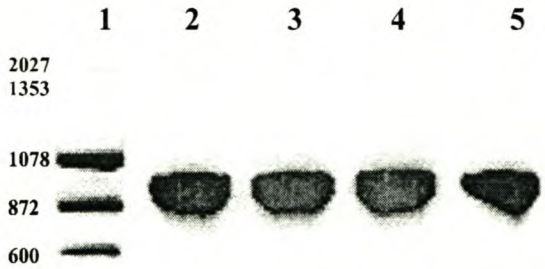
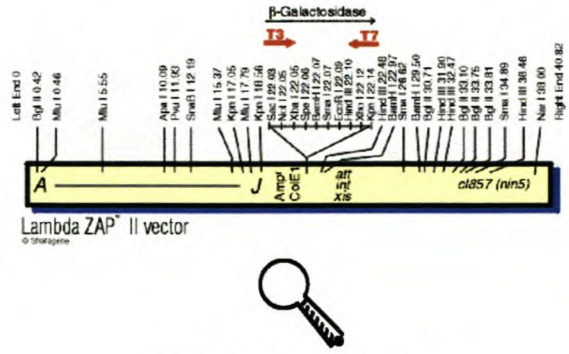


Fig. 14. PCR amplification of the WL1M636 clone DNA. The reactions were analysed by 0.7% agarose gel electrophoresis, followed by ethidium bromide staining. Lane 1, λDNA-Hind III and ϕX174 DNA-Hae III molecular weight markers. Lanes 2 to 5, PCR product.

7.3 WL1M636 PCR PRODUCT SEQUENCING ANALYSIS

The WL1M636 PCR product was sequenced in both the forward and reverse directions and the ORFs were identified using the ORF Finder program[4]. A BLAST search was performed to screen public databanks for homologues of the ORFs. As expected a gene was identified (position 50 – 251 of insert) which had homology to the *secD* gene (score 77.4 and E-value 3e-14) of *Rickettsia prowazekii*. Figure 15 shows the PCR product contained within the λ ZAPII vector, as well as the primer sequences and the *secD* sequence used to isolate the λGEM 11 clone (see 6.2).

[4] http://www.ncbi.nlm.nih.gov/gorf/



Multiple Cloning Site

GTAATACGACTCACTATAGGGCGAATTGGGTACCGGGCCCCCCCTCGATCGAAGAAAAAATTAAGATGGGCTGTA
GAGTCTGGATTTAAAAATGCTATGTCTACGATATTTGATTCAAATATTACTACGCTGATTGTTGCAGCAATTATG
TTTGTAATGGGTAGCGGCCCAGTTAGTGGCTTTGCTATTACGCTATCTATAGGTATTTTATGTTCTATGTTTTCT
GCAATTACATTGACAAAAATGCTTATAGATTTGTACATTAGGTTCTTTAATTTAAAGGTGTTAAATATATAGTAA
TTTGTTATAAAGGTTAAATTTTTATTTATTTAAAAATTGTTTTTTAATCATAATATTGCAGTGAATATAGTAACG
GAGATCATTGTTTGTTATAATTAGTAAGTTAGATAATGTTGATTGTGTGAAATACCTTAAGAAATGTATTTGATA
GTAAGAAAGAAAAAGGTTTAATTTTTATCTGTAGGTCTTGATAGCAGTTATGTGTAGTACTAGTTTTAAATTATC
AGAAAATTGTTGAGGAATTTTATTGTAATTTGGGGTATAATAGTGTTAATAATTATTTAATTCTGATGTAAAATT
TTTATCATAGGATAATCTATCTAATATTATAATGAAACTTTTATGATGACTTATCTAAGGTATATTCTTTAATCA
AGGAAAGAAAAAGTGGAGTTTTTGTCTATAGAATTATTCATAAAGCTATGTTGATAGTAAGTAAATAAGTTTAAT
GTACTAAACATGGTATTTGCTAGTAATATAATAATTTGATTATAGGTCGACGGTATCGATAAGCTTGATATCGAA
TTCCTGCAGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTTCCCTTTA
GTGAGGGTTA

Fig. 15. DNA sequence of WL1M636 PCR product contained within the λZAPII vector.

T7 primer sequence: 5'- GTA ATA CGA CTC ACT ATA GGG C -3'

WL1F primer sequence: 5'- GCT CTA GAA CTA GTG GAT CCC -3'

T3 primer sequence: 5'- TAA CCC TCA CTA AAG GGA -3'

Last 348 bp of PCR product which is referred to later in this thesis (see Addendum 6, page 167)

secD sequence used to identify λGEM11 clone

WL1M636 PCR product sequence

7.4 λGEM11 LIBRARY SCREENING ANALYSIS

The WL1M636 PCR product was readioactivelly labelled. Table 11 shows the amount of [α-^{32}P] dATP incorporation for the probes used.

Table 11. The percentage [α-^{32}P] dATP incorporation using the WL1M636 PCR product

Screening	Cerenkov Count	% Incorporation
Primary Screening	58260	87
Secondary Screening	115811	70

The first screening of the five plates, T1-T5, yielded a single plaque i.e the probe hybridised to a single plaque (Fig.16a).

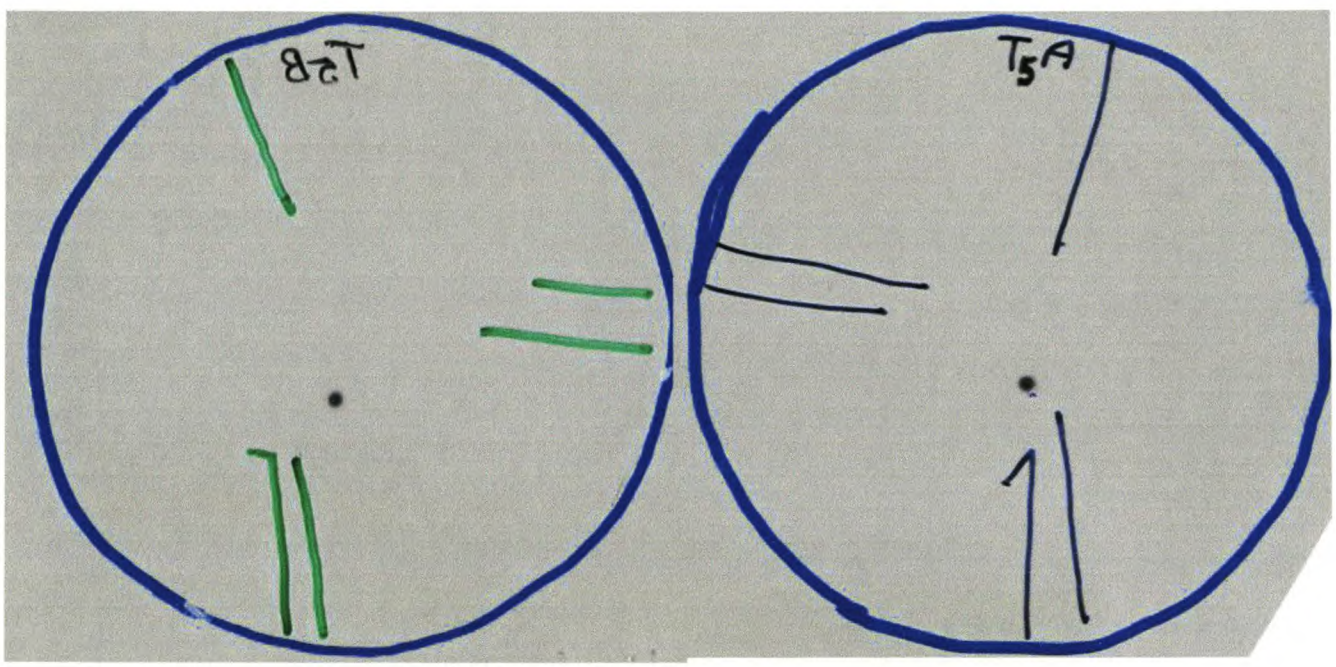


Fig. 16a. The autorads of the first screening

This plaque was isolated and a phage lysate was prepared. The subsequent screening using this phage lysate produced plates where the probe hybridised to all of the plaques (Fig. 16b). A phage lysate of these plates was made.

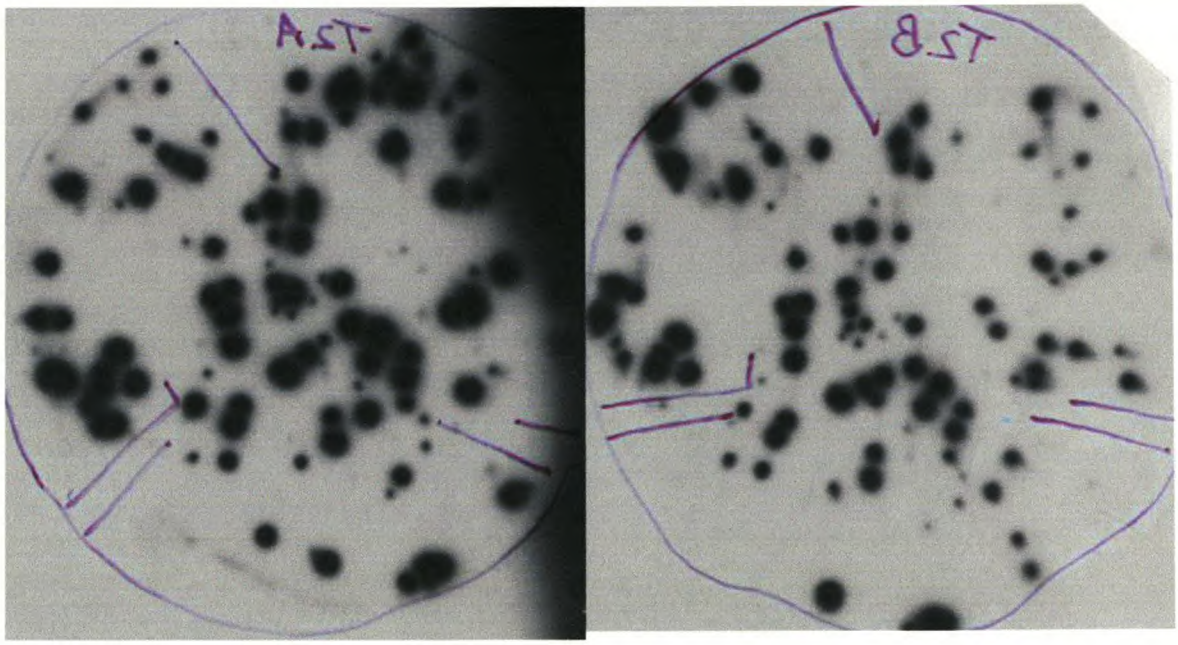


Fig. 16b. The autorads of the secondary screening

7.5 λGEM11 CLONE TITRE INCREASE ANALYSIS

The first titre increase of the phage lysate yielded in a titre of 6.033×10^2 pfu/ml. A second titre increase was performed using the phage lysate obtained from these plates, as a titre of 2×10^7 phage to 10^{10} is needed to obtain sufficient DNA for further experimentation. The second titre increase yielded a titre of 2.27×10^8 pfu/ml.

7.6 PHAGE DNA ANALYSIS

The DNA isolation yielded a phage DNA fragment of 42 kb (Fig. 17, lane 2). The phage DNA was digested with Eco RI and Xho I as these enzymes have restriction sites within the MCS of the λGEM11 vector (Fig. 17, lane 4). The restriction digestion yielded three fragments, which meant that the *C. ruminantium* insert had no restriction sites for Eco RI and Xho I. The first band (42 kb) shows the undigested λGEM11 DNA; the second band shows the 20 kb left arm; the third band shows the approximately 12 kb insert and the fourth band (very faint) shows the 9 kb right arm. The three fragments combined (20 kb left arm, 12 kb insert and 9 kb right arm) added to 41 kb, which is approximately the size of the miniprep i.e 42 kb.

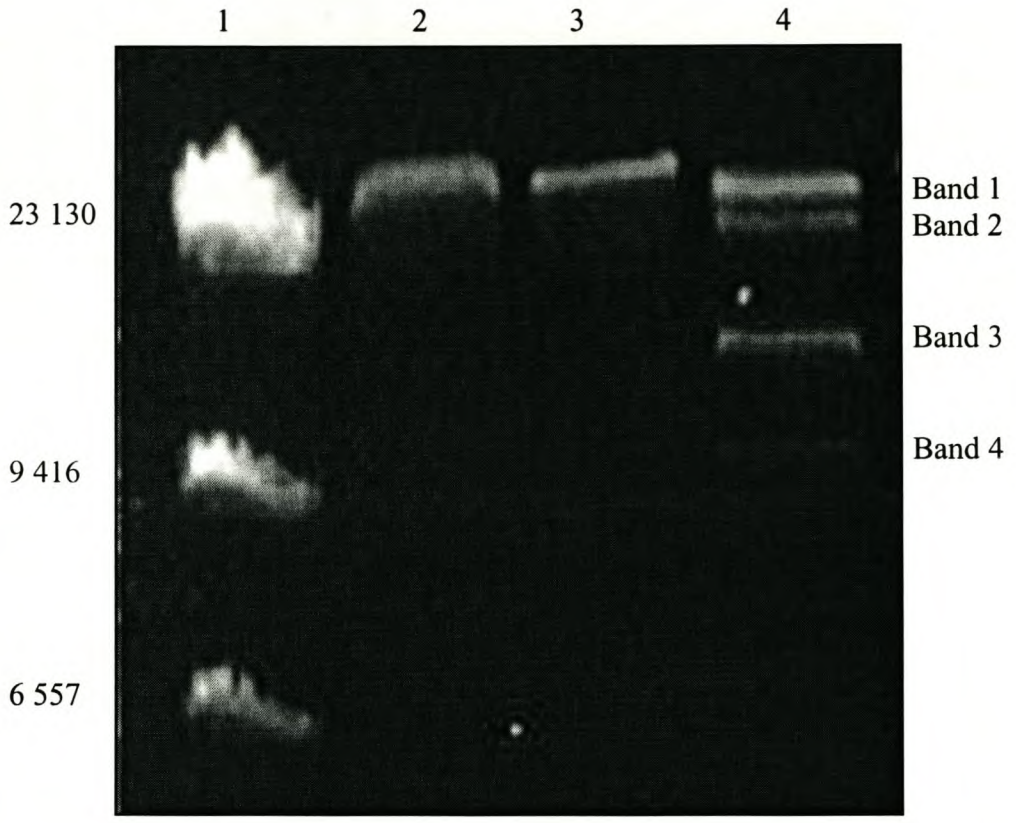


Fig. 17. The restriction digestions of the phage miniprep DNA (1μg) were analysed by 0.7% agarose gel electrophoresis, followed by ethidium bromide staining. Lane 1, ϕX174 DNA-Hae III molecular weight marker. Lane 2, phage miniprep DNA obtained from the DNA Quick prep (University of Zurich) method. Lane 3, phage miniprep DNA digested with Eco RI. Lane 4, phage miniprep DNA digested with Xho I.

Figure 18 is a diagram depicting the *C. ruminantium* DNA as it is when still contained within the λGEM11 vector. The DNA concentration was determined and a total yield of 1.62 mg DNA was obtained with an $OD_{260nm/280nm} = 1.91$.

Fig. 18. A diagram depicting the *C. ruminantium* DNA contained within the λGEM11 vector.

7.7 *C. RUMINANTIUM* DNA ANALYSIS

The amplification of the *C. ruminantium* insert located within the λGEM11 phage DNA gave a PCR product of 12 kb (Fig. 19). This is equal in size to the insert observed with the restriction analysis as discussed in 7.6.

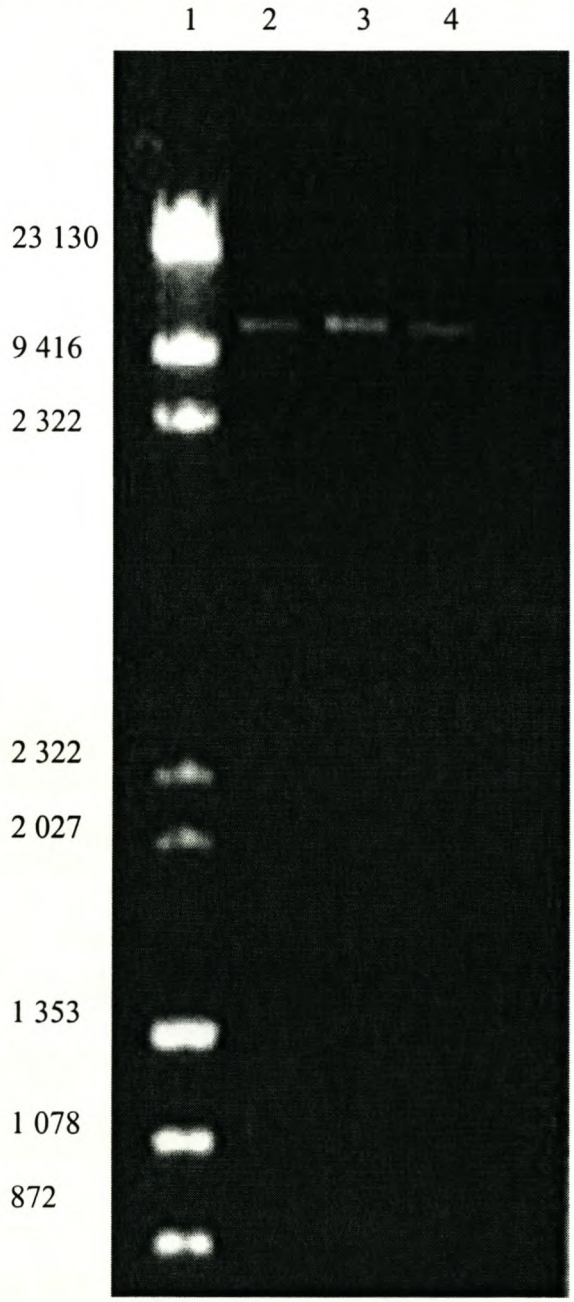


Fig. 19. Amplified *C. ruminantium* DNA PCR product analysed by 0.7% agarose gel electrophoresis, followed by ethidium bromide staining. Lane 1, λDNA-Hind III molecular weight marker. Lanes 2, 3 and 4, PCR product obtained from the phage miniprep DNA using the *TaKaRa LA Taq* ™ (TaKaRa Biomedicals) PCR method.

A few PCR amplifications were carried out to obtain sufficient DNA. The PCR products were digested with Hind III (Fig. 20) to verify that each individual reaction amplified the same product. This restriction enzyme was chosen as the *C. ruminantium* DNA had no Eco RI or Xho I restriction sites as observed in 7.6. The digestion profile of all the PCR products were the same which meant that the amplification reaction produced the same product in each of the separate reactions, providing enough material for further experiments. Six fragments were obtained after the digestion of the PCR product. The fragments (band 1 - 500 bp, band 2 - 800 bp, band 3 – 1 400 bp, band 4 – 2 300 bp, band 5 – 3 000 bp and band 6 – 4 000bp) added to 12 kb which is equal to the size of the PCR product obtained.

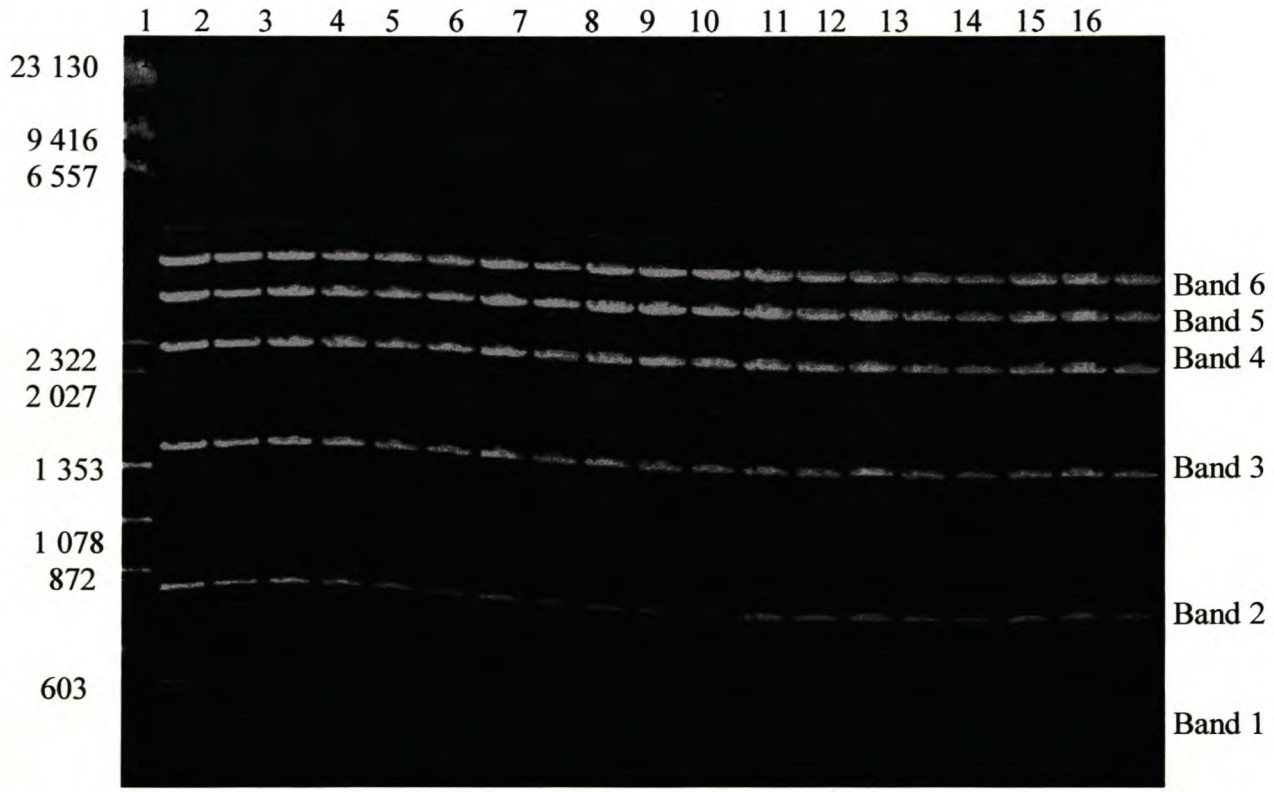


Fig. 20. The restriction digestion reactions of the PCR product were analysed by 0.7% agarose gel electrophoresis, followed by ethidium bromide staining. Lane 1, λDNA-Hind III and ϕX174 DNA-Hae III molecular weight markers. The remaining lanes show the *C. ruminantium* amplified DNA digested with Hind III.

7.8 *C. RUMINANTIUM* PCR PRODUCT SCREENING ANALYSIS

The *C. ruminantium* PCR product was digested with Hind III and Bgl II and the digested products were run on a 0.7% agarose/TBE gel (Fig. 21). These two restriction enzymes were chosen as the *C. ruminantium* DNA had no Eco RI or Xho I restriction sites as observed in 7.6. Five bands were observed for each digestion, Hind III (band 1 - 872 bp, band 2 – 1 150 bp, band 4 – 2 322 bp, band 5 – 3 500 bp and band 6 – 4 000 bp) and Bgl II (band 1 - 872 bp, band 3 – 1 500 bp, band 4 – 2 322 bp, band 5 – 3 500 bp and band 6 – 4 000 bp). These fragments added up to 11 844 bp and 12 194 bp respectively. These sizes are approximately the same as the 12 kb insert which was obtained in 7.7.

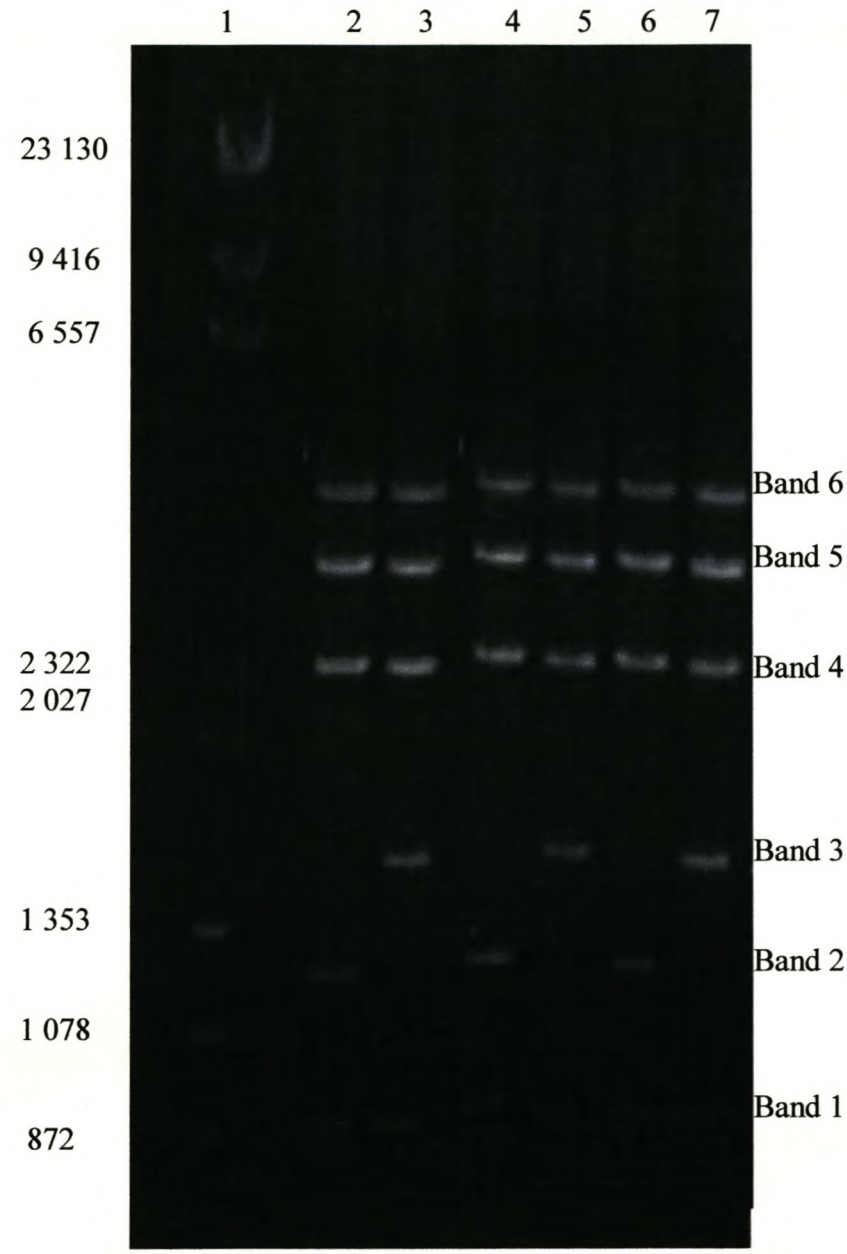


Fig. 21. The restriction digestion reactions were analysed by 0.7 % agarose gel electrophoresis, followed by ethidium bromide staining. Lane 1, λDNA-Hind III and ϕX174 DNA-Hae III molecular weight markers. Lanes 2, 4 and 6, *C. ruminantium* DNA digested with Bgl II. Lanes 3, 5 and 7, *C. ruminantium* DNA digested with Hind III.

The digested products were transfered to nylon membranes and these membranes were screened with radioactively labelled *Mycoplasma*-, bovine- and *Cowdria* probes (Fig. 22) at 50 ng/µl in 100 µl.

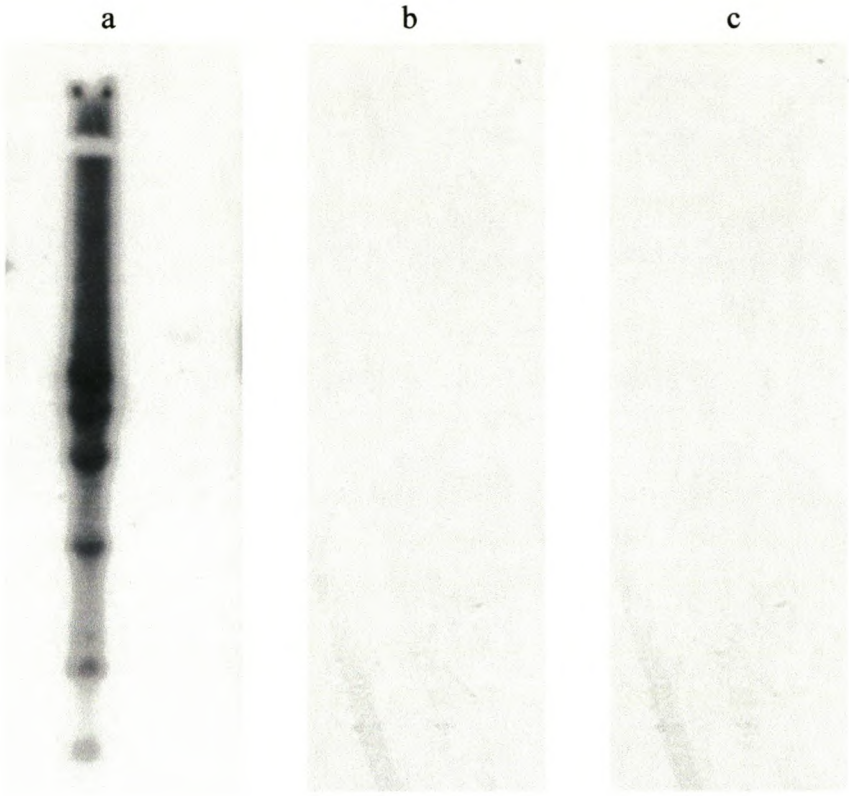


Fig. 22. The autorads for the screening of the *C. ruminantium* DNA digestions with the individual radioactive probes: (a) *C. ruminantium* DNA digestions screened with the *C. ruminantium* probe. (b) *C. ruminantium* DNA digestions screened with the bovine probe. (c) *C. ruminantium* DNA digestions screened with *Mycoplasma* the probe.

Only the *Cowdria* probes were able to hybridise to the digested PCR product suggesting that the *C. ruminantium* DNA were clean and not contaminated with other DNA.

7.9 *C. RUMINANTIUM* CLONING DNA ANALYSIS

The *C. ruminantium* DNA was nebulised to fragment the DNA. The ends were filled in with Klenow (AEC Amersham) and the product was electrophoresed on a 0.7 % agarose gel (Fig. 23).

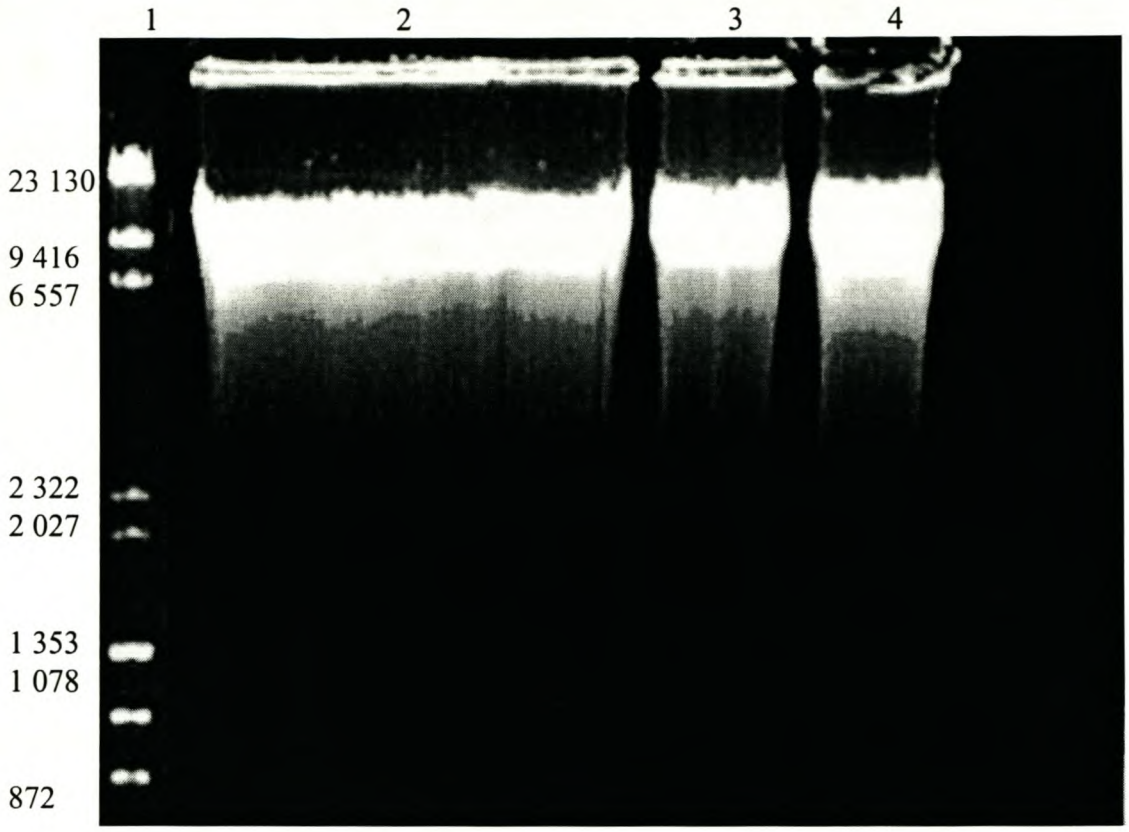


Fig. 23. The nebulised DNA was visualised on a 0.7% agarose gel stained with ethidium bromide. Lane 1, ϕX174 DNA-Hae III molecular weight marker. Lanes 2 to 4, nebulised DNA.

A narrow band was obtained with a size range from between 6 000 – 15 000 bp. This showed that the nebulisation was not as successful, as a much broader band was expected.
The DNA band was gel-purified and the DNA concentration was determined. A total yield of 61.8 µg DNA obtained; $OD_{260nm/280nm} = 1.84$.

7.10 pUC_{18} VECTOR ANALYSIS

The pUC_{18} vector (5 µg) was cut with Bam HI, dephosphorylated and blunt ended. A band of approximately 2.5 kb was obtained (Fig. 24) which is equal to the size of the pUC_{18} vector i.e. 2 686 bp.

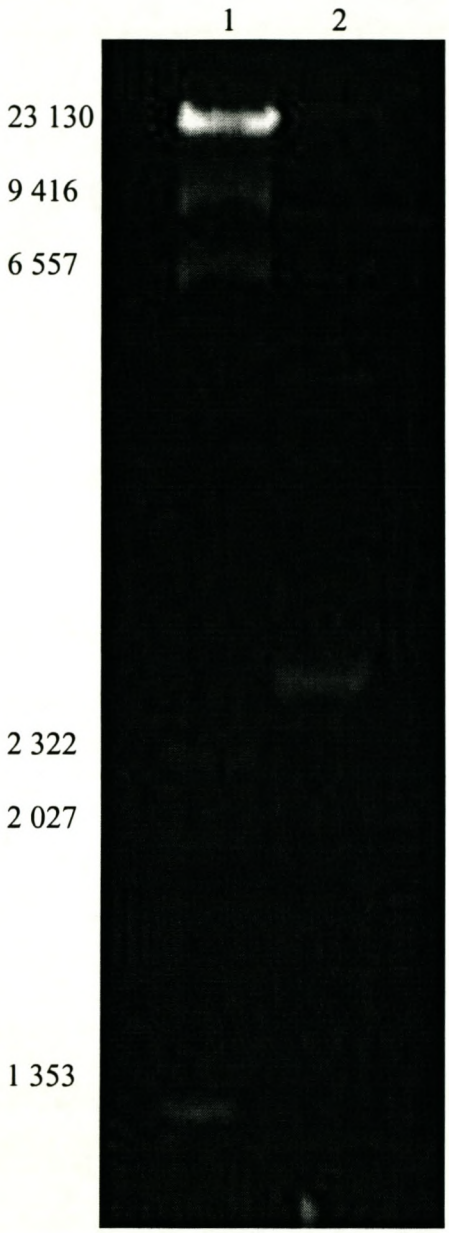


Fig. 24. Agarose gel eletrophoresis of the pUC18 vector. Lane 1, λDNA-Hind III molecular weight marker. Lane 2, digested pUC18 vector.

The DNA concentration was determined and a total yield of 26.4 μg was obtained, $OD_{260nm/280nm} = 1.83$. A total of 84% of DNA was recovered.

7.11 *C. RUMINANTIUM* DNA CLONING ANALYSIS

The *C. ruminantium* DNA was cloned into both the pUC_{18} and pMOS*Blue* vectors. A hundred and forty three clones were obtained. The clones were digested with Eco RI and Xba I as these enzymes have restriction sites within the MCS of both the pUC_{18} and pMOS*Blue* vector. Sixty percent of the clones had inserts. Figures 25 a-c represents clones obtained by cloning into the pMos*Blue* vector whereas Figures 25 d-f represents clones obtained by cloning into the pUC_{18} vector.

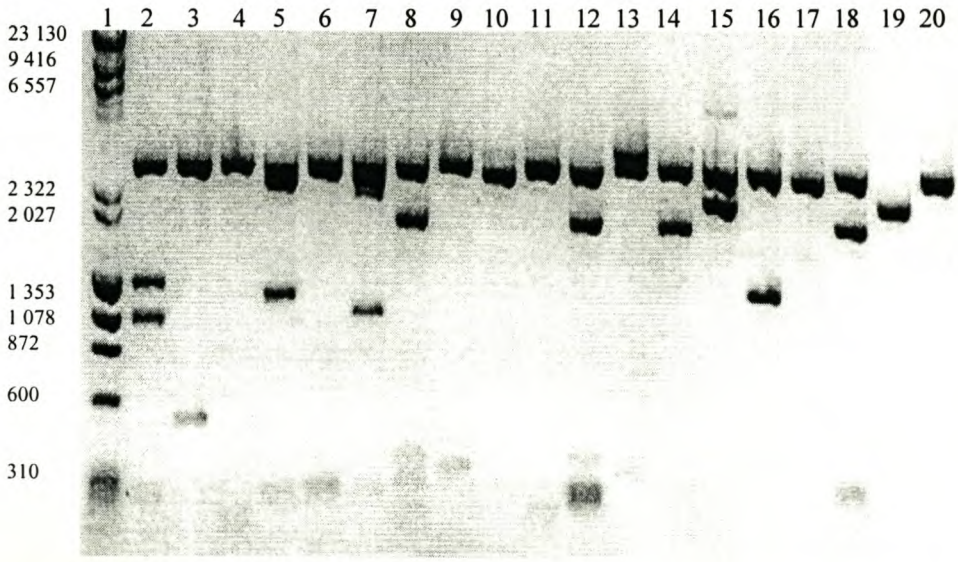


Fig. 25a. The restriction digestion reactions of the individual clones were analysed on a 0.7% agarose gel and visualised by ethidium bromide staining. Lane 1, λDNA-Hind III and ϕX174 DNA-Hae III molecular weight markers. Lanes 2 to 20, pMOS*Blue* clones digested with Eco RI and Xba I.

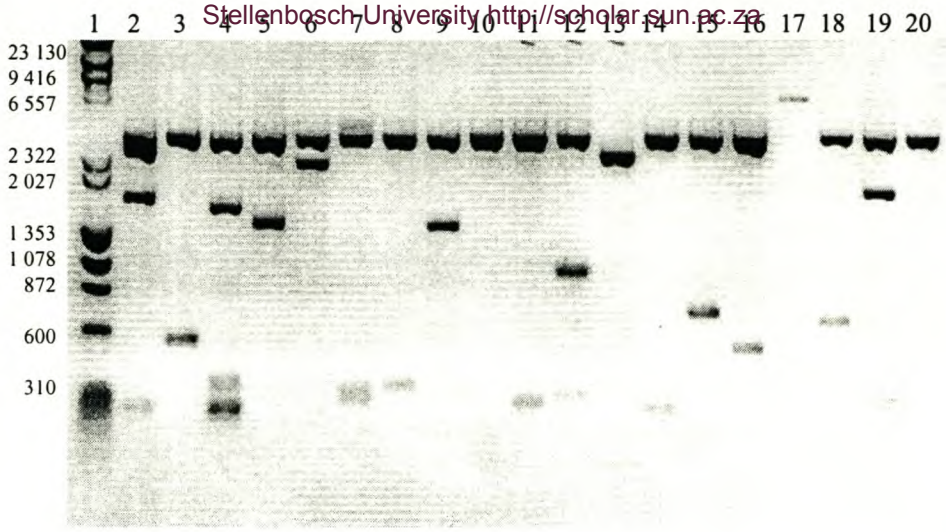


Fig. 25b. The restriction digestion reactions of the individual clones were analysed on a 0.7% agarose gel and visualised by ethidium bromide staining. Lane 1, λDNA-Hind III and ϕX174 DNA-Hae III molecular weight markers. Lanes 2 to 20, pMOS*Blue* clones digested with Eco RI and Xba I.

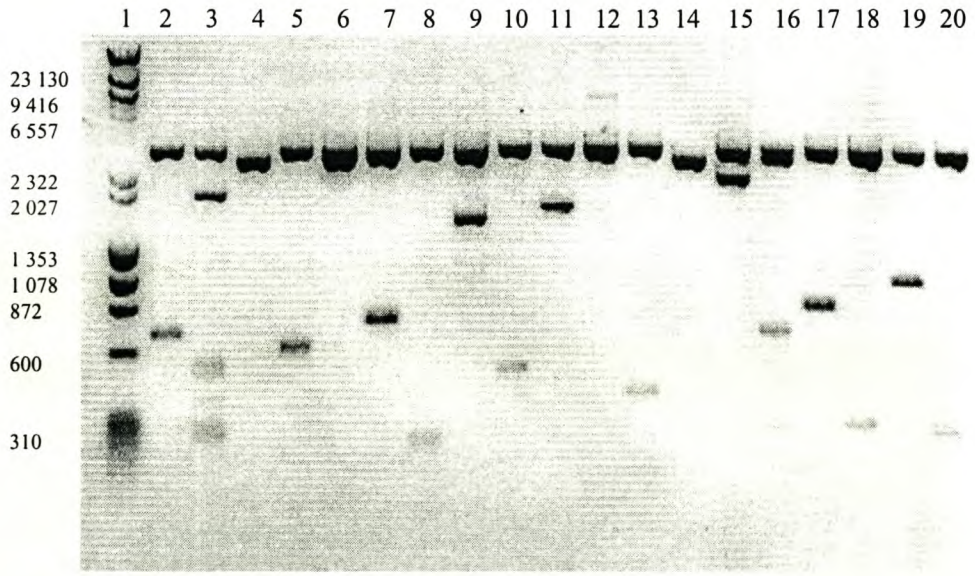


Fig. 25c. The restriction digestion reactions of the individual clones were analysed on a 0.7% agarose gel and visualised by ethidium bromide staining. Lane 1, λDNA-Hind III and ϕX174 DNA-Hae III molecular weight markers. Lanes 2 to 20, pMOS*Blue* clones digested with Eco RI and Xba I.

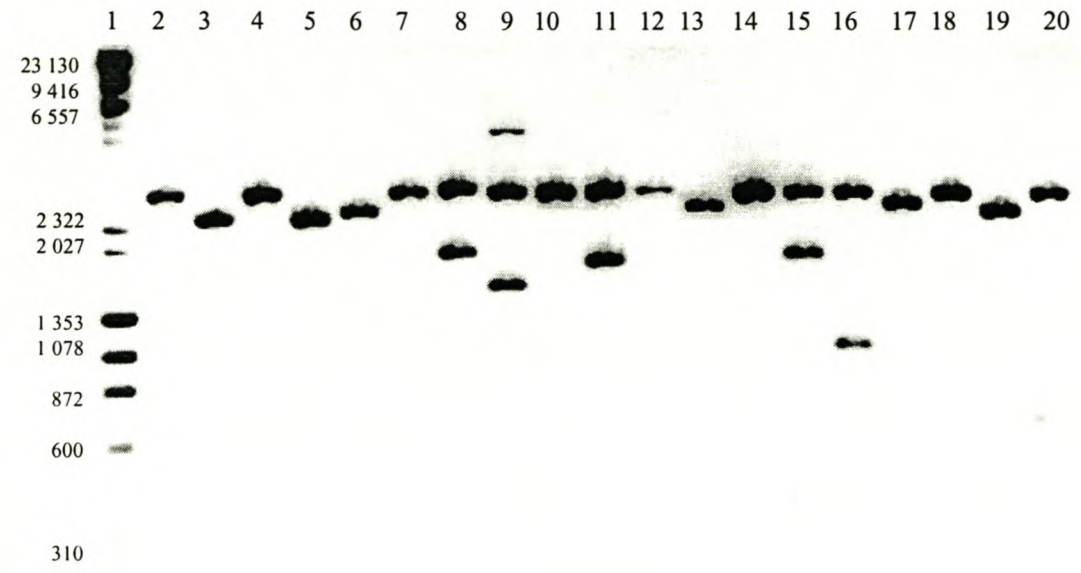


Fig. 25d. The restriction digestion reactions of the individual clones were analysed on a 0.7% agarose gel and visualised by ethidium bromide staining. Lane 1, λDNA-Hind III and ϕX174 DNA-Hae III molecular weight markers. Lanes 2 to 20, pUC_{18} clones digested with Eco RI and Xba I.

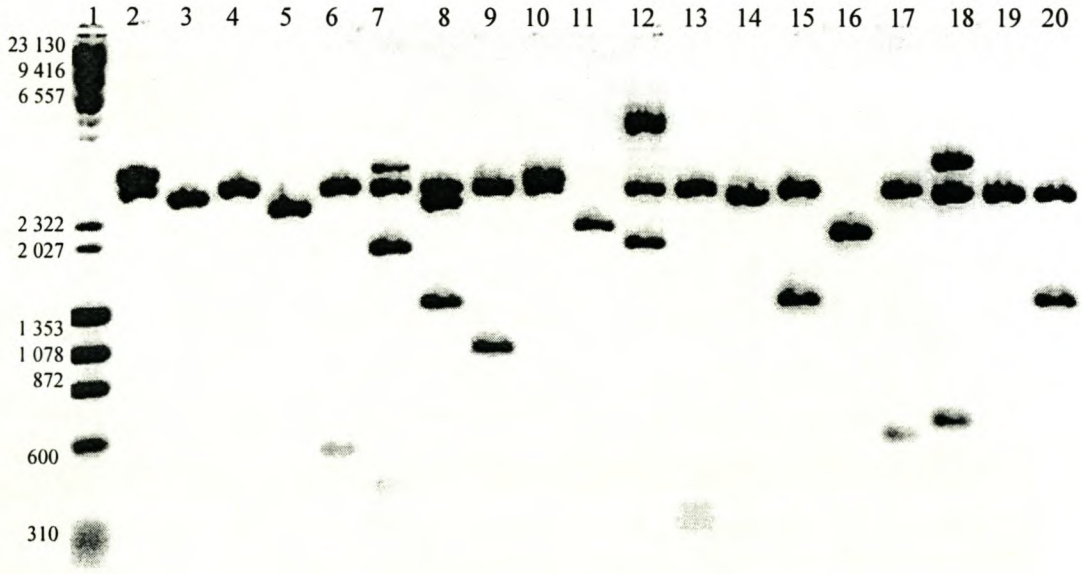


Fig. 25e. The restriction digestion reactions of the individual clones were analysed on a 0.7% agarose gel and visualised by ethidium bromide staining. Lane 1, λDNA-Hind III and ϕX174 DNA-Hae III molecular weight markers. Lanes 2 to 20, pUC_{18} clones digested with Eco RI and Xba I.

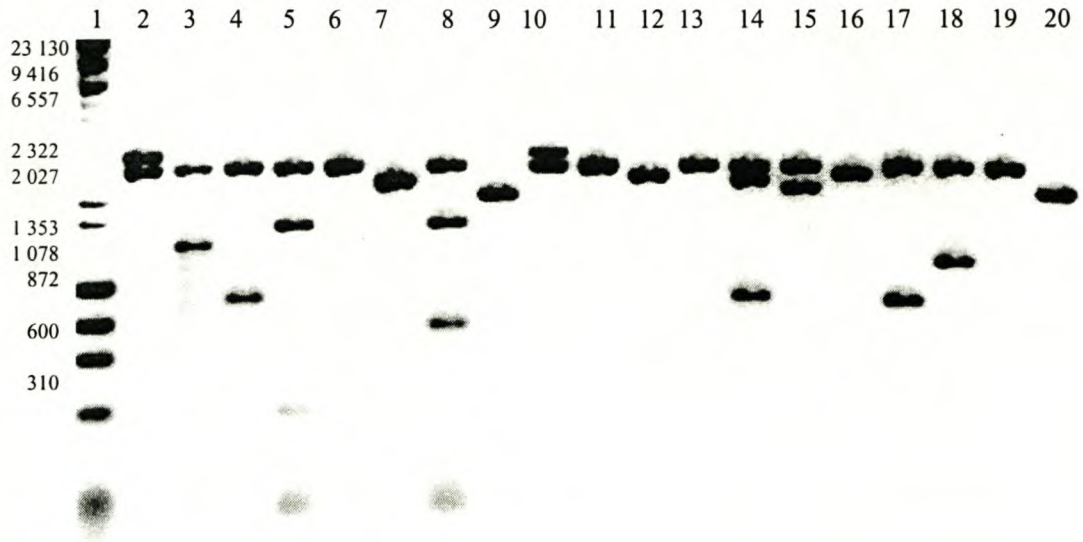


Fig. 25f. The restriction digestion reactions of the individual clones were analysed on a 0.7% agarose gel and visualised by ethidium bromide staining. Lane 1, λDNA-Hind III and ϕX174 DNA-Hae III molecular weight markers. Lanes 2 to 20, pUC_{18} clones digested with Eco RI and Xba I.

The inserts cloned into the two vectors ranged from 300 – 3 000 bp in length with the average insert length being 1 550 bp. Thus, although the electrophoresis of the nebulised fragments indicated that the bulk of these fragments were in the range of 6 000 to 15 000 bp (Fig. 23), smaller fragments were cloned. The reason why this is found may be because the smaller fragments occurred in such low concentrations that they were not visible on the electrophoresis. However, the sequences that will be presented hereafter, give clear evidence that smaller fragments of the *C. ruminantium* DNA were created during the nebulisation process.

7.12 SEQUENCING ANALYSIS OF pUC_{18} AND pMOS*Blue* CLONES

DNA sequence of a total length of 10 515 bp with a maximum read length of 30 000 bp was identified from 136 templates using the Staden Genome Assembly Program (Staden, 1998), including gap4 and nip4. Two contiguous sequences (contigs) were identified with contig 1 spanning 8 382 bp compiled from 92 sequence readings (Addendum 1, see page 116), and contig 2 spanning 2 135 bp compiled from 44 sequence readings (Addendum 2, see page 134).

Further analysis obtained two clones, clones 41 and 125, which were present in both contigs. Analysis of the DNA sequence of clone 125 showed that it started at position 6 739 and was 771 bp long. The size of the clone when analysed via restriction analysis was approximately 2 500 bp. This meant that the gap between the two contigs was approximately 856 bp. Analysis of the DNA sequence of clone 41 showed that it started at position 7 524 and was 776 bp long. The size of the clone when analysed via restriction analysis was approximately 858 bp. This meant that the gap between the two contigs was approximately 850 bp. Therefore the average size of the gap between the two contigs was approximately 850 bp. A diagrammatic presentation of the contigs within the *Cowdria* genome is shown in Figure 26.

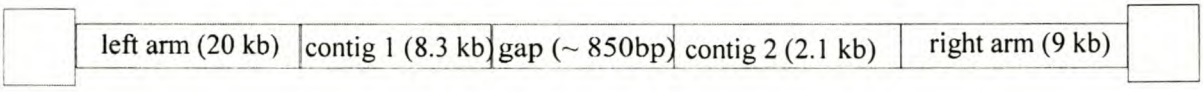


Fig. 26. A shematic representation of the two contigs representing the *Cowdria* DNA within the λGEM11 vector.

The two contigs obtained from the pUC_{18} and pMOS*Blue* clones were screened against the sequences obtained from the random sequencing of *C. ruminantium* λZAPII library in order to close the gap between the two contigs identified (Addendum 3, see page 143 and Addendum 4, see page 156). This made it possible to extend contiguous sequence 1 with 360 nucleotides at the 5' end but did not join the two contiguous sequences.

7.13 GAP CLOSURE

The gap between contigs 1 and 2 was closed by designing specific primers using the Intergrated DNA Technologies Inc. Oligoanalyser 2.5 [5] computer program. Typically, 20-mers with a GC content between 45 and 55 %, minimal secondary structure and no significant inter-primer complimentarity were selected. Primer WL2TP1_gapF was designed complementary to the 3' end of contig 1, and primer WL2TP1_gapR was designed complementary to the 5' end of contig 2:

WL2TP1_gapF: 5'- CAC TTA CAC CAA TGC CAC AC -3' primer: length = 20 -mer, GC content = 50.0 %, melting temperature = 55.2 %, molecular weight = 5975.0 g/mole,

[5] http://www.idtdna.com/program/oligocalc/oligocalc.asp

extinction coefficient = 186200 L/ (mole·cm), nmole/OD_{260nm} = 5.37 and µg/ OD_{260nm} = 32.09.

WL2TP1_gapR: 5'- TTA CCG CCA CCC TAA CAT ATA G -3' primer: length = 22 -mer, GC content = 45.5 %, melting temperature = 55.1 %, molecular weight = 6623.4 g/mole, extinction coefficient = 210300 L/ (mole·cm), nmole/OD_{260nm} = 4.76 and µg/ OD_{260nm} = 31.49.

These primers were used to amplify *Cowdria* genomic DNA as well as the phage miniprep DNA. A PCR product of ~ 700 bp was obtained (Fig. 27).

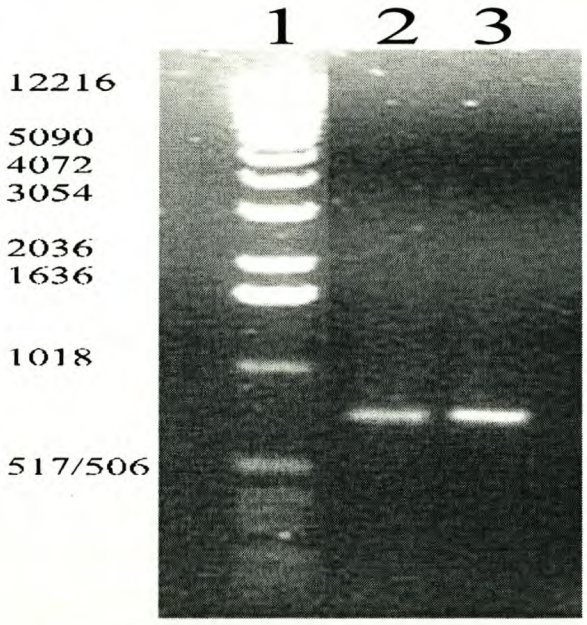


Fig. 27. PCR product obtained from the *Cowdria* genomic DNA and phage miniprep DNA visualised by electrophoresis on a 0.7% agarose gel stained with ethidium bromide. Lane 1, molecular weight marker X (Roche). Lane 2, PCR product obtained from the *Cowdria* genomic DNA. Lane 3, PCR product obtained from the phage miniprep DNA.

The PCR products obtained were sequenced as well as clones 41 and 125 (Addendum 5, see page 165). These clones were obtained from the cloning of the *C. ruminantium* DNA into the pUC_{18} and pMOS*Blue* vectors. The sequences of these clones were present in both contigs (contig 1 and 2) and would therefore span the gap between the two contigs.

The sequence closed the gap between the two contigs and a single contiguous sequence of 10 950 bp was obtained with a GC content of 26% (Addendum 6, see page 167).

The ORFs of this single contiguous sequence were identified using the ORF Finder program[4]. A total of 21 ORFs were identified with nine ORFs (ORF1- ORF9) showing homology with genes encoding known proteins (Fig. 28).

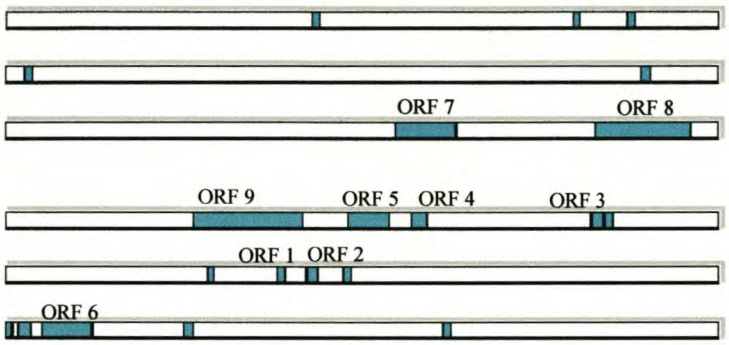


Fig. 28. Schematic diagram showing the positions of the ORFs frames in the single contiguous sequence with special reference to the ORFs showing homology to genes encoding known proteins (ORF1-ORF9).

Comparative sequence analysis was conducted using the BLAST[5] program. Public databanks, such as the National Centre for Biotechnological Information (NCBI) and nucleic acid (BLASTN) databases was screened for homologues in the sequence in order to identify possible putative *C. ruminantium* genes (Altschul *et al.*, 1990). A total of 21 putative protein-coding sequences were obtained. Twelve ORFs showed no significant homology with genes encoding known proteins. Two ORFs were less than 200 bp and had homology to genes encoding proteins of bacterial origin with unknown functions (Table 12). The remaining ORFs have homologies with existing genes as shown in Table 12.

[4] http://www.ncbi.nlm.nih.gov/gorf/
[5] http://www.ncbi.nlm.nih.gov/BLAST/

Table 12. Open reading frames identified and their homologies.

ORF	Position	Length (bp)	Closest Homology	Score	E-value
1	4 183-4 311	129	Unknown (*Pasteurella multocida*)	31	3.2
2	4 651-4 821	171	Unknown (*Rickettsia conorii*)	41	0.002
3	9 020-9 208	189	NADP-dependent alcohol dehydrogenase (*Acinetobacter* sp. M-1)	32	1.7
4	6 233-6 490	258	Hypothetical protein APE1082 (*Aeropyrum pernix* strain K_1)	48	2e-05
5	5 276-5 926	651	Dethiobiotin synthetase (*Caulobacter crescentus*)	107	1e-22
6	579-1 358	780	Prolipoprotein diacylglycerol transferase (*Rickettsia conorii*)	255	2e-67
7	6 009-6 938	930	Cell division protein FtsY (*Rickettsia conorii*)	288	6e-77
8	9 069-10 847	1 479	Putative NADH-ubiquinone oxidoreductase subunit (*Sinorhizobium meliloti*)	335	5e-91
9	2 903-4 585	1 683	Oligopeptide ABC transporter, permease protein (*Thermotoga maritime*)	38	0.23

Long repeats (100-200 bp) have been identified in both contigs (contig 1; 7 764-7 844 & 8 022-8 104 and contig 2; 2 418-2 445 & 2 661-2 689). Large repetitive sequences have been identified in a number of bacteria including *Mycoplasma* and *Rickettsia*. Repetitive DNA sequences are located either within or just before the start of a gene, or may be located within the promoter region. The function of these repeats is unknown but it is thought that they play a role in the evolutionary mechanisms that allow bacteria to adapt to environmental changes. These mechanisms include the generation of antigenic diversity and the regulation of gene expression, where the length of the repeat dictates transcription of a particular gene.

CHAPTER EIGHT
CONCLUSION

The aim of this study was to identify genes which could be used in DNA vaccines that could protect ruminants against *C. ruminantium* infection. A bacteriophage clone containing the *secD* gene (score 77.4 and E-value 3e-14) was isolated from a λGEM11 *C. ruminantium* DNA library by radioactively labelling the WL1M636 PCR product. The phage DNA (42 kb) was isolated from the bacteriophage clone and the *C. ruminantium* insert (12 kb) was amplified from the phage DNA and screened with *Mycoplasma,* bovine and *Cowdria* probes. The amplified DNA was fragmented by nebulisation and fragments of 6 000 – 15 000 bp were size selected and their ends repaired. The resulting blunt-ended fragments were subcloned into pUC_{18} and pMOS*Blue* vectors. A total of 153 clones were obtained and screened by restriction analysis to identify *C. ruminantium* clones containing inserts. A total of 93 *C. ruminantium* clones which ranged in size from 300 – 3 000 bp, were sequenced in both directions using the Sanger dideoxy chain termination method (Sanger *et al.*, 1977). Sequences were read using an Automated Sequencer, the ABI™ 377 (Applied Biosystems). Gel readings were assembled and edited using the Staden Genome Assembly Program (Staden, 1998). Two continuous sequences, contiguous sequence (contig) 1 of 8 382 bp and contig 2 of 2 135 bp, were identified with a short sequence of unidentified bases in between. Oligonucleotide primers were designed to amplify the DNA sequences between the two continuous sequences. A single contig of 10 950 bp was obtained. The entire sequence was analysed and open reading frames (ORFs) were identified using the ORF Finder program[1] and a total of 21 putative protein-coding sequences were obtained. BLAST[2] (Basic Local Alignment Search Tool) was used to screen public databanks for homologues of the sequences in order to identify putative *C. ruminantium* genes (Altschul *et al.*, 1990).

Five open reading frames of more than 300 bp were identified which showed homology to genes encoding specific proteins in bacteria. The ORFs had homology to genes which encoded the following proteins:

ORF 5: dethiobiotin synthetase

ORF 6: prolipoprotein diacylglygerol transferase

[1] http://www.ncbi.nlm.nih.gov/gorf/
[2] http://www.ncbi.nlm.nih.gov/BLAST/

ORF 7: cell division protein FtsY
ORF 8: putative NADH-ubiquinone oxidoreductase subunit
ORF 9: oligopeptide ABC transporter permease protein

Dethiobiotin synthethase belongs to the class of ATP-dependent carboxylases and catalyses the carboxylation of 7,8-diaminoperlargonic acid, leading to the formation of the ureido ring of biotin. Biotin is the vitamin essential for many biological carboxylation reactions such as the conversion of acetyl-coenzyme A (CoA) to malonyl-CoA in fatty acid biosynthesis. Biotin is essential for fatty acid biosynthesis, which is required for the formation of cell membranes.

Prolipoprotein diacylglycerol transferase is the precursor to lipoprotein diacylglycerol transferase. Lipoprotein diacylglycerol transferase catalyses the transfer of diacylglycerol from one molecule to another. Diacylglycerol acts as a second messenger in signal transduction and is therefore essential for cell communication.

FtsY plays an important role in protein secretion. Targeting of many polytopic proteins to the inner membrane of prokaryotes occurs via an essential signal recognition particle-like (SRP) pathway. The SRP pathway operates in the same manner as the protein secretory (Sec) pathway and differs only in that it is specific for hydrophobic signal anchor sequences. FtsY is a member of the GTPase superfamily. GTPases are extremely important in regulating membrane signalling pathways in all cells. FtsY is a homolog of the eukaryotic SRP particle receptor alpha-subunit that binds to membranes and plays a central role in membrane protein biogenesis. The secretory proteins SecY and SecE of the Sec pathway, forms a pore (SecYE) through the membrane and cotranslational targeting to the SecYE translocon is mediated by FtsY.

NADH-ubiquinone oxidoreductase catalyses the oxidation of NADH by ubiquinone which is coupled specifically to proton translocation across the membrane. The NADH-ubiquinone oxidoreductase is required for proton translocation, maintaining a constant pH in the cell, which is essential for cell survival.

The ATP-binding cassette (ABC) transporter superfamily is one of the most widespread of all gene families and currently has in excess of 1 100 members in organisms ranging from Archea to man. It forms a transport pathway in which transport can be inwards across the

membrane or outwards to the cell exterior. The ABC transporter protein is essential for transporting proteins across membranes thereby transporting proteins between cell compartments and removing harmful proteins out of the cells.

Although a PCR product containing the *secD* gene was used to isolate the clone from the λGEM11 library, no homologous sequence to this gene was found in the contiguous sequence identified in this thesis. This was a very puzzling result, as it was assumed that the regions flanking the *secD* gene as they had been cloned into the λZAPII vector originally (the preparation of which was not part of this thesis) would be identical to those contained in the *C. ruminantium* genome. However, only the last 348 bp of the WL1M636 PCR product was identified in the contiguous sequence (Addendum 6, see page 167) and thereafter the sequence of the WL1M636 PCR product differed from that of the contiguous sequence. Hybridisation of the WL1M636 probe to the λGEM11 library was therefore due to the homology of the last 348 bp sequence contained within the WL1M636 probe and not due to the hybridisation of the probe to the *secD* gene sequence which would explain why the *secD* gene could not be identified within the contiguous sequence. Clearly, an error must have occurred during the preparation of the WL1M636 clone during the original cloning of the *secD* gene. This WL1M636 clone was prepared by blunt end cloning of a *secD* PCR product into the λZAPII vector. The only plausible explanation for this result is that during this process, an additional piece of DNA, i.e. that corresponding to the last 348 bp of the WL1M636 PCR product, must also have been ligated into the λZAPII vector. Although this result has considerable consequences for this study, this problem could not have been envisaged at the beginning of the study. Clearly, the approach should rather have been to design primers to amplify only the *secD* sequence for use as a probe. This strategy should also be employed in future strategies to screen the λGEM11 library for the *secD* sequence.

The five ORFs show significant homology with genes encoding proteins essential for the normal physiological functioning of an organism and therefore essential to the survival of the organism. The ORFs therefore have potential to be used in a vaccine. Future strategies would be to clone these ORFs into the pCMViUB vector, a genetic vaccine vector, which is planned as a next stage of this project. These constructs will be used to immunise sheep by intramuscular injection and gene gun inoculation. The immunisation will be done three times at three-week intervals, and the sheep will be challenged five weeks after the third inoculation with a lethal dose of the pathogenic *C. ruminantium* Welgevonden isolate. Lymphocyte

proliferation assays will be performed on PBMCs collected from sheep before immunisation and one week before challenge. Cells will be stimulated by *Cowdria* lysate (positive antigen) and uninfected bovine EC lysate (negative antigen). On day three of the of the culture period, proliferation will be estimated by measurement of [^{3}H] thymidine uptake. Results will be expressed as stimulation index (SI) (counts per minute (cpm) of positive antigen divided by cpm of negative antigen) averaged from triplicate plates.

The DNA constructs prepared using these ORFs can also be used to genetically engineer the organism by mutating the genes for these ORFs. The mutated genes would produce non-functional proteins, which would interfere with the normal cell processes. The mutated organisms could be used to immunise susceptible animals thereby exposing the animals to a non-viable *C. ruminantium* strain, which would induce an immune response without causing the disease and thereby confer protection against challenge with virulent *C. ruminantium* isolates.

Through the sequencing of this portion of the *Cowdria* genome and the identification of these ORFs this study has laid the groundwork for future studies into the development of effective vaccines against *Cowdria ruminantium*.

CHAPTER NINE
REFERENCES

Adams L.B., Hibbs J.B. Jr., Taintor R.R., and Krahenbuhl J.L., **1990**. Microbiostatic effect of murine activated macrophages for *Toxoplasma gondii*: role of synthesis of inorganic nitrogen oxides from L-arginine. *Journal of Immunology*, 144: 2725-2729.

Adams L.B., Franzblau S.G., Vavrin Z., Hibbs J.B., Jr., and Krahenbuhl J.L., **1991**. L-Arginine-dependent macrophage effector functions inhibit metabolic activity of *Mycobacterium leprae*. *Journal of Immunology*, 147: 1642-1646.

Alarcon J.B., Waine G.W., and McManus D.P., **1999**. DNA vaccines: Technology and application as anti-parasite and anti-microbial agents. *Advances in Parasitology,* 42: 343-410.

Alderink F.J., and McCauley E.H., **1988**. The Probability of spread of *Amblyomma hebraeum* in the Caribbean. *Preventive Veterinary Medicine*, 6: 285-294.

Alexander R.A., **1931**. Heartwater. The present state of our knowledge of the disease. *17th Report of the Director of Veterinary Services and Animal Industry*. Union of South Africa, 89-149.

Allsopp B.A., Allsopp M.T.E.P., du Plessis J.L., and Visser E.S., **1996**. Uncharacterised *Ehrlichia* spp. may contribute to clinical heartwater. *Annals of the New York Academy of Sciences*, 791: 17-24.

Allsopp M.T.E.P., Visser E.S., du Plessis J.L., Vogel S.W., and Allsopp B.A., **1997**. Different organisms associated with heartwater as shown by analysis of 16S ribosomal RNA gene sequences. *Veterinary Parasitology,* 71: 283-300.

Allsopp M.T.E.P., Hattingh C.M., Vogel S.W., and Allsopp B.A., **1999**. Evaluation of 16S, *map1* and pCS20 probes for the detection of *Cowdria* and *Ehrlichia* species. *Epidemiology and Infection*, 122: 323-328.

Allsopp B.A., **2000**. Biannual Heartwater Research Programme Review. Agricultural Research Council. Onderstepoort Veterinary Institute.

Allsopp M.E.T.P., and Allsopp B.A., **2001**. A novel *Ehrlichia* detected in dogs in South Africa. *Onderstepoort Journal of Veterinary Research*. Submitted.

Altschul S.F., Gish W., Miller W., Myers E.W., and Lipman D.J., 1990. Basic local alignment search tool. *The Journal of Molecular Biology*, 215: 403-410.

Andrew H.R., and Norval R.A.I., **1989**. The carrier state of sheep, cattle and African buffalo recovered from heartwater. *Veterinary Parasitology*, 34: 261-266.

Anon, **1996**. Livestock census performed by the Directorate of Animal Health, Pretoria, South Africa.

Arendt W.J., **1988**. Range expansion of the cattle egret (*Bubulcus ibis*) in the greater Caribbean basin. Colon. *Waterbirds*, 11: 252-262.

Barbet A.F., Semu S.M., Chigagure N., Kelly P.J., Jongejan F., and Mahan S.M., **1994**. Size variation of the major immunodominant protein of *Cowdria ruminantium*. *Clinical and Diagnostic Laboratory Immunology*, 1: 744-746.

Barré N., Uilenberg G., Morel P.C., and Camus E., **1987**. Danger of introducing heartwater on to the American mainland: potential role of indigenous and exotic *Amblyomma* ticks. *The Onderstepoort Journal of Veterinary Research*, 54: 405-417.

Barré N., Garris G.I., Borel G., and Camus E., **1988**. Host and population dynamics of *Amblyomma variegatum* (Acari: Ixodidae) on Guadeloupe, French West Indies. *Journal of Medical Entomology*, 25: 111-115.

Bekker C.P.J., Vink D., Pereira C.M.L., Wapenaar W., Langa A., and Jongejan F., **2000**. Heartwater (Cowdria ruminantium infection) as a cause of post-restocking mortality of goats in Mozambique. *Clinical and Diagnostic Laboratory Immunology*, Submitted.

Bell-Sakyi L., Koney E.B., Dogbey O., and Sumption K.J., **1996**. Heartwater in Ghana: implications for control of ticks. *Tropical Animal Health and Production*, 59S-64S.

Bezuidenhout, J.D., **1981**. The development of a new heartwater vaccine, using the *Amblyomma hebraeum* nymphs infected with *Cowdria ruminantium*. In: Whitehead (G.B.), Gibson (J.D.), Eds. Proceedings of an International Conference on Tick Biology and Control, Grahamstown, SA., Rhodes University. 33-39.

Bezuidenhout J.D., **1985a**. The epizootiology and control of heartwater and other tick-borne diseases of cattle in South Africa. *The Onderstepoort Journal of Veterinary Research*, 52: 211-214.

Bezuidenhout J.D., Paterson C.L., and Barnard B.J.M., **1985b**. *In vitro* cultivation of *Cowdria ruminantium*. *The Onderstepoort Journal of Veterinary Research*, 52: 113-120.

Bezuidenhout J.D., and Oberem P.T., **1985c**. New vaccines against heartwater. *South African Journal of Science*, 81: 2-3.

Bezuidenhout J.D., **1987**. Natural transmission of Heartwater. *The Onderstepoort Journal of Veterinary Research*, 54: 349-351.

Birnie E.F., Burridge M.J., Camus E., and Barré N., **1985**. Heartwater in the Caribbean: isolation of *Cowdria ruminantium* from Antigua. *The Veterinary Record,* 116: 121-123.

Booccvar K.S., Maybodi M., Washington M.K., Hibbs J.B. Jr., and Kurlander R.L., **1994**. Nitric oxide produced during murine listeriosis is protective. *Infection and Immunity*, 62: 1089-1100.

Bourdoulous S., Bensaid A., Martinez D., Sheikboudou C., Trap I., Strosberg A., and Couraud P., **1995**. Infection of bovine brain microvessel endothelial cells with *Cowdria ruminantium* elicits IL-1β, -6 and –8 mRNA production and expression of unusual MHC class II DQα transcript. *Journal of Immunology*. 154: 4032-4038.

Bowie M.V., Reddy G.R., Semu S.M., Mahan S.M., and Barbet A.F., **1999**. Potential value of major antigenic protein 2 for serological diagnosis of heartwater and related Ehrlichial infections. *Clinical and Diagnostic Laboratory Immunology*, 6: 209-215.

Brayton K.A., Fehrsen J., de Villiers E.P., van Kleef M., and Allsopp B.A., **1997**. Construction and initial analysis of a representative λZAPII expression library of the intracellular rickettsia *Cowdria ruminantium*: cloning of *map*1 and three other *Cowdria* genes. *Veterinary Parasitology*, 72: 185-199.

Brayton K.A., van der Walt M., Vogel S.W., and Allsopp B.A., **1998**. A partially protective clone from *Cowdria ruminantium* identified by using a Salmonella vaccine deliverysystem. *Annals of the New York Academy of Sciences*, 849: 247-252.

Brayton K.A., de Villiers E.P., Fehrsen J., Nxomani C., Collins N.E., and Allsopp B.A., **1999**. *Cowdria ruminantium* DNA is unstable in a supercos1 library. *The Onderstepoort Journal of Veterinary Research*, 66: 111-117.

Burridge M.J., **1985**. Heartwater invades the Caribbean. *Parasitology Today*. 1: 1-3.

Busse R., Fleming I., and Schini V.B., **1995**. Nitric oxide formation in vascular wall: regulation and functional implications. *Microbiology and Immunology,* 196: 7-8.

Byrne G.I., and Turco J., **1988**. Interferon and nonviral pathogens. *Immunology Series* Vol 42. (Byrne G.I., and Turco J., eds.) Marcel Dekker. 250-317.

Byrom B. and Yunker C.E., **1990**. Improved culture conditions of *Cowdria ruminantium* (*Rickettsiales*), the agent of heartwater disease of domestic ruminants. *Cytotechnology*, 4: 285-290.

Byrom B., Yunker C.E., Donovan P.L., and Smith G.E., **1991**. *In vitro* isolation of *Cowdria ruminantium* from plasma of infected ruminants. *Veterinary Microbiology*, 26: 263-268.

Byrom B., Mahan S.M., and Barbet A.F., **1993**. The development of antibody to *Cowdria ruminantium* in mice and its role in heartwater disease. *Revue d'elevage et de Medecine Veterinaire des Pays Tropicaux,* 46: 197-201

Camus E., and Barré N., **1987**. Epidemiology of heartwater in Guadeloupe and the Caribbean. *The Onderstepoort Journal of Veterinary Research*, 54: 419-426.

Camus E., and Barré N., **1988**. Heartwater. A Review. *Office of International Epizootics*, Paris, 147.

Camus E., and Barré N., **1992**. The role of *Amblyomma variegatum* in the transmission of heartwater with special reference to Guadeloupe. *Annals of the New York Academy of Sciences*, 653: 33-41.

Camus E., **1996**. Heartwater (Cowdriosis): A review (2nd Ed), Office International des Epizooties (OIE).

Carl M., Robbins F., Hartzman R.J., and Dasch A., **1987**. Lysis of cells infected with typhus group rickettsiae by human cytotoxic T cell clone. *Journal of Immunology*, 139: 4203-4207.

Clark R., **1962**. The pathological physiology of heartwater [Cowdria (Rickettsia) *ruminantium* Cowdry 1926]. *The Onderstepoort Journal of Veterinary Research*, 29: 25-33.

Corn J.L., Barre N, Thiebot B., Creekmore T.E., Garris G.I., and Nettles V.F., **1993**. Potential role off cattle egrets, *Bubulcus ibis* (Ciconiformes: Ardeidae), in the dissemination of *Amblyomma variegatum* (Acari: Ixodidae) in the Eastern Caribbean. *Journal of Medical Entomology*, 30: 1029-1037.

Coutinho G.C., Durieu-Trautmann O., Strosberg A.D., and Courand P.O., **1991**. Catecholamines stimulate the IFN-γ-induced class II MHC expression on bovine brain capillary endothelial cells. *Journal of Immunology*. 147: 2525-2529.

Cowdry E.V., **1925**. Studies on the etiology of heartwater. I. Observation of a rickettsia, *Rickettsia ruminantium* (n.sp.) in the tissue of infected animals. *The Journal of Experimental Medicine*, 42: 231-252.

Cowdry E.V., **1926**. Cytological studies on heartwater. I. The observations of *Rickettsia ruminantium* in the tissues of infected animals. In: 11^{th}-12^{th} Report of the Director of Education and Veterinary Research Union of South Africa. 161-177.

Crosby G.T., **1972**. Spread of the cattle egret in the Western Hemisphere. *Bird-banding*. 43: 205-212.

Dame J.B., Mahan S.M., and Yowell C.A., **1992**. Phylogenetic relationship of *Cowdria ruminantium*, agent of heartwater, to *Anaplasma marginale* and other members of the order *Rickettsiales* determined on the basis of 16S rRNA sequence. *International Journal of Systematic Bacteriology*, 42: 270-274.

de Villiers E.P., Brayton K.A., Zweygarth E.P., and Allsopp B.A., **2000**. Genome size and genetic map of *Cowdria ruminantium*. *Microbiology*, 146: 2627-2634.

de Villiers E.P., **2001**. Genome Analysis of *Cowdria ruminantium*. In fulfullment of a PhD Thesis.

de Vries N., Mahan S.M., Ushewokunze-Obatolu U., Norval R.A.I., and Jongejan F., **1993**. Correlation between antibodies to *Cowdria ruminantium* (Rickettsiales) in cattle and the distribution in of *Amblyomma* vector ticks in Zimbabwe. *Experimental and Applied Acarology*, 17: 799-810.

Dinsmore J.J., **1973**. Foraging success of cattle egrets, *Bubulcus ibis*. *American Midl. Nature*, 89: 242-246.

du Plessis J.L., **1970a**. Pathogenesis of heartwater. I. *Cowdria ruminantium* in the lymph nodes of ruminants. *The Onderstepoort Journal of Veterinary Research*, 37: 89-96.

du Plessis J.L., **1970b**. Immunity in heartwater. I. Preliminary note on the role of serum antibodies. *The Onderstepoort Journal of Veterinary Research*, 37: 147-150.

du Plessis J.L., and Kümm N.A.L., **1971**. The passage of *Cowdria ruminantium* in mice. *Journal of the South African Veterinary Medical Association*, 42: 217-221.

du Plessis J.L., **1975a**. Electron microscopy of *Cowdria ruminantium* infected reticulo-endothelial cells of the mammalian host. *The Onderstepoort Journal of Veterinary Research*, 42: 1-14.

du Plessis J.L., **1975b**. Histopathological studies of the pathogenesis of heartwater as manifested in mice infected with a strain of *Cowdria ruminantium*. *Pretoria SA, University of Pretoria.*

du Plessis J.L., **1981**. The application of the indirect fluorescent antibody test to the serology of heartwater, In: Whitehead GB, Gibson JD (eds.), Tick biology and control, Tick Research Unit, Rhodes University. 47-52.

du Plessis J.L., Jansen B.C., and Prozesky L., **1983**. Heartwater in Angora goats I. Immunity subsequent to artificial infection and treatment. *The Onderstepoort Journal of Veterinary Research*, 50: 137-143.

du Plessis J.L., Bezuidenhout J.D., and Lüdemann C.J.F., **1984a**. The immunisation of calves against heartwater: subsequent immunity both in the absence and presence of natural tick challenge. *The Onderstepoort Journal of Veterinary Research*, 51: 193-196.

du Plessis J.L., **1984b**. Colostrum-derived antibodies to *Cowdria ruminantium* in the serum of calves and lambs. *The Onderstepoort Journal of Veterinary Research*, 51: 275-276.

du Plessis J.L., **1985**. A method for determining the *Cowdria ruminantium* infection rate in *Amblyomma hebraeum*: effects in mice injected with tick homogenates. *The Onderstepoort Journal of Veterinary Research*, 52: 55-61.

du Plessis J.L., Malan L., and Kowalski Z.E., **1987a**. The pathogenesis of heartwater. *The Onderstepoort Journal of Veterinary Research*, 54: 313-318.

du Plessis J.L., and Malan L., **1987b**. The application of the indirect fluorescent antibody test in research on heartwater. *The Onderstepoort Journal of Veterinary Research*, 54: 319-325.

du Plessis J.L., Camus E., Oberem P.T., and Malan L., **1987c**. Heartwater serology: some problems with the interpretation of results. *The Onderstepoort Journal of Veterinary Research*, 54: 327-329.

du Plessis J.L., and Malan L., **1987d**. The non-specific resistance of cattle to heartwater. *The Onderstepoort Journal of Veterinary Research*, 54: 333-336.

du Plessis J.L., van Gas L., Olivier J.A., and Bezuidenhout J.D., **1989**. The heterogenecity of *Cowdria ruminantium* stocks: cross-immunity and serology in sheep and pathogenicity in mice. *The Onderstepoort Journal of Veterinary Research*, 56: 195-201.

du Plessis J.L., **1990**. Increased pathogenicity of an *Ehrlichia*-like agent after passage through *Amblyomma hebraeum*: a preliminary report. *The Onderstepoort Journal of Veterinary Research,* 57: 233-237.

du Plessis J.L., Berche P., and van Gas L., **1991**. T cell-mediated immunity to *Cowdria ruminantium* in mice: the protective role of LYT-2+ T cells. *The Onderstepoort Journal of Veterinary Research*, 58: 171-179.

du Plessis J.L., Bezuidenhout J.L., Brett M.S., Camus E., Jongejan F., Mahan S.M., and Martinez D., **1993a**. The sero-diagnosis of heartwater: a comparison of five tests. *Revue d'elevage et de Medecine Veterinaire des Pays Tropicaux,* 46: 123-129.

du Plessis J.L., **1993b**. The relationship between Cowdria and Ehrlichia: change in the bahaviour of ehrlichial agents passaged through *Amblyomma hebraeum*. *Revue d'elevage et de Medecine Veterinaire des Pays Tropicau,* 46: 131-143.

Ellis R.W., **1999**. New technologies for making vaccines. *Vaccine*, 17: 1596-1604.

Estrada-Pena A., Castella J., and Morel P.C., **1994**. Cuticular hydrocarbon composition, phenotypic variability, and geographic relationships in allopatric populations of *Amblyomma variegatum* (Acari: Ixodidae) from Africa and the Caribbean. *Journal of Medical Entomology*, 31: 534-544.

Flach E.J., Woodford J.D., Morzaria S.P., Dolan T.T., and Shambwana I., **1990**. Identification of *Babesia bovis* and *Cowdria ruminantium* on the island of Ungunja, Zanzibar. *The Veterinary Record,* 126: 57-59.

Furchgott R.F., and Zawadski J.V., **1980**. The obligatory role of endothelial cells in the relaxation of arterial smooth muscle by acetylcholine. *Nature*, 288: 373-376.

Goff W.L., Johnson W.C., Wyatt C.R., and Cluff C.W., **1996**. Assessment of bovine mononuclear phagocytes and neutrophils for induced L-arginine-dependent nitric oxide production. *Veterinary Immunology and Immunopathology*, 55: 45-62.

Haig, D.A., **1952**. Tick-borne rickettsiosis in South Africa. *Advances in Veterinary Science*, 2: 307-325.

Hancock J., and Kushlan J., **1984**. The herons handbook. Harper and Row, New York.

Hanson B., **1991**. Susceptibility of *Rickettsia tsustsugamushi* Gilliam to gamma interferon in cultured mouse cells. *Infection and Immunity*, 59: 4125-4133.

Heatwole H., **1965**. Some aspects of the association of cattle egrets with cattle. *Animal Behaviour,* 13: 79-83.

Helland D.E., and Hey A.W., **1990**. A short review of the biotechnological methods of relevance to modern vaccine development. *Scandinavian Journal of Infectious diseases*, Supplement 76: 32-38.

Henry E.T., Norman B.B., Fly D.E., Wichmann R.W., and York S.M., **1983**. Effects and use of a modified live *Anaplasma marginale* vaccine in beef heifers in California. *Journal of the American Veterinary Medical Association*, 183: 66-69.

Holt and Krieg, **1979**. *Bergey's Manual of Systematic Bacteriology*. Volume I. Section 9.

Ilemobade A.A. and Leeflang P., **1977**. Epidemiology of heartwater in Nigeria. *Revue d'elevage et de Medecine Veterinaire des Pays Tropicau,* 30: 149-155.

Jackson C., and Neitz W.O., **1932**. On the etiology of heartwater, 49-70. 18th Report. Director of Veterinary Services and Animal Industry, Onderstepoort. Union of South Africa.

Janeway C.A., Jones B., and Hayday A., **1993**. Specificity and function of T cells bearing γδ receptors. *Immunology Today*, 9: 93-96.

Jerrels T.R., Turco J., Winkler H.H., and Spitalny G.L., **1986**. Neutralisation of lymphokine-mediated antirickettsial activity of fibroblasts and macrophages with monoclonal antibody specific for murine interferon gamma. *Infection and Immunity*, 51: 355-359.

Jongejan F., Morzaria S.P., Shariff O.A., and Abdalla H.M., **1984**. Isolation and transmission of *Cowdria ruminantium* (causal agent of heartwater disease) in Blue Nile Province, Sudan. *Veterinary Research Communications*, 8: 141-145.

Jongejan F., Uilenberg G., Franssen F.F.J., Gueye A., and Meuwe Huys J., **1988**. Antigenic differences between stocks of *Cowdria ruminantium*. *Research in Veterinary Science*, 44: 186-188.

Jongejan F., and Thielemans M.J.C., **1989a**. Identification of an immunodominant antigenically conserved 32-kilodalton protein from *Cowdria ruminantium*. *Infection and Immunity*, 57:3243-3246.

Jongejan F., Wassink L.A., Thielemans M.J.C., Perie N.M., and Uilenberg G., **1989b**. Serotypes in *Cowdria ruminantium* and their relationship with the *Ehrlichia phagocytophila* determined by immunofluorescence. *Veterinary Microbiology,* 21: 31-41.

Jongejan F., Thielemans M.J.C., de Groot M., van Kooten J.S., and van der Zeijst B.A.M., **1991a**. Competitive enzyme-linked immunosorbent assay for heartwater using monoclonal antibodies to a *Cowdria ruminantium* specific 32 kilodalton protein. *Veterinary Microbiology*, 28: 199-211.

Jongejan F., Thielemans M.J.C., Brière C., and Uilenberg G., **1991b**. Antigenic diversity of *Cowdria ruminantium* isolates determined by cross-immunity. *Research in Veterinary Science*, 51: 24-28.

Jongejan F., **1991c**. Protective immunity to heartwater (*Cowdria ruminantium* infection) is acquired after vaccination with *in vitro*-attenuated rickettsiae. *Infection and Immunity*, 59: 729-731.

Jongejan F., Zandbergen T.A., van de Wiel P.A., de Groot M., and Uilenberg G., **1991d**. The tick-borne rickettsia Cowdria ruminantium has a Chlamydia-like development cycle. *The Onderstepoort Journal of Veterinary Research,* 58: 227-237.

Jongejan F., **1992**. Experimental transmission of *Cowdria ruminantium* (Rickettsiales) by the American reptile tick *Amblyomma dissimile* Koch, 1844. *Experimental and Applied Acarology*, 15: 117-121.

Jongejan F., de Vries N., Nieuwenhuijs J., van Vliet A.H.M., and Wassink L.A., **1993**. The immunodominant 32-kilodalton protein of *Cowdria ruminantium* is conserved within the genus *Ehrlichia*. *Revue d'elevage et de Medecine Veterinaire des Pays Tropicaux,* 46: 145-152.

Jongejan F., and Uilenberg G., **1994**. Ticks and control methods. *Revue scientifique et technique*, 13: 1201-1226.

Jongejan F., and Bekker P.J., **1999**. *Cowdria ruminantium*: recent developments in diagnostic methods, molecular characterisation and vaccines. In rickettsae and rickettsial diseases at the turn of the third millennium, pp. 373-386. Edited by Raoult D., and Brouqui P., Paris: Elsevier.

Katz J.B., Barbet A.F., Mahan S.M., Kumbula D., Lockhart J.M., Keel M.K., Dawson J.E., Olson J.G., and Ewing S.A., **1996**. A recombinant antigen from the heartwater agent (CR) reactive with antibodies in some southeastern United States white-tailed deer (*Odocoileus virginianus*), but not cattle, sera. *Journal of Wildlife Diseases,* 32: 424-430.

Katz J.B., DeWald R., Dawson J.E., Camus E., Martinez D., and Mondry R., **1997**. Development and evaluation of a recombinant antigen, monoclonal antibody-based competitive ELISA for heartwater serodiagnosis. *Journal of Veterinary Diagnostic Investigation,* 9: 130-135.

Kaufmann S.H.E., **1993**. Immunity to intracellular bacteria. *Annual Review of Immunology,* 11: 129-163.

Kelly P.J., Matthewman L.A., Mahan S.M., Semu S., Peter T., Mason P.R., Brouqui P., and Raoult D., **1994**. Serological evidence for antigenic relationships between *Ehrlichia canis* and *Cowdria ruminantium*. *Research in Veterinary Science*, 56: 170-174.

Kocan K.M., and Bezuidenhout J.D., **1987**. Morphology and development of *Cowdria ruminantium* in *Amblyomma* ticks. *The Onderstepoort Journal of Veterinary Research*, 54: 177-182.

Kocan K.M., Bezuidenhout J.D., and Hart A., **1987a**. Demonstration and ultrastructural features of *Cowdria ruminantium* in midgut epithelial cells and salivary glands of nymphal *Amblyomma hebraeum. The Onderstepoort Journal of Veterinary Research*, 54: 87-92.

Kocan K.M., Morzaria S.P., Voigt W.P., Kiarie J., and Irvin A.P., **1987b**. Demonstration of colonies of *Cowdria ruminantium* in midgut epithelial cells of *Amblyomma variegatum. American Journal of Veterinary Research,* 48: 356-360.

Kock N.D., van Vliet A.H.M., Charlton K., and Jongejan F., **1995**. Detection of *Cowdria ruminantium* in blood and bone marrow samples from clinically normal, free-ranging Zimbabwean wild ungulates. *Journal of Clinical Microbiology*, 33: 2501-2504.

Lally N.C., Nicoll S., Paxton E.A., Cary C.M., and Sumption K.J., **1995**. The *Cowdria ruminantium* groE operon. *Microbiology*, 141: 2091-2100.

Lefkowitz D.L., Mills K.C., Moguilevsky N., Bollen A., Vaz A., and Lefkowitz S.S., **1993**. Regulation of macrophage by human recombinant myeloperoxidase. *Immunology Letters,* 36: 43-50.

Lieser M.J., Kozol R.A., and Tennenerg S.D., **1995**. Interferon-gamma primes neutrophil-mediated gastric surface cell cytotoxicity. *American Journal of Physiology,* 268: G843-G848.

Liew F.Y., Li Y., Severa A., Millott S., Schmidt J., Salter M., and Moncada S., **1991**. Regulation of lymphocyte functions by nitric oxide. *Current Opinion in Immunology,* 7: 396-399.

Logan L.L., Endris R.G., Birnie E.F., and Mebus C.A., **1985**. Research objectives to improve U.S. diagnostic capabilities for heartwater disease. Proceedings of 89th Annual Meeting, United States Animal Health Association, Milwaukee, Wisconsin, USA.

Logan, L. L., Holland C. J., Mebus C.A., and Ristic M., **1986**. Serological relationship between *Cowdria ruminantium* and certain *Ehrlichia spp. The Veterinary Record,* 119: 458-489.

Logan, L. L., Holland C. J., Mebus C.A., and Ristic M., **1987**. The development of *Cowdria ruminantium* in neutrophils. *The Onderstepoort Journal of Veterinary Research*, 54: 197-204.

Lounsbury C.P., **1900**. Tick-heartwater experiment. *Agricultural Journal of the Cape of Good Hope*, 16: 682-687.

MacKenzie P.K.I., and van Rooyen R.E., **1981**. The isolation and culture of *Cowdria ruminantium* in albino mice. Proceedings of the International Congress on Tick Biology and Control, 27-29 January, Grahamstown, South Africa. 33-39.

MacKenzie P.K.L., and McHardy N., **1987**. *Cowdria ruminantium* infection in the mouse: A review. *The Onderstepoort Journal of Veterinary Research,* 54: 267-269.

Mahan S.M., Waghela S.D., McGuire T.C., Rurangirwa F.R., Wassink L.A., and Barbet A.F., **1992**. A cloned DNA probe for *Cowdria ruminantium* hybridises with eight heartwater strains and detects infected sheep. *Journal of Clinical Microbiology*, 30: 981-986.

Mahan S.M., Tebele N., Mukwedeya D., Semu S., Nyathi C.B., Wassink L.A., Kelly P.J., Peter T., and Barbet A.F., **1993**. An immunoblotting diagnostic assay for heartwater based on the immunodominant 32-kilodalton protein *Cowdria ruminantium* detects false positives in field sera. *Journal of Clinical Microbiology*, 31: 2729-2737.

Mahan S.M., McGuire T.C., Semu S.M., Bowie M.V., Jongejan F., Rurangirwa F.R., and Barbet A.E., **1994a**. Molecular cloning of a gene encoding the immunogenic 21 kDa protein of *Cowdria ruminantium*. *Microbiology*, 140: 2135-2142.

Mahan S.M., Smith G.E., and Byrom B., **1994b**. Concanavalin A-stimulated bovine T cell supernatants inhibit the growth of *Cowdria ruminantium* in bovine endothelial cells in vitro. *Infection and Immunity*, 62: 747-750.

Mahan S.M., **1995a**. Review of the molecular biology of *Cowdria ruminantium*. *Veterinary Parasitology*, 57: 51-56.

Mahan S.M., Andrew H.R., Tebele N., Burridge M.J., and Barbet A.F., **1995b**. Immunisation of sheep with inactivated *Cowdria ruminantium* provides protection against heartwater. *Research in Veterinary Science*, 58: 46-49.

Mahan S.M., Sileghem M., Smith G.E., and Byrom B., **1996**. Neutralisation of bovine concanavalin-A T cell supernatant-mediated *Cowdria ruminantium* activity with antibodies specific to interferon gamma but not to tumor necrosis factor. *Parasite Immunology,* 18: 317-324.

Mahan S.M., Kumbula D., Burridge M.J., and Barbet A.F., **1998a**. The inactivated *Cowdria ruminantium* vaccine for heartwater protects against heterologous strains and against laboratory and field tick challenge. *Vaccine*, 16: 1203-1211.

Mahan S.M., Peter T.F., Simbi B.H., Kocan K.M., Camus E., Barbet A.F., and Burridge M.J., **1998b**. PCR detection of *Cowdria ruminantium* infection in ticks and animals from heartwater-endemic regions of Zimbabwe. *Annals of the New York Academy of Science*, 849: 85-87.

Mahan S.M., Semu S.M., Peter T.F., and Jongejan F., **1998c**. Evaluation of the MAP1B ELISA for cowdriosis with field sera from livestock in Zimbabwe. *Annals of the New York Academy of Science*, 849: 85-87.

Mahan S.M., Peter T.F., Perry B.D., Modisa L., Axsel G., Bryson N., van der Merwe L., van der Riet P., Pistorius J., Dlamini P., Makala L., Munyama G., Barbet A.F., and Burridge M.J., **1998d**. Characterisation of heartwater challenge at field sites in six SADC countries. P. 30. In Proceedings of the 9th International Conference of the Association of Institutes of Tropical Veterinary Medicine (AITVM).

Mahan S.M., Allsopp B., Kocan K.M., Palmer G.H., and Jongejan F., **1999**. Vaccine strategies for *Cowdria ruminantium* infections and their application to other Ehrlichial infections. *Parasitology Today*, 15: 290-294.

Martinez D., Coisne S., and Sheikboudou C., **1993**. Detection of antibodies to *Cowdria ruminantium* in the serum of domestic ruminants by ELISA. *Revue d'elevage et de Medecine Veterinaire des Pays Tropicaux,* 46: 115-120.

Martinez D., Maillard J.C., Coisne S., Sheikboudou C., and Bensaid A., **1994**. Protection of goats against heartwater acquired by immunisation with inactivated elementary bodies of *Cowdria ruminantium*. *Veterinary Immunology and Immunopathology,* 41: 153-163.

Martinez D., Perez J.M., Sheikboudou C., Debus A., and Bensaid A., **1996**. Comparative efficacy of Freund's and Montanide ISA50 adjuvants for the immunisation of goats against heartwater with inactivated *Cowdria ruminantium*. *Veterinary Parasitology*, 67: 175-184

Matsuyama S., Fujita Y., Sagara K., and Mizushima S., **1992**. Overproduction, purification and characterisation of SecD and SecF, integral membrane components of the protein translocation machinery of *Escherichia coli*. *Biochimica et Biophysica Acta*, 1122: 77-84.

Matsuyama S., Fujita Y., and Mizushima S., **1993**. SecD is involved in the release of translocated secretory proteins from the cytoplasmic membrane of *Escherichia coli*. *The EMBO Journal*, 12: 265-270.

McBride J.W., Yu Xj., Walker D.H., **1999**. Molecular cloning of the gene for a conserved major immunoreactive 28- kilodalton protein of *Ehrlichia canis*: a potential serodiagnostic antigen. *Clinical and Diagnostic Laboratory Immunology,* 6: 392-399.

McKeever D.J., **1993**. Progress towards a vaccine against *Theilaria parva*: relevance for heartwater research. *Revue d'elevage et de Medecine Veterinaire des Pays Tropicaux,* 46: 231-235.

Minjuaw B., and McLeod A., **2001**. Epidemiology of tick-borne diseases. Their effects on the livelihood of the poor in east and southern Africa and India. Poster: Heartwater in sub-Saharan Africa. Department for International Development, Animal Health Programme.

Moncada S., Palmer R.M.J., and Higgs E.A., **1988**. The discovery of nitric oxide as the endogenous nitrovasodilator. *Hypertension*, 12: 365-372.

Moncada S., Palmer R.M.J., and Higgs E.A., **1990**. Nitric oxide: the endogenous regulator of vascular tone p1-6. In: J.B. Warren (ed.), The endothelium: an introduction to current research. Wiley-Liss, New York.

Moncada S., and Palmer R.M.J., **1993**. The L-arginine-nitric oxide pathway. *The New England Journal of Medicine,* 329: 2002-2012.

Muller-Kobold A., Martinez D., Camus E., and Jongejan F., **1992**. Distribution of heartwater in the Caribbean determined on the basis of detection of antibodies to the conserved 32 kilodalton protein of *Cowdria ruminantium*. *Journal of Clinical Microbiology*, 30: 1870-1873.

Mullis K.B., and Faloona F.A., **1987**. Specific synthesis of DNA *in vitro* via a polymerase catalysed chain reaction. *Methods in Enzymology,* 155: 335-350.

Musoke A.J., Palmer G.H., McElwain T.F., Nene V., and McKleever D., **1996**. Prospects for subunit vaccines against tick-borne diseases. *British Veterinary Journal*, 152: 621-639.

Musoke A.J., McKleever D., and Nene V., **1997**. Subunit vaccines for the control of tick-borne diseases: implications for the future. *Parasitologia*, 39: 131-137.

Mutunga M., Preston P.M., and Sumpton K.J., **1998**. Nitric oxide is produced by *Cowdria ruminantium*-infected bovine pulmonary endothelial cells *in vitro* and is stimulated by gamma interferon. *Infection and Immunity*, 66: 2115-2121.

Mwangi D.M., Mahan S.M., Nyanjui J.K., Taracha E.L.N., and McKeever D.J., **1998a**. Immunisation of cattle by infection with *Cowdria ruminantium* elicits T lymphocytes that recognise autologous, infected endothelial cells and monocytes. *Infection and Immunity*, 66: 1855-1860.

Mwangi D.M., McKeever D.J., Nyanjui J.K., Barbet A.F., and Mahan S.M., **1998b**. Major antigenic proteins 1 and 2 of *Cowdria ruminantium* are targets for T-lymphocytes responses of immune cattle. *Annals of the New York Academy of Sciences,* 849: 372-374.

Neitz A.W.H., Viljoen G.J., Bezuidenhout J.D., Oberem P.T., van Wyngaardt W., and Vermeulen N.M.J., **1986**. The detection of antibodies to *Cowdria ruminantium* in serum and *Cowdria ruminantium* antigen in *Amblyomma hebraeum* by means of enzyme-linked immunosorbent assays. *The Onderstepoort Journal of Veterinary Research*, 53: 39-41.

Neitz W.O., and Alexander R.A., **1941**. The immunisation of calves against heartwater. *Journal of the South African Veterinary Medical Association*, 12: 103-111.

Neitz W.O., **1964**. Tick-borne diseases as a hazard in the rearing of calves in South Africa. *Bulletin Office of International Epizootics*, 62: 607-625.

Neitz W.O., **1968**. Heartwater. *Bulletin Office of International Epizootics*, 70: 329-336.

Norval R.A.I., Andrew H.R., and Yunker C.E., **1990**. Infection rates with *Cowdria ruminantium* of nymphs and adults of the bont tick collected in the field in Zimbabwe. *Veterinary Parasitology*, 36: 277-282.

Norval R.A.I., Perry B.D., Meltzer M.I., Kruska R.L., and Booth T.H., **1994**. Factors affecting the distribution of the ticks *Amblyomma hebraeum* and *Amblyomma variegatum* in Zimbabwe: implications of reduced acaricide usage. *Experimental and Applied Acarology,* 18: 383-407.

Nyika A., Mahan S.M., Burridge M.J., McGuire T.C., Rurangirwa F., and Barbet A.F., **1998**. A DNA vaccine protects mice against the rickettsial agent *Cowdria ruminantium*. *Parasite Immunology,* 20: 111-119.

Oberem P.T., and Bezuidenhout J.D., **1987a**. Heartwater in hosts other than domestic ruminants. *The Onderstepoort Journal of Veterinary Research*, 54: 271-275.

Oberem P.T., and Bezuidenhout J.D., **1987b**. The production of a heartwater vaccine. *The Onderstepoort Journal of Veterinary Research*, 54: 485-488.

Ogra P.L., Faden H.S., Abraham R., Duffy L.C., Sun M., and Minor P.D., **1991**. Effect of prior immunity on the shedding of virulent revertant virus in feces after oral immunisation with live attenuated poliovirus vaccines. *Journal of Infectious Diseases*, 164: 191-194.

Owen N.C., Littlejohn A., Kruger J.M., and Erasmus B.J., **1973**. Physiopathological features of heartwater in sheep. *Journal of the South African Veterinary Medical Association,* 44: 397-403.

Perez J.M., Martinez D., Sheikboudou C., Jongejan F., and Bensaid A., **1998**. Characterisation of variable immunodominant antigens of *Cowdria ruminantium* by ELISA and immunoblots. *Parasite Immunology,* 20: 613-622.

Peter T.F., Deem S.L., Barbet A.F., Norval R.A.I., Simbi B.H., Kelly P.J., and Mahan S.M., **1995**. Development and evaluation of a PCR assay for detection of low levels of *Cowdria ruminantium* infection in *Amblyomma* ticks not detected by DNA probe. *Journal of Clinical Microbiology*, 33: 166-172.

Peter T.F., Bryson N.R., Perry B.D., O'Callaghan C.J., Medley G.F., Smith G.E., Mlambo G., Horak I.G., Burridge M.J., and Mahan S.M., **1999**. *Cowdria ruminantium* infection in ticks in the Kruger National Park. *The Veterinary Record*, 145: 304-307.

Peter T.F., Barbet A.F., Alleman A.R., Simbi B.H., Burridge M.J., and Mahan S.M., **2000**. Detection of the agent of heartwater, *Cowdria ruminantium*, in *Amblyomma* ticks by PCR: validation and application of the assay to field ticks. *Journal of Clinical Microbiology,* 38: 1539-1544.

Pienaar J.G., **1970**. Electron microscopy of *Cowdria* (Rickettsia) *ruminantium* (Cowdry, 1926) in the endothelial cells of the vertebrate host. *The Onderstepoort Journal of Veterinary Research*, 37: 67-78.

Pogliano K.J., and Beckwith J., **1994**. Genetic and molecular characterisation of the *Escherichia coli secD* operon and its products. *Journal of Bacteriology*, 176: 804-814.

Potgieter F.T., and van Rensburg L., **1983**. Infectivity, virulence and immunogenicity of *Anaplasma centrale* live vaccine. *The Onderstepoort Journal of Veterinary Research*, 50: 29-31.

Prescott L.M., Harvey J.P., and Klein D.A., **1993**. *Microbiology 2nd edition*, Wm. C Brown Publishers, Kane K (Editor). 441-443.

Preston P.M., and Jongejan F., **1999a**. Protective immune mechanisms to ticks and tick-borne diseases of ruminants. *Parasitology Today*, 15: 255-258.

Preston P.M., and Jongejan F., **1999b**. Protective immune mechanisms to ticks and tick-borne diseases of ruminants. *Parasitology Today*, Poster.

Prichard R., **1997**. Application of molecular biology in veterinary parasitology. *Veterinary Parasitology*, 71: 155-175.

Promega Technical Manual, **1995**. *Genomic cloning manual.*

Provost A., and Bezuidenhout J.D., **1987**. The historical background and global importance of heartwater. *The Onderstepoort Journal of Veterinary Research*, 54: 165-169.

Prozesky L., and du Plessis J.L., **1984**. The pathology of heartwater. A pathological study of sheep, goats and mice infected with *Cowdria ruminantium. Proceedings of XIIIth World Congres on Diseases of Cattle,* 1: 547-549.

Prozesky L., and du Plessis J.L., **1985a**. The pathology of heartwater I. A Study of mice infected with the Welgevonden strain of *Cowdria ruminantium. The Onderstepoort Journal of Veterinary Research*, 52: 71-79.

Prozesky L., and du Plessis J.L., **1985b**. The pathology of heartwater II. A Study of the lung lesions in sheep and goats infected with the Ball3 strain of *Cowdria ruminantium. The Onderstepoort Journal of Veterinary Research*, 52: 81-85.

Prozesky L., Bezuidenhout J.D., and Paterson C.L., **1986**. Heartwater. An *in vitro* study of the ultrastructure of *Cowdria ruminantium. The Onderstepoort Journal of Veterinary Research*, 53: 153-159.

Prozesky L., **1987a**. The pathology of heartwater III. A Review. *The Onderstepoort Journal of Veterinary Research*, 54: 281-286.

Prozesky L., **1987b**. The diagnosis of heartwater at post mortem in ruminants and the conformation of *Cowdria ruminantium* in mice. *The Onderstepoort Journal of Veterinary Research*, 54: 301-303.

Prozesky L., and du Plessis J.L., **1987c**. Heartwater. The development and life cycle of *Cowdria ruminantium* in the vertebrate host, ticks and cultured endothelial cells. *The Onderstepoort Journal of Veterinary Research*, 54: 193-196.

Purchase H.S., **1945**. A simple and rapid method for demonstrating *Rickettsia ruminantium* (Cowdry, 1925) in heartwater brains. *The Veterinary Record,* 57: 413-414.

Radi R., Beckman J.S., Bush K.M., and Freeman B.A., **1991**. Peroxynitrite oxidation of sulfhydryls: the cytotoxic potential of superoxide and nitric oxide. *Journal of Biological Chemistry,* 266: 4244-4250.

Reddy G.R., **1995**. Determining the sequence of parasite DNA. *Parasitology Today*, 11: 37-42.

Reddy G.R., Sulsona C.R., Harrison R.H., Mahan S.M., Burridge M.J., and Barbet A.F., **1996**. Sequence heterogeneity of the Major Antigen Protein 1 genes from *Cowdria ruminantium* isolates from different geographical areas. *Clinical and Diagnostic Laboratory Immunology*, 3: 417-422.

Ristic M., and Huxoll D.L., **1984**. Tribe II. Ehrlicheae Philip. 948^{AL}. In: Krieg N.R., and Holt J.G., (eds.). Bergey's Manual of Systematic Bacteriology, vol. 1. Williams and Wilkins, Baltimore, Maryland. Pp. 704-711.

Rogers R.J., Dimmock C.K., de Vos A.J., and Rodwell B.J., **1988**. Bovine leucosis virus contamination of a vaccine produced *in vivo* against bovine babesiosis and anaplasmosis. *The Australian Veterinary Journal*, 65: 285-287.

Rossouw M., Neitz A.W.H., de Waal D.J., du Plessis J.L., van Gas L., and Brett S., **1990**. Identification of antigenic proteins of *Cowdria ruminantium*. *The Onderstepoort Journal of Veterinary Research*, 57: 215-221.

Saiki R.K., Gelfand D.H., Stoffel S., Scharf S.J., Higuchi R., Horn G.T., Mullis K.B., and Ehrlich H.A., **1988**. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, 239: 487-491.

Sanger E., Nicklen S., and Coulson A.R., 1977. DNA sequencing with chain-terminating inhibitors.

Semu S.M., Mahan S.M., Yunker C.E., and Burridge M.J., **1992**. Development and persistence of *Cowdria ruminantium* specific antibodies following experimental infection of cattle, as detected by the indirect fluorescent antibody test. *Veterinary Immunology and Immunopathology,* 33: 339-352.

Sher A., and Coffman R.L., **1992**. Regulation of immunity to parasites by T cells and T cell-derived cytokines. *Annual Review of Immunology,* 10: 385-409.

Smith G.E., Anderson E.C., Burridge M.J., Peter T.F., and Mahan S.M., **1998**. Growth of *Cowdria ruminantium* in tissue culture endothelial cell lines from wild African mammals. *Journal of Wildlife Diseases*, 34: 297-304.

Spreull J., **1922**. Heartwater. *Joernaal van die Departement Landbou*, 11: 1-12.

Staden R., Beal K.F., and Bonfield J.K., **1998**. The Staden Package. Computer methods in Molecular Biology, Eds. Stephen Misener and Steve Krawetz. The Humana Press Inc., Totowa, NJ 07512.

Stratagene Instruction manual, **1993**. *Undigested Lambda ZAP® II cloning kit.*

Sulsona C.R., Mahan S.M., and Barbet A.F., **1999**. The *map1* gene of *Cowdria ruminantium* is a member of a muligene family containing both conserved and variable genes. *Biochemical and Biophysical Research Communications*, 257: 300-305.

Totté P., de Gee A.L.W., and Werenne J., **1993a**. Role of interferons in infectious diseases in the bovine species; effect of viruses and rickettsias. *Revue d'elevage et de Medecine Veterinaire des Pays Tropicaux,* 46: 83-86.

Totté P., Blankaert D., Zilimwabagabo P., and Werenne J., **1993b**. Inhibition of *Cowdria ruminantium* infectious yield by interferons alpha and gamma in endothelial cells. *Revue d'elevage et de Medecine Veterinaire des Pays Tropicaux,* 46: 189-194.

Totté P., Jongejan F., de Gee A.L., and Werenne J., **1994a**. Production of interferon alpha in *Cowdria ruminantium*-infected cattle and its effect on infected endothelial cell cultures. *Infection and Immunity*, 62: 2600-2604.

Totté P., **1994b**. In: *Animal cell technology: Basic and applied aspects* (Vol 6) (Kobayashi, T. *et al.*, eds.), pp 595-599, Kluwer academic publishers.

Totté P., Vachiéry N., Martinez D., Trap I., Ballingall K.T., MacHugh N.D., Bensaid A., and Werenne J., **1996**. Recombinant bovine interferon gamma inhibits growth of *Cowdria ruminantium* but fails to induce MHC class II following infection of endothelial cells. *Veterinary Immunology and Immunopathology*, 53: 61-71.

Totté P., McKeever D., Martinez D., and Bensaid A., **1997**. Analysis of T cell responses in cattle immunised against heartwater by vaccination with killed elementary bodies of *Cowdria ruminantium*. *Infection and Immunity*, 65: 236-241.

Totté P., McKeever D., Jongejan F., Barbet A., Mahan S.M., Mwangi D., and Bensaid A., **1998**. Analysis of cellular responses to native and recombinant proteins of *Cowdria ruminantium*. *Annals of the New York Academy of Sciences,* 849: 155-160.

Totté P., Bensaid A., Mahan S.M., Martinez D., and McKeever D.J., **1999**. Immune responses to *Cowdria ruminantium* infections. *Parasitology Today*, 15: 286-290.

Uilenberg G., **1982**. Experimental transmission of *Cowdria ruminantium* by the Gulf coast tick *Amblyomma maculatum*: Danger of introducing heartwater and benign Africa theileriosis onto the American mainland. *American Journal of Veterinary Research,* 43: 1279-1282.

Uilenberg G., **1983**. Heartwater (*Cowdria ruminantium* infection): Current status. *Advances in Veterinary Science and Comparative Medicine*, 27: 427-480.

Uilenberg G., Barré N., Camus E., Burridge M.J., and Garris G.I., **1984**. Heartwater in the Caribbean. *Preventive Veterinary Medicine*, 2: 255-267.

Uilenberg G., Camus E., and Barré N., **1985**. Quelques observations sur une souche de *Cowdria ruminantium* isolée en Guadeloupe (Antilles francaises). *Revue d'elevage et de Medecine Veterinaire des Pays Tropicaux*, 38: 34-42.

Uilenberg G., **1990a**. Methods currently used for the control of heartwater: their validity and proposals for future control strategies. *Parasitologia*, 32: 55-62.

Uilenberg G., **1990b**. Present and future possibilities for control of cowdriosis and anaplasmosis. *The Veterinary Quarterly*, 12: 39-45.

Uilenberg G., Dobbelaere D.A.E., de Gee A.L.W., and Koch H.T., **1993a**. Progress in research on tick-borne diseases: Theileriosis and Heartwater. *The Veterinary Quarterly,* 15: 48-54.

Uilenberg G., and Camus E., **1993b**. Heartwater (Cowdriosis). In: Z Woldehiwet and M Ristic (ed), Rickettsial and chlamydial diseases of domestic animals. Pergamon Press, Oxford. 293-332.

Umansky V., Becur M., Schirrmacher V., and Rocha M., **1997**. Activated endothelial cells induce apoptosis in lymphoma cells: role of nitric oxide. *International Journal of Oncology,* 10: 465-471.

Vachiéry N., Trap I., Totté P., Martinez D., and Bensaid A., **1998**. Inhibition of MHC class I and class II cell surface expression on bovine endothelial cells upon infection with *Cowdria ruminantium*. *Veterinary Immunology and Immunopathology,* 61: 37-48.

van Amstel S.R., and Oberem P.T., **1987**. The treatment of heartwater. *The Onderstepoort Journal of Veterinary Research*, 54: 475-479.

van de Pypekamp H.E., and Prozesky L., **1987**. Heartwater. An overview of the clinical signs, susceptibility and differential diagnoses of the disease in domestic ruminants. *The Onderstepoort Journal of Veterinary Research*, 54: 263-266.

van der Merwe L., **1987**. The infection and treatment method of vaccination against heartwater. *The Onderstepoort Journal of Veterinary Research*, 54: 489-491.

van Kleef M., Neitz A.W.H., and de Waal D.T., **1993**. Isolation and characterisation of antigenic proteins of *Cowdria ruminantium*. *Revue d'elevage et de Medecine Veterinaire des Pays Tropicaux*, 46: 157-164.

van Vliet A.H.M., Jongejan F., and van der Zeijst B.A.M., **1992**. Phylogenetic position of *Cowdria ruminatium* (*Rickettsiales*) determined by analysis of amplified 16S ribosomal DNA sequence. *International Journal of Systematic Bacteriology*, 42: 494-498.

van Vliet A.H.M., Jongejan F., van Kleef M., and van der Zeijst B.A.M., **1994**. Molecular cloning, sequence analysis and expression of the gene encoding the immunodominant 32-kilodalton protein of *Cowdria ruminantium*. *Infection and Immunity*, 62: 1451-1456.

van Vliet A.H.M., van der Zeijst B.A.M., Camus E., Mahan S.M., Martinez D., and Jongejan F., **1995**. Use of a specific immunogenic region on the *Cowdria ruminantium* MAP1 protein in a serological assay. *Journal of Clinical Microbiology*, 33: 2405-2410.

Vora M., Yssel H., de Vries J.E., and Karasek M.A., **1994**. Antigen presentation by human dermal microvascular endothelial cells: immunoregulatory effect of IFN-γ and IL-10. *Journal of Immunology*, 152: 5734-5741.

Waghela S.D., Rurangirwa F.R., Mahan S.M., Yunker C.E., Crawford T.B., Barbet A.F., Burridge M.J., and McGuire T.C., **1991**. A cloned DNA probe identifies *Cowdria ruminantium* in *Amblyomma variegatum* ticks. *Journal of Clinical Microbiology*, 29: 2571-2577.

Waine G.J., and McManus D.P., **1995**. Nucleic acids: Vaccines of the future. *Parasitology Today*, 11: 113-116.

Walker J.B., and Olwage A., **1987**. The tick vectors of *Cowdria ruminantium* (Ixodoidea, Ixodidae genus *Amblyomma*) and their distribution. *The Onderstepoort Journal of Veterinary Research*, 54: 353-379.

Weaver C.T., and Unanue E.R., **1990**. The costimulatory function of antigen-presenting cells. *Immunology Today*, 11: 49-55.

Weiner D.B., and Kennedy R.C., **1999**. Genetic vaccines. *Scientific American Feature Article*.

Weeks-Levy C., Tatem J.M., DiMichele S.J., Waterfield W., Georgiu A.F. and Mento S.J., **1991**. Identification and characterisation of a new base substitution in the vaccine strain of Sabin 3 poliovirus. *Virology,* 185: 934-937.

Woodman J.P., Dimier I.J., and Bout D.T., **1991**. Human endothelial cells are activated by IFNγ to inhibit *Toxoplasma gondii* replication - inhibition is due to a different mechanism from that existing in mouse macrophages and human fibroblasts. *Journal of Immunology*, 147: 2019-2023.

Young E., and Basson P.A., **1973**. Heartwater in the eland. *Journal of the South African Veterinary Association*, 44: 185-186.

Yunker C.E., Kocan K.M., Norval R.A.I., and Burridge M.J., **1987**. Distinctive staining of colonies of *Cowdria ruminantium* in midguts of *Amblyomma hebraeum*. *The Onderstepoort Journal of Veterinary Research*, 54: 183-185.

Yunker C.E., Mahan S.M., Waghela S.D., McGuire T.C., Rurangirwa F.R., Barbet A.F., and Wassink L.A., **1993**. Detection of *Cowdria ruminantium* by means of a DNA probe, pCS20 in infected bont ticks, *Amblyomma hebraeum*, the major vector in South Africa. *Epidemiology and Infection,* 110: 95-104.

Yunker C.E., and Allsopp B.A., **1994**. Modern biotechnological methods for diagnosis of cowdriosis – present and future trends, Proc FAO Expert Consultation on Use of Applicable Biotechnology. *Methods for diagnosis of Haemoparasites*, 4-6 October, Merida, Mexico, FAO, Rome. 143-155.

Yunker C.E., **1995**. Current status of *in vitro* cultivation of *Cowdria ruminantium*. *Veterinary Parasitology*, 57: 205-211.

ADDENDA

List of Solutions

Church and Gilbert HYB	7% SDS, 0.5 M NaPi pH 7,5 (42 ml 1 M Na_2HPO_4, 8 ml 1 M NaH_2PO_4)
LB	10 g tryptone, 5 g yeast extract, 5 g NaCl, 1 L water
LB plates	10 g tryptone, 5 g yeast extract, 5 g NaCl, 15 g agar,1 L water
LB top agarose	1.0 g tryptone, 0.5 g NaCl, 0.7 g agarose, 100 ml water
SM Buffer	0.01 % gelatin, 50 mM Tris-HCl pH 7.5, 100 mM NaCl, 8 mM $MgSO_4$
2x SSC	30 mM Na_3 citrate, 300 mM NaCl pH 7.0
STE buffer	1.5 % SDS, 0.3 M Tris pH 9, 0.15 M EDTA
TAE	4.84 g Tris, 2 ml 0.5 M Na_2EDTA pH 8.0, 1.142 ml glacial acetic acid, 1 L water
TBE	10.8 g Tris, 5.5 g boric acid, 4 ml 0.5 M Na_2EDTA pH 8.0, 1 L water
TE buffer	10 mM Tris-HCl pH 8.0, 1 mM EDTA

Addendum 1. DNA sequence of contig one. Column 1: Position where clone sequence begins. Column 2: Clone number and direction of primer used (F = forward, R = reverse)

Primer SP6GEM11: 5'-CCA TTT AGG TGA CAC TAT AG-3'.

Primer WL2TP1_gapF: 5'- CAC TTA CAC CAA TGC CAC AC -3'

```
                        10        20        30        40        50        60        70
7       152F    CCATTTAGGTGACACTATAGAAGAGCTCGAGGATCTTCGGCTACCTGAAAGAAGCGATGTTAAAACTTTTATTTTCATT
-43     T118R   CCATTTAGGTGACACTATAGAAGAGCTCGAGGATCTTCGGCTACCTGAAAGAAGCGATGTTAAAACTTTTATTTTCATT
-23     T105R   CCATTTAGGTGACACTATAGAAGAGCTCGAGGATCTTCGGCTACCTGAAAGAAGCGATGTTAAAACTTTTATTTTCATT
9       17R     CCATTTAGGTGACACTATAGAAGAGCTCGAGGATCTTCGGCTACCTGAAAGAAGCGATGTTAAAACTTTTATTTTCATT
78      T2R     CCATTTAGGTGACACTATAGAAGAGCTCGAGGATCTTCGGCTACCTGAAAGAAGCGATGTTAAAACTTTTATTTTCATT
87      T34F    CCATTTAGGTGACACTATAGAAGAGCTCGAGGATCTTCGGCTACCTGAAAGAAGCGATGTTAAAACTTTTATTTTCATT
82      T31R    CCATTTAGGTGACACTATAGAAGAGCTCGAGGATCTTCGGCTACCTGAAAGAAGCGATGTTAAAACTTTTATTTTCATT
55      T131R   CCATTTAGGTGACACTATAGAAGAGCTCGAGGATCTTCGGCTACCTGAAAGAAGCGATGTTAAAACTTTTATTTTCATT
48      T123R   CCATTTAGGTGACACTATAGAAGAGCTCGAGGATCTTCGGCTACCTGAAAGAAGCGATGTTAAAACTTTTATTTTCATT
62      T141_F  CCATTTAGGTGACACTATAGAAGAGCTCGAGGATCTTCGGCTACCTGAAAGAAGCGATGTTAAAACTTTTATTTTCATT
11      84R     CCATTTAGGTGACACTATAGAAGAGCTCGAGGATCTTCGGCTACCTGAAAGAAGCGATGTTAAAACTTTTATTTTCATT
12      86R     CCATTTAGGTGACACTATAGAAGAGCTCGAGGATCTTCGGCTACCTGAAAGAAGCGATGTTAAAACTTTTATTTTCATT
42      T116_F  CCATTTAGGTGACACTATAGAAGAGCTCGAGGATCTTCGGCTACCTGAAAGAAGCGATGTTAAAACTTTTATTTTCATT
107     T53R    CCATTTAGGTGACACTATAGAAGAGCTCGAGGATCTTCGGCTACCTGAAAGAAGCGATGTTAAAACTTTTATTTTCATT
109     T54R    CCATTTAGGTGACACTATAGAAGAGCTCGAGGATCTTCGGCTACCTGAAAGAAGCGATGTTAAAACTTTTATTTTCATT
112     T58_F   CCATTTAGGTGACACTATAGAAGAGCTCGAGGATCTTCGGCTACCTGAAAGAAGCGATGTTAAAACTTTTATTTTCATT
113     T5F     CCATTTAGGTGACACTATAGAAGAGCTCGAGGATCTTCGGCTACCTGAAAGAAGCGATGTTAAAACTTTTATTTTCATT
115     T62R    CCATTTAGGTGACACTATAGAAGAGCTCGAGGATCTTCGGCTACCTGAAAGAAGCGATGTTAAAACTTTTATTTTCATT
117     T6R     CCATTTAGGTGACACTATAGAAGAGCTCGAGGATCTTCGGCTACCTGAAAGAAGCGATGTTAAAACTTTTATTTTCATT
119     T72R    CCATTTAGGTGACACTATAGAAGAGCTCGAGGATCTTCGGCTACCTGAAAGAAGCGATGTTAAAACTTTTATTTTCATT
128     T86F    CCATTTAGGTGACACTATAGAAGAGCTCGAGGATCTTCGGCTACCTGAAAGAAGCGATGTTAAAACTTTTATTTTCATT
131     T88_F   CCATTTAGGTGACACTATAGAAGAGCTCGAGGATCTTCGGCTACCTGAAAGAAGCGATGTTAAAACTTTTATTTTCATT
35      T111R   CACACTATAGAAGAGCTCGAGGATCTTCGGCTACCTGAAAGAAGCGATGTTAAAACTTTTATTTTCATT
-36     T111_F  CACACTATAGAAGAGCTCGAGGATCTTCGGCTACCTGAAAGAAGCGATGTTAAAACTTTTATTTTCATT
57      T135R   ACTATAGAAGAGCTCGAGGATCTTCGGCTACCTGAAAGAAGCGATGTTAAAACTTTTATTTTCATT
-40     T115R                 GCTCGAGGATCTTCGGCTACCTGAAAGAAGCGATGTTAAAACTTTTATTTTCATT

CONSENSUS       CCATTTAGGTGACACTATAGAAGAGCTCGAGGATCTTCGGCTACCTGAAAGAAGCGATGTTAAAACTTTTATTTTCATT
```

Primer SP6GEM11: 5'-CCA TTT AGG TGA CAC TAT AG-3'.

```
                80        90        100       110       120       130       140       150
7       152F    TATAGCCGAAACATATTAATGCTAGTATATTTTACCGATTATATTACCTTTTTAATAAAAATACAACTTCACATTTATG
-43     T118R   TATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATATTACCTTTTTAATAAAAATACAACTTCACATTTATG
-23     T105R   TATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATATTACCTTTTTAATAAAAATACAACTTCACATTTATG
9       17R     TATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATATTACCTTTTTAATAAAAATACAACTTCACATTTATG
78      T2R     TATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATATTACCTTTTTAATAAAAATACAACTTCACATTTATG
87      T34F    TATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATATTACCTTTTTAATAAAAATACAACTTCACATTTATG
82      T31R    TATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATATTACCTTTTTAATAAAAATACAACTTCACATTTATG
55      T131R   TATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATATTACCTTTTTAATAAAAATACAACTTCACATTTATG
48      T123R   TATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATATTACCTTTTTAATAAAAATACAACTTCACATTTATG
62      T141_F  TATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATATTACCTTTTTAATAAAAATACAACTTCACATTTATG
11      84R     TATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATATTACCTTTTTAATAAAAATACAACTTCACATTTATG
12      86R     TATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATATTACCTTTTTAATAAAAATACAACTTCACATTTATG
42      T116_F  TATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATATTACCTTTTTAATAAAAATACAACTTCACATTTATG
107     T53R    TATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATATTACCTTTTTAATAAAAATACAACTTCACATTTATG
109     T54R    TATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATATTACCTTTTTAATAAAAATACAACTTCACATTTATG
112     T58_F   TATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATATTACCTTTTTAATAAAAATACAACTTCACATTTATG
113     T5F     TATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATATTACCTTTTTAATAAAAATACAACTTCACATTTATG
115     T62R    TATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATATTACCTTTTTAATAAAAATACAACTTCACATTTATG
117     T6R     TATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATATTACCTTTTTAATAAAAATACAACTTCACATTTATG
119     T72R    TATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATATTACCTTTTTAATAAAAATACAACTTCACATTTATG
128     T86F    TATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATATTACCTTTTTAATAAAAATACAACTTCACATTTATG
13      T88_F   TATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATATTACCTTTTTAATAAAAATACAACTTCACATTTATG
35      T111R   TATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATATTACCTTTTTAATAAAAATACAACTTCACATTTATG
-36     T111_F  TATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATATTACCTTTTTAATAAAAATACAACTTCACATTTATG
57      T135R   TATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATATTACCTTTTTAATAAAAATACAACTTCACATTTATG
-40     T115R   TATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATATTACCTTTTTAATAAAAATACAACTTCACATTTATG
-66     T17F                                                              ACAACTTCACATTTATG

CONSENSUS       TATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATATTACCTTTTTAATAAAAATACAACTTCACATTTATG
```

```
                 160       170       180       190       200       210       220       230
7        152F    GAATCATTACAATAGAAAAA
-43      T118R   GAATCATTACAATAGAAAAATAAAAAACATTAAGATA
-23      T105R   GAATCATTACAATAGAAAAATAAAAAACATTAAGATAACTAGTTTAACTAGTTATAAAT
9        17R     GAATCATTACAATAGAAAAATAAAAAACATTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACATA
78       T2R     GAATCATTACAATAGAAAAATAAAAAACATTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACATA
87       T34F    GAATCATTACAATAGAAAAATAAAAAACATTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACATA
82       T31R    GAATCATTACAATAGAAAAATAAAAAACATTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACATA
55       T131R   GAATCATTACAATAGAAAAATAAAAAACATTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACATA
48       T123R   GAATCATTACAATAGAAAAATAAAAAACATTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACATA
62       T141_F  GAATCATTACAATAGAAAAATAAAAAACATTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACATA
11       84R     GAATCATTACAATAGAAAAATAAAAAACATTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACATA
12       86R     GAATCATTACAATAGAAAAATAAAAAACATTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACATA
42       T116_F  GAATCATTACAATAGAAAAATAAAAAACATTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACATA
107      T53R    GAATCATTACAATAGAAAAATAAAAAACATTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACATA
109      T54R    GAATCATTACAATAGAAAAATAAAAAACATTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACATA
112      T58_F   GAATCATTACAATAGAAAAATAAAAAACATTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACATA
113      T5F     GAATCATTACAATAGAAAAATAAAAAACATTAAGATAACTAGTTTAACTAGTTATAAATCCTCCCATACATGTTACATA
115      T62R    GAATCATTACAATAGAAAAATAAAAAACATTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACATA
117      T6R     GAATCATTACAATAGAAAAATAAAAAACATTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACATA
119      T72R    GAATCATTACAATAGAAAAATAAAAAACATTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACATA
128      T86F    GAATCATTACAATAGAAAAATAAAAAACATTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACATA
131      T88_F   GAATCATTACAACAGAAAAATAAAAAACATTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACATA
35       T111R   GAATCATTACAATAGAAAAATAAAAAACATTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACATA
-36      T111_F  GAATCATTACAATAGAAAAATAAAAAACATTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACATA
57       T135R   GAATCATTACAATAGAAAAATAAAAAACATTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACATA
-40      T115R   GAATCATTACAATAGAAAAATAAAAAACATTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACAT
-66      T17F    GAATCATTACAATAGAAAAATAAAAAACATTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACATA

CONSENSUS        GAATCATTACAATAGAAAAATAAAAAACATTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACATA


                   240       250       260       270       280       290       300       310
9        17R     ACTATTTCACTACTAAATATTATGAGACATTCTAGATAACCTTAATATACAGATACCTAAAATAACCATAATGAAGGAC
78       T2R     ACTATTTCACTACTAAATATTATGAGACATTCTAGATAACCTTAATATACAGATACCTAAAATAACCATAATGAAGGAC
87       T34F    ACTATTTCACTACTAAATATTATGAGACATTCTAGATAACCTTAATATACAGATACCTAAAATAACCATAATGAAGGAC
82       T31R    ACTATTTCACTACTAAATATTATGAGACATTCTAGATAACCTTAATATACAGATACCTAAAATAACCATAATGAAGGAC
55       T131R   ACTATTTCACTACTAAATATTATGAGACATTCTAGATAACCTTAATATACAGATACCTAAAATAACCATAATGAAGGAC
48       T123R   ACTATTTCACTACTAAATATTATGAGACATTCTAGATAACCTTAATATACAGATACCTAAAATAACCATAATGAAGGAC
62       T141_F  ACTATTTCACTACTAAATATTATGAGACATTCTAGATAACCTTAATATACAGATACCTAAAATAACCATAATGAAGGAC
11       84R     ACTATTTCGCTACTAAATATTATGAGACATTCTAGATAACCTTAATATACAGATACCTAAAATAACCATAATGAAGGAC
12       86R     ACTATTTCACTACTAAATATTATGAGACATTCTAGATAACCTTAATATACAGATACCTAAAATAACCATAATGAAGGAC
42       T116_F  ACTATTTCACTACTAAATATTATGAGACATTCTAGATAACCTTAATATACAGATACCTAAAATAACCATAATGAAGGAC
107      T53R    ACTATTTCACTACTAAATATTATGAGACATTCTAGATAACCTTAATATACAGATACCTAAAATAACCATAATGAAGGAC
109      T54R    ACTATTTCACTACTAAATATTATGAGACATTCTAGATAACCTTAATATACAGATACCTAAAATAACCATAATGAAGGAC
112      T58_F   ACTATTTCACTACTAAATATTATGAGACATTCTAGATAACCTTAATATACAGATACCTAAAATAACCATAATGAAGGAC
113      T5F     ACTATTTCACTACTAAATATTATGAGACATTCTAGATAACCTTAATATACAGATACCTAAAATAACCATAATGAAGGAC
115      T62R    ACTATTTCACTACTAAATATTATGAGACATTCTAGATAACCTTAATATACAGATACCTAAAATAACCATAATGAAGGAC
117      T6R     ACTATTTCACTACTAAATATTATGAGACATTCTAGATAACCTTAATATACAGATACCTAAAATAACCATAATGAAGGAC
119      T72R    ACTATTTCACTACTAAATATTATGAGACATTCTAGATAACCTTAATATACAGATACCTAAAATAACCATAATGAAGGAC
128      T86F    ACTATTTCACTACTAAATATTATGAGACATTCTAGATAACCTTAATATACAGATACCTAAAATAACCATAATGAAGGAC
131      T88_F   ACTATTTCACTACTAAATATTATGAGACATTCTAGATAACCTTAATATACAGATACCTAAAATAACCATAATGAAGGAC
35       T111R   ACTATTTCACTAGTGAATATTATGAGACATTCTAGATAACCTTAATATACAGATACCTAAAATAACCATAATGAAGGAC
-36      T111_F  ACTATTTCACTAGTGAATATTATGAGACATTCTAGATAACCTTAATATACAGATACCTAAAATAACCATAATGAAGGAC
57       T135R   ACTATTTCACTACTAAATATTATGAGACATTCTAGATAACCTTAATATACAGATACCTAAAATAACCATAATGAAGGAC
-66      T17F    ACTATTTCACTACTAAATATTATGAGACATTCTAGATAACCTTAATATACAGATACCTAAAATAACCATAATGAAGGAC
-106     T53F                                                    TAATATACAGATACCTAAAATAACCATAATGAAGGGC

CONSENSUS        ACTATTTCACTACTAAATATTATGAGACATTCTAGATAACCTTAATATACAGATACCTAAAATAACCATAATGAAGGAC
```

```
                    320        330        340        350        360        370        380        390
9         17R       AATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTCTCTTACAAATTCAATAAAAA
78        T2R       AATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTCTCTTACAAATTCAATAAAAA
87        T34F      AATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTCTCTTACAAATTCAATAAAAA
82        T31R      AATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTCTCTTACAAATTCAATAAAAA
55        T131R     AATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTCTCTTACAAATTCAATAAAAA
48        T123R     AATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTCTCTTACAAATTCAATAAAAA
62        T141_F    AATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTCTCTTACAAATTCAATAAAAA
11        84R       AATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTCTCTTACAAATTCAATAAAAA
12        86R       AATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTCTCTTACAAATTCAATAAAAA
42        T116_F    AATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTCTCTTACAAATTCAATAAAAA
107       T53R      AATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTCTCTTACAAATTCAATAAAAA
109       T54R      AATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTCTCTTACAAATCCAATAAAAA
112       T58_F     AATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTCGAACATCTGGTTCTCTTACAAATTCAATAAAAA
113       T5F       AATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTTTTTTACAAATTCAATAAAAA
115       T62R      AATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTCTCTTACAAATTCAATAAAAA
117       T6R       AATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTCTCTTACAAATTCAATAAAAA
119       T72R      AATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTCTCTTACAAATTCAATAAAAA
128       T86F      AATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTCGAACATCTGGTTCTCTTACAAATTCAATAAAAA
131       T88_F     AATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTCTCTTACAAATTCAATAAAAA
35        T111R     AATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTCTCTTACAAATTCAATAAAAA
-36       T111_F    AATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTCTCTTACAAATTCAATAAAAA
57        T135R     AATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTCTCTTACAAATTCAATAAAAA
-66       T17F      AATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTCTCTTACAAATTCAATAAAAA
-106      T53F      AATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTCTCTTACAAATTCAATAAAAA
-77       T2F                                                                          CAATAAAAA

CONSENSUS          AATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTCTCTTACAAATTCAATAAAAA


                         400        410        420        430        440        450        460        470
78        T2R       AACGTACTATTCCATACCATATCATAAAAATAGAAAATAACATACCTTGATATGATTTTACCTTAGTAAAAAAAAATAA
87        T34F      AACGTACTATTCCATACCATATCATAAAAATAGAAAATAACATACCTTGATATGATTTTACCTTAGTAAAAAAAAATAA
82        T31R      AACGTACTATTCCATACCATATCATAAAAATAGAAAATAACATACCTTGATATGATTTTACCTTAGTAAAAAAAAATAA
55        T131R     AACGTACTATTCCATACCATATCATAAAAATAGAAAATAACATACCTTGATATGATTTTACCTTAGTAAAAAAAAATAA
48        T123R     AACGTACTATTCCATACCATATCATAAAAATAGAAAATAACATACCTTGATATGATTTTACCTTAGTAAAAAAAAATAA
62        T141_F    AACGTACTATTCCATACCATATCATAAAAATAGAAAATAACATACCTTGATACGATTTTACCTTAGTAAAAAAAAATAA
11        84R       AACGTACTATTCCATACCATATCATAAAAATAGAAAATAACATACCTTGATATGATTTTACCTTAGTAAAAAAAAATAA
12        86R       AACGTACTATTCCATACCATATCATAAAAATAGAAAATAACATACCTTGATATGATTTTACCTTAGTAAAAAAAAATAA
42        T116_F    AACGTACTACTCCATACCATATCATAAAAATAGAAAATAACATACCTTGATATGATTTTACCTTAGTAAAAAAAAATAA
107       T53R      AACGTACTATTCCATACCATATCATAAAAATAGAAAATAACATACCTTGATATGATTTTACCTTAGTAAAAAAAAATAA
109       T54R      AACGTACTATTCCATACCATATCATAAAAATAGAAAATAACATACCTTGATATGATTTTACCTTAGTAAAAAAAAATAA
112       T58_F     AACGTACTATTCCATACCATATCATAAAAATAGAAAATAACATACCTTGATATGATTTTACCTTAGTAAAAAAAAATAA
113       T5F       AACGTACTATTCCATACCATATCATAAAAATAGAAAACAACATACCTTGATATGATTTTACCTTAGTAAAAAAAAATAA
115       T62R      AACGTACTATTCCATACCATATCATAAAAATAGAAAATAACATACCTTGATATGATTTTACCTTAGTAAAAAAAAATAA
117       T6R       AACGTACTATTCCATACCATATCATAAAAATAGAAAATAACATACCTTGATATGATTTTACCTTAGTAAAAAAAAATAA
119       T72R      AACGTACTATTCCATACCATATCATAAAAATAGAAAATAACATACCTTGATATGATTTTACCTTAGTAAAAAAAAATAA
128       T86F      AACGTACTATTCCATACCATATCATAAAAATAGAAAATAACATACCTTGATATGATTTTACCTTAGTAAAAAAAAATAA
131       T88_F     AACGTACTATTCCATACCATATCATAAAAATAGAAAATAACATACCTTGATATGATTTTACCTTAGTAAAAAAAAATAA
35        T111R     AACGTACTATTCCATACCATATCATAAAAATAGAAAATAACATACCTTGATATGATTTTACCTTAGTAAAAAAAAATAA
-36       T111_F    AACGTACTATTCCATACCATATCATAAAAATAGAAAATAACATACCTTGATATGATTTTACCTTAGTAAAAAAAAATAA
57        T135R     AACGTACTATTCCATACCATATCATAAAAATAGAAAATAACATACCTTGATATGATTTTACCTTAGTAAAAAAA**TAA
-66       T17F      AACGTACTATTCCATACCATATCATAAAAATAGAAAATAACATACCTTGATATGATTTTACCTTAGTAAAAAAAAATAA
-106      T53F      AACGTACTATTCCATACCATATCATAAAAATAGAAAATAACATACCTTGATATGATTTTACCTTAGTAAAAAAAAATAA
-77       T2F       AACGTACTATTCCATACCATATCATAAAAATAGAAAATAACATACCTTGATATGATTTTACCTTAGTAAAAAAAAATAA

CONSENSUS          AACGTACTATTCCATACCATATCATAAAAATAGAAAATAACATACCTTGATATGATTTTACCTTAGTAAAAAAAAATAA
```

```
                        480        490        500        510        520        530        540        550
78      T2R     TAAATTCATTACTACAAACAATAATAATCCTTCAAAAA
87      T34F    TAAATTCATTACTACAAACAATAATAATCCTTCAAAAAAAGCTTCATATAACTGACTAGGATGCCTATAAAAAA
82      T31R    TAAATTCATTACTACAAACAATAATAATCCTTCAAAAAAAGCTTCATATAACTGACTAGGATGCCTATAAAAAAA
55      T131R   TAAATTCATTACTACAAACAATAATAATCCTTCAAAAAAAGCTTCATATAACTGGCTAGGATGCCTATAAAAAAAATCT
48      T123R   TAAATTCATTACTACAAACAATAATAATCCTTCAAAAAAAGCTTCATATAACTGACTAGGATGCCTATAAAAAAAATCT
62      T141_F  TAAATTCATTACTACAAACAATAATAATCCTTCAAAAAAAGCTTCATATAACTGACTAGGATGCCTATAAAAAAAATCT
11      84R     TAAATTCATTACTACAAACAATAATAATCCTTCAAAAAAAGCTTCATATAACTGACTAGGATGCCTATAAAAAAAATCT
12      86R     TAAATTCATTACTACAAACAATAATAATCCTTCAAAAAAAGCTTCATATAACTGACTAGGATGCCTATAAAAAAAATCT
42      T116_F  TAAATTCATTACTACAAACAATAATAATCCTTCAAAAAAAGCTTCATATAACTGACTAGGATGCCTATAAAAAAAATCT
107     T53R    TAAATTCATTACTACAAACAATAATAATCCTTCAAAAAAAGCTTCATATAACTGACTAGGATGCCTATAAAAAAAATCT
109     T54R    TAAATTCATTACTACAAACAATAATAATCCTTCAAAAAAAGCTTCATATAACTGACTAGGATGCCTATAAAAAAAATCT
112     T58_F   TAAATTCATTACTACAAACAATAATAATCCTTCAAAAAAAGCTTCATATAACTGACTAGGATGCCCATAAAAAAAATCT
113     T5F     TAAATTCATTACTACAAACAATAATAATCCTTCAAAAAAAGCTTCATATAACTGACTAGGATGCCTATAAAAAAAAACT
115     T62R    TAAATTCATTACTACAAACAATAATAATCCTTCAAAAAAAGCTTCATATAACTGACTAGGATGCCTATAAAAAAAATCT
117     T6R     TAAATTCATTACTACAAACAATAATAATCCTTCAAAAAAAGCTTCATATAACTGACTAGGATGCCTATAAAAAAAATCT
119     T72R    TAAATTCATTACTACAAACAATAATAATCCTTCAAAAAAAGCTTCATATAACTGACTAGGATGCCTATAAAAAAAATCT
128     T86F    TAAATTCATTACTACAAACAATAATAATCCTTCAAAAAAAGCTTCATATAACTGACTAGGATGCCCATAAAAAAAATCT
131     T88_F   TAAATTCATTACTACAAACAATAATAATCCTTCAAAAAAAGCTTCATATAACTGACTAGGATGCCTATAAAAAAAATCT
35      T111R   TAAATTCATTACTACAAACAATAATAATCCTTCAAAAAAAGCTTCGTATAACTGACTAGGATGCCTATAAAAAAAATCT
-36     T111_F  TAAATTCATTACTACAAACAATAATAATCCTTCAAAAAAAGCTTCGTATAACTGACTAGGATGCCTATAAAAAAAATCT
57      T135R   TAAATTCATTACTACAAACAATAATAATCCTTCAAAAAAAGCTTCATATAACTGACTAGGATGCCTATAAAAAAAATCT
-66     T17F    TAAATTCATTACTACAAACAATAATAATCCTTCAAAAAAAGCTTCATATAACTGACTAGGATGCCTATAAAAAAAATCT
-106    T53F    TAAATTCATTACTACAAACAATAATAATCCTTCAAAAAAAGCTTCATATAACTGACTAGGATGCCTATAAAAAAAATCT
-77     T2F     TAAATTCATTACTACAAACAATAATAATCCTTCAAAAAAAGCTTCATATAACTGACTAGGATGCCTATAAAAAAAATCT

CONSENSUS       TAAATTCATTACTACAAACAATAATAATCCTTCAAAAAAAGCTTCATATAACTGACTAGGATGCCTATAAAAAAAATCT


                        560        570        580        590        600        610        620        630
55      T131R   CCACTATTTTGAAAAATCATTCCAAACCGTGTATTTGTAACTTTACCATATAACTCACCATTAATAAAATTCGCAATAC
48      T123R   CCACTATTTTGAAAAATCATTCCAAACCGTGTATTTGTAACTTTACCATATAACTCACCATTAATAAAATTCGCAATAC
62      T141_F  CCACTATTTTGAAAAGTCATTCCAAACCGTGTATTTGTAACTTTACCATATAACTCACCATTAATAAAATTCGCAATAC
11      84R     CCACTATTTTGAAAAATCATTCCAAACCGTGTATTTGTAACTTTACCATATAACTCACCATTAATAAAATTCGCAATAC
12      86R     CCACTATTTTGAAAAATCATTCCAAACCGTGTATTTGTAACTTTACCATATAACTCACCATTAATAAAATTCGCAATAC
42      T116_F  CCACTATTTTGAAAAATCATTCCAAACCGTGTATTTGTAACTTTACCATATAACTCACCATTAATAAAATTCGCAATAC
107     T53R    CCACTATTTTGAAAAATCATTCCAAACCGTGTATTTGTAACTTTACCATATAACTCACCATTAATAAAATTCGCAATAC
109     T54R    CCACTATTTTGAAAAATCATTCCAAACCGTGTATTTGTAACTTTACCATATAACTCACCATTAATAAAATTCGCAATAC
112     T58_F   CCACTATTTTGAAAAATCATTCCAAACCGTGTATTTGTAACTTTACCATATAACTCACCATTAATAAAATTCGCAATAC
113     T5F     CCACTATTTTGAAA
115     T62R    CCACTATTTTGAAAAATCATTCCAAACCGTGTATTTGTAACTTTACCATATAACTCACCATTAATAAAATTCGCAATAC
117     T6R     CCACTATTTTGAAAAATCATTCCAAACCGTGTATTTGTAACTTTACCATATAACTCACCATTAATAAAATTCGCAATAC
119     T72R    CCACTATTTTGAAAAATCATTCCAAACCGTGTATTTGTAACTTTACCATATAACTCACCATTAATAAAATTCGCAATAC
128     T86F    CCACTATTTTGAAA
131     T88_F   CCACTATTTTGAAAAATCATTCCAAACCGTGTATTTGTAACTTTACCATATAACTCACCATTAATAAAATTCGCAATAC
35      T111R   CCACTATTTTGAAAAAT
-36     T111_F  CCACTATTTTGAAAAAT
57      T135R   CCACTATTTTGAAAAATCATTCCAAACCGTGTATTTGTAACTTTACCATATAACTCACCATTAATAAAATTCGCAATAC
-66     T17F    CCACTATTTTGAAAAATCATTCCAAACCGTGTATTTGTGACTTTACCATATAACTCACCATTAATAAAATTCGCAATAC
-106    T53F    CCACTATTTTGAAAAATCATTCCAAACCGTGTATTTGTAACTTTACCATATAACTCACCATTAATAAAATTCGCAATAC
-77     T2F     CCACTATTTTGAAAAATCATTCCAAACCGTGTATTTGTAACTTTACCATATAACTCACCATTAATAAAATTCGCAATAC
-41     T116R                                                           CTCACCATTAATAAAATTCGCAATAC

CONSENSUS       CCACTATTTTGAAAAATCATTCCAAACCGTGTATTTGTAACTTTACCATATAACTCACCATTAATAAAATTCGCAATAC


                        640        650        660        670        680        690        700
55      T131R   GTCCTAAAAATATTC
48      T123R   GTCCTAAAAATATTCCGA
62      T141_F  GTCCTAAAAATATTCCGACAGGTACAGCACACAAACACAAATCAATTGCTGATAAA
11      84R     GTCCTANAAATATTCCGACAGGTACAGCACACAAACACAAATCAATTGCTGATAAAAACTTTATT*TATACTTTTTACA
12      86R     GTCCTAAAAATATTCCGACAGGTACAGCACACAAACACAAATCAATTGCTGATAAAAACTTTATTNTATACTTTTTACA
42      T116_F  GTCCTAAAAATATTCCGACAGGTACAGCACACAAACACAAATCAATTGCTGATAAAAACTTTATTTTATACTTTTTACA
107     T53R    GTCCTAAAAA
109     T54R    GTCCTAAAAATATTCCGA
112     T58_F   GTCCTAAAAATATTCCGACAGGTACAGCACACAAACACAAATCAATTGCTGATAAAAACTNTATTTTATACTTTNTACA
115     T62R    GTCCTAAAAATATTCCGA
117     T6R     GTCCTAAAAA
119     T72R    GTCCTAAAAATATTCCGAC
131     T88_F   GTCCTAAAAATATTCCGACAGGTACAGCACACAAACACAAATCAATTGCTGATAAAAACTNTATTNTATACTTTNTACA
57      T135R   GTCCTAAAAATATTCCGAC
-66     T17F    GTCCTAAAAATATTCCGACAGGTACAGCACACAAACACAAATCAATTGCTGATAAAAACTTTATTTTATACTTTTTACA
-106    T53F    GTCCTAAAAATATTCTGACAGGTACAGCACACAAACACAAATCAATTGCTGATAAAAACTTTATTTTATACTTTTTACA
-77     T2F     GTCCTAAAAATATTCCGACAGGTACAGCACACAAACACAAATCAATTGCTGATAAAAACTTTATTTTATACTTTTTACA
-41     T116R   GTCCTAAAAATATTCCGACAGGTACAGCACACAAACACAAATCAATTGCTGATAAAAACTTTATTTTATACTTTTTACA

CONSENSUS       GTCCTAAAAATATTCCGACAGGTACAGCACACAAACACAAATCAATTGCTGATAAAAACTTTATTTTATACTTTTTACA
```

```
                710       720       730       740       750       760       770       780
11      84R     AAAA
12      86R     AAAA
42      T116_F  AAAAATATACATAGTACAAAAAAGCCCTAAACTTGCTCCATGAAATGACATA
112     T58_F   AAAAATATACATAGTACAAA
131     T88_F   AAAAATATACATAGTACAAAAAAGCCCTAAACTTGCTCCATGAAATGACATACCACCTTTCCATAACTTAAACATTTC
-66     T17F    AAAAATATACATAGTACAAAAAAGCCCTAAACTTGCTCCATGAA
-106    T53F    AAAAATATACATAGTACAAAAAAGCCCTAAACTTGCTCCATGAAATGACATACCACCTTTCCATAACTTAAACATTTCA
-77     T2F     AAAAATATACATAGTACAAAAAAGCCCTAAACTTGCTCCATGAAATGACATACCACCTTTCCATAACTTAAACATTTCA
-41     T116R   AAAAATATACATAGTACAAAAAAGCCCTAAACTTGCTCCATGAAATGACATACCACCTTTCCATAACTTAAACATTTCA
-88     T34R                                                                          ATTTCA

CONSENSUS       AAAAATATACATAGTACAAAAAAGCCCTAAACTTGCTCCATGAAATGACATACCACCTTTCCATAACTTAAACATTTCA*


                790       800       810       820       830       840       850       860
-106    T53F    ATAGGAAAACTCATATAGAAATTTAAGTTATAAAATAATATAT
-77     T2F     ATAGGAAAACTCATATAGAAATTTAAGTTATAAAATAATATATATCCTATCCTTCCACCAAGAATCATTCCAGTAACCC
-41     T116R   ATAGGAAAACTCATATAGAAATTTAAGTTATAAAATAATATATATCCTATCCTTCCACCAAGAATCATTCCAGTAACCC
-88     T34R    ATAGGAAAACTCATATAGAAATTTAAGTTATAAAATAATATATATCCTATCCTTCCACCAAGAATCATTCCAGTAACCC

CONSENSUS       ATAGGAAAACTCATATAGAAATTTAAGTTATAAAATAATATATATCCTATCCTTCCACCAAGAATCATTCCAGTAACCC


                870       880       890       900       910       920       930       940
-77     T2F     ACCATGAAATTATCGACTTATAACTTTCTGGAGTAAAAACCTTGTACTTATCTATTTTCTGCACATACCAATAAGCAAA
-41     T116R   ACCATGAAATTATCGACTTATAACTTTCTGGAGTAAAAACCTTGTACTTATCTATTTTCTGCACATACCAATAAGCAAA
-88     T34R    ACCATGAAATTATCGACTTATAACTTTCTGGAGTAAAAACCTTGTACTTATCTATTTTCTGCACATACCAATAAGCAAA

CONSENSUS       ACCATGAAATTATCGACTTATAACTTTCTGGAGTAAAAACCTTGTACTTATCTATTTTCTGCACATACCAATAAGCAAA


                950       960       970       980       990       1000      1010      1020
-77     T2F     TAAAATACCTATTATATATGCTAAAGAATACCATCTGATCTCTAACAACCCTATCTTTAAAGCCACTGGATCTATATTC
-41     T116R   TAAAATACCTATTATATATGCTAAAGAATACCATCTGATCTCTAACAACCCTATCTTTAAAGCCACTGGATCTATATTC
-88     T34R    TAAAATACCTATTATATATGCTAAAGAATACCATCTGATCTCTAACAACCCTATCTTTAAAGCCACTGGATCTATATTC
-114    T5R                                                                           TATTC

CONSENSUS       TAAAATACCTATTATATATGCTAAAGAATACCATCTGATCTCTAACAACCCTATCTTTAAAGCCACTGGATCTATATTC


                1030      1040      1050      1060      1070      1080      1090      1100
-77     T2F     ATAAAATATAATTACCTCTTTAATA
-41     T116R   ATAAAATATAATTACCTCTTTAATAATTTTAATAAATAACTATATACTTCCATAAAAGCATTAATACAGTTTTTTCTAT
-88     T34R    ATAAAATATAATTACCTCTTTAATAATTTTAATAAATAACTATATACTTCCATAAAAGCATTAATACAGTTTTTTCTAT
-114    T5R     ATAAAATATAATTACCTCTTTAATAATTTTAATAAATAACTATATACTTCCATAAAAGCATTAATACAGTTTTTTCTAT

CONSENSUS       ATAAAATATAATTACCTCTTTAATAATTTTAATAAATAACTATATACTTCCATAAAAGCATTAATACAGTTTTTTCTAT


                1110      1120      1130      1140      1150      1160      1170      1180
-41     T116R   ATTCTCAACTATAAATAATAGAGAATTTGGCATCTTCTAAAATTTAAAAACATATCTGGCAATACATATTCTTAAAAAT
-88     T34R    ATTCTCAACTATAAACAATAGAGAATTTGGCATCTTCTAAAATTTAAAAACATATCTGGCAATACATATTCTTAAAAAT
-114    T5R     ATTCTCAACTATAAACAATAGAGAATTTGGCATCTTTTAAAATTTAAAAACATATCTGGCAATACATATTCTTAAAAAT

CONSENSUS       ATTCTCAACTATAAACAATAGAGAATTTGGCATCTTCTAAAATTTAAAAACATATCTGGCAATACATATTCTTAAAAAT


                1190      1200      1210      1220      1230      1240      1250      1260
-41     T116R   AACGTCATTTCACATATCTATAAAAAGTAAAAATATCA
-88     T34R    AACGTCATTTCACATATCTATAAAAAGTAAAAATATCAACCAATAATAACCATGTTGATTTAGTACAATTTATTAAACT
-114    T5R     AACGTCATTTCACATATCTATAAAAAGTAAAAATATCAACCAATAATAACCATGTTGATTTAGTACAATTTATTAAACT

CONSENSUS       AACGTCATTTCACATATCTATAAAAAGTAAAAATATCAACCAATAATAACCATGTTGATTTAGTACAATTTATTAAACT


                1270      1280      1290      1300      1310      1320      1330      1340
-88     T34R    AATTCCATTTTAACAAAATGTTTAATGTTCTAATTTAGATTCACATAAG
-114    T5R     AATTCCATTTTAACAAAATGTTTAATGTTCTAATTTAGATTCACATAAGATTATACATACAAGCTTTCATTAACTATAA
-49     T123_F     TCCATTTTAACANAATGTTTAATGTTCTAATTTAGATTCACATAAGATTATACATACAAGCTTTCATTAATTATAA

CONSENSUS       AATTCCATTTTAACAAAATGTTTAATGTTCTAATTTAGATTCACATAAGATTATACATACAAGCTTTCATTAACTATAA
```

```
                      1350       1360       1370       1380       1390       1400       1410      1420
-114    T5R     ATAGAATAAAAAATATAACGTAATACCAAGTTGCGTTACACTTTATTGCACAGTCTATAATTTATCAATCATAAACAAA
-49     T123_F  ATAGAATAAAAAATATAACGTAATACCAAGTTGCGTTACACTTTATTGCACAGTCTATAATTTATCAATCATAAACAAA

CONSENSUS       ATAGAATAAAAAATATAACGTAATACCAAGTTGCGTTACACTTTATTGCACAGTCTATAATTTATCAATCATAAACAAA


                     1430       1440       1450       1460       1470       1480       1490     1500
-114    T5R     TTAAAAAGAAAGTTTAGTATTAGATTAACAAAGCCTGTAGTGATTCTAACATATATTCA
-49     T123_F  TTAAAAAGAAAGTTTAGTATTAGATTAACAAAGCCTGTAGTGATTCTAACATATATTCACAATACAATATATGTATAAT
-111    T58R                    TATTAGATTAACAAAGCCTGTAGTGATTCTAACATATATTCCCAATACAATATATGTATAAT
-110    T54_F                                                                      TATGTATAAT

CONSENSUS       TTAAAAAGAAAGTTTAGTATTAGATTAACAAAGCCTGTAGTGATTCTAACATATATTCACAATACAATATATGTATAAT


                       1510       1520       1530       1540       1550       1560       1570
-49     T123_F  GTATCATTAAATACAAAAATAATAAGTAAAAACTCAATTTGACACACTCATTATTTAATATAATTAGATATATAACATA
-111    T58R    GTATCATTAAATACAAAAATAATAAGTAAAAACTCAATTTGACACACTCATTATTTAATATAATTAGATATATAACATA
-110    T54_F   GTATCATTAAATACAAAAATAATAAGTAAAAACTCAATTTGACACACTCATTATTTAATATAATTAGATATATAACATA

CONSENSUS       GTATCATTAAATACAAAAATAATAAGTAAAAACTCAATTTGACACACTCATTATTTAATATAATTAGATATATAACATA


                1580       1590       1600       1610       1620       1630       1640       1650
-49     T123_F  TGTTTTCACTATACAAAATTATCAGTACTATTTTTAATATTAAATTGATACATAATAATTAGTTATGGTATTAGTATGA
-111    T58R    TGTTTTCACTATACAAAATTATCAGTACTATTTTTAATATTAAATTGATACATAATAATTAGTTATGGTATTAGTATGA
-110    T54_F   TGTTTTCACTATACANAATTATCAGTACTATTTTTAATATTAAATTGATACATAATAATTAGTTATGGTATTAGTATGA

CONSENSUS       TGTTTTCACTATACAAAATTATCAGTACTATTTTTAATATTAAATTGATACATAATAATTAGTTATGGTATTAGTATGA


                1660       1670       1680       1690       1700       1710       1720       1730
-49     T123_F  TAAAACATTAGGTTTTATTCTATCATAGAAACTATACAAACCTTAATATCCACAATCCTTTACTATAGAAATTTACATA
-111    T58R    TAAAACATTAGGTTTTATTCTATCATAGAAACTATACAAACCTTAATATCCACAATCCTTTACTATAGAAATTTACATA
-110    T54_F   TAAAACATTAGGTTTTATTCTATCATAGAAACTATACAAACCTTAATATCCACAATCCTTTACTATAGAAATTTACATA
-133    T95R                                                                   TAGAAATTTACATA
-108    T54F                                                                      ATTTACATA

CONSENSUS       TAAAACATTAGGTTTTATTCTATCATAGAAACTATACAAACCTTAATATCCACAATCCTTTACTATAGAAATTTACATA


                 1740       1750       1760       1770       1780       1790       1800       1810
-49     T123_F  TTCTTAAAAGAAAATAGCAAGATATTATACATCTATTATCGCCAACATCAAAACTTGTCTAGCATAAAACCTTTGTACC
-111    T58R    TTCTTAAAAGAAAATAGCAAGATATTATACATCTATTATCGCCAACATCAAAACTTGTCTAGCATAAAACCTTTGTACC
-110    T54_F   TTCTTAAAAGAAAATAGCAAGATATTATACATCTATTATCGCCAACATCAAAACTTGTCTAGCATAAAACCTTTGTACC
-133    T95R    TTCTTAAAAGAAAATAGCAAGATATTATACATCTATTATCGCCAACATCAAAACTTGTCTAGCATAAAACCTTTGTACC
-108    T54F    TTCTTAAAAGAAAATAGCAAGATATTATACATCTATTATCGCCAACATCAAAACTTGTCTAGCATAAAACCTTTGTACC

CONSENSUS       TTCTTAAAAGAAAATAGCAAGATATTATACATCTATTATCGCCAACATCAAAACTTGTCTAGCATAAAACCTTTGTACC


                   1820       1830       1840       1850       1860       1870       1880       1890
-49     T123_F  ATTATTATATAAATAATCAACAACAAGGTACTACCTAGTACCTACAAATTCTATTATAGATTCATCTTAAGCTATAATG
-111    T58R    ATTATTATATAAATAATCAACAACAAGGTACTACCTAGTACCTACAAATTCTATTATAGATTCATCTTAAGCTATAATG
-110    T54_F   ATTATTATATAAATAATCAACAACAAGGTACTACCTAGTACCTACAAATTCTATTATAGATTCATCTTAAGCTATAATG
-133    T95R    ATTATTATATAAATAATCAACAACAAGGTACTACCTAGTACCTACAAATTCTATTATAGATTCACCTTAAGCTATAATG
-108    T54F    ATTATTATATAAATAATCAACAACAAGGTACTACCTAGTACCTACAAATTCTATTATAGATTCATCTTAAGCTATAATG
-2       141R                                                                          TAATG

CONSENSUS       ATTATTATATAAATAATCAACAACAAGGTACTACCTAGTACCTACAAATTCTATTATAGATTCATCTTAAGCTATAATG


                   1900       1910       1920       1930       1940       1950       1960       1970
-49     T123_F  TATTTATCAAGAAATATCCTATTTAATAATAGATATTAATCATCTCACACTTATGAAATTACACTTCACTAATTTCACA
-111    T58R    TATTTATCAAGAAATATCCTATTTAATAATAGATATTAATCATCTCACACTTATGAAATTACACTTCACTAATTTCACA
-110    T54_F   TATTTATCAAGAAATATCCTATTTAATAATAGATATTAATCATCTCACACTTATGAAATTACACTTCACTAATTTCACA
-133    T95R    TATTTATCAAGAAATATCCTATTTAATAATAGATATTAATCATCTCACACTTATGAAATTACACTTCACTAATTTCACA
-108    T54F    TATTTATCAAGAAATATCCTATTTAATAATAGATATTAATCATCTCACACTTATGAAATTACACTTCACTAATTTCACA
-2       141R   TATTTATCAAGAAATATCCTATTTAATAATAGATATTAATCATCTCACACTTATGAAATTACACTTCACTAATTTCACA
28      T108_F                  CTATTTAATAATAGATATTAATCATCTCACACTTATGAAATTACACTTCACTAATTTCACA

CONSENSUS       TATTTATCAAGAAATATCCTATTTAATAATAGATATTAATCATCTCACACTTATGAAATTACACTTCACTAATTTCACA
```

```
                    1980       1990       2000       2010       2020       2030       2040       2050
-49      T123_F  ATATGTAGTTTTTAAATTAATATTTAAATCTCGAAAGAAATTTAGGTAAT
-111     T58R    ATATGTAGTTTTTAAATTAATATTTAAATCTCGAAAGAAATTTAGGTAATAAAATTTATTAGTTATATTAAACATTTTA
-110     T54_F   ATATGTAGTTTTTAAATTAATATTTAAATCTCGAAAGAAATTTAGGTAATAAAATTTATTAGTTATATTAAACATTTTA
-133     T95R    ATATGTAGTTTTTAAATTAATATTTAAATCTCGAAAGAAATTTAGGTAATAAAATTTATTAGTTATACTAAACATTTTA
-108     T54F    ATATGTAGTTTTTAAATTAATATTTAAATCTCGAAAGAAATTTAGGTAATAAAATTTATTAGTTATATTAAACATTTTA
-2       141R    ATATGTAGTTTTTAAATTAATATTTAAATCTCGAAAGAAATTTAGGTAATAAAATTTATTAGTTATATTAAACATTTTA
28       T108_F  ATATGTAGTTTTTAAATTAATATTTAAATCTCGAAAGAAATTTAGGTAATAAAATTTATTAGTTATATTAAACATTTTA

CONSENSUS        ATATGTAGTTTTTAAATTAATATTTAAATCTCGAAAGAAATTTAGGTAATAAAATTTATTAGTTATATTAAACATTTTA


                        2060       2070       2080       2090       2100       2110       2120       2130
-111     T58R    TAGTAATAAATAGTATTTTTACTATTTATTATACGTCA
-110     T54_F   TAGTAATAAATAGTATTTTTACTATTTATTATATGTCATACAACACATTATTAAACTAACAAGTTTTGCCAAATCCTAA
-133     T95R    TAGTAATAAATAGTATTTTTACTATTTATTATATGTCATACAACACATTATTAAACTAACAAGTTTTGCCAAATCCTAA
-108     T54F    TAGTAATAAATAGTATTTTTACTATTTATTATATGTCATACAACACATTATTAAACTAACAAGTTTTGCCAAATCCTAA
-2       141R    TAGTAATAAATAGTATTTTTACTATTTATTATATGTCATACAACACATTATTAAACTAACAAGTTTTGCCAAATCCTAA
28       T108_F  TAGTAATAAATAGTATTTTTACTATTTATTATATGTCATACAACACATTATTAAACTAACAAGTTTTGCCAAATCCTAA
-81      T31F                                                CACATTATTAAACTAACAAGTTTTGCCAAATCCTAA

CONSENSUS        TAGTAATAAATAGTATTTTTACTATTTATTATATGTCATACAACACATTATTAAACTAACAAGTTTTGCCAAATCCTAA


                    2140       2150       2160       2170       2180       2190       2200       2210
-110     T54_F   TATGTTCTCCAAAAAGCACCACTAGCAACACTTATTTTACTAATGATACTACCTAGCATTAATAAATTGATAATTGTAA
-133     T95R    TATGTTCTCCAAAAAGCACCACT
-108     T54F    TATGTTCTCCAAAAAGCACCACTAGCAACACTTATTTTACTAATGATACTACCTAGCATTAATAAATTGATAATTGTAA
-2       141R    TATGTTCTCCAAAAAGCACCACTAGCAACACTTATTTTACTAATGATACTACCTAGCATTAATAAATTGATAATTGTAA
28       T108_F  TATGTTCTCCAAAAAGCACCACTAGCAACACTTATTTTACTAATGATACTACCTAGCATTAATAAATTGATAATTGTAA
-81      T31F    TATGTTCTCCAAAAAGCACCACTAGCAACACTTATTTTACTAATGATACTACCTAGCATTAATAAATTGATAATTGTAA
63       T145_F                                      TTTACTAATGATACTACCTAGCATTAATAAATTGATAATTGTAA
-3       145R                                        TTTACTAATGATACTACCTAGCATTAATAAATTGATAATTGTAA
-136     T98R                                        TTTACTAATGATACTACCTAGCATTAATAAATTGATAATTGTAA

CONSENSUS        TATGTTCTCCAAAAAGCACCACTAGCAACACTTATTTTACTAATGATACTACCTAGCATTAATAAATTGATAATTGTAA


                     2220       2230       2240       2250       2260       2270       2280       2290
-110     T54_F   CATCTAGAACTAACTCTCATATATCTTAAATATAAACAGTAAG
-108     T54F    CATCTAGAACTAACTCTCATATATCTTAAATATAAACAGTAAG
-2       141R    CATCTAGAACTAACTCTCATATATCTTAAATATAAACAGTAAGACATTATCTATCACCCTAAATTCTATCATAATTACA
28       T108_F  CATCTAGAACTAACTCTCATATATCTTAAATATAAACAGTAAGACATTATCTATCACCCTAAATTCTATCATAATTACA
-81      T31F    CATCTAGAACTAACTCTCATATATCTTAAATATAAACAGTAAGACATTATCTATCACCCTAAATTCTATCATAATTACA
63       T145_F  CATCTAGAACTAACTCTCATATATCTTAAATATAAACAGTAAGACATTATCTATCACCCTAAATTCTATCATAATTACA
-3       145R    CATCTAGAACTAACTCTCATATATCTTAAATATAAACAGTAAGACATTATCTATCACCCTAAATTCTATCATAATTACA
-136     T98R    CATCTAGAACTAACTCTCATATATCTTAAATATAAACAGTAAGACATTATCTATCACCCTAAATTCTATCATAATTACA

CONSENSUS        CATCTAGAACTAACTCTCATATATCTTAAATATAAACAGTAAGACATTATCTATCACCCTAAATTCTATCATAATTACA


                   2300       2310       2320       2330       2340       2350       2360       2370
-2       141R    GCATTTGTTTTATTTTATATAAAGACACCACTTTTTATACTTTAATAAGTCCATAAAAGTATAATTAACAGAACACTTG
28       T108_F  GCATTTGTTTTATTTTATATAAAGACACCACTTTTTATACTTTAATAAGTCCATAAAAGTATAATTAACAGAACACTTG
-81      T31F    GCATTTGTTTTATTTTATATAAAGACACCACTTTTTATACTTTAATAAGTCCATAAAAGTATAATTAACAGAACACTTG
63       T145_F  GCATTTGTTTTATTTTATATAAAGACACCACTTTTTATACTTTAATAAGTCCATAAAAGTATAATTAACAGAACACTTG
-3       145R    GCATTTGTTTTATTTTATATAAAGACACCACTTTTTATACTTTAATAAGTCCATAAAAGTATAATTAACAGAACACTTG
-136     T98R    GCATTTGTTTTATTTTATATAAAGACACCACTTTTTATACTTTAATAAGTCCATAAAAGTATAATTAACAGAACACTTG

CONSENSUS        GCATTTGTTTTATTTTATATAAAGACACCACTTTTTATACTTTAATAAGTCCATAAAAGTATAATTAACAGAACACTTG


                     2380       2390       2400       2410       2420       2430       2440       2450
-2       141R    TTTACATATTGCTTGCACACTTTATGGTATAAAACTTCATTAATTTTACAACCTGTAGCACCTACATTAACTCTTCAAT
28       T108_F  TTTACATATTGCTTGCACACTTTATGGTATAAAACTTCATTAATTTTACAACCTGTAGCACCTACATTAACTCTTCAAT
-81      T31F    TTTACATATTGCTTGCACACTTTATGGTATAAAACTTCATTAATTTTACAACCTGTAGCACCTACATTAACTCTTCAAT
63       T145_F  TTTACATATTGCTTGCACACTTTATGGTATAAAACTTCATTAATTTTACAACCTGTAGCACCTACATTAACTCTTCAAT
-3       145R    TTTACATATTGCTTGCACACTTTATGGTATAAAACTTCATTAATTTTACAACCTGTAGCACCTACATTAACTCTTCAAT
-136     T98R    TTTACATATTGCTTGCACACTTTATGGTATAAAACTTCATTAATTTTACAACCTGTAGCACCTACATTAACTCTTCAAT

CONSENSUS        TTTACATATTGCTTGCACACTTTATGGTATAAAACTTCATTAATTTTACAACCTGTAGCACCTACATTAACTCTTCAAT
```

```
                    2460       2470       2480       2490       2500       2510       2520       2530
-2      141R    CTTGAAAAATTAG
28      T108_F  CTTGAAAAATTAGACAACAAACTACCACCAACTATACTAAATGTTTTATAAACACAAATAACATTTTTTCTATTATTGG
-81     T31F    CTTGAAAAATTAGACAACAAACTACCACCAACTATACTAAATGTTTTATAAACACAAATAACATTTTTTCTATTATTGG
63      T145_F  CTTGAAAAATTAGACAACAAACTACCACCAACTATACTAAATGTTTTATAAACACAAATAACATTTTTTCTATTATTGG
-3      145R    CTTGAAAAATTAGACAACAAACTACCACCAACTATACTAAATGTTTTATAAACACAAATAACATTTTTTCTATTATTGG
-136    T98R    CTTGAAAAATTAGACAACAAACTACCACCAACTATACTAAATGTTTTATAAACACAAATAACATTTTTTCTATTATTGG
-130    T88R                GACAACAAACTATCACCAACTATACTAAATGTTTTATAAACACAAATAACATTTTTTCTATTATTGG
-127    T84_F                                          TGTTTTATAAACACAAATAACATTTTTTCTATTATCGG
-4      149R                                                                TTTCTATTATTGG

CONSENSUS       CTTGAAAAATTAGACAACAAACTACCACCAACTATACTAAATGTTTTATAAACACAAATAACATTTTTTCTATTATTGG


                        2540       2550       2560       2570       2580       2590       2600       2610
28      T108_F  ATATTACATCACATATCACTAAATTAATAAGTTTTTCCATAAACTTTAACTTGATGTTTCTCTAGATGTATTACTAGCA
-81     T31F    ATATTACATCACATATCACTAAATTAATAAGTTTTTCCATAAACTTTAACTTGATGTTTCTCT
63      T145_F  ATATTACATCACATATCACTAAATTAATAAGTTTTTCCATAAACT
-3      145R    ATATTACATCACATATCACTAAATTAATAAGTTTTTCCATAAACT
-136    T98R    ATATTACATCACATATCACTAAATTAATAAGTTTTTCCATAAACT
-130    T88R    ATATTACATCACATATCACTAAATTAATAAGTTTTCCCATAAACTTTAACGTGATGTTTCTCTAGATGTATTACTAGCA
-127    T84_F   ATATTACATCACATATCACTAAATTAATAAGTTTTCCCATAAACTTTAACGTGATGTTTCTCTAGATGTATTACTAGCA
-4      149R    ATATTACATCACATATCACTAAATTAATAAGTTTTCCCATAAACTTCAACGTGATGTTTCTCTAGATGTATTACTAGCA
59      T137_F                                          ATAAACTTTAACGTGATGTTTCTCTAGATGTATTACTAGCA

CONSENSUS       ATATTACATCACATATCACTAAATTAATAAGTTTTTCCATAAACTTTAACGTGATGTTTCTCTAGATGTATTACTAGCA


                          2620       2630       2640       2650       2660       2670       2680       2690
28      T108_F  GTACCTGTTACAACTGCCTCTTCTATGTCACTATCTAGCACTTCAGCAACTTCACAACCTGTAGCACTTACATTAACTC
-130    T88R    GTGCCTGTTACAACTGCTTCTTCTATGTCACTATCTAGCAGTTCAGTAATTTCACAACCTGTAGCACCTACATTAACTC
-127    T84_F   GTGCCTGTTACAACTGCTTCTTCTATGTCACTATCTAGCAGTTCAGTAATTTCACAACCTGTAGCACCTACATTAACTC
-4      149R    GTGCCTGTTACAACTGCTTCTTCTATGTCACTATCTAGCAGTTCAGTAATTTCACAACCTGTAGCACCTACATTAACTC
59      T137_F  GTGCCTGTTACAACTGCTTCTTCTATGTCACTATCTAGCAGTTCAGTAATTTCACAACCTGTAGCACCTACATTAACTC
-116    T6F                                 TGTCACTATCTAGCAGTTCAGTAATTTCACAACCTGTAGCACCTACATTAACTC

CONSENSUS       GTGCCTGTTACAACTGCTTCTTCTATGTCACTATCTAGCAGTTCAGTAATTTCACAACCTGTAGCACCTACATTAACTC


                        2700       2710       2720       2730       2740       2750       2760     2770
28      T108_F  TTCAATCTTGAAAA
-130    T88R    TTGAATCTTGAACAGCAGGATCAGGTAATAAAGTATCATCATCTACCCTAGGAACTGTATCAATATCCCCAGTACTTTT
-127    T84_F   TTGAATCTTGAACAGCAGGATCAGGTAATAAAGTATCATCATCTACCCTAGGAACTGTATCAATATCCCCAGTACTTTT
-4      149R    TTGAATCTTGAACAGCAGGATCAGGTAATAAAGTATCATCATCTACCCTAGGAACTGTATCAATATCCCCAGTACTTTT
59      T137_F  TTGAATCTTGAACAGCAGGATCAGGTAATAAAGTATCATCATCTACCCTAGGAACTGTATCAATATCCCCAGTACTTTT
-116    T6F     TTGAATCTTGAACAGCAGGATCAGGTAATAAAGTATCATCATCTACCCTAGGAACTGTATCAATACCCCCAGTACTTTT

CONSENSUS       TTGAATCTTGAACAGCAGGATCAGGTAATAAAGTATCATCATCTACCCTAGGAACTGTATCAATATCCCCAGTACTTTT


                      2780       2790       2800       2810       2820       2830       2840
-130    T88R    TTCCTCTTGTAGATTATCTAATGCCTGCTGACCAAACACACCAGCTTTACAGGCATTATAAATAGGCATATTAACACTT
-127    T84_F   TTCCTCTTGTAGATTATCTAATGCCTGCTGACCAGACACATCAGCTTTACAGGCATTATAAATAGGCATATTAACACTT
-4      149R    TTCCTCTTGTAGATTATCTAATGCCTGCTGACCAAACACACCAGCTTTACAGGCATTATAAATAGGCATATTAACACTT
59      T137_F  TTCCTCTTGTAGATTATCTAATGCCTGCTGACCAAACACACCAGCTTTACAGGCATTATAAATAGGCATATTAACACTT
-116    T6F     TTCCTCTTGTAGATTATCTAATGCCTGCTGACCAAACACACCAGCTTTACAGGCATTATAAATAGGCATATTAACACTT
-47     T122_F                        TGCCTGCTGACCAAACACACCAGCTTTACAGGCATTATAAATAGGCATATTAACACTT

CONSENSUS       TTCCTCTTGTAGATTATCTAATGCCTGCTGACCAAACACACCAGCTTTACAGGCATTATAAATAGGCATATTAACACTT


                    2850       2860       2870       2880       2890       2900       2910       2920
-130    T88R    TCAACAAACATAAATTCATATTTGCCAATTTTGATATATTCAATGTCATCAAACATTTTTAGGCGATCTTTTATAGATA
-127    T84_F   TCAACAAACATAAATTCATATTTGCCAATTTTGATATATTCAATGTCATCAAACATTTTTAGGTGATCTTTTATAGATA
-4      149R    TCAACAAACATAAATTCATATTTGCCAATTTTGATATATTCAATGTCATCAAACATTTTTAGGTGATCTTTTATAGATA
59      T137_F  TCAACAAACATAAATTCATATTTGCCAATTTTGATATATTCAATGTCATCAAACATTTTTAGGTGATCTTTTATAGATA
-116    T6F     TCAACAAACATAAATTCATATTTGCCAATTTTGATATATTCAATGTCATCAAACATTTTTAGGTGATCTTTTATAGATA
-47     T122_F  TCAACAAACATAAATTCATATTTGCCAATTTTGATATATTCAATGTCATCAAACATTTTTAGGTGATCTTTTATAGATA
37      T112R                                                                   TCTTTTATAGATA

CONSENSUS       TCAACAAACATAAATTCATATTTGCCAATTTTGATATATTCAATGTCATCAAACATTTTTAGGTGATCTTTTATAGATA
```

```
                2930       2940       2950       2960       2970       2980       2990       3000
-130    T88R    AGTAAGCTGTAGTACTAAGACCATCTAATTTTGATACTATTTTTCGATCTTTCCTACGCAGAAAATGACATATTAAAGT
-127    T84_F   AGTAAGCTGTAGTACTAAGACCATCTAATTTTGATACTATTTTTCGATCTTTCCTACGCAGAAAATGACATATTAAAGT
-4      149R    AGTAAGCTGTAGTACTAAGACCATCTAATTTTGATACTATTTTTCGATCTTTCCTACGCAGAAAATGACATATTAAAGT
59      T137_F  AGTAAGCTGTAGTACTAAGACCATCTAATTTTGATACTATTTTTCGATCTTTCCTACGCAGAAAATGACATATTAAAGT
-116    T6F     AGTAAGCTGTAGTACTAAGACCATCTAATTTTGATACTATTTTTCGATCTTTCCTACGCAGAAAATGACATATTAAAGT
-47     T122_F  AGTAAGCTGTAGTACTAAGACCATCTAATTTTGATACTATTTTTCGATCTTTCCTACGCAGAAAATGACATATTAAAGT
37      T112R   AGTAAGCTGTAGTACTAAGACCATCTAATTTTGATACTATTTTTCGATCTTTCCTACGCGGAAAATGACATATTAAAGT
-44     T11R                                                                AGAAAATGACATATTAAAGT

CONSENSUS       AGTAAGCTGTAGTACTAAGACCATCTAATTTTGATACTATTTTTCGATCTTTCCTACGCAGAAAATGACATATTAAAGT


                 3010       3020       3030       3040       3050       3060       3070       3080
-130    T88R    AATAGCTATCATTGTTCCAGTTATTCCAGCTATTACACTAGATAATACAAACTCATGACCAAATTTAGCATCGCAACAC
-127    T84_F   AATAGCTATCATTGTTCCAGTTATTCCAGCTATTACACTAGATAATACAAACTCATGACCAAATTTAGCATCGCAACAC
-4      149R    AATAGCTATCATTGTTCCAGTTATTCCAGCTATTACACTAGATAATACAA
59      T137_F  AATAGCTATCATTGTTCCAGTTATTCCAGCTATTACACTAGATAATACAAACTCATGACCAAATTTAGCATCGCAACAC
-116    T6F     AATAGCTATCATTGTTCCAGTTATTCCAGCTATTACACTAGATAATACAAACTCATGACCAAATTTAGCATCGCAACAC
-47     T122_F  AATAGCTATCATTGTTCCAGTTATTCCAGCTATTACACTAGATAATACAAACTCATGACCAAATTTAGCATCGCAACAC
37      T112R   AATAGCTATCATTGTTCCAGTTATTCCAGCTATTACACTAGATAATACAAACTCATGACCAAATTTAGCATCGCAACAC
-44     T11R    AATAGCTATCATTGTTCCAGTTATTCCAGCTATTACACTAGATAATACAAACTCATGACCAAATTTAGCATCGCAACAC

CONSENSUS       AATAGCTATCATTGTTCCAGTTATTCCAGCTATTACACTAGATAATACAAACTCATGACCAAATTTAGCATCGCAACAC


                    3090       3100       3110       3120       3130       3140       3150       3160
-130    T88R    AACATCAATACATTTGTGGCCAT
-127    T84_F   AACATCAATACATTTGTGGCCATAAAAATACCTGCAAATCCAATAATAAACGCTGCACAGACTCTTGATATTTTATTCT
59      T137_F  AACATCAATACATTTGTGGCCATAAAAATACCTGCAAATCCAATAATAAACGCTGCACAGACTCTTGATATTTTATTCT
-116    T6F     AACATCAATACATTTGTGGCCATAAAAATACCTGCAAATCCAATAATAAACG
-47     T122_F  AACATCAATACATTTGTGGCCATAAAAATACCTGCAAATCCAATAATAAACGCTGCACAGACTCTTGATATTTTATTCT
37      T112R   AACATCAATACATTTGTGGCCATAAAAATACCTGCAAATCCAATAATAAACGCTGCACAGACTCTTGATATTTTATTCT
-44     T11R    AACATCAATACATTTGTGGCCATAAAAATACCTGCAAATCCAATAATAAACGCTGCACAGACTCTTGATATTTTATCCT

CONSENSUS       AACATCAATACATTTGTGGCCATAAAAATACCTGCAAATCCAATAATAAACGCTGCACAGACTCTTGATATTTTATTCT


                       3170       3180       3190       3200       3210       3220       3230       3240
-127    T84_F   TATACGATAGTAAATACAATCCTCCTAGAAACATTATAAA
59      T137_F  TATACGATAGTAAATACAATCCTCCTAGAAACATTATAAAATACGATACAACAGAACAGTATAAAGCTATTCTATCAAA
-47     T122_F  TATACGATAGTAAATACAATCCTCCTAGAAACATTATAAAATACGATACAACAGAACAGTATAAAGCTATTCTATCAAA
37      T112R   TATACGATAGTAAATACAATCCTCCTAGAAACATTATAAAAT
-44     T11R    TATACGATAGTAAATACAATCCTCCTAGAAACATTATAAAATACGATACAACAGAACAGTATAAAGCTATTCTATCAAA

CONSENSUS       TATACGATAGTAAATACAATCCTCCTAGAAACATTATAAAATACGATACAACAGAACAGTATAAAGCTATTCTATCAAA


                        3250       3260       3270       3280       3290       3300       3310       3320
59      T137_F  TTTAAAAAAATGACATGA
-47     T122_F  TTTAAAAAAATGACATGATAAAGCCACAAGTTTTAAAAGATACACAGACCCTAGTATTAGAAAAGTAGTACATCTTAAC
-44     T11R    TTTAAAAAAATGACATGATAAAGCCACAAGTTTTAAAAGATACACAGATCCTAGTATTAGAAAAGTAGTACATCTTAAC
34      T110_F            AATGACATGATAAAGCCACAAGTTTTAAAAGATACACAGATCCTAGTATTAGAAAAGTAGTACATCTTAAC

CONSENSUS       TTTAAAAAAATGACATGATAAAGCCACAAGTTTTAAAAGATACACAGATCCTAGTATTAGAAAAGTAGTACATCTTAAC


                     3330       3340       3350       3360       3370       3380       3390       3400
-47     T122_F  AGTAATTGCCTTTTTGTAAGCTCATTGTTACCTGCCATAAGAAACAAATTGTGTTTTGTACTTACTTTCTTTC
-44     T11R    AGTAATTGCCTTTTTGTAAGCTCATCGTTACCTGCCATAAGAAACAAATTGTGTTTTGTACTTACTTTCTTTCCTAACT
34      T110_F  AGTAATTGCCTTTTTGTAAGCTCATTGTTACCTGCCATAAGAAACAAATTGTGTTTTGTACTTACTTTCTTTCCTAACT
120     T73F                                      CCATAAGAAACAAATTGTGTTTTGTACTTACTTTCTTTCCTAACT
-121    T73R                                      CCATAAGAAACAAATTGTGTTTTGTACTTACTTTCTTTCCTAACT

CONSENSUS       AGTAATTGCCTTTTTGTAAGCTCATTGTTACCTGCCATAAGAAACAAATTGTGTTTTGTACTTACTTTCTTTCCTAACT


                     3410       3420       3430       3440       3450       3460       3470       3480
-44     T11R    CAGTAAGCTGTAAATTTCCTTTGCCATCATTGACCATAAAACTCTGTTTTTCAGCTTCCTGTTTTATATGTGATAAAAT
34      T110_F  CAGTAAGCTGTAAATTTCCTTTGCCATCATTGACCATAAAACTCTGTTTTTCAGCTTCCTGTTTTATATGTGATAAAAT
120     T73F    CAGTAAGCTGTAAATTTCCTTTGCCATCATTGACCATAAAACTCTGTTTTTCAGCTTCCTGTTTTATATGTGATAAAAT
-121    T73R    CAGTAAGCTGTAAATTTCCTTTGCCATCATTGACCATAAAACTCTGTTTTTCAGCTTCCTGTTTTATATGTGATAAAAT
31      T10F                          CATCATTGACCATAAAACTCTGTTTTTCAGCTTCCTGTTTTATATGTGATAAAAT

CONSENSUS       CAGTAAGCTGTAAATTTCCTTTGCCATCATTGACCATAAAACTCTGTTTTTCAGCTTCCTGTTTTATATGTGATAAAAT
```

```
                      3490       3500       3510       3520       3530       3540       3550       3560
-44      T11R     TCTCTCTCTTTTCGTAGAAGGAACATCATCATTTAGAAGATCCTCAGAAACAGTACGTATACCTTGATTATATTCTCTA
34       T110_F   TCTCTCTCTTTTCGTAGAAGGAACATCATCATTTAGAAGATCCTCAGAAACAGTACGTATACCTTGATTATATTCTCTA
120      T73F     TCTCTCTCTTTTCGTAGAAGGAACATCATCATTTAGAAGATCCTCAGAAACAGTACGTATACCTTGATTATATTCTCTA
-121     T73R     TCTCTCTCTTTTCGTAGAAGGAACATCATCATTTAGAAGATCCTCAGAAACAGTACGTATACCTTGATTATATTCTCTA
31       T10F     TCTCTCTCTTTTCGTAGAAGGAACATCATCATTTAGAAGATCCTCAGAAACAGTACGTATACCTTGATTATATTCTCTA
101      T49F                                                                  GTATACCTTGATTATATTCTCTA

CONSENSUS         TCTCTCTCTTTTCGTAGAAGGAACATCATCATTTAGAAGATCCTCAGAAACAGTACGTATACCTTGATTATATTCTCTA


                      3570       3580       3590       3600       3610       3620       3630
-44      T11R     TACTGCCATGAAAGTAAATTCATAC
34       T110_F   TACTGCCATGAAAGTAAATTCATACTACTTGCTTCTTCTTGAACACTTCTATTCCCAACATAAGATAATAAAAGAACAA
120      T73F     TACTGCCATGAAAGTAAATTCATACTACTTGCTTCTTCTTGAACACTTCTATTCCCAACATAAGATAATAAAAGAACAA
-121     T73R     TACTGCCATGAAAGTAAATTCATACTACTTGCTTCTTCTTGAACACTTCTATTCCCAACATAAGATAATAAAAGAACAA
31       T10F     TACTGCCATGAAAGTAAATTCATACTACTTGCTTCTTCTTGAACACTTCTATTCCCAACATAAGATAATAAAAGAACAA
101      T49F     TACTGCCATGAAAGTAAATTCATACTACTTGCTTCTTCTTGAACACTTCTATTCCCAACATAAGATAATAAAAGAACAA
-27      T108R           ATGAAAGTAAATTCATACTACTTGCTTCTTCTTGAACACTTCTATTCCCAACATAAGATAATAAAAGAACAA

CONSENSUS         TACTGCCATGAAAGTAAATTCATACTACTTGCTTCTTCTTGAACACTTCTATTCCCAACATAAGATAATAAAAGAACAA


                    3640     3650       3660       3670       3680       3690       3700       3710
34       T110_F   TAGGAATAACAATAGTAGGAGCAAATACTACAGTAAAACAACATCTTAATATTTTCTCACGTATAGTTAATAAAGACCA
120      T73F     TAGGAATAACAATAGTAGGAGCAAATACTACAGTAAAACAACATCTTAATATTTTCT
-121     T73R     TAGGAATAACAATAGTAGGAGCAAATACTACAGTAAAACAACATCTTAATATTTTCT
31       T10F     TAGGAATAACAATAGTAGGAGCAAATACTACAGTAAAACAACATCTTAATATTTTCTCACGTATAGTTAATAAAGACCA
101      T49F     TAGGAATAACAATAGTAGGAGCAAATACTACAGTAAAACAACATCTTAATATTTTCTCACGTATAGTTAATAAAGACCA
-27      T108R    TAGGAATAACAATAGTAGGAGCAAATACTACAGTAAAACAACATCTTAATATTTTCTCACGTATAGTTAATAAAGACCA

CONSENSUS         TAGGAATAACAATAGTAGGAGCAAATACTACAGTAAAACAACATCTTAATATTTTCTCACGTATAGTTAATAAAGACCA


                    3710     3720       3730       3740       3750       3760       3770       3780
34       T110_F   CCTTTCAGCATATAACACAGAATCATATATTTCATGACTTAATGGTTTACTACTGACATCATATTTTTCAACTCGAAAA
31       T10F     CCTTTCAGCATATAACACAGAATCATATATTTCATGACTTAATGGTTTACTACTGACATCATATTTTTCAACTCGAAAA
101      T49F     CCTTTCAGCATATAACACAGAATCATATATTTCATGACTTAATGGTTTACTACTGACATCATATTTTTCAACGCGAAAA
-27      T108R    CCTTTCAGCATATAACACAGAATCATATATTTCATGACTTAATGGTTTACTACTGACATCATATTTTTCAACTCGAAAA
-58      T137R                                                                  TCATATTTTTCAACTCGAAAA
CONSENSUS         CCTTTCAGCATATAACACAGAATCATATATTTCATGACTTAATGGTTTACTACTGACATCATATTTTTCAACTCGAAAA


                    3790       3800       3810       3820       3830       3840       3850       3860
34       T110_F   ACTCGATCAGTACCAGCTTGTGTTTCTAACTCATCAACAAAATCTCTTTGACTTTTTAAAATTATATCATATATATCAC
31       T10F     ACTCGATCAGTACCAGCTTGTGTTTCTAACTCATCAACAAAATCTCTTTGACTTTTTAAAATTATATCATATATATCAC
101      T49F     ACTCGATCAGTACCAGCTTGTGTTTCTAACTCATCAACAAAATCTCTTTGACTTTTTAAAATTATATCATATATATCAC
-27      T108R    ACTCGATCAGTACCAGCTTGTGTTTCTAACTCATCAACAAAATCTCTTTGACTTTTTAAAATTATATCATATATATCAC
-58      T137R    ACTCGATCAGTACCAGCTTGTGTTTCTAACTCATCAACAAAATCTCTTTGACTTTTTAAAATTATATCATATATATCAC

CONSENSUS         ACTCGATCAGTACCAGCTTGTGTTTCTAACTCATCAACAAAATCTCTTTGACTTTTTAAAATTATATCATATATATCAC


                     3870       3880       3890       3900       3910       3920       3930       3940
34       T110_F   TTTCAGGAGTGTGCTCTTTAGCACCATCACATGTAGAAATAATAAAAACAGTTTTAGGATTAGAATAAGCTTTACCTTC
31       T10F     TTTCAGGAGTGTGCTCTTTAGCACCATCACATGTAGAAATAATAAAAACAGTTTTAGGATTAGAATAAGCTTTACCTTC
101      T49F     TTTCAGGAGTGTGCTCTTTAGCACCATCACATGTAGAAATAATAAAAACAGTTTTAGGATTAGAATAAGCTTTACCTTC
-27      T108R    TTTCAGGAGTGTGCTCTTTAGCACCATCACATGTAGAAATAATAAAAACAGTTTTAGGATTAGAATAAGCTTTACCTTC
-58      T137R    TTTCAGGAGTGTGCTCTTTAGCACCATCACATGTAGAAATAATAAAAACAGTTTTAGGATTAGAATAAGCTTTACCTTC

CONSENSUS         TTTCAGGAGTGTGCTCTTTAGCACCATCACATGTAGAAATAATAAAAACAGTTTTAGGATTAGAATAAGCTTTACCTTC


                     3950       3960       3970       3980       3990       4000       4010       4020
34       T110_F   ACGATAAGTACATTCAGAAATACCAATATCATAACTGGAATGCACTACTGGCAAAAAC
31       T10F     ACGATAAGTACATTCAGAAATACCAATATCATAACTGGAATGCACTACTGGCAAAAACCCTAGCGAACGTATAAATGAT
101      T49F     ACGATAAGTACATTCAGAAATACCAATATCATAACTGGAATGCACTACTGGCAAAAACCCTAGCGAACGTATAAATGAT
-27      T108R    ACGATAAGTACATTCAGAAATACCAATATCATAACTGGAATGCACTACTGGCAAAAACCCTAGCGAACGTATAAATGAT
-58      T137R    ACGATAAGTACATTCAGAAATACCAATATCATAACTGGAATGCACTACTGGCAAAAACCCTAGCGAACGTATAAATGAT

CONSENSUS         ACGATAAGTACATTCAGAAATACCAATATCATAACTGGAATGCACTACTGGCAAAAACCCTAGCGAACGTATAAATGAT
```

```
                    4030       4040       4050       4060       4070       4080       4090       4100
31        T10F      ACAGTCTTAC
101       T49F      ACAGTCTTACGTATATCGTCTTTATTAGTAGAAGTATGTATTAATACCCTTTGTAAAAAAGATTCATCTACCGCATCAT
-27       T108R     ACAGTCTTACGTACATCGTCTTTATTAGTAGAAGTATGTATTAATACCCTTTGTAAAAAAGATTCATCTACCGCATCAT
-58       T137R     ACAGTCTTACGTATATCGTCTTTATTAGTAGAAGTATGTATTAATACCCTTTGTAAAAAAGATTCATCTACCGCATCAT
-20       T103R                                                ATACCCTTTGTAAAAAAGATTCATCTACCGCATCAT

CONSENSUS           ACAGTCTTACGTATATCGTCTTTATTAGTAGAAGTATGTATTAATACCCTTTGTAAAAAAGATTCATCTACCGCATCAT


                    4110       4120       4130       4140       4150       4160       4170       4180
101       T49F      CTTTATTTTGAATAACAAAAGCAGCAGGAGTAAGATTTCTTATTACCTTTCCTGAACAAAACATACCAATACTTGGATT
-27       T108R     CTTTATTTTGAATAACAAAAGCAGCAGGAGTAAGATTTCTTATTACCTTTCCTGAACAAAACATACCAA
-58       T137R     CTTTATTTTGAATAACAAAAGCAGCAGGAGTAAGATTTCTTATTACCTTTCCTGAACAAAACATACCAATACTTGGATT
-20       T103R     CTTTATTTTGAATAACAAAAGCAGCAGGAGTAAGATTTCTTATTACCTTTCCTGAACAAAACATACCAATACTTGGATT

CONSENSUS           CTTTATTTTGAATAACAAAAGCAGCAGGAGTAAGATTTCTTATTACCTTTCCTGAACAAAACATACCAATACTTGGATT


                    4190       4200       4210       4220       4230       4240       4250       4260
101       T49F      ATAAAATAT
-58       T137R     ATAAAATATGGTATATCTTAAATTGTCAAGCCATTCCTTATAAAACTTATACTTAGCAGGATCATACATTAGTTTTTAA
-20       T103R     ATAAAATATGGTATATCTTAAATTGTCAAGCCATTCCTTATAAAACTTATACTTAGCAGGATCATACATTAGTTTTTAA

CONSENSUS           ATAAAATATGGTATATCTTAAATTGTCAAGCCATTCCTTATAAAACTTATACTTAGCAGGATCATACATTAGTTTTTAA


                          4270       4280       4290       4300       4310       4320       4330       4340
-58       T137R     CTTATTACCTATTTACAATCACAATACTACATAATTACTATTTTACATTAATTTTTTAAATATTATAAGTAAAAGTTTT
-20       T103R     CTTATTACCTATTTACAATCACAATACT

CONSENSUS           CTTATTACCTATTTACAATCACAATACTACATAATTACTATTTTACATTAATTTTTTAAATATTATAAGTAAAAGTTTT


                          4350       4360       4370       4380       4390       4400       4410
-58       T137R     ACTACAATAAGGGCAAACTATTTCTTGACCATTTCTGATAGTCAAATATATTTTCGGATGTTCAGTATAATCAACATCA
93        T43F                                           TGATAGTCAAATATATtTTTCGGATGTTCAGTATAATCAACATCA
-94       T43R                                           TGATAGTCAAATATATtTTTCGGATGTTCAGTATAATCAACATCA

CONSENSUS           ACTACAATAAGGGCAAACTATTTCTTGACCATTTCTGATAGTCAAATATATTTTCGGATGTTCAGTATAATCAACATCA


                    4420     4430       4440       4450       4460       4470       4480       4490
-58       T137R     CTTCC
93        T43F      CTTCCTTCCCCATTACAAGATACTATTTGCTCTTCAAGTACATCTATTTTTTTTATCATGATTAGACATAAATATAATTT
-94       T43R      CTTCCTTCCCCATTACAAGATACTATTTGCTCTTCAAGTACATCTATTTTTTTTATCATGATTAGACATAAATATAATTT

CONSENSUS           CTTCCTTCCCCATTACAAGATACTATTTGCTCTTCAAGTACATCTATTTTTTTTATCATGATTAGACATAAATATAATTT


                    4500       4510       4520       4530       4540       4550       4560       4570
93        T43F      TTAATATTAATAAAACACAACTCTTATAACATCATAATAAATAACAAGCTACCTTTCTTACCATTCAAATATATATTAT
-94       T43R      TTAATATTAATAAAACACAACTCTTATAACATCATAATAAATAACAAGCTACCTTTCTTACCATTCAAATATATATTAT
39        T113_F                                         AATAAATAACAAGCTACCTTTCCTACCATTCAAATATATATTAT
-118      T72F                                                                   ATTCAAATACATATTAT
-8        153R                                                                    TTCAAATATATATTAT

CONSENSUS           TTAATATTAATAAAACACAACTCTTATAACATCATAATAAATAACAAGCTACCTTTCTTACCATTCAAATATATATTAT


                    4580       4590       4600       4610       4620       4630       4640       4650
93        T43F      TATACACTTTGTATAATCAAATTATTATATTACTAGCAAATACATGTTTACTACATTAAACTTATTTACTACTATCAAC
-94       T43R      TATACACTTTGTATAATCAAATTATTATATTACTAGCAAATACATGTTTACTACATTAAACTTATTTACTACTATCAAC
39        T113_F    TATACACTTTGTATAATCAAATTATTATATTACTAGCAAATACATGTTTACTACATTAAACTTATTTACTACTATCAAC
-118      T72F      TATACACTTTGTATAATCAAATTATTATATTAGTAGCAAATACATGTTTACTACATTAAACTTATTTACTACTATCAAC
-8        153R      TATACACTTTGTATAATCAAATTATTATATTACTAGCAAATACATGTTTACTACATTAAACTTATTTACTACTATCAAC

CONSENSUS           TATACACTTTGTATAATCAAATTATTATATTACTAGCAAATACATGTTTACTACATTAAACTTATTTACTACTATCAAC
```

```
                4660      4670      4680      4690      4700      4710      4720      4730
93      T43F    ATAGCTTTATGAATAATTCTATAGACAAAAACTCCACTTTTTCTTTCCTTGATTAAAGAATATACCTTAGATAAGTCAT
-94     T43R    ATAGCTTTATGAATAATTCTATAGACAAAAACTCCACTTTTTCTTTCCTTGATTAAAGAATATACCTTAGATAAGTCAT
39      T113_F  ATAGCTTTATGAATAATTCTATAGACAAAAACTCCACTTTTTCTTTCCTTGATTAAAGAATATACCTTAGATAAGTCAT
-118T   72F     ATAGCTTTATGAATAATTCTATAGACAAAAACTCCACTTTTTCTTTCCTTGATTAAAGAATATACCTTAGATAAGTCAT
-8      153R    ATAGCTTTATGAATAATTCTATAGACAAAAACTCCACTTTTTCTTTCCTTGATTAAAGAATATACCTTAGATAAGTCAT

CONSENSUS       ATAGCTTTATGAATAATTCTATAGACAAAAACTCCACTTTTTCTTTCCTTGATTAAAGAATATACCTTAGATAAGTCAT


                4740      4750      4760      4770      4780      4790      4800      4810
93      T43F    CATAAAAGTTTCATTATAATATTAGATAGATTATCCTATGATAAAAATTTTACATCAGAATTAAATAATTATTAACACT
-94     T43R    CATAAAAGTTTCATTATAATATTAGATAGATTATCCTATGATAAAAATTTTACATCAGAATTAAATAATTATTAACACT
39      T113_F  CATAAAAGTTTCATTATAATATTAGATAGATTATCCTATGATAAAAATTTTACATCAGAATTAAATAATTATTAACACT
-118    T72F    CATAAAAGTTTCATTATAATATTAGATAGATTATCCTATGATAAAAATTTTACATCAGAATTAAATAATTATTAACACT
-8      153R    CATAAAAGTTTCATTATAATATTAGATAGATTATCCTATGATAAAAATTTTACATCAGAATTAAATAATTATTAACCCT
-26     T107R                                                   TTACATCAGAATTAAATAATTATTAACACT
104     T50F                                                    TTACATCAGAATTAAATAATTATTAACACT
-105    T50R                                                    TTACATCAGAATTAAATAATTATTAACACT

CONSENSUS       CATAAAAGTTTCATTATAATATTAGATAGATTATCCTATGATAAAAATTTTACATCAGAATTAAATAATTATTAACACT


                4820      4830      4840      4850      4860      4870      4880      4890
93      T43F    ATTATACCCCAAATTACAATAAAATTCCTCAACAATTTTCTGATAATTTAAAACTAGTACTACACATAACTGCTATCAA
-94     T43R    ATTATACCCCAAATTACAATAAAATTCCTCAACAATTTTCTGATAATTTAAAACTAGTACTACACATAACTGCTATCAA
39      T113_F  ATTATACCCCAAATTACAATAAAATTCCTCAACAATTTTCTGATAATTTAAAACTAGTACTACACATAACTGCTATCAA
-118    T72F    ATTATACCCCAAATTACAATAAAATTCCTCAACAATTTTCTGATAATTTAAAACTAGTACTACACATAACTGCTATCAA
-8      153R    ATTATAC
-26     T107R   ATTATACCCCAAATTACAATAAAATTCCTCAACAATTTTCTGATAATTTAAAACTAGTACTACACATAACTGCTATCAA
104     T50F    ATTATACCCCAAATTACAATAAAATTCCTCAACAATTTTCTGATAATTTAAAACTAGTACTACACATAACTGCTATCAA
-105    T50R    ATTATACCCCAAATTACAATAAAATTCCTCAACAATTTTCTGATAATTTAAAACTAGTACTACACATAACTGCTATCAA
-33     T110R                                   CAATTTTCTGATAATTTAAAACTAGTACTACACATAACTGCTATCAA
-56     T132R                                                           TACACATAACTGCTATCAA

CONSENSUS       ATTATACCCCAAATTACAATAAAATTCCTCAACAATTTTCTGATAATTTAAAACTAGTACTACACATAACTGCTATCAA


                4890      4900      4910      4920      4930      4940      4950      4960
93      T43F    GACCTACAGATAAAAATTAAACCTTTTTCTTTCTTACTATCATTTTGGCTATAGAAAA
-94T    43R     GACCTACAGATAAAAATTAAACCTTTTTCTTTCTTACTATCATTTTGTCTATAGAATA
39      T113_F  GACCTACAGATAAAAATTAAACCTTTTTCTTTCTTACTATCATTTTGTCTATAGAATAGTACTTTCCCATTTGTTAGCA
-118    T72F    GACCTACAGATAAAA
-26     T107R   GACCTACAGATAAAAATTAAACCTTTTTCTTTCTTACTATCATTTTGTCTATAGAATAGTACTTTCCCATTTGTTAGCA
104     T50F    GACCTACAGATAAAAATTAAACCTTTTTCTTTCTTACTATCATTTTGTCTATAGAATAGTACTTTCCCATTTGTTAGCA
-105    T50R    GACCTACAGATAAAAATTAAACCTTTTTCTTTCTTACTATCATTTTGTCTATAGAATAGTACTTTCCCATTTGTTAGCA
-33     T110R   GACCTACAGATAAAAATTAAACCTTTTTCTTTCTTACTATCATTTTGTCTATAGAATAGTACTTTCCCATTTGTTAGCA
-56     T132R   GACCTACAGATAAAAATTAAACCTTTTTCTTTCTTACTATCATTTTGTCTATAGAATAGTACTTTCCCATTTGTTAGCA

CONSENSUS       GACCTACAGATAAAAATTAAACCTTTTTCTTTCTTACTATCATTTTGTCTATAGAATAGTACTTTCCCATTTGTTAGCA


                4970      4980      4990      4500      5010      5020      5030      5040
39      T113_F  TTATCTTTAACATATTGCTGAACATATACTAATTTACCTGTATACTCATAAATAAATTTCACTATATCATTAATATTAG
-26     T107R   TTATCTTTAACATATTGCTGAACATATACTAATTTACCTGTATACTCATAAATAAATTTCACTATATCATTAATATTAG
104     T50F    TTATCTTTAACATATTGCTGAACATATACTAATTTACCTGTATACTCATAAATAAATTTCACTATATCATTAATATTAG
-105    T50R    TTATCTTTAACATATTGCTGAACATATACTAATTTACCTGTATACTCATAAATAAATTTCACTATATCATTAATATTAG
-33     T110R   TTATCTTTAACATATTGCTGAACATATACTAATTTACCTGTATACTCATAAATAAATTTCACTATATCATTAATATTAG
-56     T132R   TTATCTTTAACATATTGCTGAACATATACTAATTTACCTGTATACTCATAAATAAATTTCACTATATCATTAATATTAG
64      T150_F                                  ACCTGTATACTCATAAATAAATTTCACTATATCATTAATATTAG

CONSENSUS       TTATCTTTAACATATTGCTGAACATATACTAATTTACCTGTATACTCATAAATAAATTTCACTATATCATTAATATTAG


                5050      5060      5070      5080      5090      5100      5110      5120
39      T113_F  TATCAGTTTTATTGTTAACACTTAGCACAATCTTCACATCGATATCACCTAACACTTTAAGTACCGTCAATGTATGACT
-26     T107R   TATCAGTTTTATTGTTAACACTTAGCACAATCTTCACATCGATATCACCTAACACTTTAAGTACCGTCAATGTATGACT
104     T50F    TATCAGTTTTATTGTTAACACTTAGCACAATCTTCACATCGATATCACCTAACACTTTAAGTACCGTCAATGTATGACT
-105    T50R    TATCAGTTTTATTGTTAACACTTAGCACAATCTTCACATCGATATCACCTAACACTTTAAGTACCGTCAATGTATGACT
-33     T110R   TATCAGTTTTATTGTTAACACTTAGCACAATCTTCACATCGATATCACCTAACACTTTAAGTACCGTCAATGTATGACT
-56     T132R   TATCAGTTTTATTGTTAACACTTAGCACAATCTTCACATCGATATCACCTAACACTTTAAGTACCGTCAATGTATGACT
64      T150_F  TATCAGTTTTATTGTTAACACTTAGCACAATCTTCACATCGATATCACCTAACACTTTAAGTACCGTCAATGTATGACT

CONSENSUS       TATCAGTTTTATTGTTAACACTTAGCACAATCTTCACATCGATATCACCTAACACTTTAAGTACCGTCAATGTATGACT
```

```
                        5130       5140       5150       5160       5170       5180       5190
39      T113_F  AATGCTCCCTAAGTAAGAACCAATAACTAAAATTACTTTTATTTTTAAATCTTGTATTAATTCCAAGCAAGTCTTCTGA
-26     T107R   AATGCTCCCTAAGTAAGAACCAATAACTAAAATTACTTTTATTTTTAAATCTTGTATTAATTCCAAGCAAGTCTTCTGA
104     T50F    AATGCTCCCTAAGTAAGAACCAATAACTAAAATTACTTTTATTTTTAAATCTTGTATTAATTCCAAGCAAGTCTTCTGA
-105    T50R    AATGCTCCCTAAGTAAGAACCAATAACTAAAATTACTTTTATTTTTAAATCTTGTATTAATTCCAAGCAAGTCTTCTGA
-33     T110R   AATGCTCCCTAAGTAAGAACCAATAACTAAAATTACTTTTATTTTTAAATCTTGTATTAATTCCAAGCAAGTCTTCTGA
-56     T132R   AATGCTCCCTAAGTAAGAACCAA
64      T150_F  AATGCTCCCTAAGTAAGAACCAATAACTAAAATTACTTTTATTTTTAAATCTTGTATTAATTCCAAGCAAGTCTTCTGA

CONSENSUS       AATGCTCCCTAAGTAAGAACCAATAACTAAAATTACTTTTATTTTTAAATCTTGTATTAATTCCAAGCAAGTCTTCTGA


                5200       5210       5220       5230       5240       5250       5260       5270
39      T113_F  TCAGTAATAGGAGACATAACACCACCAACACCTTCTATTAATAGATAATCTATATCTTGATTAATATGTGTATAACAGA
-26     T107R   TCAGTAATAGGAGACATAACACCACCAACACCTTCTATTAATAGATAATCTATATCTTGATTAATATGTGTATAACAGA
104     T50F    TCAGTAATAGGAGACATAACACCACCAACACCTTCTATTAATAGATAATCTATATCTTGATTAATATGTGTATAACAGA
-105    T50R    TCAGTAATAGGAGACATAACACCACCAACACCTTCTATTAATAGATAATCTATATCTTGATTAATATGTGTATAACAGA
-33     T110R   TCAGTAATAGGAGACATAACACCACCAACACCTTCTATTAATAGATAATCTATATCTTGATTAATATGTGTATAACAGA
64      T150_F  TCAGTAATAGGAGACATAACACCACCAACACCTTCTATTAATAGATAATCTATATCTTGATTAATATGTGTATAACAGA
38      T113R                                                   AGATAATCTATATCTTGATTAATATGTGTATAACAGA

CONSENSUS       TCAGTAATAGGAGACATAACACCACCAACACCTTCTATTAATAGATAATCTATATCTTGATTAATATGTGTATAACAGA


                5280       5290       5300       5310       5320       5330       5340       5350
39      T113_F  ATTTAAAAATTTCATTATAATCTAATTTAATATTT
-26     T107R   ATTTAAA
104     T50F    ATTTAAA
-105    T50R    ATTTAAA
-33     T110R   ATTTAAAAATTTCATTATAATCTAATTTAATATTTTCCAGCCTAGCAGCTATATTAGGAGCATGAGGATAAGATAGCCT
64      T150_F  ATTTAAAAATTTCATTATAATCTAATTTAATATTTTCCAGCCTAGCAGCTATATTAGGAGCATGAGGATAAGATAGCCT
-38     T113R   ATTTAAAAATTTCATTATAATCTAATTTAATATTTTCCAGCCTAGCAGCTATATTAGGAGCATGAGGATAAGATAGCCT
-5      150R                                                                    TGAGGATAAGATAGCCT

CONSENSUS       ATTTAAAAATTTCATTATAATCTAATTTAATATTTTCCAGCCTAGCAGCTATATTAGGAGCATGAGGATAAGATAGCCT


                 5360       5370       5380       5390       5400       5410       5420       5430
-33     T110R   CCAAGGAGAAATTCTATTGATGTTATGAGTATTGCAATCCACATCTAAACTACATAATATTTTGCTTGTATCATTACTC
64      T150_F  CCAAGGAGAAATTCTATTGATGTTATGAGTATTGCAATCCACATCTAAACTACATAATATTTTGCTTGTATCATTACTC
-38     T113R   CCAAGGAGAAATTCTATTGATGTTATGAGTATTGCAATCCACATCTAAACTACATAATATTTTGCTTGTATCATTACTC
-5      150R    CCAAGGAGAAATTCTATTGATGTTATGAGTATTGCAATCCACATCTAAACTACATAATATTTTGCTTGTATCATTACTC

CONSENSUS       CCAAGGAGAAATTCTATTGATGTTATGAGTATTGCAATCCACATCTAAACTACATAATATTTTGCTTGTATCATTACTC


                  5440       5450       5460       5470       5480       5490       5500       5510
-33     T110R   ATAATTTCCACATCATTCCACCCACTTATTATA
64      T150_F  ATAACTTCCACATCATTCCACCCACTTATTATAGGTTTTATTGCATGTACTGTTTTACAATTTTTTCGTAAATGCCAGC
-38     T113R   ATAATTTCCACATCATTCCACCCACTTATTATAGGTTTTATTGCATGTACTGTTTTACAATTTTTTCGTAAATGCCAGC
-5      150R    ATAACTTCCACATCATTCCACCCACTTATTATAGGTTTTATTGCATGTACTGTTTTACAATTTTTTCGTAAATGCCAGC
-10     49R                                                                                 GC

CONSENSUS       ATAACTTCCACATCATTCCACCCACTTATTATAGGTTTTATTGCATGTACTGTTTTACAATTTTTTCGTAAATGCCAGC


                    5520       5530       5540       5550       5560       5570       5580       5590
64      T150_F  ACAGTACTGTAGTAATAAAAGTTTTACCTATACCAGTACCACATGATGTAATAAAATAAGCTGACATTTAATACAGTGA
-38     T113R   ACAGTACTGTAGTAATAAAAGTTTTACCTATACCAGTACCACATGATGTAATAAAATAAGCTGACATTTAATACAGTGA
-5      150R    ACAGTACTGTAGTAATAAAAGTTTTACCTATACCAGTACCACATGATGTAATAAAATAAGCTGACATTTAATACAGTGA
-10     49R     ACAGTACTGTAGTAATAAAAGTTTTACCTATACCAGTACCACATGATGTAATAAAATAAGCTGACATTTAATACAGTGA

CONSENSUS       ACAGTACTGTAGTAATAAAAGTTTTACCTATACCAGTACCACATGATGTAATAAAATAAGCTGACATTTAATACAGTGA


                    5600       5610       5620       5630       5640       5650       5660       5670
64      T150_F  ACCTTATAATTTACTTAATACTAACTGATAAACACTGAACTCACAACAAAACACTAAGGAAAATTGAGCTATGA
-38     T113R   ACCTTATAATTTACTTAATACTAACTGATAAACACTGAACTCACAACAAAACACTAAGGAAAATTGAGCTATGACAGAT
-5      150R    ACCTTATAATTTACTTAATACTAACTGATAAACACTGAACTCACAACAAAACACTAAGGAAAATTGAGCTATGACAGAT
-10     49R     ACCTTATAATTTACTTAATACTAACTGATAAACACTGAACTCACAACAAAACACTAAGGAAAATTGAGCTATGACAGAT

CONSENSUS       ACCTTATAATTTACTTAATACTAACTGATAAACACTGAACTCACAACAAAACACTAAGGAAAATTGAGCTATGACAGAT
```

```
                      5680       5690       5700       5710       5720       5730       5740       5750
-38      T113R   AACCAAGGATTTTTCAGCAATATCAAAAAAAGTCTATTTAAAACTTCATCAAAATTAAGTGATGGAATTAAAAAAATTT
-5       150R    AACCAAGGATTTTTCAGCAATATCAAAAAAAGTCTATTTAAAACTTCATCAAAATTAAGTGATGGAATTAAAAAAATTT
-10      49R     AACCAAGGATTTTTCAGCAATATCAAAAAAAGTCTATTTAAAACTTCATCAAAATTAAGTGATGGAATTAAAAAAATTT
60       T139R                                                                          AAAAATTT

CONSENSUS        AACCAAGGATTTTTCAGCAATATCAAAAAAAGTCTATTTAAAACTTCATCAAAATTAAGTGATGGAATTAAAAAAATTT


                            5760       5770       5780       5790       5800       5810       5820       5830
-38      T113R   TTTCCAATAGTAAAAAAATCAATCAAGAAACTTTAGAAGAATTAAAGGAATTACTGATTACAGCAGATATTGGATATGA
-5       150R    TTTCCAATAGTAAAAAAATCAATCAAGAAACCTTAGAAGAATTAAAGGAATTACTGATTACAGCAGATATTGGATATGA
-10      49R     TTTCCAATAGTAAAAAAATCAATCAAGAAACTTTAGAAGAATTAAAGG
60       T139R   TTTCCAATAGTAAAAAAATCAATCAAGAGACTTTAGAAGAATTAAAGGAATTACTGATTACAGCAGATATTGGATATGA

CONSENSUS        TTTCCAATAGTAAAAAAATCAATCAAGAAACTTTAGAAGAATTAAAGGAATTACTGATTACAGCAGATATTGGATATGA


                          5840       5850       5860       5870       5880       5890       5900
-38      113R    AAATGCATCATTATTAATAAAAAA
-5       150R    AAATGCATCATTATTAATAAAAAAAATTGCAGAAGCAAAATTTGATGAAGTAACTGATCATACAATAAAGCAAAAACTA
60       T139R   AAATGCATCATTATTAATAAAAAAAATTGCAGAAGCAAAATTTGATGAAGTAACTGATCATACAATAAAGCAAAAACTA

CONSENSUS        AAATGCATCATTATTAATAAAAAAAATTGCAGAAGCAAAATTTGATGAAGTAACTGATCATACAATAAAGCAAAAACTA


                  5910    5920       5930       5940       5950       5960       5970       5980
-5       150R    GCAGAAGAGA
60       T139R   GCAGAAGAGATAGAAAATATTTTATTACAAGTTGAAAAACCTTTTTCTATAATAAAAAAACCACATGTAATTATGATAT

CONSENSUS        GCAGAAGAGATAGAAAATATTTTATTACAAGTTGAAAAACCTTTTTCTATAATAAAAAAACCACATGTAATTATGATAT


                  5990       6000       6010       6020       6030       6040       6050       6060
60       T139R   GTGGAACTAATGGAAACGGTAAAACAACAACAGTAGGTAAATTAGCATATAAATTTAAGAATAATGGAAAAAGTGTATT
45       T120R                                                          TAAATTTAAGAATAATGGAAAAAGTGTATT
-46      T120_F                                                         TAAATTTAAGAATAATGGAAAAAGTGTATT

CONSENSUS        GTGGAACTAATGGAAACGGTAAAACAACAACAGTAGGTAAATTAGCATATAAATTTAAGAATAATGGAAAAAGTGTATT


                  6070       6080       6090       6100       6110       6120       6130       6140
60       T139R   AGTTGCAGCGTGTGACACATTTCGTGCAGCAGCTACAGAACAACTTACAGTATGGTCACAAAAAGTAGATTTTCCCATT
45       T120R   AGTTGCAGCGTGTGACACATTTCGTGCAGCAGCTACAGAACAACTTACAGTATGGTCACAAAAAGTAGATTTTCCCATT
-46      T120_F  AGTTGCAGCGTGTGACACATTTCGTGCAGCAGCTACAGAACAACTTACAGTATGGTCACAAAAAGTAGATTTTCCCATT
24       T106R       GCAGCGTGTGACACATTTCGTGCAGCAGCTACAGAACAACTTACAGTATGGTCACAAAAAGTAGATTTTCCCATT

CONSENSUS        AGTTGCAGCGTGTGACACATTTCGTGCAGCAGCTACAGAACAACTTACAGTATGGTCACAAAAAGTAGATTTTCCCATT


                   6150       6160       6170       6180       6190       6200       6210       6220
60       T139R   GTAACAAGTACACAAGGATCAGACGCAGCAAGTGTAGCATATCAAGCCATGCAGCAAGCTTTAAAAAATGAAATAGATA
45       T120R   GTAACAAGTACACAAGGATCAGATGCAGCAAGTGTAGCATATCAAGCCATGCAGCAAGCTTTAAAAAATGAAATAGATA
-46      T120_F  GTAACAAGTACACAAGGATCAGATGCAGCAAGTGTAGCATATCAAGCCATGCAGCAAGCTTTAAAAAATGAAATAGATA
24       T106R   GTAACAAGTACACAAGGATCAGATGCAGCAAGTGTAGCACATCAAGCCATGCAGCAAGCTTTAAAAAATGAAATAGATA
-61      T139_F                                                                         AAATAGATA

CONSENSUS        GTAACAAGTACACAAGGATCAGATGCAGCAAGTGTAGCATATCAAGCCATGCAGCAAGCTTTAAAAAATGAAATAGATA


                     6230       6240       6250       6260       6270       6280       6290       6300
60       T139R   TTTTACTTGTTGATACTGCAGGAAGATTACACAACTATAAAAATCTAATGGAAGAATTAGCAAAAATTAAACGAATTAT
45       T120R   TTTTACTTGTTGATACTGCAGGAAGATTACACAACTATAAAAATCTAATGGAAGAATTAGCAAAAATTAAACGAATTAT
-46      T120_F  TTTTACTTGTTGATACTGCAGGAAGATTACACAACTATAAAAATCTAATGGAAGAATTAGCAAAAATTAAACGAATTAT
24       T106R   TTTTACTTGTTGATACTGCAGGAAGATTACACAACTATAAAAATCTAATGGAAGAATTAGCAAAAATTAAACGAATTAT
-61      T139_F  TTTTACTTGTTGATACTGCAGGAAGATTACACAACTATAAAAATCTAATGGAAGAATTAGCAAAAATTAAACGAATTAT

CONSENSUS        TTTTACTTGTTGATACTGCAGGAAGATTACACAACTATAAAAATCTAATGGAAGAATTAGCAAAAATTAAACGAATTAT
```

```
                    6310       6320       6330       6340       6350       6360       6370       6380
60       T139R   AGGTAAACACGATAATGAAGCACCACATGATGTAGTATTAATACTAGACGCAACAACTGGGCAAAACG
45       T120R   AGGTAAACACGATAATGAAGCACCACATGATGTAGTATTAATACTAGACGCAACAACTGGGCAAAACGCTCTTAATCAA
-46      T120_F  AGGTAAACACGATAATGAAGCACCACATGATGTAGTATTAATACTAGACGCAACAACTGGGCAAAACGCTCTTAATCAA
24       T106R   AGGTAAACACGATAATGAAGCACCACATGATGTAGTATTAATACTAGACGCAACAACTGGGCAAAACGCTCTTAATCAA
-61      T139_F  AGGTANACACGATAATGAAGCACCACATGATGTAGTATTAATACTAGACGCAACAACTGGGCAAAACGCTCTTAATCAA
-126     T84F                              ACATGATGTAGTATTAATACTAGACGCAACAGGTGGGCAAAACGCTTTTAATCAA

CONSENSUS        AGGTAAACACGATAATGAAGCACCACATGATGTAGTATTAATACTAGACGCAACAACTGGGCAAAACGCTCTTAATCAA


                      6390       6400       6410       6420       6430       6440       6450       6460
45       T120R   GTCGAAGTTTTTTTACAATTTGTAAACATTAGTGGACTTGTTATCACAAAACTGGATGGTACAGCAAAAGGTGGAGTAG
-46      T120_F  GTCGAAGTTTTTTTACAATTTGTAAACATTAGTGGACTTGTTATCACAAAACTGGATGGTACAGCAAAAGGTGGAGTAG
24       T106R   GTCGAAGTTTTTTTACAATTTGTAAACATTAGTGGACTTGTTATCACAAAACTGGATGGTACAGCAAAAGGTGGAGTAG
-61      T139_F  GTCGAAGTTTTTTTACAATTTGTAAACATTAGTGGACTTGTTATCACAAAACTGGATGGTACAGCAAAAGGTGGAGTAG
-126     T84F    GTCGAAGTTTTTTTACAATTTGTAAACATTAGTGGACTTGTTATCACAAAACTGGATGGTACAGCAAAAGGTGGAGTAG
84       T32R       GAAGTTTTTTTACAATTTGTAAACATTAGTGGACTTGTTATCACAAAACTGGATGGTACAGCAAAAGGTGGAGTAG

CONSENSUS        GTCGAAGTTTTTTTACAATTTGTAAACATTAGTGGACTTGTTATCACAAAACTGGATGGTACAGCAAAAGGTGGAGTAG


                      6470       6480       6490       6500       6510       6520       6530       6540
45       T120R   TAATAAGAATAGCACAAAAATATAAATTG
-46      T120_F  TAATAAGAATAGCACAAAAATATAAATTG
24       T106R   TAATAAGAATAGCACAAAAATATAAATTGAATATTCATGCAATAGGAATAGGAGAACAGGTAGAAGATCTTAAAGATTT
-61      T139_F  TAATAAGAATAGCACAAAAATATAAATTGAATATTCATGCGATAGGAATAGGAGAACAGGTAGAAGATCTTAAAGATTT
-126     T84F    TAATAAGAATAGCCCAAAAATATAAATTGAATATTCATGCAATAGGAATAGGAGAACAGGTAGAAGATCTTAAAGATTT
84       T32R    TAATAAGAATAGCACAAAAATATAAATTGAATATTCATGCAATAGGAATAGGAGAACAGGTAGAAGATCTTAAAGATTT
-85      T33F                          ATATAAATTGAATATTCATGCAATAGGAATAGGAGAACAGGTAGAAGATCTTAAAGATTT
86       T33R                          ATATAAATTGAATATTCATGCAATAGGAATAGGAGAACAGGTAGAAGATCTTAAAGATTT

CONSENSUS        TAATAAGAATAGCACAAAAATATAAATTGAATATTCATGCAATAGGAATAGGAGAACAGGTAGAAGATCTTAAAGATTT


                    6550       6560       6570       6580       6590       6600       6610       6620
24       T106R   CTCCGCTAAAGAGTTCACTGCTGGACTTTTAAATATGGATAACATATTGTAATTTAATAGCAACTTTGCTTGTAAAGCA
-61      T139_F  CTCCGCTAAAGAGTTCACTGCTGGACTTTTAAATATGGATAACATATTGTAATTTAATAGCAACTTTGCTTGTAAAGCA
-126     T84F    CTCCGCTAAAGAGTTCACTGCTGGACTTTTAAATATGGATAACATATTGTAATTTAATAGCAACTTTGCTTGTAAAGCA
84       T32R    CTCCGCTAAAGAGTTCACTGCTGGACTTTTAAATATGGATAACATATTGTAATTTAATAGCAACTTTGCTTGTAAAGCA
-85      T33F    CTCCGCTAAAGAGTTCACTGCTGGACTTTTAAATATGGATAACATATTGTAATTTAATAGCAACTTTGCTTGTAAAGCA
86       T33R    CTCCGCTAAAGAGTTCACTGCTGGACTTTTAAATATGGATAACATATTGTAATTTAATAGCAACTTTGCTTGTAAAGCA
54       T128_F                          CTTTTAAATATGGATAACATATTGTAATTTAATAGCAACTTTGCTTGTAAAGCA
68       T18R                            CTTTTAAATATGGATAACATATTGTAATTTAATAGCAACTTTGCTTGTAAAGCA

CONSENSUS        CTCCGCTAAAGAGTTCACTGCTGGACTTTTAAATATGGATAACATATTGTAATTTAATAGCAACTTTGCTTGTAAAGCA


                    6630       6640       6650       6710       6720       6730       6740     6750
24       T106R   GTATTCTTCTTGCATTTACTGTTTAATTTAATATTAACTCATTGCTATTAGACATTTATTTTTTCCTAATTCAATTGAT
-61      T139_F  GTATTCTTCTTGCATTTACTGTTTAGTTTAATATTAACTCATTGCTATTAGACATTTATTTTTTCCTAATTCAATTGAT
-126     T84F    GTATTCTTCTTGCATTTACTGTTTAGTTTAATATTAACTCATTGCTATTAGACATTTATTTTTTCCTAATTCAATTGAT
84       T32R    GTATTCTTCTTGCATTTACTGTTTAGTTTAATGTTAACTCATTGCTATTAGACATTTATTTTTTCCTAATTCAATTGAT
-85      T33F    GTATTCTTCTTGCATTTACTGTTTAGTTTAATATTAACTCATTGCTATTAGACATTTATTTTTTCCTAATTCAATTGAT
86       T33R    GTATTCTTCTTGCATTTACTGTTTAGTTTAATATTAACTCATTGCTATTAGACATTTATTTTTTCCTAATTCAATTGAT
54       T128_F  GTATTCTTCTTGCATTTACTGTTTAGTTTAATATTAACTCATTGCTATTAGACATTTATTTTTTCCTAATTCAATTGAT
68       T18R    GTATTCTTCTTGCATTTACTGTTTAGTTTAATATTAACTCATTGCTATTAGACATTTATTTTTTCCTAATTCAATTGAT

CONSENSUS        GTATTCTTCTTGCATTTACTGTTTAGTTTAATATTAACTCATTGCTATTAGACATTTATTTTTTCCTAATTCAATTGAT


                      6760       6770       6780       6790       6800       6810       6820
24       T106R   GTAATATAAAA
-61      T139_F  GTAATATAAAAATTTCTCGTTGCTACTAATTTTTAGTTCTCTTGTAACAAAAATTCTACTGAGTTGATACAATCATAAT
-126     T84F    GTAATATAAAAATTTCTCGTTGCTACTAATTTTTAGTTCTCTTGTAACAAAAATTCTACTGAGTTGAT
84       T32R    GTAATATAAAAATTTCTCGTTGCTACTAATTTTTAGTTCTCTTGTAACAAAAATTCTACTGAGTTGATACAATCATAAT
-85      T33F    GTAATATAAAAATTTCTCGTTGCTACTAATTTTTAGTTCTCTTGTAACAA
86       T33R    GTAATATAAAAATTTCTCGTTGCTACTAATTTTTAGTTCTCTTGTAACAA
54       T128_F  GTAATATAAAAATTTCTCGTTGCTACTAATTTTTAGTTCTCTTGTAACAAAAATTCTACTGAGTTGATACAATCATAAT
68       T18R    GTAATATAAAAATTTCTCGTTGCTACTAATTTTTAGTTCTCTTGTAACAAAAATTCTACTGAGTTGATACAATCATAAT
22       T104_F  GTAATATAAAAATTTCTCGTTGCTACTAATTTTTAGTTCTCTTGTAACAAAAATTCTACTGAGTTGATACAATCATAAT
51       T125_F                          TACTAATTTTTAGTTCTCTTGTAACAAAAATTCTACTGAGTTGATACAATCATAAT

CONSENSUS        GTAATATAAAAATTTCTCGTTGCTACTAATTTTTAGTTCTCTTGTAACAAAAATTCTACTGAGTTGATACAATCATAAT
```

```
                6830       6840       6850       6860       6870       6880       6890       6900
-61     T139_F  ATAACAATAAGTTTCTATTCATACAGAAAAAATTTTTTTCTGTACACTACTTAAAAATAAAAAAAAGTACTATACAAGC
84      T32R    ATAACAATAAGTTTCTATTCATACAGAAAAAATTTTTTTCTGTACACTACTTAAAAATAAAAAAA*GTACTATACAAGC
54      T128_F  ATAACAATAAGTTTCTATTCATACAGAAAAAATTTTTTTCTGTACACTACTTAAAAATAAAAAAAAGTACTATACAAGC
68      T18R    ATAACAATAAGTTTCTATTCATACAGAAAAAATTTTTTTCTGTACACTACTTAAAAATAAAAAAAAGTACTATACAAGC
22      T104_F  ATAACAATAAGTTTCTATTCATACAGAAAAAATTTTTTTCTGTACACTACTTAAAAATAAAAAAAAGTACTATACAAGC
51      T125_F  ATAACAATAAGTTTCTATTCATACAGAAAAAATTTTTTTCTGTACACTACTTAAAAATAAAAAAAAGTACTATACAAGC

CONSENSUS       ATAACAATAAGTTTCTATTCATACAGAAAAAATTTTTTTCTGTACACTACTTAAAAATAAAAAAAAGTACTATACAAGC


                 6910       6920       6930       6940       6950       6960       6970       6980
-61     T139_F  AATACTAAAAACACATCAGTTTATATCATAGTAAGTAAATAAGATAGAAACTTTGTAATATTTGCAATAAAATCAATAA
84      T32R    AATACTAAAAACACATCAGTTTATATCATAGTAAGGAAATAAGATAGAAACTTTGTAATATTTGCAATAAAATCAATAA
54      T128_F  AATACTAAAAACACATCAGTTTATATCATAGTAAGTAAATAAGATAGAAACTTTGTAATATTTGCAATAAAATCAATAA
68      T18R    AATACTAAAAACACATCAGTTTATATCATAGTAAGTAAATAAGATAGAAACTTTGTAATATTTGCAATAAAATCAATAA
22      T104_F  AATACTAAAAACACATCAGTTTATATCATAGTAAGTAAATAAGATAGAAACTTTGTAATATTTGCAATAAAATCAATAA
51      T125_F  AATACTAAAAACACATCAGTTTATATCATAGTAAGTAAATAAGATAGAAACTTTGTAATATTTGCAATAAAATCAATAA

CONSENSUS       AATACTAAAAACACATCAGTTTATATCATAGTAAGTAAATAAGATAGAAACTTTGTAATATTTGCAATAAAATCAATAA


                   6990       7000       7010       7020       7030       7040       7050       7060
-61     T139_F  TAT
84      T32R    TATATTATGTTAAACTAACATACAGGAAATATTACGTTTTCACTCTTTTAAAAAAGAAATTCATATTCATTTATACTAT
54      T128_F  TATATTATGTTAAACTAACATACAGGAAATATTACGTTTTCACTCTTTTAAAAAAGAAATTCATATTCATTTATACTAT
68      T18R    TATATTATGTTAAACTAACATACAGGAAATATTACGTTTTCACTCTTTTAAAAAAGAAATTCATATTCATTTATACTAT
22      T104_F  TATATTATGTTAAACTAACATACAGGAAATATTACGTTTTCACTCTTTTAAAAAAGAAATTCATATTCATTTATACTAT
51      T125_F  TATATTATGTTAAACTAACATACAGGAAATATTACGTTTTCACTCTTTTAAAAAAGAAATTCATATTCATTTATACTAT
99      T47R                    CATACAGGAAATATTACGTTTTCACTCTTTTAAAAAAGAAATTCATATTCATTTATACTAT
-21     T104R                                                                   ATTTATACTAT

CONSENSUS       TATATTATGTTAAACTAACATACAGGAAATATTACGTTTTCACTCTTTTAAAAAAGAAATTCATATTCATTTATACTAT


                     7070       7080       7090       7100       7110       7120       7130       7140
84      T32R    AAGTAGTTTAGAGTATTTACATGAAAGATCATG
54      T128_F  AAGTAGTTTAGAGTATTTACATGAAAGATCATGATATAATAGAGTATATTATCAAGTAATCTATTAAATCTGGTTGAAT
68      T18R    AAGTAGTTTAGAGTATTTACATGAAAGATCATGATATAATAGAGTATATTATCAAGTAATCTATTAAATCTGGTTGAAT
22      T104_F  AAGTAGTTTAGAGTATTTACATGAAAGATCATGATATAATAGAGTATATTATCAAGTAATCTATTAAATCTGGTTGAAT
51      T125_F  AAGTAGTTTAGAGTATTTACATGAAAGATCATGATATAATAGAGTATATTACCAAGTAATCTATTAAATCTGGTTGAAT
99      T47R    AAGTAGTTTAGAGTATTTACATGAAAGATCATGATATAATAGAGTATATTATCAAGTAATCTATTAAATCTGGTTGAAT
-21     T104R   AAGTAGTTTAGAGTATTTACATGAAAGATCATGATATAATAGAGTATATTATCAAGTAATCTATTAAATCTGGTTGAAT

CONSENSUS       AAGTAGTTTAGAGTATTTACATGAAAGATCATGATATAATAGAGTATATTATCAAGTAATCTATTAAATCTGGTTGAAT


                        7150       7160       7170       7180       7190       7200       7210       7220
54      T128_F  ATATAATAATATTTTCAGAATATTAGATACCTACACTATTAGTTTATCATTTAACATCCTAACAAAATTAATTTTAACA
68      T18R    ATATAATAATATTTT
22      T104_F  ATATAATAATATTTTCAGAATATTAGATACCTACACTATTAGTTTATCATTTAACATCCTAACAAAATTAATTTTAACA
51      T125_F  ATATAATAATATTTTCAGAATATTAGATACCTACACTATTAGTTTATCATTTAACATCCTAACAAAATTAATTTTAACA
99      T47R    ATATAATAATATTTTCAGAATATTAGATACCTACACTATTAGTTTATCATTTAACATCCTAACAAAATTAATTTTAACA
-21     T104R   ATATAATAATATTTTCAGAATATTAGATACCTACACTATTAGTTTATCATTTAACATCCTAACAAAATTAATTTTAACA
17      T101_F                          ATTAGATACCTACACTATTAGTTTATCATTTAACATCCTAACAAAATTAATTTTAACA

CONSENSUS       ATATAATAATATTTTCAGAATATTAGATACCTACACTATTAGTTTATCATTTAACATCCTAACAAAATTAATTTTAACA


                     7230       7240       7250       7260       7270       7280       7290       7300
54      T128_F  TTGAATCTATGATATTTTTACATATTAAGTCTTAATTTACACAACATGTAAAACACATTTTGTATAAATTAAGAAAGTA
22      T104_F  TTGAATCTATGATATTTTTACATATTAAGTCTTAATTTACACAACATGTAAAACACATTTTGTATAAATTAAGAAAGTA
51      T125_F  TTGAATCTATGATATTTTTACATATTAAGTCTTAATTTACACAACATGTAAAACACATTTTGTATAAATTAAGAAAGTA
99      T47R    TTGAATCTATGATATTTTTACATATTAAGTCTTAATTTACACAACATGTAAAACACATTTTGTATAAATTAAGAAAGTA
-21     T104R   TTGAATCTATGATATTTTTACATATTAAGTCTTAATTTACACAACATGTAAAACACATTTTGTATAAATTAAGAAAGTA
17      T101_F  TTGAATCTATGATATTTTTACATATTAAGTCTTAATTTACACAACATGTAAAACACATTTTGTATAAATTAAGAAAGTA

CONSENSUS       TTGAATCTATGATATTTTTACATATTAAGTCTTAATTTACACAACATGTAAAACACATTTTGTATAAATTAAGAAAGTA
```

```
                        7310       7320       7330       7340       7350       7360       7370       7380
54      T128_F  CATAAATTTAGGTAACTATATTAAAATCTCATTATAACAAA
22      T104_F  CATAAATTTAGGTAACTATATTAAAATCTCATTATAACAAACAATAACCTAGCGATTATTATATATACCTACACACAAA
51      T125_F  CATAAATTTAGGTAACTATATTAAAATCTCATTATAACAAACAATAACCTAGCGATTATTATATATACCTACACACAAA
99      T47R    CATAAATTTAGGTAACTATATTAAAATCTCATTATAACAAACAATAACCTAGCGATTATTATATATACCTACACACAAA
-21     T104R   CATAAATTTAGGTAACTATATTAAAATCTCATTATAACAAACAATAACCTAGCGATTATTATATATACCTACACACAAA
17      T101_F  CATAAATTTAGGTAACTATATTAAAATCTCATTATAACAAACAATAACCTAGCGATTATTATATATACCTACACACAAA

CONSENSUS       CATAAATTTAGGTAACTATATTAAAATCTCATTATAACAAACAATAACCTAGCGATTATTATATATACCTACACACAAA


                      7390       7400       7410       7420       7430       7440       7450
22      T104_F  ATAGGTAACACCAGTTTTTTAGTAGATATGGGAAATCTACTTTATAAATTCAGAATTTAAGTAAAACTGTTATATTTAA
51      T125_F  ATAGGTAACACCAGTTTTTTAGTAGATATGGGAAATCTACTTTATAAATTCAGAATTTAAGTAAAACTGTTATATTTAA
99      T47R    ATAGGTAACACCAGTTTTTTAGTAGATATGGGAAATCTACTTTATAAATTCAGAATTTAAGTAAAACTGTTATATTTAA
-21     T104R   ATAGGTAACACCAGTTTTTTAGTAGATATGGGAAATCTACTTTATAAATTCAGAATTTAAGTAAAACTGTTATATTTAA
17      T101_F  ATAGGTAACACCAGTTTTTTAGTAGATATGGGAAATCTACTTTATAAATTCAGAATTTAAGTAAAACTGTTATATTTAA

CONSENSUS       ATAGGTAACACCAGTTTTTTAGTAGATATGGGAAATCTACTTTATAAATTCAGAATTTAAGTAAAACTGTTATATTTAA


                    7460       7470       7480       7490       7500       7510       7520       7530
22      T104_F  GTAACATGTAGTAAACTAATAAACATAATCACATTAAAATACAGTTATAATATATCAT
51      T125_F  GTAACATGTAGTAAACTAATAAACATAATCACATTAAAATACAGTTATAATATATCATCTATAAATATTTGATACCAAT
99      T47R    GTAACATGTAGTAAACTAATAAACATAATCACATTAAAATACAGTTATAATATATCATCTATAAATATTTGATACCAAT
-21     T104R   GTAACATGTAGTAAACTAATAAACATAATCACATTAAAATACAGTTATAATATATCATCTATAAATATTTGATACCAAT
17      T101_F  GTAACATGTAGTAAACTAATAAACATAATCACATTAAAATACAGTTATAATATATCATCTATAAATATTTGATACCAAT

CONSENSUS       GTAACATGTAGTAAACTAATAAACATAATCACATTAAAATACAGTTATAATATATCATCTATAAATATTTGATACCAAT


                  7540       7550       7560       7570       7580       7590       7600       7610
51      T125_F  AAGT
99      T47R    AAGTATTGCTAACTATTATTTGAAAAATATCAGGCATTTTTGTAAATACAATAATCCATAAATTAAATCCAATTTAAAT
-21     T104R   AAGTATTGCTAACTATTATTTGAAAAATATCAGGCATTTTTGTAAATACAATAATCCATAAATTAAATCCAATTTAAAT
17      T101_F  AAGTATTGCTAACTATTATTTGAAAAATATCAGGCATTTTTGTAAATACAATAATCCATAAATTAAATCCAATTTAAAT
91      T41F                    TTTGAAAAATATCAGGCATTTTTGTAAATACAATAATCCATAAATTAAATCCAATTTAAAT

CONSENSUS       AAGTATTGCTAACTATTATTTGAAAAATATCAGGCATTTTTGTAAATACAATAATCCATAAATTAAATCCAATTTAAAT


                       7620       7630       7640       7650       7660       7670       7680       7690
99      T47R    AAACAACAAAC
-21     T104R   AAACAACAAACTTATACGAAATTAAAAATAATTATATTTGTAATCTATAAACACTCACATACAAATCAGTAACACAATC
17      T101_F  AAACAACAAACTTATACGAAATTAAAAATAATTATATTTGTAATCTATAAACACTCACATACAAATCAGTAACACAATC
91      T41F    AAACAACAAACTTATACGAAATTAAAAATAATTATATTTGTAATCTATAAACACTCACATACAAATCAGTAACACAATC
-100    T47_F                                                                              AATC

CONSENSUS       AAACAACAAACTTATACGAAATTAAAAATAATTATATTTGTAATCTATAAACACTCACATACAAATCAGTAACACAATC


                     7700       7710       7720       7730       7740       7750       7760
-21     T104R   CTATTTATCTTGTTTACATA
17      T101_F  CTATTTATCTTGTTTACATAAGATATAAAGCTTGTTTTGTAAGTTAAAAAAACTTATAAGTGAAGGAGTAGTTATATTA
91      T41F    CTATTTATCTTGTTTACATAAGATATAAAGCTTGTTTTGTAAGTTAAAAAAACTTATAAGTGAAGGAGTAGTTATATTA
-100    T47_F   CTATTTATCTTGTTTACATAAGATATAAAGCTTGTTTTGTAAGTTAAAAAAACTTATAAGTGAAGGAGTAGTTATATTA

CONSENSUS       CTATTTATCTTGTTTACATAAGATATAAAGCTTGTTTTGTAAGTTAAAAAAACTTATAAGTGAAGGAGTAGTTATATTA


                  7770     7780       7790       7800       7810       7820       7830       7840   7850
17      T101_F  ATCAAGTATTCCATTACTCATTTATTCATAATGACTCTAGAAAATCTCCTACATTACAAGCAAGATATATAGAAAAATA
91      T41F    ATCAAGTATTCCATTACTCATTTATTCATAATGACTCTAGAAAATCTCCTACATTACAAGCAAGATATATAGAAAAATA
-100    T47_F   ATCAAGTATTCCATTACTCATTTATTCATAATGACTCTAGAAAATCTCCTACATTACAAGCAAGATATATAGAAAAATA

CONSENSUS       ATCAAGTATTCCATTACTCATTTATTCATAATGACTCTAGAAAATCTCCTACATTACAAGCAAGATATATAGAAAAATA


                        7860       7870       7880       7890       7900       7910       7920
17      T101_F  TAAACACCCAGCTAAATTTTTTTCTTATTCACAATGAANACCTTTAACTTATTTCAAAATTAAAAACTATTAAAGTATA
91      T41F    TAAACACCCAGCTAAATTTTTTTCTTATTCACAATGAAAACCTTTAACTTATTTCAAAATTAAAAACTATTAAAGTATA
-100    T47_F   TAAACACCCAGCTAAATTTTTTTCTTATTCACAATGAAAACCTTTAACTTATTTCAAAATTAAAAACTATTAAAGTATA
-98     T47F              GCTAAATTTTTTTCTTATTCACAATGAAAACCTTTAACTTATTTCAAAATTAAAAACTATTAAAGTATA

CONSENSUS       TAAACACCCAGCTAAATTTTTTTCTTATTCACAATGAAAACCTTTAACTTATTTCAAAATTAAAAACTATTAAAGTATA
```

```
                7930       7940       7950       7960       7970       7980       7990       8000
17      T101_F  CTCCTACAATAATGTAGCACTAATCTCTTATCCATGAAAATTCTATCACCATTTTACAAACTTAAACACTTACACCAA
91      T41F    CTCCTACAATAATGTAGCACTAATCTCTTATCCATGAAAATTCTATCACCATTTTACAAACTTAAACACTTACACCAAT
-100    T47_F   CTCCTACAATAATGTAGCACTAATCTCTTATCCATGAAAATTCTATCACCATTTTACAAACTTAAACACTTACACCAAT
-98     T47F    CTCCTACAATAATGTAGCACTAATCTCTTATCCATGAAAATTCTATCACCATTTTACAAACTTAAACACTTACACCAAT

CONSENSUS       CTCCTACAATAATGTAGCACTAATCTCTTATCCATGAAAATTCTATCACCATTTTACAAACTTAAACACTTACACCAAT


                  8010       8020       8030       8040       8050       8060       8070
91      T41F    GCCACACAACACTAAATATTAATCAAATATTCTTCAATTAACTTATTCATAATGACTCTAGAAAATCTCCTACA
-100    T47_F   GCCACACAACACTAAATATTAATCAAATATTCTTCAATTAACTTATTCATAATGACTCTAGAAAATCTCCTACA
-98     T47F    GCCACACAACAGTAAATATTAATCAAATATTCTTCAATTAACTTATTCATAATGACTCTAGAAAATCTCCTACA

CONSENSUS       GCCACACAACACTAAATATTAATCAAATATTCTTCAATTAACTTATTCATAATGACTCTAGAAAATCTCCTACA
```

Primer WL2TP1_gapF: 5'- CAC TTA CAC CAA TGC CAC AC -3'

```
                8080    8090       8100       8110       8120       8130       8140       8150
91      T41F    TTACAAGCAAGATATATAGAAAAATATCAACACCCAGCTAAATTTTTTTTCTTATTCACAATGAAAACTTTTAACTTAT
-100    T47_F   TTACAAGCAAGATATATAGAAAAATATCAACACCCAGCTAAATTTTTTTTCTTATTCACAATGAAAACTTTTAACTTAT
-98     T47F    TTACAAGCAAGATATATAGAAAAATATCAACCCCCAGCTAAATTTTTTTTCTTATTCACAATGAAAACTTTTAACTTAT

CONSENSUS       TTACAAGCAAGATATATAGAAAAATATCAACACCCAGCTAAATTTTTTTTCTTATTCACAATGAAAACTTTTAACTTAT


                     8160       8170       8180       8190       8200       8210       8220       8230
91      T41F    TTCAAAATTAAAAACTATTAATTCTTATATAATATAAAAACTACTCTTTCTACTAATATCCTATACAAACACTTTAATA
-100    T47_F   TTCAAAATTAAAAACTATTAATTCTTATATAATATAAAAACTACTCTTTCTACTAATATCCTATACAAACACTTTAGTA
-98     T47F    TTCAAAATTAAAAACTATTAATTCTTATATAATATAAAAACTACTCTTTCTACTAATATCCTATACAAACACTTTAGTA

CONSENSUS       TTCAAAATTAAAAACTATTAATTCTTATATAATATAAAAACTACTCTTTCTACTAATATCCTATACAAACACTTTAGTA


                        8240       8250       8260       8270       8280       8290       8300       8310
91      T41F    TTTTATAGTACTACGTAAAAAATCCAGGTTGATCAACATCAGGCTACTTTATCTTTACTATACAACGTGTNCTAATGAT
-100    T47_F   TTTTATAGTACTACGTAAAAAATCCAGGTTGATCAACATCAGGCTACTTTATCTTTACTATACAACGTGTACTAATGAT
-98     T47F    TTTTATAGTACTACGTAAAAAATCCAGGTTGATCAACATCAGGCTACTTTATCTTTACTATACAACGTGTACTAATGAT

CONSENSUS       TTTTATAGTACTACGTAAAAAATCCAGGTTGATCAACATCAGGCTACTTTATCTTTACTATACAACGTGTACTAATGAT


                           8320       8330       8340       8350       8360       8370       8380
91      T41F    AAATAGCAA
-100    T47_F   AAATAGCAACAAACACAATGTTAGATTATATTTATCAGTAACATAGCACTTCTTATAAAAAATCTCTTAAAATACATAA
-98     T47F    AAATAGCAACAAACCCAATGTTAGATTATATTTATCAGTAACATAGCACTTCTTATAAAAAATCTCTTAAAATACATAA

CONSENSUS       AAATAGCAACAAACACAATGTTAGATTATATTTATCAGTAACATAGCACTTCTTATAAAAAATCTCTTAAAATACATAA


                8340
-100    T47_F   GCATACCATATC
-98     T47F    GCATACCATATC

CONSENSUS       GCATACCATATC
```

Addendum 2. DNA sequence of contig two. Column 1: Position where clone sequence begins. Column 2: Clone number and direction of primer used (F = forward, R = reverse)

Primer WL2TP1_gapR: 5'- TTA CCG CCA CCC TAA CAT ATA G -3'

Primer T7GEM11: 5'-CTA ATA CGA CTC ACT ATA GG-3'

```
                        10         20         30         40         50         60         70
14      T100R   CTACGAATAATGTTTACACAATTAAATATTGATTAGCTACTTAAATAATAAAAAACAATTTATCATTACAATTATATAT
-1      125R                                                          TTTATCATTCCAATTATATAT

CONSENSUS       CTACGAATAATGTTTACACAATTAAATATTGATTAGCTACTTAAATAATAAAAAACAATTTATCATTACAATTATATAT

                 80         90         100        110        120        130        140        150
14      T100R   GTATATAAAAAAACAGTACCGTCTTTATACATCTAAATAATATATATTGTAACTACTCATCAGATCATTAAATTATTTA
-1      125R    GTATATAAAAAAACAGTACCGTCTTTATACATCTAAATAATATATATTGTAACTACTCATCAGATCATTAAATTATTTA

CONSENSUS       GTATATAAAAAAACAGTACCGTCTTTATACATCTAAATAATATATATTGTAACTACTCATCAGATCATTAAATTATTTA

                 160        170        180        190        200        210        220        230
14      T100R   ACTATACAACAAGTAATAATAATCTATAAATCAAGACTCTTTACTTTAAAAAACTTTATTGATATCTATATATACTATA
-1      125R    ACTATACAACAAGTAATAATAATCTATAAATCAAGACTCTTTACTTTAAAAAACTTTATTGATATCTATATATACTATA
70      T1R                                AATCAAGACTCTTTACTTTAAAAAACTTTATTGATATCTATATATACTATA
-69     T1F                                AATCAAGACTCTTTACTTTAAAAAACTTTATTGATATCTATATATACTATA

CONSENSUS       ACTATACAACAAGTAATAATAATCTATAAATCAAGACTCTTTACTTTAAAAAACTTTATTGATATCTATATATACTATA

                  240        250        260        270        280        290        300        310
14      T100R   TGTTAGGGTGGCGGTAAAAAGGCTTATGCACTTTTTCAAAAATTCATTATCAGATTTTTCTATAATCTGGTTAGTTTTC
-1      125R    TGTTAGGGTGCCGGTAAAAAGGCTTATGCACTTTTTCAAAAATTCATTATCAGATTTTTCTATAATCTGGTTAGTTTTC
70      T1R     TGTTAGGGTGGCGGTAAAAAGGCTTATGCACTTTTTCAAAAATTCATTATCAGATTTTTCTATAA
-69     T1F     TGTTAGGGTGGCGGTAAAAAGGCTTATGCACTTTTTCAAAAATTCATTATCAGATTTTTCTATAA
-50     T125R         GGTGGCGGTAAAAAGGCTTATGCACTTTTTCAAAAATTCATTATCAGATTTTTCTATAATCTGGTTAGTGTTC
124     T82F                CGGTAAAAAGGCTTATGCACTTTTTCAAAAATTCATTATCAGATTTTTCTATAATCTGGTTAGTTTTC

CONSENSUS       TGTTAGGGTGGCGGTAAAAAGGCTTATGCACTTTTTCAAAAATTCATTATCAGATTTTTCTATAATCTGGTTAGTTTTC
```

Primer WL2TP1_gapR: 5'- TTA CCG CCA CCC TAA CAT ATA G -3'

```
                   320        330        340        350        360        370        380        390
14      T100R   CTACCATATATATCTTCAGTATTGATATTATTATTAAAAAAAGTTAGATACTCACAAGAAACAATAAGTTTTATTGCTT
-1      125R    CTACCATATATATCTTCAGTATTGATATTATTATTAAAAAAAGTTAGATACTCACAAGAAACAATAAGTTTTATTGCTT
-50     T125R   CTACCATATATATCTTCAGTATTGATATTATTATTAAAAAAAGTTAGATACTCACAAGAAACAATAAGTTTTATTGCTT
124     T82F    CTACCATATATATCTTCAGTATTGATATTATTATTAAAAAAAGTTAGATACTCACAAGAAACAATAAGTTTTATTGCTT
-52     T126R        ATATATATCTTCAGTATTGATATTATTATTAAAAAAAGTTAGATACTCACAAGAAACAATAAGTTTTATTGCTT

CONSENSUS       CTACCATATATATCTTCAGTATTGATATTATTATTAAAAAAAGTTAGATACTCACAAGAAACAATAAGTTTTATTGCTT

                   400        410        420        430        440        450        460        470
14      T100R   CATGCATGCTATTTTACGCAGCATTTAATATTTACACCCAAACAACCGATAACATATCAGCCTTACATGTTGTATTATG
-1      125R    CATGCATGCTATTTTACGCAGCATTTAATATTTACACCCAAACAACCGATAACATATCAGCCTTACATGTTGTATTATG
-50     T125R   CATGCATGCTATTTTACGCAGCATTTAATATTTACACCCAAACAACCGATAACATATCAGCCTTACATGTTGTATTATG
124     T82F    CATGCATGCTATTTTACGCAGCATTTAATATTTACACCCAAACAACCGATAACATATCAGTCTTACATGTTGTATTATG
-52     T126R   CATGCATGCTATTTTACGCAGCATTTAATATTTACACCCAAACAACCGATAACATATCAGCCTTACATGTTGTATTATG
132     T90R                                                                    TTGTATTATG
103     T4R                                                                      GTATTATG

CONSENSUS       CATGCATGCTATTTTACGCAGCATTTAATATTTACACCCAAACAACCGATAACATATCAGCCTTACATGTTGTATTATG
```

```
                      480        490        500        510        520        530        540        550
14       T100R  TGAGTTTACTCCTAAAATACAGCTAACATTAAAAACAGAATATATAGGCATAATATTCAGCATACTAGCTTCATTTTTG
-1       125R   TGAGTTTACTCCTAAAATACAGCTAACATTAAAAACAGAATATATAGGCATAATATTCAGCATACTAGCTTCATTTTTG
-50      T125R  TGAGTTTACTCCTAAAATACAGCTAACATTAAAAACAGAATATATAGGCATAATATTCAGCATACTAGCTTCATTTTTG
124      T82F   TGAGTTTACTCCTAAAATACAGCTAACATTAAAAACAGAATATATAGGCATAATATTCAGCATACTAGCTTCATTTTTG
-52      T126R  TGAGTTTACTCCTAAAATACAGCTAACATTAAAAACAGAATATATAGGCATAATATTCAGCATACTAGCTTCATTTTTG
132      T90R   TGAGTTTACTCCTAAAATACAGCTAACATTAAAAACAGAATATATAGGCATAATATTCAGCATACTAGCTTCATTTTTG
103      T4R    TGAGTTTACTCCTAAAATACAGCTAACATTAAAAACAGAATATATAGGCATAATATTCAGCATACTAGCTTCATTTTTG
135      T96_F                              CATTAAAAACAGAATATATAGGCATAATATTCAGCATACTAGCTTCATTTTTG
19       T102_F                                                                          TTTG

CONSENSUS      TGAGTTTACTCCTAAAATACAGCTAACATTAAAAACAGAATATATAGGCATAATATTCAGCATACTAGCTTCATTTTTG


                    560        570        580        590        600        610        620        630
14       T100R  TGGATAGTAACTACTATTTATACTATTAGTTATATGAGACACAATGATAAAAATAACAAAAAGCAATCTATATTTTATG
-1       125R   TGGATAGTAACTACTATTTATACTATTAGTTATATGAGACACAATGATAAAAATAACAAAAAGCAATCTATATTTTATG
-50      T125R  TGGATAGTAACTACTATTTATACTATTAGTTATATGAGACACAATGATAAAAATAACAAAAAGCAATCTATATTTTATG
124      T82F   TGGATAGTAACTACTATTTATACTATTAGTTATATGAGACACAATGATAAAAATAACAAAAAGCAATCTATATTTTATG
-52      T126R  TGGATAGTAACTACTATTTATACTATTAGTTATATGAGACACAATGATAAAAATAACAAAAAGCAATCTATATTTTATG
132      T90R   TGGATAGTAACTACTATTTATACTATTAGTTATATGAGACACAATGATAAAAATAACAAAAAGCAATCTATATTTTATG
103      T4R    TGGATAGTAACTACTATTTATACTATTAGTTATATGAGACACAATGATAAAAATAACAAAAAGCAATCTATATTTTATG
135      T96_F  TGGATAGTAACTACTATTTATACTATTAGTTATATGAGACACAATGATAAAAATAACAAAAAGCAATCTATATTTTATG
19       T102_F TGGATAGTAGCTACTATTTATACTATTAGTTATATGAGACACAATGATAAAAATAACAAAAAGCAATCTATATTTTATG

CONSENSUS      TGGATAGTAACTACTATTTATACTATTAGTTATATGAGACACAATGATAAAAATAACAAAAAGCAATCTATATTTTATG


                      640        650        660        670        680        690        700
14       T100R  CGTGTTTTGCAGCAAGTATTGGATGTACAATGGGAAATGCCTTTTCAGGAAATCTATTAACATTATTCA
-1       125R   CGTGTTTTGCAGCAAGTATTGGATGTACAATGGGAATTGCTTTTTCAGGAAATCTATTAACATTATTCATATTTTATGA
-50      T125R  CGTGTTTTGCAGCAAGTATTGGATGTACAATGGGAATTGCTTTTTCAGGAAATCTATTAACATTATTCATATTTTATGA
124      T82F   CGTGTTTTGCAGCAAGTATTGGATGTACAATGGGAATTGCTTTTTCAGGAAATCTATTAACATTATTCATATTTTATGA
-52      T126R  CGTGTTTTGCAGCAAGTATTGGATGTACAATGGGAATTGCTTTTTCAGGAAATCTATTAACATTATTCATATTTTATGA
132      T90R   CGTGTTTTGCAGCAAGTATTGGATGTACAATGGGAATTGCTTTTTCAGGAAATCTATTAACATTATTCATATTTTATGA
103      T4R    CGTGTTTTGCAGCAAGTATTGGATGTACAATGGGAATTGCTTTTTCAGGAAATCTATTAACATTATTCATATTTTATGA
135      T96_F  CGTGTTTTGCAGCAAGTATTGGATGTACAATGGGAATTGCTTTTTCAGGAAATCTATTAACATTATTCATATTTTATGA
19       T102_F CGTGTTTTGCAGCAAGTATTGGATGTACAATGGGAATTGCTTTTTCAGGAAATCTATTAACATTATTCATATTTTATGA
-18      T102R                GTATTGGATGTACAATGGGAATTGCTTTTTCAGGAAATCTATTAACATTATTCATATTTTATGA
73       T22F                        GTACAATGGGAATTGCTTTTTCAGGAAATCTATTAACATTATTCATATTTTATGA
97       T46R                                                    CATTATTCATATTTTATGA

CONSENSUS      CGTGTTTTGCAGCAAGTATTGGATGTACAATGGGAATTGCTTTTTCAGGAAATCTATTAACATTATTCATATTTTATGA


                    710        720        730        740        750        760        770        780
-1       125R   GTTGCTAACTATTAGCACTTACCCTCTTGTTACATATTATGCAAATCACGAATCACAAATTTCTGGTAGGTATTATATG
-50      T125R  GTTGCTAACTATTAGCACTTACCCTCTTGTTACATATTATGCAAATCACGAATCACAAATTTCTGGTAGGTATTATATG
124      T82F   GTTGCTAACTATTAGCACTTACCCTCTTGTTACATATTATGCAAATCACGAATCACAAATTTCTGGTAGGTATTATATG
-52      T126R  GTTGCTAACTATTAGCACTTACCCTCTTGTTACATATTATGCAAATCACGAATCACAAATTTCTGGTAGGTATTATATG
132      T90R   GTTGCTAACTATTAGCACTTACCCTCTTGTTACATATTATGCAAATCACGAATCACAAATTTCTGGTAGGTATTATATG
103      T4R    GTTGCTAACTATTAGCACTTACCCTCTTGTTACATATTATGCAAATCACGAATCACAAATTTCTGGTAGGTATTATATG
135      T96_F  GTTGCTAACTATTAGCACTTACCCTCTTGTTACATATTATGCAAATCACGAATCACAAATTTCTGGTAGGTATTATATG
19       T102_F GTTGCTAACTATTAGCACTTACCCTCTTGTTACATATTATGCAAATCACGAATCACAAATTTCTGGTAGGTATTATATG
-18      T102R  GTTGCTAACTATTAGCACTTACCCTCTTGTTACATATTATGCAAATCACGAATCACAAATTTCTGGTAGGTATTATATG
73       T22F   GTTGCTAACTATTAGCACTTACCCTCTTGTTACATATTATGCAAATTACGAATCACAAATTTCTGGTAGGTATTATATG
97       T46R   GTTGCTAACTATTAGCACTTACCCTCTTGTTACATATTATGCAAATCACGAATCACAAATTTCTGGTAGGTATTATATG
30       T109_F                                                          ACAAATTTCTGGTAGGTATTATATG
89       T40F                                                                            ATG

CONSENSUS      GTTGCTAACTATTAGCACTTACCCTCTTGTTACATATTATGCAAATCACGAATCACAAATTTCTGGTAGGTATTATATG
```

```
                 790       800       810       820       830       840       850       860
-1       125R    GGAATACTGCTGGGAACTTCAATGTTATTATTTTTACCAGCAATCATCATTACTTAT
-50      T125R   GGAATACTGCTGGGAACTTCAATGTTATTATTTTTACCAGCAATCATCATTACTTAT
124      T82F    GGAATACTGCTGGGA
-52      T126R   GGAATACTGCTGGG
132      T90R    GGAATACTGCTGGGAACTTCAATGTTATTATTTTTACCAGCAATCATCATTACTTATAACATTAGCGGCACCTTAGATT
103      T4R     GGAATACTGCTAGGAACTTCAATGTTATTATTTTTACCAGCAATCATCATTACTTATAACATTAGCGGCACCTTAGATT
135      T96_F   GGAATACTGCTGGGAACTTCAATGTTATTATTTTTACCAGCAATCATCATTACTTATAACATTAGCGGCACCTTAGATT
19       T102_F  GGAATACTGCTGGGAACTTCAATGTTATTATTTTTACCAGCAATCATCATTACTTATAACATTAGCGGCACCTTAGATT
-18      T102R   GGAATACTGCTGGGAACTTCAATGTTATTATTTTTACCAGCAATCATCATTACTTATAACATTAGCGGCACCTTAGATT
73       T22F    GGAATACTGCTGGGAACTTCAATGTTATTATTTTTACCAGCAATCATCATTACTTATAACATTAGCGGCACCTTAGATT
97       T46R    GGAATACTGCTGGGAACTTCAATGTTATTATTTTTACCAGCAATCATCATTACTTATAACATTAGCGGCACCTTAGATT
30       T109_F  GGAATACTGCTGGGAACTTCAATGTTATTATTTTTACCAGCAATCATCATTACTTATAACATTAGCGGCACCTTAGATT
89       T40F    GGAATACTGCTGGGAACTTCAATGTTATTATTTTTACCAGCAATCATCATTACTTATAACATTAGCGGCACCTTAGATT
122      T81F                                                                        CTTAGATT

CONSENSUS        GGAATACTGCTGGGAACTTCAATGTTATTATTTTTACCAGCAATCATCATTACTTATAACATTAGCGGCACCTTAGATT

                 870       880       890       900       910       920       930       940
132      T90R    TTACAAAAGGTGGCATATTACCATCAAGCATCTCTAGCGTTTTTTTAATGAGCTTATTATTCCTATTCATTTACAGTAT
103      T4R     TTACAAAAGGTGGCATATTACCATCAAGCATCTCTAGCGTTTTTTTAATGAGCTTATTATTCCTATTCATTTACAGTAT
135      T96_F   TTACAAAAGGTGGCATATTACCATCAAGCATCTCTAGCGTTTTTTTAATGAGCTTATTATTCCTATTCATTTACAGTAT
19       T102_F  TTACAAAAGGTGGCATATTACCATCAAGCATCTCTAGCGTTTTTTTAATGAGCTTATTATTCCTATTCATTTACAGTAT
-18      T102R   TTACAAAAGGTGGCATATTACCATCAAGCATCTCTAGCGTTTTTTTAATGAGCTTATTATTCCTATTCATTTACAGTAT
73       T22F    TTACAAAAGGTGGCATATTACCATCAAGCATCTCTAGCGTTTTTTTAATGAGCTTATTATTCCTATTCATTTACAGTAT
97       T46R    TTACAAAAGGTGGCATATTACCATCAAGCATCTCTAGCGTTTTTTTAATGAGCTTATTATTCCTATTCATTTACAGTAT
30       T109_F  TTACAAAAGGTGGCATATTACCATCAAGCATCTCTAGCGTTTTTTTAATGAGCTTATTATTCCTATTCATTTACAGTAT
89       T40F    TTACAAAAGGTGGCATATTACCATCAAGCATCTCTAGCGTTTTTTTAATGAGCTTATTATTCCTATTCATTTACAGTAT
122      T81F    TTACAAAAGGTGGCATATTACCATCAAGCATCTCTAGCGTTTTTTTAATGAGCTTATTATTCCTATTCATTTACAGTAT

CONSENSUS        TTACAAAAGGTGGCATATTACCATCAAGCATCTCTAGCGTTTTTTTAATGAGCTTATTATTCCTATTCATTTACAGTAT

                 950       960       970       980       990      1000      1010      1020
132      T90R    AGGAAAAACTGCATTAATGCCTATACATTCATGGTTACCAAGAGCTATGGTTGCCCCTACTCCAGTAAGTGCA
103      T4R     AGGAAAAACTGCATTAATGCCTATACATTCA
135      T96_F   AGGAAAAACTGCATTAATGCCTATACATTCATGGTTACCAAGAGCTATGGTTGCCCCTACTCCAGTAAGTGCACTATTA
19       T102_F  AGGAAAAACTGCATTAATGCCTATACATTCATGGTTACCAAGAGCTATGGTTGCCCCTACTCCAGTAAGTGCACTATTA
-18      T102R   AGGAAAAACTGCATTAATGCCTATACATTCATGGTTACCAAGAGCTATGGTTGCCCCTACTCCAGTAAGTGCACTATTA
73       T22F    AGGAAAAACTGCATTAATGCCTATACATTCATGGTTACCAAGAGCTATGGTTGCCCCTACTCCAGTAAGTGCACTATTA
97       T46R    AGGAAAAACTGCATTAATGCCTATACATTCATGGTTACCAAGAGCTATGGTTGCCCCTACTCCAGTAAGTGCACTATTA
30       T109_F  AGGAAAAACTGCATTAATGCCTATACATTCATGGTTACCAAGAGCTATGGTTGCCCCTACTCCAGTAAGTGCACTATTA
89       T40F    AGGAAAAACTGCATTAATGCCTATACATTCATGGTTACCAAGAGCTATGGTTGCCCCTACTCCAGTAAGTGCACTATTA
122      T81F    AGGAAAAACTGCATTAATGCCTATACATTCATGGTTACCAAGAGCTATGGTTGCCCCTACTCCAGTAAGTGCACTATTA
71       T21F                                                CTATGGTTGCCCCTACTCCAGTAAGTGCACTATTA

CONSENSUS        AGGAAAAACTGCATTAATGCCTATACATTCATGGTTACCAAGAGCTGCTGTTGCCCCTACTCCAGTAAGTGCACTATTA

                1030      1040      1050      1060      1070      1080      1090      1100
135      T96_F   CATGCAGTAGCAGTTGTCAAATCTGGAGTTTTTGCTTTAATAAAGGTTTTATTGTATATAATCGGGTTAGAAAGACTAC
19       T102_F  CATGCAGTAGCAGTTGTCAAATCTGGAGTTTTTGCTTTAATAAAGGTTTTATTGTATATAATCGGGTTAGAAAGACTAC
-18      T102R   CATGCAGTAGCAGTTGTCAAATCTGGAGTTTTTGCTTTAATAAAGGTTTTATTGTATATAATCGGGTTAGAAAGACTAC
73       T22F    CATGCAGTAGCAGTTGTCAAATCTGGAGTTTTTGCTTTAATAAAGGTTTTATTGTATATAACCGGGTTAGAAAGACTAC
97       T46R    CATGCAGTAGCAGTTGTCAAATCTGGAGTTTTTGCTTTAATAAAGGTTTTATTGTATATAATCGGGTTAGAAAGACTAC
30       T109_F  CATGCAGTAGCAGTTGTCAAATCCGGAGTTTTTGCTTTAATAAAGGTTTTATTGTATATAATCGGGTTAGAAAGACTAC
89       T40F    CATGCAGTAGCAGTTGTCAAATCTGGAGTTTTTGCTTTAATAAAGGTTTTATTGTATATAATCGGGTTAGAAAGACTAC
122      T81F    CATGCAGTAGCAGTTGTCAAATCTGGAGTTTTTGCTTTAATAAAGGTTTTATTGTATATAATCGGGTTAGAAAGACTAC
71       T21F    CATGCAGTAGCAGTTGTCAAATCTGGAGTTTTTGCTTTAATAAAGGTTTTATTGTATATAATCGGGTTAGAAAGACTAC

CONSENSUS        CATGCAGTAGCAGTTGTCAAATCTGGAGTTTTTGCTTTAATAAAGGTTTTATTGTATATAATCGGGTTAGAAAGACTAC

                1110      1120      1130      1140      1150      1160      1170      1180
135      T96_F   AAGCGTTAGTTCAAACACATAACAATATATTGATGTACGCTTCAGCAATTACAATTCTCTTAGCATCCTTTATAGCAAT
19       T102_F  AAGCGTTAGTTCAAACACATAACAATATATTGATGTACGCTTCAGCAATTACAATTCTCTTAGCATCCTTTATAGCAAT
-18      T102R   AAGCGTTAGTTCAAACACATAACAATATATTGATGTACGCTTCAGCAATTACAATTCTCTTAGCATCCTTTATAGCAAT
73       T22F    AAGCGTTAGTTCAAACACATAACAATATATTGATGTACGCTTCAGCAATTACAATTCTCTTAGCATCCTTTATAGCAAT
97       T46R    AAGCGTTAGTTCAAACACATAACAATATATTGATGTACGCTTCAGCAATTACAATTCTCTTAGCATCCTTTATAGCAAT
30       T109_F  AAGCGTTAGTTCAAACACATAACAATATATTGATGTACGCTTCAGCAATTACAATTCTCTTAGCATCCTTTATAGCAAT
89       T40F    AAGCGTTAGTTCAAACACATAACAATATATTGATGTACGCTTCAGCAATTACAATTCTCTTAGCATCCTTTATAGCAAT
122      T81F    AAGCGTTAGTTCAAACACATAACAATATATTGATGTACGCTTCAGCAATTACAATTCTCTTAGCATCCTTTATAGCAAT
71       T21F    AAGCGTTAGTTCAAACACATAACAATATATTGATGTACGCTTCAGCAATTACAATTCTCTTAGCATCCTTTATAGCAAT

CONSENSUS        AAGCGTTAGTTCAAACACATAACAATATATTGATGTACGCTTCAGCAATTACAATTCTCTTAGCATCCTTTATAGCAAT
```

```
                    1190       1200       1210       1220       1230       1240       1250       1260
135      T96_F      AAAACAAACAAACTTAAAGAAGTTATTAGCATATTCAACA
19       T102_F     AAAACAAACAAACTTAAAGAAGTTATTAGCATATTCAACAATTTCAC
-18      T102R      AAAACAAACAAACTTAAAGAAGTTATTAGCATATTCAACAATTTCACAGCTTTCTTATATAACAATAGCAGTATCATTG
73       T22F       AAAACAAACAAACTTAAAGAAGTTATTAGCATATTCAACAATTTCACAGCTTTCTTATATAACAATAGCAGTATCATTG
97       T46R       AAAACAAACAAACTTAAAGAAGTTATTAGCATATTCAACAATTTCACAGCTTTCTTATATAACAATAGCAGTATCATTG
30       T109_F     AAAACAAACAAACTTAAAGAAGTTATTAGCATATTCAACAATTTCACAGCTTTCTTATATAACAATAGCAGTATCATTG
89       T40F       AAAACAAACAAACTTAAAGAAGTTATTAGCATATTCAACAATTTCACAGCT
122      T81F       AAAACAAACAAACTTAAAGAAGTTATTAGCATATTCAACAATTTCACAGCTTTCTTATATAACAATAGCAGTATCATTG
71       T21F       AAAACAAACAAACTTAAAGAAGTTATTAGCATATTCAACAATTTCACAGCTTTCTTATATAACAATAGCAGTATCATTG

CONSENSUS           AAAACAAACAAACTTAAAGAAGTTATTAGCATATTCAACAATTTCACAGCTTTCTTATATAACAATAGCAGTATCATTG

                    1270       1280       1290       1300       1310       1320       1330       1340
-18      T102R      TATACAGAACGTGCCGGTTGATATTTCTATATTTCAAATGATATCGCATGCATTT
73       T22F       TATACAGAACGTGCCGGTTGATATTTCTATATTTCAAATGATATCGCATGCATT
97       T46R       TATACAGAACGTGCCGGTTGATATTTCTATATTTCAAATGATATCGCATGCATTTGCAAAAATAACA
30       T109_F     TATACAGAACGTGCCGGTTGATATTTCTATATTTCAAATGATATCGCATGCATTTGCAAAAATAACATTATTCTTTACT
122      T81F       TATACAGAACGTGCCGGTTGATATTTCTATATTTCAAATGATATCGCATGCATTTGCAAAAATAACATTATTCTTTACT
71       T21F       TATACAGAACGTGCCGGTTGATATTTCTATATTTCAAATGATATCGCACGCATTTGCAAAAATAACATTATTCTTTACT

CONSENSUS           TATACAGAACGTGCCGGTTGATATTTCTATATTTCAAATGATATCGCATGCATTTGCAAAAATAACATTATTCTTTACT

                    1350       1360       1370       1380       1390       1400       1410       1420
30       T109_F     GCAGGAGCAATATATACCAAAACAGGAAAAAAATACTTAAATGAACTTCAAGGTATTGGTAAATCCATGCCAATAACAA
122      T81F       GCAGGAGCAATATATACCAAAACAGGAAAAAAATACTTAAAT
71       T21F       GCAGGAGCAATATATACCAAAACAGGAAAAAAATACTTAAATGAACTTCAAGGTATTGGTAAATCCATGCCAATAACAA
-83      T32F                                                                                    AA
-25      T106_F                                                                                  AA
-134     T96R                                                                                    AA
15       T100_F                                                                                  AA
-16      T101R                                                                                   AA
53       T128R                                                                                   AA
-123     T81R                                                                                    AA

CONSENSUS           GCAGGAGCAATATATACCAAAACAGGAAAAAAATACTTAAATGAACTTCAAGGTATTGGTAAATCCATGCCAATAACAA

                    1430       1440       1450       1460       1470       1480       1490     1500
30       T109_F     TGACAGCGTTCTCTATANGAGCTGCTGCAATGATAGGCATTCCTCCTGCTGTAACATTTTGGGGAAAATTTTTTATTAT
71       T21F       TGACAGCGTTCTCTATAGGAGCTGCTGCAATGATAGGCATTCCTCCTGCTGTAACATTTTGGGGAAAATTTTTTATTAT
-83      T32F       TGACAGCGTTCTCTATAGGAGCTGCTGCAATGATAGGCATTCCTCCTGCTGTAACATTTTGGGGAAAATTTTTTATTAT
-25      T106_F     TGACAGCGTTCTCTATAGGAGTTGCTGCAATGATAGGCATTCCTCCTGCTGTAACATTTTGGGGAAAATTTTTTATTAT
-134     T96R       TGACAGCGTTCTCTATAGGAGCTGCTGCAATGATAGGCATTCCTCCTGCTGTAACATTTTGGGGAAAATTTTTTATTAT
-15      T100_F     TGACAGCGTTCTCTATAGGAGCTGCTGCAATGATAGGCATTCCTCCTGCTGTAACATTTTGGGGAAAATTTTTTATTAT
-16      T101R      TGACAGCGTTCTCTATAGGAGCTGCTGCAATGATAGGCATTCTTCCTGCTGTAACATTTTGGGGAAAATTTTTTATTAT
-53      T128R      TGACAGCGTTCTCTATAGGAGCTGCTGCAATGATAGGCATTCCTCCTGCTGTAACATTTTGGGGAAAATTTTTTATTAT
-123     T81R       TGACAGCGTTCTCTATAGGAGCTGCTGCAATGATAGGCATTCCTCCTGCTGTAACATTTTGGGNAAAATTTTTTATTAT
-90      T40R                 TCTCTATAGGAGCTGCTGCAATGATAGGCATTCCTCCTGCTGTACATTTTTGGGGAAAATTTTTTATTAT
-125     T82R                   TCTATAGGAGCTGCTGCAATGATAGGCATTCCTCCTGCTGTAACATTTTGGGGAAAATTTTTTATTAT
-74      T22R                   TTTATAGGAGCTGTTGCAATGATAGGCATTCCTCCTGCTGTAACATTTGGGGGAAAATTTTTTATTAT
-102     T4F                     TTATAGGAGCTGCTGCAATGATAGGCATTCCTCCTGGTGTAACATTTTGGGGAAAATTTTTTATTAT
-67      T18F                        GGAAGCTGCTGCAATGATAGGCATTCCTCGTGCTGTAACATTTTGGGGAAAATTTTTTATTAT
-29      T109R                       GGAGCTGCTGCAATAATAGGCATTCCTCCTGCTGTAACATTTTGGGGAAAATTTTTTATTAT
-96      T46F                          AGCTGCTCCAATGATAGGCATTCCTCCTGGTGTAACATTTGGGGGAAAATTTTTTATTAT
 95      T44R                                                                  GGAAAATTTTTTATTAT
13       87R                                                                       TTTTTTATTAT
-129     T87_F                                                                     TTTTTTATTAT
-92      T41R                                                                         TTATTAT

CONSENSUS           TGACAGCGTTCTCTATAGGAGCTGCTGCAATGATAGGCATTCCTCCTGCTGTAACATTTTGGGGAAAATTTTTTATTAT
```

```
                          1510        1520        1530        1540        1550        1560        1570
30        T109_F  ATCAGAATCATTAAATCAAAATATTATGTCTGTAGTACTAGTATTAATAGCAAGCACCATACTAAATACAATATATTTT
71        T21F    ATCAGAATCATTAAATCAAAATATTATGTCTGTAGTACTAGTATTAATAGCAAGCACCATACTAAATACAATATATTTT
-83       T32F    ATCAGAATCATTAAATCAAAATATTATGTCTGTAGTACTAGTATTAATAGCAAGCACCATACTAAATACAATATATTTT
-25       T106_F  ATCAGAATCATTAAATCAAAATATTATGTCTGTAGTACTAGTATTAATAGCAAGCACCATACTAAATACAATATATTTT
-13       4T96R   ATCAGAATCATTAAATCAAAATATTATGTCTGTAGTACTAGTATTAATAGCAAGCACCATACTAAATACAATATATTTT
-15       T100_F  ATCAGAATCATTAAATCAAAATATTATGTCTGTAGTACTAGTATTAATAGCAAGCACCATACTAAATACAATATATTTT
-16       T101R   ATCAGAATCATTAAATCAAAATATTATGTCTGTAGTACTAGTATTAATAGCAAGCACCATACTAAATACAATATATTTT
-53       T128R   ATCAGAATCATTAAATCAAAATATTATGTCTGTAGTACTAGTATTAATAGCAAGCACCATACTAAATACAATATATTTT
-123      T81R    ATCAGAATCATTAAATCAAAATATTATGTCTGTAGTACTAGTATTAATAGCAAGCACCATACTAAATACAATATATTTT
-90       T40R    ATCAGAATCATTAAATCANAATATTATGTCTGTAGTACTAGTATTAATAGCAAGCACCATACTAAATACAATATATTTT
-125      T82R    ATCAGAATCATTAAATCAAAATATTATGTCTGTAGTACTAGTATTAATAGCAAGCACCATACTAAATACAATATATTTT
-74       T22R    ATCAGAATCATTAAATCAAAATATTATGTCTGTAGTACTAGTATTAATAGCAAGCACCATACTAAATACAATATATTTT
-102      T4F     ATCAGAATCATTAAATCAAAATATTATGTCTGTAGTACTAGTATTAATAGCAAGCGCCATACTAAATACAATATATTTT
-67       T18F    ATCAAAATCATTAAATCAAAATATTATGTCTGTAGTACTAGTATTAATAGCAAGCGCCATACTAAATACAATATATTTT
-29       T109R   ATCAGAATCATTAAATCAAAATATTATGTCTGTAGTACTAGTATTAATAGCAAGCACCATACTAAATACAATATATTTT
-96       T46F    ATCAGAATCATTAAATCAAAATATTATGTCTGTAGTACTAGTATTAATAGCAAGCGCCATACTAAATACAATATATTTT
95        T44R    ATCAGAATCATTAAATCAAAATATTATGTCTGTAGTACTAGTATTAATAGCAAGCACCATACTAAATACAATATATTTT
13        87R     ATCAGAATCATTAAATCAAAATATTATGTCTGTAGTACTAGTATTAATAGCAAGCACCATACTAAATACAATATATTTT
-129      T87_F   ATCAGAATCATTAAATCAAAATATTATGTCTGTAGTACTAGTATTAATAGCAAGCACCATACTAAATACAATATATTTT
-92       T41R    ATCAGAATCATTAAATCAAAATATTACGTCTGTAGTACTAGTATTAATAGCAAGCACCATACTAAATACAATATATTTT
61        51R         GAATCATTAAATCAAAATATTATGTCTGTAGTACTAGTATTAATAGCAAGCACCATACTAAATACAATATATTTT
-65       T151_F      GAATCATTAAATCAAAATATTATGTCTGTAGTACTAGTATTAATAGCAAGCACCATACTAAATACAATATATTTT
-72       T21R                                                    TTAATAGCAAGCACCATACTAAATACAATATATTTT

CONSENSUS         ATCAGAATCATTAAATCAAAATATTATGTCTGTAGTACTAGTATTAATAGCAAGCACCATACTAAATACAATATATTTT


                  1580      1590      1600      1610      1620      1630      1640      1650
30        T109_F  ATCCCTATTATATACAATGCGTT
71        T21F    ATCCCTATTATATACAATGCGTT
-83       32F     ATCCCTATTATATACAATGCGTTTTATGTCCCATGTAACACTAATAAATATGCTGAAGCTCCTATACCCATGCTAATTG
-25       T106_F  ATCCCTATTATATACAATGCGTTTTATGTCCCATGTAACACTAATAAATATGCTGAAGCTCCTATACCCATGCTAATTG
-134      T96R    ATCCCTATTATATACAATGCGTTTTATGTCCCATGTAACACTAATAAATATGCTGAAGCTCCTATACCCATGCTAATTG
-15       T100_F  ATCCCTATTATATACAATGCGTTTTATGTCCCATGTAACACTAATAAATATGCTGAAGCTCCTATACCCATGCTAATTG
-16       101R    ATCCCTATTATATACAATGCGTTTTATGTCCCATGTAACACTAATAAATATGCTGAAGCTCCTATACCCATGCTAATTG
-53       T128R   ATCCCTATTATATACAATGCGTTTTATGTCCCATGTAACACTAATAAATATGCTGAAGCTCCTATACCCATGCTAATTG
-123      T81R    ATCCCTATTATATACAATGCGTTTTATGTCCCATGTAACACTAATAAATATGCTGAAGCTCCTATACCCATGCTAATTG
-90       T40R    ATCCCTATTATATACAATGCGTTTTATGTCCCATGTAACACTAATAAATATGCTGAAGCTCCTATACCCATGCTAATTG
-125      T82R    ATCCCTATTATATACAATGCGTTTTACGTCCCATGTAACACTAATAAATATGCTGAAGCTCCTATACCCATGCTAATTG
-74       T22R    ATCCCTATTATATACAATGCGTTTTATGTCCCATGTAACACTAATAAATATGCTGAAGCTCCTATACCCATGCTAATTG
-102      T4F     ATCCCTATTATATACAATGCGTTTTATGTCCCATGTAACACTAATAAATATGCTGAAGCTCCTATACCCATGCTAATTG
-67       18F     ATCCCTATTATATACAATGCGTTTTATGTCCCATGTAACACTAATAAATATGCTGAAGCTCCTATACCCATGCTAATTG
-29       T109R   ATCCCTATTATATACAATGCGTTTTATGTCCCATGTAACACTAATAAATATGCTGAAGCTCCTATACCCATGCTAATTG
-96       46F     ATCCCTATTATATACAATGCGTTTTATGTCCCATGTAACACTAATAAATATGCTGAAGCTCCTATACCCATGCTAATTG
95        T44R    ATCCCTATTATATACAATGCGTTTTATGTCCCATGTAACACTAATAAATATGCTGAAGCTCCTATACCCATGCTAATTG
13        87R     ATCCCTATTATATACAATGCGTTTTATGTCCCATGTAACACTAATAAATATGCTGAAGCTCCTATACCCATGCTAATTG
-129      T87_F   ATCCCTATTATATACAATGCGTTTTATGTCCCATGTAACACTAATAAATATGCTGAAGCTCCTATACCCATGCTAATTG
-92       41R     ATCCCTATTAAATACAATGCGTTTTATGTCCCATGTAACACTAATAAATATGCTGAAGCTCCTATACCCATGNTAATTG
6         151R    ATCCCTATTATATACAATGCGTTTTATGTCCCATGTAACACTAATAAATATGCTGAAGCTCCTATACCCATGCTAATTG
-65       T151_F  ATCCCTATTATATACAATGCGTTTTATGTCCCATGTAACACTAATAAATATGCTGAAGCTCCTATACCCATGCTAATTG
-72       T21R    ATCCCTATTATATACAATGCGTTTTATGTCCCATGTAACACTAATAAATATGCTGAAGCTCCTATACCCGTGCTAATTG
-79       T30F                                                                          TGCTAATTG
80        T30R                                                                          TGCTAATTG

CONSENSUS         ATCCCTATTATATACAATGCGTTTTATGTCCCATGTAACACTAATAAATATGCTGAAGCTCCTATACCCATGCTAATTG
```

```
                1660       1670       1680       1690       1700       1710       1720       1730
-83     T32F    CCATTTCTATAACAACAATATGTACAATATTATTATTTTTATATCCTGACGTAATATTCAATATAATAAATCACCTCAA
-25     T106_F  CCATTTCTATAACAACAATATGTACAATATTATTATTTTTATATCCTGACGTAATATTCAATATAATAAATCACCTCAA
-134    T96R    CCATTTCTATAACAACAATATGTACAATATTATTATTTTTATATCCTGACGTAATATTCAATATAATAAATCACCTCAA
-15     T100_F  CCATTTCTATAACAACAATATGTACAATATTATTATTTTTATATCCTGACGTAATATTCAATATAATAAATCACCTCAA
-16     T101R   CCATTTCTATAACAACAATATGTACAATATTATTATTTTTATATCCTGACGTAATATTCAATATAATAAATCACCTCAA
-53     T128R   CCATTTCTATAACAACAATATGTACAATATTATTATTTTTATATCCTGACGTAATATTCAATATAATAAATCACCTCAA
-123    T81R    CCATTTCTATAACAACAATATGTACAATATTATTATTTTTATATCCTGACGTAATATTCAATATAATAAATCACCTCAA
-90     T40R    CCATTTCTATAACAACAATATGTACAATATTATTATTTTTATATCCTGACGTAATATTCAATATAATAAATCACCTCAA
-125    T82R    CCATTTCTATAACAACAATATGTACAATATTATTATTTTTATATCCTGACGTAATATTCAATATAATAAATCACCTCAA
-74     T22R    CCATTTCTATAACAACAATATGTACAATATTATTATTTTTATATCCTGACGTAATATTCAATATAATAAATCACCTCAA
-102    T4F     CCATTTCTATAACAACAATATGTACAATATTATTATTTTTATATCCTGACGTAATATTCAATATAATAAATCACCTCAA
-67     T18F    CCATTTCTATAACAACAATATGTACAATATTATTATTTTTATATCCTGACGTAATATTCAATATAATAAATCACCTCAA
-29     T109R   CCATTTCTATAACAACAATATGTACAATATTATTATTTTTATATCCTGACGTAATATTCAATATAATAAATCACCTCAA
-96     T46F    CCATTTCTATAACAACAATATGTACAATATTATTATTTTTATATCCTGACGTAATATTCAATATAATAAATCACCTCAA
95      T44R    CCATTTCTATAACAACAATATGTACAATATTATTATTTTTATATCCTGACGTAATATTCAATATAATAAATCACCTCAA
13      87R     CCATTTCTATAACAACAATATGTACAATATTATTATTTTTATATCCTGACGTAATATTCAATATAATAAATCACCTCAA
-129    T87_F   CCATTTCTATAACAACAATATGTACAATATTATTATTTTTATATCCTGACGTAATATTCAATATAATAAATCACCTCAA
-92     T41R    CCATTTCTATAACAACAATATGTACAATATTATTATTTTTCTGACGTAATATTCAATATAATAAATCACCTCAA
6       151R    CCAT
-65     T151_F  CCATTTCTATAACAACAATATGTACAATATTATTATTTTTATATCCTGACGTAATATTCAATATAATAAATCACCTCAA
-72     T21R    CCATTTCTATAACAACAATATGTACAATATTATTATTTTTATATCCTGACGTAATATTCAATATAATAAATCACCTCAA
-79     T30F    CCATTTCTATAACAACAATATGTACAATATTATTATTTTTATATCCTGACGTAATATTCAATATAATAAATCACCTCAA
80      T30R    CCATTTCTATAACAACAATATGTACAATATTATTATTTTTATATCCTGACGTAATATTCAATATAATAAATCACCTCAA

CONSENSUS       CCATTTCTATAACAACAATATGTACAATATTATTATTTTTATATCCTGACGTAATATTCAATATAATAAATCACCTCAA


                1740       1750       1760       1770       1780       1790       1800       1810
-83     T32F    ATAAGATAATGTATATAAATACATTGGTTAATACCATATAAAATATTCTATTTGGAGTACAAAATAAGAATTAGTTAAA
-25     T106_F  ATAAGATAATGTATATAAATACATTGGTTAATACCATATAAAATATTCTATTTGGAGTACAAAATAAGAATTAGTTAAA
-134    T96R    ATAAGATAATGTATATAAATACATTGGTTAATACCATATAAAATATTCTATTTGGAGTACAAAATAAGAATTAGTTAAA
-15     T100_F  ATAAGATAATGTATATAAATACATTGGTTAATACCATATAAAATATTCTATTTGGAGTACAAAATAAGAATTAGTTAAA
-16     T101R   ATAAGATAATGTATATAAATACATTGGTTAATACCATATAAAATATTCTATTTGGAGTACAAAATAAGAATTAGTTAAA
-53     T128R   ATAAGATAATGTATATAAATACATTGGTTAATACCATATAAAATATTCTATTTGGAGTACAAAATAAGAATTAGTTAAA
-123    T81R    ATAAGATAATGTATATAAATACATTGGTTAATACCATATAAAATATTCTATTTGGAGTACAAAATAAGAATTAGTTAAA
-90     T40R    ATAAGATAATGTATATAAATACATTGGTTAATACCATATAAAATATTCTATTTGGAGTACAAAATAAGAATTAGTTAAA
-125    T82R    ATAAGATAATGTATATAAATACATTGGTTAATACCATATAAAATATTCTATTTGGAGTACAAAATAAGAATTAGTTAAA
-74     T22R    ATAAGATAATGTATATAAATACATTGGTTAATACCATATAAAATATTCTATTTGGAGTACAAAATAAGAATTAGTTAAA
-102    T4F     ATAAGATAATGTATATAAATACATTGGTTAATACCATATAAAATATTCTATTTGGAGTACAAAATAAGAATTAGTTAAA
-67     T18F    ATAAGATAATGTATATAAATACATTGGTTAATACCATATAAAATATTCTATTTGGAGTACAAAATAAGAATTAGTTAAA
-29     T109R   ATAAGATAATGTATATAAATACATTGGTTAATACCATATAAAATATTCTATTTGGAGTACAAAATAAGAATTAGTTAAA
-96     T46F    ATAAGATAATGTATATAAATACATTGGTTAATACCATATAAAATATTCTATTTGGAGTACAAAATAAGAATTAGTTAAA
95      T44R    ATAAGATAATGTATATAAATACATTGGTTAATACCATATAAAATATTCTATTTGGAGTACAAAATAAGAATTAGTTAAA
13      87R     ATAAGATAATGTATATAAATACATTGGTTAATACCATATAAAATATTCTATTTGGAGTACAAAATAAGAATTAGTTAAA
-129    T87_F   ATAAGATAATGTATATAAATACATTGGTTAATACCATATAAAATATTCTATTTGGAGTACAAAATAAGAATTAGTTAAA
-92     T41R    ATAAGATAATGTATATAAATACATTGGTTAATACCATATAAAATATTCTATTTGGAGTACAAAATAAGAATTAGTTAAA
-65     T151_F  ATAAGATAATGTATATAAATACATTGGTTAATACCATATAAAATATTCTATTTGGAGTACAAAATAAGAATTAGTTAAA
-72     T21R    ATAAGATAATGTATATAAATACATTGGTTAATACCATATAAAATATTCTATTTGGAGTACAAAATAAGAATTAGTTAAA
-79     T30F    ATAAGATAATGTATATAAATACATTGGTTAATACCATATAAAATATTCTATTTGGAGTACAAAATAAGAATTAGTTAAA
80      T30R    ATAAGATAATGTATATAAATACATTGGTTAATACCATATAAAATATTCTATTTGGAGTACAAAATAAGAATTAGTTAAA
-32     T10R                                    CATATAAAATATTCTATTTGGAGTACAAAATAAGAATTAGTTAAA

CONSENSUS       ATAAGATAATGTATATAAATACATTGGTTAATACCATATAAAATATTCTATTTGGAGTACAAAATAAGAATTAGTTAAA
```

```
                1820      1830      1840      1850      1860      1870      1880      1890
-83     T32F    TATTATTTTGAATTGATAACGTATTATTTACACTATATATTAATCAAGTCATTAATTATAAGTATAGATTCTACATTCT
-25     T106_F  TATTATTTTGAATTGATAACGTATTATTTACACTATATATTAATCAAGTCATTAATTATAAGTATAGATTCTACATTCT
-134    T96R    TATTATTTTGAATTGATAACGTATTATTTACACTATATATTAATCAAGTCATTAATTATAAGTATAGATTCTACATTCT
-15     T100_F  TATTATTTTGAATTGATAACGTATTATTTACACTATATATTAATCAAGTCATTAATTATAAGTATAGATTCTACATTCT
-16     T101R   TATTATTTTGAATTGATAACGTATTATTTACACTATATATTAATCAAGTCATTAATTATAAGTATAGATTCTACATTCT
-53     T128R   TATTATTTTGAATTGATAACGTATTATTTACACTATATATTAATCAAGTCATTAATTATAAGTATAGATTCTACATTCT
-123    T81R    TATTATTTTGAATTGATAACGTATTATTTACACTATATATTAATCAAGTCATTAATTATAAGTATAGATTCTACATTCT
-90     T40R    TATTATTTTGAATTGATAACGTATTATTTACACTATATATTAATCAAGTCATTAATTATAAGTATAGATTCTACATTCT
-125    T82R    TATTATTTTGAATTGATAACGTATTATTTACACTATATATTAATCAAGTCATTAATTATAAGTATAGATTCTACATTCT
-74     T22R    TATTATTTTGAATTGATAACGTATTATTTACACTATATATTAATCAAGTCATTAATTATAAGTATAGATTCTACATTCT
-102    T4F     TATTATTTTGAATTGATAACGTATTATTTACACTATATATTAATCAAGTCATTAATTATAAGTATAGATTCTACATTCT
-67     T18F    TATTATTTTGAATTGATAACGTATTATTTACACTATATATTAATCAAGTCATTAATTATAAGTATAGATTCTACATTCT
-29     T109R   TATTATTTTGAATTGATAACGTATTATTTACACTATATATTAATCAAGTCATTAATTATAAGTATAGATTCTACATTCT
-96     T46F    TATTATTTTGAATTGATAACGTATTATTTACACTATATATTAATCAAGTCATTAATTATAAGTATAGATTCTACATTCT
95      T44R    TATTATTTTGAATTGATAACGTATTATTTACACTATATATTAATCAAGTCATTAATTATAAGTATAGATTCTACATTCT
13      87R     TATTATTTTGAATTGATAACGTATTATTTACACTATATATTAATCAAGTCATTAATTATAAGTATAGATTCTACATTCT
-129    T87_F   TATTATTTTGAATTGATAACGTATTATTTACACTATATATTAATCAAGTCATTAATTATAAGTATAGATTCTACATTCT
-92     T41R    TATTATTTTGAATTGATAACGTATTATTTACACTATATATTAATCAAGTCATTAATTATAAGTATAGATTCTACATTCT
-65     T151_F  TATTATTTTGAATTGATAACGTATTATTTACACTATATATTAATCAAGTCATTAATTATAAGTATAGATTCTACATTCT
-72     T21R    TATTATTTTGAATTGATAACGTATTATTTACACTATATATTAATCAAGTCATTAATTATAAGTATAGATTCTACATTCT
-79     T30F    TATTATTTTGAATTGATAACGTATTATTTACACTATATATTAATCAAGTCATTAATTATAAGTATAGATTCTACATTCT
80      T30R    TATTATTTTGAATTGATAACGTATTATTTACACTATATATTAATCAAGTCATTAATTATAAGTATAGATTCTACATTCT
-32     T10R    TATTATTTTGAATTGATAACGTATTATTTACACTATATATTAATCAAGTCATTAATTATAAGTATAGATTCTACATTCT

CONSENSUS       TATTATTTTGAATTGATAACGTATTATTTACACTATATATTAATCAAGTCATTAATTATAAGTATAGATTCTACATTCT


                1900      1910      1920      1930      1940      1950      1960      1970
-83     T32F    ACATACTACACCTAAACAACTACAATTAAATACTTCAATTTGAAATATAAATACAATATTAAAATAGTCGTAAACTAGT
-25     T106_F  ACATACTACACCTAAACAACTACAATTAAATACTTCAATTTGAAATATAAATACAATATTAAAATAGTCGTAAACTAGT
-134    T96R    ACATACTACACCTAAACAACTACAATTAAATACTTCAATTTGAAATATAAATACAATATTAAAATAGTCGTAAACTAGT
-15     T100_F  ACATACTACACCTAAACAACTACAATTAAATACTTCAATTTGAAATATAAATACAATATTAAAATAGTCGTAAACTAGT
-16     T101R   ACATACTACACCTAAACAACTACAATTAAATACTTCAATTTGAAATATAAATACAATATTAAAATAGTCGTAAACTAGT
-53     T128R   ACATACTACACCTAAACAACTACAATTAAATACTTCAATTTGAAATATAAATACAATATTAAAATAGTCGTAAACTAGT
-123    T81R    ACATACTACACCTAAACAACTACAATTAAATACTTCAATTTGAAATATAAATACAATATTAAAATAGTCGTAAACTAGT
-90     T40R    ACATACTACACCTAAACAACTACAATTAAATACTTCAATTTGAAATATAAATACAATATTAAAATAGTCGTAAACTAGT
-125    T82R    ACATACTACACCTAAACAACTACAATTAAATACTTCAATTTGAAATATAAATACAATATTAAAATAGTCGTAAACTAGT
-74     T22R    ACATACTACACCTAAACAACTACAATTAAATACTTCAATTTGAAATATAAATACAATATTAAAATAGTCGTAAACTAGT
-102    T4F     ACATACTACACCTAAACAACTACAATTAAATACTTCAATTTGAAATATAAATACAATATTAAAATAGTCGTAAACTAGT
-67     T18F    ACATACTACACCTAAACAACTACAATTAAATACTTCAATTTGAAATATAAATACAATATTAAAATAGTCGTAAACTAGT
-29     T109R   ACATACTACACCTAAACAACTACAATTAAATACTTCAATTTGAAATATAAATACAATATTAAAATAGTCGTAAACTAGT
-96     T46F    ACATACTACACCTAAACAACTACAATTAAATACTTCAATTTGAAATATAAATACAATATTAAAATAGTCGTAAACTAGT
95      T44R    ACATACTACACCTAAACAACTACAATTAAATACTTCAATTTGAAATATAAATACAATATTAAAATAGTCGTAAACTAGT
13      87R     ACATACTACACCTAAACAACTACAATTAAATACTTCAATTTGAAATATAAATACAATATTAAAATAGTCGTAAACTAGT
-129    T87_F   ACATACTACACCTAAACAACTACAATTAAATACTTCAATTTGAAATATAAATACAATATTAAAATAGTCGTAAACTAGT
-92     T41R    ACATACTACACCTAAACAACTACAATTAAATACTTCAATTTGAAATATAAATACAATATTAAAATAGTCGTAAACTAGT
-65     T151_F  ACATACTACACCTAAACAACTACAATTAAATACTTCAATTTGAAATATAAATACAATATTAAAATAGTCGTAAACTAGT
-72     T21R    ACATACTACACCTAAACAACTACAATTAAATACTTCAATTTGAAATATAAATACAATATTAAAATAGTCGTAAACTAGT
-79     T30F    ACATACTACACCTAAACAACTACAATTAAATACTTCAATTTGAAATATAAATACAATATTAAAATAGTCGTAAACTAGT
80      T30R    ACATACTACACCTAAACAACTACAATTAAATACTTCAATTTGAAATATAAATACAATATTAAAATAGTCGTAAACTAGT
-32     T10R    ACATACTACACCTAAACAACTACAATTAAATACTTCAATTTGAAATATAAATACCATATTAAAATAGTCGTAAACTAGT

CONSENSUS       ACATACTACACCTAAACAACTACAATTAAATACTTCAATTTGAAATATAAATACAATATTAAAATAGTCGTAAACTAGT
```

```
                    1980       1990       2000       2010       2020       2030       2040       2050
-83      T32F    AGCTTATCACTTAAATTATAATACCAACCAAAATCATTAATTATTCTTTTTCACATATTACTTAAAGTAAAAAATAGAA
-25      T106_F  AGCTTATCACTTAAATTATAATACCAACCAAAATCATTAATTATTCTTTTTCACATATTACTTAAAGTAAAAAATAGAA
-134     T96R    AGCTTATCACTTAAATTATAATACCAACCAAAATCATTAATTATTCTTTTTCACATATTACTTAAAGTAAAAAATAGAA
-15      T100_F  AGCTTATCACTTAAATTATAATACCAACCAAAATCATTAATTATTCTTTTTCACATATTACTTAAAGTAAAAAATAGAA
-16      T101R   AGCTTATCACTTAAATTATAATACCAACCAAAATCATTAATTATTCTTTTTCACATATTACTTAAAGTAAAAAATAGAA
-53      T128R   AGCTTATCACTTAAATTATAATACCAACCAAAATCATTAATTATTCTTTTTCACATATTACTTAAAGTAAAAAATAGAA
-123     T81R    AGCTTATCACTTAAATTATAATACCAACCAAAATCATTAATTATTCTTTTTCACATATTACTTAAAGTAAAAAATAGAA
-90      T40R    AGCTTATCACTTAAATTATAATACCAACCAAAATCATTAATTATTCTTTTTCACATATTACTTAAAGTAAAAAATAGAA
-125     T82R    AGCTTATCACTTAAATTATAATACCAACCAAAATCATTAATTATTCTTTTTCACATATTACTTAAAGTAAAAAATAGAA
-74      T22R    AGCTTATCACTTAAATTATAATACCAACCAAAATCATTAATTATTCTTTTTCACATATTACTTAAAGTAAAAAATAGAA
-102     T4F     AGCTTATCACTTAAATTATAATACCAACCAAAATCATTAATTATTCTTTTTCACATATTACTTAAAGTAAAAAATAGAA
-67      T18F    AGCTTATCACTTAAATTATAATACCAACCAAAATCATTAATTATTCTTTTTCACATATTACTTAAAGTAAAAAATAGAA
-29      T109R   AGCTTATCACTTAAATTATAATACCAACCAAAATCATTAATTATTCTTTTTCACATATTACTTAAAGTAAAAAATAGAA
-96      T46F    AGCTTATCACTTAAATTATAATACCAACCAAAATCATTAATTATTCTTTTTCACATATTACTTAAAGTAAAAAATAGAA
95       T44R    AGCTTATCACTTAAATTATAATACCAACCAAAATCATTAATTATTCTTTTTCACATATTACTTAAAGTAAAAAATAGAA
13       87R     AGCTTATCACTTAAATTATAATACCAACCAAAATCATTAATTATTCTTTTTCACATATTACTTAAAGTAAAAAATAGAA
-129     T87_F   AGCTTATCACTTAAATTATAATACCAACCAAAATCATTAATTATTCTTTTTCACATATTACTTAAAGTAAAAAATAGAA
-92      T41R    AGCTTATCACTTAAATTATAATACCAACCAAAATCATTAATTATTCTTTTTCACATATTACTTAAAGTAAAAAATAGAA
-65      T151_F  AGCTTATCACTTAAATTATAATACCAACCAAAATCATTAATTATTCTCTTTCACATATTACTTAAAGTAAAAAATAGAA
-72      T21R    AGCTTATCACTTAAATTATAATACCAACCAAAATCATTAATTATTCTTTTTCACATATTACTTAAAGTAAAAAATAGAA
-79      T30F    AGCTTATCACTTAAATTATAATACCAACCAAAATCATTAATTATTCTTTTTCACATATTACTTAAAGTAAAAAATAGAA
80       T30R    AGCTTATCACTTAAATTATAATACCAACCAAAATCATTAATTATTCTTTTTCACATATTACTTAAAGTAAAAAATAGAA
-32      T10R    AGCTTATCACTTAAATTATAATACCAACCAAAATCATTAATTATTCTTTTTCACATATTACTTAAAGTAAAAAATAGAA
-75      T25F    AGCTTATCACTTAAATTATAATACCAACCAAAATCATTAATTATTCTTTTTCACATATTACTTAAAGTAAAAAATAGAA
76       T25R    AGCTTATCACTTAAATTATAATACCAACCAAAATCATTAATTATTCTTTTTCACATATTACTTAAAGTAAAAAATAGAA

CONSENSUS        AGCTTATCACTTAAATTATAATACCAACCAAAATCATTAATTATTCTTTTTCACATATTACTTAAAGTAAAAAATAGAA


                    2060       2070       2080       2090       2100       2110       2120       2130
-83      T32F    TATAAGATTAAAATATAACCCTAAGTTAGGAGTACTTTATCTACACTAGATCCTCGAGCTCCCTATAGTGAGTCGTATT
-25      T106_F  TATAAGATTAAAATATAACCCTAAGTTAGGAGTACTTTATCTACACTAGATCCTCGAGCTCCCTATAGTGAGTCGTATT
-134     T96R    TATAAGATTAAAATATAACCCTAAGTTAGGAGTACTTTATCTACACTAGATCCTCGAGCTCCCTATAGTGAGTCGTATT
-15      T100_F  TATAAGATTAAAATATAACCCTAAGTTAGGAGTACTTTATCTACACTAGATCCTCGAGCTCCCTATAGTGAGTCGTATT
-16      T101R   TATAAGATTAAAATATAACCCTAAGTTAGGAGTACTTTATCTACACTAGATCCTCGAGCTCCCTATAGTGAGTCGTATT
-53      T128R   TATAAGATTAAAATATAACCCTAAGTTAGGAGTACTTTATCTACACTAGATCCTCGAGCTCCCTATAGTGAGTCGTATT
-123     T81R    TATAAGATTAAAATATAACCCTAAGTTAGGAGTACTTTATCTACACTAGATCCTCGAGCTCCCTATAGTGAGTCGTATT
-90      T40R    TATAAGATTAAAATATAACCCTAAGTTAGGAGTACTTTATCTACACTAGATCCTCGAGCTCCCTATAGTGAGTCGTATT
-125     T82R    TATAAGATTAAAATATAACCCTAAGTTAGGAGTACTTTATCTACACTAGATCCTCGAGCTCCCTATAGTGAGTCGTATT
-74      T22R    TATAAGATTAAAATATAACCCTAAGTTAGGAGTACTTTATCTACACTAGATCCTCGAGCTCCCTATAGTGAGTCGTATT
-102     T4F     TATAAGATTAAAATATAACCCTAAGTTAGGAGTACTTTATCTACACTAGATCCTCGAGCTCCCTATAGTGAGTCGTATT
-67      T18F    TATAAGATTAAAATATAACCCTAAGTTAGGAGTACTTTATCTACACTAGATCCTCGAGCTCCCTATAGTGAGTCGTATT
-29      T109R   TATAAGATTAAAATATAACCCTAAGTTAGGAGTACTTTATCTACACTAGATCCTCGAGCTCCCTATAGTGAGTCATATT
-96      T46F    TATAAGATTAAAATATAACCCTAAGTTAGGAGTACTTTATCTACACTAGATCCTCGAGCTCCCTATAGTGAGTCGTATT
95       T44R    TATAAGATTAAAATATAACCCTAAGTTAGGAGTACTTTATCTACACTAGATCCTCGAGCTCCCTATAGTGAGTCGTATT
13       87R     TATAAGATTAAAATATAACCCTAAGTTAGGAGTACTTTATCTACACTAGATCCTCGAGCTCCCTATAGTGAGTCGTATT
-129     T87_F   TATAAGATTAAAATATAACCCTAAGTTAGGAGTACTTTATCTACACTAGATCCTCGAGCTCCCTATAGTGAGTCGTATT
-92      T41R    TATAAGATTAAAATATAACCCTAAGTTAGGAGTACTTTATCTACACTAGATCCTCGAGCTCCCTATAGTGAGTCGTATT
-65      T151_F  TATAAGATTAAAATATAACCCTAAGTTAGGAGTACTTTATCTACACTAGATCCTCGAGCTCCCTATAGTGAGTCGTATT
-72      T21R    TATAAGATTAAAATATAACCCTAAGTTAGGAGTACTTTATCTACACTAGATCCTCGAGCTCCCTATAGTGAGTCGTATT
-79      T30F    TATAAGATTAAAATATAACCCTAAGTTAGGAGTACTTTATCTACACTAGATCCTCGAGCTCCCTATAGTGAGTCGTATT
80       T30R    TATAAGATTAAAATATAACCCTAAGTTAGGAGTACTTTATCTACACTAGATCCTCGAGCTCCCTATAGTGAGTCGTATT
-32      T10R    TATAAGATTAAAATATAACCCTAAGTTAGGAGTACTTTATCTACACTAGATCCTCGAGCTCCCTATAGTGAGTCGTATT
-75      T25F    TATAAGATTAAAATATAACCCTAAGTTAGGAGTACTTTATCTACACTAGATCCTCGAGCTCCCTATAGTGAGTCGTATT
76       T25R    TATAAGATTAAAATATAACCCTAAGTTAGGAGTACTTTATCTACACTAGATCCTCGAGCTCCCTATAGTGAGTCGTATT

CONSENSUS        TATAAGATTAAAATATAACCCTAAGTTAGGAGTACTTTATCTACACTAGATCCTCGAGCTCCCTATAGTGAGTCGTATT
```

Addendum 3. DNA sequence of contig one screened against the sequences obtained from the random sequencing of the *C. ruminantium* λZAPII library. Column 1: Position where clone sequence begins. Column 2: Clone number and direction of primer used (F = forward, R = reverse)

```
                                    10        20        30        40        50        60
-1714   WL1F2385                    TCATACTCAACATTTATGTATTACTTAGAACAAATACTGACCTTATTATAAAATCTACTAAACA
-1713   WL1F2386                    TCATACTCAACATTTATGTGTTACTTAGAACAAATACTGACCTTATTATAAAATCTACTAAACA
9096    contig1                     TCATACTCAACATTTATGTATTACTTAGAACAAATACTGACCTTATTATAAAATCTACTAAACA
1712    cowdria-15g11.p1c             ATACTCAACATTTATGTATTACTTAGAACAAATACTGACCTTATTATAAAATCTACTAAACA

CONSENSUS                           TCATACTCAACATTTATGTATTACTTAGAACAAATACTGACCTTATTATAAAATCTACTAAACA


                                         70        80        90       100       110       120
-1714   WL1F2385                    GCACAAATCTATCTTAAGTCTACACAAAATACTATAATAAATATAGCTTATTACATACTTCATA
-1713   WL1F2386                    GCACAAATCTATCTTAAGTCTACACAAAATACTATAATAAATATAGCTTATTACATACTTCATA
9096    contig1                     GCACAAATCTATCTTAAGTCTACACAAAATACTATAATAAATATAGCTTATTACATACTTCATA
1712    cowdria-15g11.p1c           GCACAAATCTATCTTAAGTCTACACAAAATACTATAATAAATATAGCTTATTACATACTTCATA

CONSENSUS                           GCACAAATCTATCTTAAGTCTACACAAAATACTATAATAAATATAGCTTATTACATACTTCATA


                                    130       140       150       160       170       180       190
-1714   WL1F2385                    TATCTATACAGACAAAAACTGTAACTCCCAAATATTACACAGGCAATAATACATATTAAATTCA
-1713   WL1F2386                    TATCTATACAGACAAAAACTGTAACTCCCAAATATTACACAGGCAATAATACATATTAAATTCA
9096    contig1                     TATCTATACAGACAAAAACTGTAACTCCCAAATATTACACAGGCAATAATACATATTAAATTCA
1712    cowdria-15g11.p1c           TATCTATACAGACAAAAACTGTAACTCCCAAATATTACACAGGCAATAATACATATTAAATTCA

CONSENSUS                           TATCTATACAGACAAAAACTGTAACTCCCAAATATTACACAGGCAATAATACATATTAAATTCA


                                         200       210       220       230       240       250
-1714   WL1F2385                    ATTATTTTCAAGAAAGATATAGTTACACACTAAATTATTACAAATTCCTCCTGATATTTTGCTT
-1713   WL1F2386                    ATTATTTTCAAGAAAGATATAGTTACACACTAAATTATTACAAATTCCTCCTGATATTTTGCTT
9096    contig1                     ATTATTTTCAAGAAAGATATAGTTACACACTAAATTATTACAAATTCCTCCTGATATTTTGCTT
1712    cowdria-15g11.p1c           ATTATTTTCAAGAAAGATATAGTTACACACTAAATTATTACAAATTCCTCCTGATATTTTGCTT

CONSENSUS                           ATTATTTTCAAGAAAGATATAGTTACACACTAAATTATTACAAATTCCTCCTGATATTTTGCTT


                                      260       270       280       290       300       310
-1714   WL1F2385                    GCAGGTACATTTCTTTATCAATACATTTTCAAGATAAGTTGCAGGTAATGTTTGAAAATTTCGT
-1713   WL1F2386                    GCAGGTACATTTCTTTATCAATACATTTTCAAGATAAGTTGCAGGTAATGTTTGAAAATTTCGT
9096    contig1                     GCAGGTACATTTCTTTATCAATACATTTTCAAGATAAGTTGCAGGTAATGTTTGAAAATTTCGT
1712    cowdria-15g11.p1c           GCAGGTACATTTCTTTATCAATACATTTTCAAGATAAGTTGCAGGTAATGTTTGAAAATTTCGT

CONSENSUS                           GCAGGTACATTTCTTTATCAATACATTTTCAAGATAAGTTGCAGGTAATGTTTGAAAATTTCGT


                                    320       330       340       350       360       370       380
-1714   WL1F2385                    TCTTTCATACCTAAATACAAACATTGCAAAACAGAAGATC
-1713   WL1F2386                    TCTTTCATACCTAAATACAAACATTGCAAAACAGAAGATC
9096    contig1                     TCTTTCATACCTAAATACAAACATTGCAAAACAGAAGATCTTCGGCTACCTGAAAGAAGCGATG
1712    cowdria-15g11.p1c           TCTTTCATACCTAAATACAAACATTGCAAAACAGAAGATCTTCGGCTACCTGAAAGAAGCGATG
1711    cowdria-42f07.p1c                                                   CGGCTACCTGAAAGAAGCGATG

CONSENSUS                           TCTTTCATACCTAAATACAAACATTGCAAAACAGAAGATCTTCGGCTACCTGAAAGAAGCGATG


                                       390       400       410       420       430       440
9096    contig1                     TTAAAACTTTTATTTTCATTTATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATAT
1712    cowdria-15g11.p1c           TTAAAACTTTTATTTTCATTTATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATAT
1711    cowdria-42f07.p1c           TTAAAACTTTTATTTTCATTTATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATAT

CONSENSUS                           TTAAAACTTTTATTTTCATTTATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATAT


                                  450       460       470       480       490       500       510
9096    contig1                     TACCTTTTTAATAAAAATACAACTTCACATTTATGGAATCATTACAATAGAAAAATAAAAAACA
1712    cowdria-15g11.p1c           TACCTTTTTAATAAAAATACAACTTCACATTTATGGAATCATTACAATAGAAAAATAAAAAACA
1711    cowdria-42f07.p1c           TACCTTTTTAATAAAAATACAACTTCACATTTATGGAATCATTACAATAGAAAAATAAAAAACA

CONSENSUS                           TACCTTTTTAATAAAAATACAACTTCACATTTATGGAATCATTACAATAGAAAAATAAAAAACA
```

```
                                      520       530       540       550       560       570
9096    contig1                 TTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACATAACTATTTCACTACT
1712    cowdria-15g11.p1c       TTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACATAACTATTTCACTACT
1711    cowdria-42f07.p1c       TTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACATAACTATTTCACTACT

CONSENSUS                       TTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACATAACTATTTCACTACT


                                  580       590       600       610       620       630       640
9096    contig1                 AAATATTATGAGACATTCTAGATAACCTTAATATACAGATACCTAAAATAACCATAATGAAGGA
1712    cowdria-15g11.p1c       AAAT
1711    cowdria-42f07.p1c       AAATATTATGAGACATTCTAGATAACCTTAATATACAGATACCTAAAATAACCATAATGAAGGA
-1709   cowdria-68a08.p1c                                 AATATACAGATACTTA-AATAACCATAATGAAGG-
-1710   cowdria-14g11.p1c                                                 AAATAACCATAATGAAGG-

CONSENSUS                       AAATATTATGAGACATTCTAGATAACCTTAATATACAGATACCTAAAATAACCATAATGAAGGA


                                650       660       670       680       690       700       710
9096    contig1                 CAATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTCTCTT
1711    cowdria-42f07.p1c       CAATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTCTCTT
-1709   cowdria-68a08.p1c       CAATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTCTCTT
-1710   cowdria-14g11.p1c       CAATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTCTCTT

CONSENSUS                       CAATAACTGTCCCATAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTCTCTT


                                     720       730       740       750       760       770
9096    contig1                 ACAAATTCAATAAAAAAACGTACTATTCCATACCATATCATAAAAATAGAAAATAACATACCTT
1711    cowdria-42f07.p1c       ACAAATTCAATAAAAAAACGTACTATTCCATACCATATCATAAAAATAGAAAATAACATACCTT
-1709   cowdria-68a08.p1c       ACAAATTCAATAAAAAAACGTACTATTCCATACCATATCATAAAAATAGAAAATAACATACCTT
-1710   cowdria-14g11.p1c       ACAAATTCAATAAAAAAACGTACTATTCCAT-CCATATCATAAAAATAGAAAATAACATACCTT

CONSENSUS                       ACAAATTCAATAAAAAAACGTACTATTCCATACCATATCATAAAAATAGAAAATAACATACCTT


                                        780       790       800       810       820       830
9096    contig1                 GATATGATTTTACCTTAGTAAAAAAAAATAATAATTCATTACTACAAACAATAATAATCCTTCA
1711    cowdria-42f07.p1c       GATATGATTTTACCTTAGTAAAAAAAAATAATAATTCATTACTACAAACAATAATAATCCTTCA
-1709   cowdria-68a08.p1c       GATATGATTTTACCTTAGTAAAAAAAAATAATAATTCATTACTACAAACAATAATAATCCTTCA
-1710   cowdria-14g11.p1c       GATATGATTTTACCTTAGTAAAAAAAAATAATAATTCATTACTACAAACAATAATAATCCTTCA

CONSENSUS                       GATATGATTTTACCTTAGTAAAAAAAAATAATAATTCATTACTACAAACAATAATAATCCTTCA


                                    840       850       860       870       880       890
9096    contig1                 AAAAAAGCTTCATATAACTGACTAGGATGCCTATAAAAAAAATCTCCACTATTTTGAAAAATCA
1711    cowdria-42f07.p1c       AAAAAAGCTTCATATAACTGACTAGGATGCCTATAAAAAAAATCTCCACTATTTTGAAAAATCA
-1709   cowdria-68a08.p1c       AAAAAAGCTTCATATAACTGACTAGGATGCCTATAAAAAAAATCTCCACTATTTTGAAAAATCA
-1710   cowdria-14g11.p1c       AAAAAAGCTTCATATAACTGACTAGGATGCCTATAAAAAAAATCTCCACTATTTTGAAAAATCA

CONSENSUS                       AAAAAAGCTTCATATAACTGACTAGGATGCCTATAAAAAAAATCTCCACTATTTTGAAAAATCA


                                900       910       920       930       940       950       960
9096    contig1                 TTCCAAACCGTGTATTTGTAACTTTACCATATAACTCACCATTAATAAAATTCGCAATACGTCC
1711    cowdria-42f07.p1c       TTCCAAACCGTGTATTTGTAACTTTACCATA
-1709   cowdria-68a08.p1c       TTCCAAACCGTGTATTTGTAACTTTACCATATAACTCACCATTAATAAAATTCGCAATACGTCC
-1710   cowdria-14g11.p1c       TTCCAAACCGTGTATTTGTAACTTTACCATATAACTCACCATTAATAAAATTCGCAATACGTCC

CONSENSUS                       TTCCAAACCGTGTATTTGTAACTTTACCATATAACTCACCATTAATAAAATTCGCAATACGTCC
                                      970       980       990      1000      1010      1020
9096    contig1                 TAAAAATATTCCGACAGGTACAGCACACAAACACAAATCAATTGCTGATAAAAACTTTATTTTA
-1709   cowdria-68a08.p1c       TAAAAATATTCCGACAGGTACAGCACACAAACACAAATCAATTGCTGATAAAAACTTTATTTTA
-1710   cowdria-14g11.p1c       TAAAAATATTCCGACAGGTACAGCACACAAACACAAATCAATTGCTGATAAAAACTTTATTTTA

CONSENSUS                       TAAAAATATTCCGACAGGTACAGCACACAAACACAAATCAATTGCTGATAAAAACTTTATTTTA


                                  1030      1040      1050      1060      1070      1080
9096    contig1                 TACTTTTTACAAAAAATATACATAGTACAAAAA*AGCCCTAAACTTGCTCCATGAAATGACATA
-1709   cowdria-68a08.p1c       TACTTTTTACAAAAAATATACATAGTACAAAAA*AGCCCTAAACTTGCTCCATGAAATGACATA
-1710   cowdria-14g11.p1c       TACTTTTTACAAAAAATATACATAGTACAAAAAGAGCCCTAAACTTGCTCCATGAAATGACATA

CONSENSUS                       TACTTTTTACAAAAAATATACATAGTACAAAAA*AGCCCTAAACTTGCTCCATGAAATGACATA
```

```
                                1090      1100      1110      1120      1130      1140      1150
9096    contig1                 CCACCTTTCCATAACTTAAACATTTCAATAGGAAAACTCATATAGAAATTTAAGTTATAAAATA
-1709   cowdria-68a08.p1c       CCACCTTTCCATAACTTAAACATTTCAATAGGAAAACTCATATAGAAATTTAAGTTATAAAATA
-1710   cowdria-14g11.p1c       CCACCTTTCCATAACTTAAACATTTCAATAGGAAAACTCATATAGAAATTTAAGTTATAAAATA

CONSENSUS                       CCACCTTTCCATAACTTAAACATTTCAATAGGAAAACTCATATAGAAATTTAAGTTATAAAATA


                                     1160      1170      1180      1190      1200      1210
9096    contig1                 ATATATATCCTATCCTTCCACCAAGAATCATTCCAGTAACCCACCATGAAATTATCGACTTATA
-1709   cowdria-68a08.p1c       ATATATATCCTATCCTTCCACCAAGAATCATTCCAGTAACCCACCATGAAATTATCGACTTATA
-1710   cowdria-14g11.p1c       ATATATATCCTATCCTTCCACCAAGAATCATTCCAGTAACCCACCATGAAATTATCGACTTATA

CONSENSUS                       ATATATATCCTATCCTTCCACCAAGAATCATTCCAGTAACCCACCATGAAATTATCGACTTATA


                                 1220      1230      1240      1250      1260      1270      1280
9096    contig1                 ACTTTCTGGAGTAAAAACCTTGTACTTATCTATTTTCTGCACATACCAATAAGCAAATAAAATA
-1709   cowdria-68a08.p1c       ACTTTCTGGAGTAAAAACCTTGTACTTATCT*TTTTATGCACAGACCAATAAGCAAATAAAATA
-1710   cowdria-14g11.p1c       ACTTTCTGGAGTAAAAACCTTGTACTTATCTATTTTCTGCACATACCAATAAGCAAATAAAATA

CONSENSUS                       ACTTTCTGGAGTAAAAACCTTGTACTTATCTATTTTCTGCACATACCAATAAGCAAATAAAATA


                                        1290      1300      1310      1320      1330      1340
9096    contig1                 CCTATTATATATGCTAAAGAAT*ACCATCGGATCGATAACAACCCTATCTTTAAAGCCACTGGA
-1709   cowdria-68a08.p1c       CCTATTATATATGTTAAAGAATCACCAT
-1710   cowdria-14g11.p1c       CCTATTATATATGCTAAAGAAT*ACCATCGTATCGA

CONSENSUS                       CCTATTATATATGCTAAAGAAT*ACCATCGGATCGATAACAACCCTATCTTTAAAGCCACTGGA


                                   1350      1360      1370      1380      1390      1400
9096    contig1                 TCTATATTCATAAAATATAATTACCTCTTTAATAATTTTAATAAATAACTATATACTTCCATAA
CONSENSUS                       TCTATATTCATAAAATATAATTACCTCTTTAATAATTTTAATAAATAACTATATACTTCCATAA


                                1410      1420      1430      1440      1450      1460      1470
9096    contig1                 AAGCATTAATACAGTTTTTTCTATATTCTCAACCTATAAACAATAGAGAATTTGGCATCTTCTA
CONSENSUS                       AAGCATTAATACAGTTTTTTCTATATTCTCAACCTATAAACAATAGAGAATTTGGCATCTTCTA


                                     1480      1490      1500      1510      1520      1530
9096    contig1                 AAATTTAAAAACATATCTGGCAATACATATTCTTAAAAATAACGTCATTTCACATATCTATAAA
CONSENSUS                       AAATTTAAAAACATATCTGGCAATACATATTCTTAAAAATAACGTCATTTCACATATCTATAAA


                                 1540      1550      1560      1570      1580      1590      1600
9096    contig1                 AAGTAAAAATATCAACCAATAATAACCATGTTGATTTAGTACAATTTATTAAACTAATTCCATT
CONSENSUS                       AAGTAAAAATATCAACCAATAATAACCATGTTGATTTAGTACAATTTATTAAACTAATTCCATT


                                1610      1620      1630      1640      1650      1660      1670
9096    contig1                 TTAACAAAATGTTTAATGTTCTAATTTAGATTCACATAAGATTATACATACAAGCTTTCATTAA
CONSENSUS                       TTAACAAAATGTTTAATGTTCTAATTTAGATTCACATAAGATTATACATACAAGCTTTCATTAA


                                     1680      1690      1700      1710      1720      1730
9096    contig1                 CTATAAATAGAATAAAAAATATAACGTAATACCAAGTTGCGTTACACTTTATTGCACAGTCTAT
CONSENSUS                       CTATAAATAGAATAAAAAATATAACGTAATACCAAGTTGCGTTACACTTTATTGCACAGTCTAT

                                 1740      1750      1760      1770      1780      1790      1800
9096    contig1                 AATTTATCAATCATAAACAAATTAAAAAGAAAGTTTAGTATTAGATTAACAAAGCCTGTAGTGA
CONSENSUS                       AATTTATCAATCATAAACAAATTAAAAAGAAAGTTTAGTATTAGATTAACAAAGCCTGTAGTGA


                                     1810      1820      1830      1840      1850      1860
9096    contig1                 TTCTAACATATATTCACAATACAATATATGTATAATGTATCATTAAATACAAAAATAATAAGTA
CONSENSUS                       TTCTAACATATATTCACAATACAATATATGTATAATGTATCATTAAATACAAAAATAATAAGTA


                                1870      1880      1890      1900      1910      1920      1930
9096    contig1                 AAAACTCAATTTGACACACTCATTATTTAATATAATTAGATATATAACATATGTTTTCACTATA
CONSENSUS                       AAAACTCAATTTGACACACTCATTATTTAATATAATTAGATATATAACATATGTTTTCACTATA
```

```
                                1940      1950      1960      1970      1980      1990
9096    contig1                 CAAAATTATCAGTACTATTTTTAATATTAAATTGATACATAATAATTAGTTATGGTATTAGTAT
CONSENSUS                       CAAAATTATCAGTACTATTTTTAATATTAAATTGATACATAATAATTAGTTATGGTATTAGTAT

                                2000      2010      2020      2030      2040      2050      2060
9096    contig1                 GATAAAACATTAGGTTTTATTCTATCATAGAAACTATACAAACCTTAATATCCACAATCCTTTA
CONSENSUS                       GATAAAACATTAGGTTTTATTCTATCATAGAAACTATACAAACCTTAATATCCACAATCCTTTA

                                2070      2080      2090      2100      2110      2120
9096    contig1                 CTATAGAAATTTACATATTCTTAAAAGAAAATAGCAAGATATTATACATCTATTATCGCCAACA
CONSENSUS                       CTATAGAAATTTACATATTCTTAAAAGAAAATAGCAAGATATTATACATCTATTATCGCCAACA

                                2130      2140      2150      2160      2170      2180      2190
9096    contig1                 TCAAAACTTGTCTAGCATAAAACCTTTGTACCATTATTATATAAATAATCAACAACAAGGTACT
CONSENSUS                       TCAAAACTTGTCTAGCATAAAACCTTTGTACCATTATTATATAAATAATCAACAACAAGGTACT

                                2200      2210      2220      2230      2240      2250
9096    contig1                 ACCTAGTACCTACAAATTCTATTATAGATTCATCTTAAGCTATAATGTATTTATCAAGAAATAT
CONSENSUS                       ACCTAGTACCTACAAATTCTATTATAGATTCATCTTAAGCTATAATGTATTTATCAAGAAATAT

                                2260      2270      2280      2290      2300      2310
9096    contig1                 CCTATTTAATAATAGATATTAATCATCTCACACTTATGAAATTACACTTCACTAATTTCACAAT
CONSENSUS                       CCTATTTAATAATAGATATTAATCATCTCACACTTATGAAATTACACTTCACTAATTTCACAAT

                                2320      2330      2340      2350      2360      2370      2380
9096    contig1                 ATGTAGTTTTTAAATTAATATTTAAATCTCGAAAGAAATTTAGGTAATAAAATTTATTAGTTAT
CONSENSUS                       ATGTAGTTTTTAAATTAATATTTAAATCTCGAAAGAAATTTAGGTAATAAAATTTATTAGTTAT

                                2390      2400      2410      2420      2430      2440
9096    contig1                 ATTAAACATTTTATAGTAATAAATAGTATTTTTACTATTTATTATATGTCATACAACACATTAT
CONSENSUS                       ATTAAACATTTTATAGTAATAAATAGTATTTTTACTATTTATTATATGTCATACAACACATTAT

                                2450      2460      2470      2480      2490      2500      2510
9096    contig1                 TAAACTAACAAGTTTTGCCAAATCCTAATATGTTCTCCAAAAAGCACCACTAGCAACACTTATT
CONSENSUS                       TAAACTAACAAGTTTTGCCAAATCCTAATATGTTCTCCAAAAAGCACCACTAGCAACACTTATT

                                2520      2530      2540      2550      2560      2570
9096    contig1                 TTACTAATGATACTACCTAGCATTAATAAATTGATAATTGTAACATCTAGAACTAACTCTCATA
CONSENSUS                       TTACTAATGATACTACCTAGCATTAATAAATTGATAATTGTAACATCTAGAACTAACTCTCATA

                                2580      2590      2600      2610      2620      2630
9096    contig1                 TATCTTAAATATAAACAGTAAGACATTATCTATCACCCTAAATTCTATCATAATTACAGCATTT
CONSENSUS                       TATCTTAAATATAAACAGTAAGACATTATCTATCACCCTAAATTCTATCATAATTACAGCATTT

                                2640      2650      2660      2670      2680      2690      2700
9096    contig1                 GTTTTATTTTATATAAAGACACCACTTTTTATACTTTAATAAGTCCATAAAAGTATAATTAACA
CONSENSUS                       GTTTTATTTTATATAAAGACACCACTTTTTATACTTTAATAAGTCCATAAAAGTATAATTAACA

                                2710      2720      2730      2740      2750      2760
9096    contig1                 GAACACTTGTTTACATATTGCTTGCACACTTTATGGTATAAAACTTCATTAATTTTACAACCTG
CONSENSUS                       GAACACTTGTTTACATATTGCTTGCACACTTTATGGTATAAAACTTCATTAATTTTACAACCTG

                                2770      2780      2790      2800      2810      2820
9096    contig1                 TAGCACCTACATTAACTCTTCAATCTTGAAAAATTAGACAACAAACTACCACCAACTATACTAA
CONSENSUS                       TAGCACCTACATTAACTCTTCAATCTTGAAAAATTAGACAACAAACTACCACCAACTATACTAA

                                2830      2840      2850      2860      2870      2880      2890
9096    contig1                 ATGTTTTATAAACACAAATAACATTTTTTCTATTATTGGATATTACATCACATATCACTAAATT
CONSENSUS                       ATGTTTTATAAACACAAATAACATTTTTTCTATTATTGGATATTACATCACATATCACTAAATT
```

```
                                2900       2910       2920       2930       2940       2950
9096    contig1                 AATAAGTTTTTCCATAAACTTTAACGTGATGTTTCTCTAGATGTATTACTAGCAGTGCCTGTTA
CONSENSUS                       AATAAGTTTTTCCATAAACTTTAACGTGATGTTTCTCTAGATGTATTACTAGCAGTGCCTGTTA

                        2960       2970       2980       2990       3000       3010       3020
9096    contig1         CAACTGCTTCTTCTATGTCACTATCTAGCAGTTCAGTAATTTCACAACCTGTAGCACCTACATT
CONSENSUS               CAACTGCTTCTTCTATGTCACTATCTAGCAGTTCAGTAATTTCACAACCTGTAGCACCTACATT

                              3030       3040       3050       3060       3070       3080
9096    contig1               AACTCTTGAATCTTGAACAGCAGGATCAGGTAATAAAGTATCATCATCTACCCTAGGAACTGTA
CONSENSUS                     AACTCTTGAATCTTGAACAGCAGGATCAGGTAATAAAGTATCATCATCTACCCTAGGAACTGTA

                          3090       3100       3110       3120       3130       3140
9096    contig1           TCAATATCCCCAGTACTTTTTTCCTCTTGTAGATTATCTAATGCCTGCTGACCAAACACACCAG
CONSENSUS                 TCAATATCCCCAGTACTTTTTTCCTCTTGTAGATTATCTAATGCCTGCTGACCAAACACACCAG

                        3150    3160       3170       3180       3190       3200       3210
9096    contig1         CTTTACAGGCATTATAAATAGGCATATTAACACTTTCAACAAACATAAATTCATATTTGCCAAT
CONSENSUS               CTTTACAGGCATTATAAATAGGCATATTAACACTTTCAACAAACATAAATTCATATTTGCCAAT

                             3220       3230       3240       3250       3260       3270
9096    contig1              TTTGATATATTCAATGTCATCAAACATTTTTAGGTGATCTTTTATAGATAAGTAAGCTGTAGTA
CONSENSUS                    TTTGATATATTCAATGTCATCAAACATTTTTAGGTGATCTTTTATAGATAAGTAAGCTGTAGTA

                        3280       3290       3300       3310       3320       3330       3340
9096    contig1         CTAAGACCATCTAATTTTGATACTATTTTTCGATCTTTCCTACGCAGAAAATGACATATTAAAG
CONSENSUS               CTAAGACCATCTAATTTTGATACTATTTTTCGATCTTTCCTACGCAGAAAATGACATATTAAAG

                              3350       3360       3370       3380       3390       3400
9096    contig1               TAATAGCTATCATTGTTCCAGTTATTCCAGCTATTACACTAGATAATACAAACTCATGACCAAA
CONSENSUS                     TAATAGCTATCATTGTTCCAGTTATTCCAGCTATTACACTAGATAATACAAACTCATGACCAAA

                          3410       3420       3430       3440       3450       3460
9096    contig1           TTTAGCATCGCAACACAACATCAATACATTTGTGGCCATAAAAATACCTGCAAATCCAATAATA
CONSENSUS                 TTTAGCATCGCAACACAACATCAATACATTTGTGGCCATAAAAATACCTGCAAATCCAATAATA

                        3470       3480       3490       3500       3510       3520       3530
9096    contig1         AACGCTGCACAGACTCTTGATATTTTATTCTTATACGATAGTAAATACAATCCTCCTAGAAACA
CONSENSUS               AACGCTGCACAGACTCTTGATATTTTATTCTTATACGATAGTAAATACAATCCTCCTAGAAACA

                              3540       3550       3560       3570       3580       3590
9096    contig1               TTATAAAATACGATACAACAGAACAGTATAAAGCTATTCTATCAAATTTAAAAAAATGACATGA
CONSENSUS                     TTATAAAATACGATACAACAGAACAGTATAAAGCTATTCTATCAAATTTAAAAAAATGACATGA

                          3600       3610       3620       3630       3640       3650
9096    contig1           TAAAGCCACAAGTTTTAAAAGATACACAGATCCTAGTATTAGAAAAGTAGTACATCTTAACAGT
CONSENSUS                 TAAAGCCACAAGTTTTAAAAGATACACAGATCCTAGTATTAGAAAAGTAGTACATCTTAACAGT

                        3660    3670       3680       3690       3700       3710       3720
9096    contig1         AATTGCCTTTTTGTAAGCTCATTGTTACCTGCCATAAGAAACAAATTGTGTTTTGTACTTACTT
CONSENSUS               AATTGCCTTTTTGTAAGCTCATTGTTACCTGCCATAAGAAACAAATTGTGTTTTGTACTTACTT

                             3730       3740       3750       3760       3770       3780
9096    contig1              TCTTTCCTAACTCAGTAAGCTGTAAATTTCCTTTGCCATCATTGACCATAAAACTCTGTTTTTC
CONSENSUS                    TCTTTCCTAACTCAGTAAGCTGTAAATTTCCTTTGCCATCATTGACCATAAAACTCTGTTTTTC

                        3790       3800       3810       3820       3830       3840       3850
9096    contig1         AGCTTCCTGTTTTATATGTGATAAAATTCTCTCTCTTTTCGTAGAAGGAACATCATCATTTAGA
CONSENSUS               AGCTTCCTGTTTTATATGTGATAAAATTCTCTCTCTTTTCGTAGAAGGAACATCATCATTTAGA

                              3860       3870       3880       3890       3900       3910
9096    contig1               AGATCCTCAGAAACAGTACGTATACCTTGATTATATTCTCTATACTGCCATGAAAGTAAATTCA
CONSENSUS                     AGATCCTCAGAAACAGTACGTATACCTTGATTATATTCTCTATACTGCCATGAAAGTAAATTCA
```

```
                                3920       3930       3940       3950       3960       3970
9096    contig1                 TACTACTTGCTTCTTCTTGAACACTTCTATTCCCAACATAAGATAATAAAAGAACAATAGGAAT
CONSENSUS                       TACTACTTGCTTCTTCTTGAACACTTCTATTCCCAACATAAGATAATAAAAGAACAATAGGAAT

                                3980     3990       4000       4010       4020       4030       4040
9096    contig1                 AACAATAGTAGGAGCAAATACTACAGTAAAACAACATCTTAATATTTTCTCACGTATAGTTAAT
CONSENSUS                       AACAATAGTAGGAGCAAATACTACAGTAAAACAACATCTTAATATTTTCTCACGTATAGTTAAT

                                    4050       4060       4070       4080       4090       4100
9096    contig1                 AAAGACCACCTTTCAGCATATAACACAGAATCATATATTTCATGACTTAATGGTTTACTACTGA
CONSENSUS                       AAAGACCACCTTTCAGCATATAACACAGAATCATATATTTCATGACTTAATGGTTTACTACTGA

                                4110       4120       4130       4140       4150       4160       4170
9096    contig1                 CATCATATTTTTCAACTCGAAAAACTCGATCAGTACCAGCTTGTGTTTCTAACTCATCAACAAA
CONSENSUS                       CATCATATTTTTCAACTCGAAAAACTCGATCAGTACCAGCTTGTGTTTCTAACTCATCAACAAA

                                      4180       4190       4200       4210       4220       4230
9096    contig1                 ATCTCTTTGACTTTTTAAAATTATATCATATATATCACTTTCAGGAGTGTGCTCTTTAGCACCA
CONSENSUS                       ATCTCTTTGACTTTTTAAAATTATATCATATATATCACTTTCAGGAGTGTGCTCTTTAGCACCA

                                  4240       4250       4260       4270       4280       4290
9096    contig1                 TCACATGTAGAAATAATAAAAACAGTTTTAGGATTAGAATAAGCTTTACCTTCACGATAAGTAC
CONSENSUS                       TCACATGTAGAAATAATAAAAACAGTTTTAGGATTAGAATAAGCTTTACCTTCACGATAAGTAC

                                4300     4310       4320       4330       4340       4350       4360
9096    contig1                 ATTCAGAAATACCAATATCATAACTGGAATGCACTACTGGCAAAAACCCTAGCGAACGTATAAA
CONSENSUS                       ATTCAGAAATACCAATATCATAACTGGAATGCACTACTGGCAAAAACCCTAGCGAACGTATAAA

                                    4370       4380       4390       4400       4410       4420
9096    contig1                 GATACAGTCTTACGTATATCGTCTTTATTAGTAGAAGTATGTATTAATACCCTTTGTAAAAAAG
CONSENSUS                       GATACAGTCTTACGTATATCGTCTTTATTAGTAGAAGTATGTATTAATACCCTTTGTAAAAAAG

                                4430       4440       4450       4460       4470       4480       4490
9096    contig1                 ATTCATCTACCGCATCATCTTTATTTTGAATAACAAAAGCAGCAGGAGTAAGATTTCTTATTAC
CONSENSUS                       ATTCATCTACCGCATCATCTTTATTTTGAATAACAAAAGCAGCAGGAGTAAGATTTCTTATTAC

                                      4500       4510       4520       4530       4540       4550
9096    contig1                 CTTTCCTGAACAAAACATACCAATACTTGGATTATAAAATATGGTATATCTTAAATTGTCAAGC
CONSENSUS                       CTTTCCTGAACAAAACATACCAATACTTGGATTATAAAATATGGTATATCTTAAATTGTCAAGC

                                  4560       4570       4580       4590       4600       4610
9096    contig1                 CATTCCTTATAAAACTTATACTTAGCAGGATCATACATTAGTTTTTAACTTATTACCTATTTAC
CONSENSUS                       CATTCCTTATAAAACTTATACTTAGCAGGATCATACATTAGTTTTTAACTTATTACCTATTTAC

                                4620     4630       4640       4650       4660       4670       4680
9096    contig1                 AATCACAATACTACATAATTACTATTTTACATTAATTTTTTAAATATTATAAGTAAAAGTTTTA
CONSENSUS                       AATCACAATACTACATAATTACTATTTTACATTAATTTTTTAAATATTATAAGTAAAAGTTTTA

                                    4690       4700       4710       4720       4730       4740
9096    contig1                 CTACAATAAGGGCAAACTATTTCTTGACCATTTCTGATAGTCAAATATATTTTCGGATGTTCAG
CONSENSUS                       CTACAATAAGGGCAAACTATTTCTTGACCATTTCTGATAGTCAAATATATTTTCGGATGTTCAG

                                4750       4760       4770       4780       4790       4800       4810
9096    contig1                 TATAATCAACATCACTTCCTTCCCCATTACAAGATACTATTTGCTCTTCAAGTACATCTATTTT
CONSENSUS                       TATAATCAACATCACTTCCTTCCCCATTACAAGATACTATTTGCTCTTCAAGTACATCTATTTT

                                      4820       4830       4840       4850       4860       4870
9096    contig1                 TTTATCATGATTAGACATAAATATAATTTTTAATATTAATAAAACACAACTCTTATAACATCAT
CONSENSUS                       TTTATCATGATTAGACATAAATATAATTTTTAATATTAATAAAACACAACTCTTATAACATCAT
```

```
                                  4880      4890      4900      4910      4920      4930
9096    contig1                   AATAAATAACAAGCTACCTTTCTTACCATTCAAATATATATTATTATACACTTTGTATAATCAA
CONSENSUS                         AATAAATAACAAGCTACCTTTCTTACCATTCAAATATATATTATTATACACTTTGTATAATCAA

                                  4940    4950      4960      4970      4980      4990      5000
9096    contig1                   ATTATTATATTACTAGCAAATACATGTTTACTACATTAAACTTATTTACTACTATCAACATAGC
CONSENSUS                         ATTATTATATTACTAGCAAATACATGTTTACTACATTAAACTTATTTACTACTATCAACATAGC

                                      5010      5020      5030      5040      5050      5060
9096    contig1                   TTTATGAATAATTCTATAGACAAAAACTCCACTTTTTCTTTCCTTGATTAAAGAATATACCTTA
CONSENSUS                         TTTATGAATAATTCTATAGACAAAAACTCCACTTTTTCTTTCCTTGATTAAAGAATATACCTTA

                                  5070      5080      5090      5100      5110      5120      5130
9096    contig1                   GATAAGTCATCATAAAAGTTTCATTATAATATTAGATAGATTATCCTATGATAAAAATTTTACA
CONSENSUS                         GATAAGTCATCATAAAAGTTTCATTATAATATTAGATAGATTATCCTATGATAAAAATTTTACA

                                        5140      5150      5160      5170      5180      5190
9096    contig1                   TCAGAATTAAATAATTATTAACACTATTATACCCCAAATTACAATAAAATTCCTCAACAATTTT
CONSENSUS                         TCAGAATTAAATAATTATTAACACTATTATACCCCAAATTACAATAAAATTCCTCAACAATTTT

                                    5200      5210      5220      5230      5240      5250
9096    contig1                   CTGATAATTTAAAACTAGTACTACACATAACTGCTATCAAGACCTACAGATAAAAATTAAACCT
CONSENSUS                         CTGATAATTTAAAACTAGTACTACACATAACTGCTATCAAGACCTACAGATAAAAATTAAACCT

                                  5260    5270      5280      5290      5300      5310      5320
9096    contig1                   TTTTCTTTCTTACTATCATTTTGTCTATAGAATAGTACTTTCCCATTTGTTAGCATTATCTTTA
CONSENSUS                         TTTTCTTTCTTACTATCATTTTGTCTATAGAATAGTACTTTCCCATTTGTTAGCATTATCTTTA

                                      5330      5340      5350      5360      5370      5380
9096    contig1                   ACATATTGCTGAACATATACTAATTTACCTGTATACTCATAAATAAATTTCACTATATCATTAA
CONSENSUS                         ACATATTGCTGAACATATACTAATTTACCTGTATACTCATAAATAAATTTCACTATATCATTAA

                                  5390      5400      5410      5420      5430      5440      5450
9096    contig1                   TATTAGTATCAGTTTTATTGTTAACACTTAGCACAATCTTCACATCGATATCACCTAACACTTT
CONSENSUS                         TATTAGTATCAGTTTTATTGTTAACACTTAGCACAATCTTCACATCGATATCACCTAACACTTT

                                        5460      5470      5480      5490      5500      5510
9096    contig1                   AAGTACCGTCAATGTATGACTAATGCTCCCTAAGTAAGAACCAATAACTAAAATTACTTTTATT
CONSENSUS                         AAGTACCGTCAATGTATGACTAATGCTCCCTAAGTAAGAACCAATAACTAAAATTACTTTTATT

                                    5520      5530      5540      5550      5560      5570
9096    contig1                   TTTAAATCTTGTATTAATTCCAAGCAAGTCTTCTGATCAGTAATAGGA*G*ACA*TAACACCAC
3250    WL1C4443-T7                                               GATCAGTAATAGGA*G*ACA*TAACACCAC
3244    cowdria-29c05.p1c                                         GATCAGTAATAGGA*G*AAA*TAACACCAC
3249    WL1H113T7                                                 GATCAGTAATAGGA*G*ACA*TAACACCAC
3245    cowdria-81b10.p1c                                           TCAGTAATAGGAAG*AC**TAACACCAC
3247    cowdria-18c05.p1c                                           TCAGTAATAGGG*GTACA*TAACACCAC
3248    cowdria-26h04.p1c                                            CAGTAATAGGA*G*ACAATAACACCAC
3246    cowdria-26f04.p1c                                                                A*TAACACTCC

CONSENSUS                         TTTAAATCTTGTATTAATTCCAAGCAAGTCTTCTGATCAGTAATAGGA*G*ACA*TAACACCAC

                                  5580    5590      5600      5610      5620      5630      5640
9096    contig1                   C*AACACCTTCTATTAATAGATAATC*TATATCTT*G*ATTAATAT*GTGTA*TAACA*GAATT
3250    WL1C4443-T7               C*AACACCTTCTATTAATAGATAATC*TATATCTT*G*ATTAATAT*GTGTA*TAACA*GAATT
3244    cowdria-29c05.p1c         C*AACACCTTCTATTAATAGATAATC*TATATCTT*G*ATTAATAT*GTGTA*TAACA*GAATT
3249    WL1H113T7                 C*AACACCTTCTATTAATAGATAATC*TATATCTT*G*ATTAATAT*GTGTA*TAACA*GAATT
3245    cowdria-81b10.p1c         C*AACACCTTCTATTAATAGATAATC*TATATCTT*G*ATTAATAT*GTGTA*TAACA*GAATT
3247    cowdria-18c05.p1c         C*AACACCTTCTATTAATAGATAATC*TATATCTT*G*ATTAATAT*GTGTA*TAACA*GAATT
3248    cowdria-26h04.p1c         CCAACCCCTTCTATTAATAGATAATC*TATATCTTAG*AT*AATAT*GTGTA*TAACA*GAATT
3246    cowdria-26f04.p1c         C*TACACCTTGTATTAATAGGTAATCATTTTTCGT*GGATTAATCCAGTGCCCTGACAAGATTT

CONSENSUS                         C*AACACCTTCTATTAATAGATAATC*TATATCTT*G*ATTAATAT*GTGTA*TAACA*GAATT
```

```
                                    5650       5660       5670       5680       5690       5700
9096    contig1                     *TAAAAATTT*CATTATAATCTAATTTAAT*ATTTTCCA*GCCTAGC*AGCTATATTAGCAGCA
3250    WL1C4443-T7                 *TAAAAATTT*CATTATAATCTAATTTAAT*ATTTTCCA*GCCTAGC*AGCTATATTAGGAGCA
3244    cowdria-29c05.p1c           *TAAAAATTT*CATTATAATCTAATTTAAT*ATTTTCCA*GCCTAGC*AGCTATATTAGGAGCA
3249    WL1H113T7                   *TAAAAATGG*CATTATAATCTAATTTAAT*ATTTTCCA*GCCTAGC*AGCTATATTAGGAGCA
3245    cowdria-81b10.p1c           *TAAAAATTT*CATTATAATCTAATTTAAT*ATTTTCCA*GCCTAGC*AGCTATATTAGGAGCA
3247    cowdria-18c05.p1c           *TAAAAATTT*CATTATAATCTAATTTAAT*ATTTTCCA*GCCTAGC*AGCTATATTAGGAGCA
3248    cowdria-26h04.p1c           *TAAAAATTT*CATTATAATCTAATTTAAT*ATTTTCCA*GCCTAGC*AGCTATATTAGGAGCA
3246    cowdria-26f04.p1c           CTAAAAATTTGCATCATAATCTAATTTAATGATTTTCCAAGCCTGACTAGCTATATCACGAGCA

CONSENSUS                           *TAAAAATTT*CATTATAATCTAATTTAAT*ATTTTCCA*GCCTAGC*AGCTATATTAGGAGCA

                                    5710       5720       5730       5740       5750       5760       5770
9096    contig1                     TGAGGATAAGATAGCCTCCAAGGAGAAATTCTATTGATGTTATGAGTATTGCAATCCACATC*T
3250    WL1C4443-T7                 TGAGGATAAGATAGCCTCCAAGGAGAAATTCTATTGATGTTATGAGTATTGCAATCCACATC*T
3244    cowdria-29c05.p1c           TGAGGATAAGATAGCCTCCAAGGAGAAATTCTATTGATGTTATGAGTATTGCAATCCACATC*T
3249    WL1H113T7                   TGAGGATAAGATAGCCTCCAAGGAGAAATTCTATGGATGTTATGAGTATGGCAATCCACATC*T
3245    cowdria-81b10.p1c           TGAGGATAAGATAGCCTCCAAGGAGAAATTCTATTGATGTTATGAGTATTGCAATCCACATCAT
3247    cowdria-18c05.p1c           TGAGGATAAGATAGCCTCCAAGGAGAAATTCTATTGATGTTATGAGTATTGCAATCCACATC*T
3248    cowdria-26h04.p1c           TGAGGATAAGATAGCCTCCAAGGAGAAATTCTATTGATGTTATGAGTATTGCAATCCACATC*T
3246    cowdria-26f04.p1c           TGAGGATAAGATAGCCTGCAAGGAGAAAATCTATAGATGTTATGAGTATTGCAATCCACCTC*T

CONSENSUS                           TGAGGATAAGATAGCCTCCAAGGAGAAATTCTATTGATGTTATGAGTATTGCAATCCACATC*T

                                          5780       5790       5800       5810       5820       5830
9096    contig1                     AAACTACATAATATTTTGCTTGTATCATTACTCATAACTTCCACATCATTCCACCCACTTATTA
3250    WL1C4443-T7                 AAACTACATAATATTTTGCTTGTATCATTACTCATAATTTCCACATCATTCCACCCACTTATTA
3244    cowdria-29c05.p1c           AAACTACATAATATTTTGCTTGTATCATTACTCATAATTTCCACATCATTCCACCCACTTATTA
3249    WL1H113T7                   AAACTACATAATATTTGGCTGGTATCATTACTCATAATTTCCACATCATTCCACCCACTTATTA
3245    cowdria-81b10.p1c           AAACTACATAATATTTAGCTTGTATCATTACTCATAATTTCCACATCATTCCACCCACTTATTA
3247    cowdria-18c05.p1c           AAACTACATAATATTTTGCTTGTATCATTACTCATAATTTCCACATCATTCCACCCACTTATTA
3248    cowdria-26h04.p1c           AAACTACATAATATTTTGCTTGTATCATTACTCATAATTTCCACATCATTCCACCCACTTATTA
3246    cowdria-26f04.p1c           AAACTACATAATGTGTAGCTTGTATCATTACTCATAATTTCCACATCATCCCACCCACTTATTA

CONSENSUS                           AAACTACATAATATTTTGCTTGTATCATTACTCATAATTTCCACATCATTCCACCCACTTATTA

                                      5840       5850       5860       5870       5880       5890     5900
9096    contig1                     TAGGTTT*TATTGCATGTACTGTTTTACAATTTTTTCGTAAATGCCAGCACA*GTACTGTAGTA
3250    WL1C4443-T7                 TAGGTTT*TATTGCATGTACTGTTTTACAATTTTTTCGTAAATGCCAGCACA*GTACTGTAGTA
3244    cowdria-29c05.p1c           TAGGTTT*TATTGCATGTACTGTTTTACAATTTTTTCGTAAATGCCAGCCCA*GTACTGTAGTA
3249    WL1H113T7                   TAGGTTTATATGGCATGTACTGTTTTACAATTTTTTCGTAAATGCCAGCCCA*GTACTGTAGTA
3245    cowdria-81b10.p1c           TAGGTTT*TATTGCATGTACTGTTTTACAATTTTTTCGTAAATGCCAGCACAAGTACTGTAGTA
3247    cowdria-18c05.p1c           TAGGTTT*TATTGCATGTACTGTTTTACAATTTTTTCGTAAATGCCAGCACA*GTACTGTAGTA
3248    cowdria-26h04.p1c           TAGGTTT*TATTGCATGTACTGTTTTACAATTTTTTCGTAAATGCCAGCACA*GTACTGTAGTA
3246    cowdria-26f04.p1c           TAGGTTT*TATTGCATGTACTGCTTCACAATTTTTTCGTAAATGCCAGCACA*GCACTGTAGTA

CONSENSUS                           TAGGTTT*TATTGCATGTACTGTTTTACAATTTTTTCGTAAATGCCAGCACA*GTACTGTAGTA

                                            5910       5920       5930       5940       5950       5960
9096    contig1                     ATAAAAGTTTTACCTATACCAGTACCACA*TGATGTAAT*AAAATAAGCTGACATTT*AAT*AC
3250    WL1C4443-T7                 ATAAAAGTTTTACCTATACCAGTACCACA*TGATGTAAT*AAAATAAGCTGACATTT*AAT*AC
3244    cowdria-29c05.p1c           ATAAAAGTTTTACCTATACCAGTACCACA*TGATGTAAT*AAAATAAGCTGACATTT*AAT*AC
3249    WL1H113T7                   ATAAAAGTTTTACCTATACCAGTCCCACA*TGAGGTAAT*AAAATAAGCTGACATTT*AAT*AC
3245    cowdria-81b10.p1c           ATAAAAGTTTTACCTATACCAGTACCACA*TGATGTAATAAAAATAAGCTGACATTT*AAT*AC
3247    cowdria-18c05.p1c           ATAAAAGTTTTACCTATACCAGTACCACA*TGATGTAAT*AAAATAAGCTGACATTT*AAT*AC
3248    cowdria-26h04.p1c           ATAAAAGTTTTACCTATACCAGTACCACA*TGATGTAAT*AAAATAAGCTGACATTT*AAT*AC
3246    cowdria-26f04.p1c           ATAAAAGTTTTACCTATACCACTCCCACAGTGATGCCAC*AAAATATGTTGACATCTCAATCAC

CONSENSUS                           ATAAAAGTTTTACCTATACCAGTACCACA*TGATGTAAT*AAAATAAGCTGACATTT*AAT*AC

                                        5970       5980       5990       6000       6010       6020
9096    contig1                     AGTGAACCTTATAATTTACTTAATACTAACTGATAAACACTGAACTCACAACAAAACACTAAGG
3250    WL1C4443-T7                 AGTGAACCTTATAATTTACTTAATACTAACTGATAAACACTGAACTCACAACAAAACACTAAGG
3244    cowdria-29c05.p1c           AGTGAACCTTATAATTTACTTAATACTAACTGATAAACACTGAACTCACAACAAAACACTAAGG
3249    WL1H113T7                   AGGGAACCTTATAATTAACTTAATACTAACTGATAAACACTGAACTCACAACAAAACACTAAGG
3245    cowdria-81b10.p1c           AGTGAACCTTATAATTTACTTAATACTAACTGATAAACACTGAACTCACAACAAAACACTAAGG
3247    cowdria-18c05.p1c           AGTGAACCTTATAATTTACTTAATACTAACTGATAAACACTGAACTCACAACAAAACACTAAGG
3248    cowdria-26h04.p1c           AGTGAACCTTATAATTTACTTAATACTAACTGATAAACACTGAACTCACAACAAAACACTAAGG
3246    cowdria-26f04.p1c           AGCGAACCTGGTAGTG*ACTTAATACTAACTGA*AAGCACTGAGCTCATA*CAAA

CONSENSUS                           AGTGAACCTTATAATTTACTTAATACTAACTGATAAACACTGAACTCACAACAAAACACTAAGG
```

```
                                6030       6040       6050       6060       6070       6080       6090
9096    contig1                 AAAATTGAGCTATGACAGATAACCAA*GGATTTTTCAGCAATATCAAAAAAAGTCTATTTAAAA
3250    WL1C4443-T7             AAAATTGAGCTATGACAGATAACCAA*GGATTTTTCAGCAATATCAAAAAAAGTCTATTTAAAA
3244    cowdria-29c05.p1c       AAAATTGAGCTATGACAGATAACCAA*GGATTTTTCAGCAATATCAAAAAAAGTCTATTTAAAA
3249    WL1H113T7               AAAATTGAGCTATGACAGATAACCAA*GGATTTTTCAGCAATATCAAAAAAAGTCTATTTAAAA
3245    cowdria-81b10.p1c       AAAATTGAGCTATGACAAATAACCAA*GGATTTTTCAGCAATATCAAAAAAAGTCTATTTAAAA
3247    cowdria-18c05.p1c       AAAATTGAGCTATGACAGATAACCAA*GGATTTTTCAGCAATATCAAAAAAAGTCTATTTAAAA
3248    cowdria-26h04.p1c       AAAATTGAGCTATGACAGATAACCAAGGGATTTTTCAGCAATATCAAAAAAAGTCTATTTAAAA

CONSENSUS                       AAAATTGAGCTATGACAGATAACCAA*GGATTTTTCAGCAATATCAAAAAAAGTCTATTTAAAA


                                      6100       6110       6120       6130       6140       6150
9096    contig1                 CTTCATCAAAATTAAGTGATGGAATTAAAAAAATTTTTTCCAATAGTAAAAAAATCAATCAAGA
3250    WL1C4443-T7             CTTCATCAAAATTAAGTGATGGAATTAAAAAAATTTTTTCCAATAGTAAAAAAATCAATCAAGA
3244    cowdria-29c05.p1c       CTTCATCAAAATTAAGTGATGGAATTAAAAAAATTT
3249    WL1H113T7               CTTCATCAAAATTAAGTGAGGGAATTAAAAAAATTTTTTCCAATGGTAAAAAAATCAATCAAGA
3245    cowdria-81b10.p1c       CTTCATCAAAATTAAGTGATGGAATTAAAAAAATTTTTTCCAATAGTAAAAAAATCAATCAAGA
3247    cowdria-18c05.p1c       CTTCA-*AAAATTAAGTGATGGGATTAAAAAAATTTTTTCCAATAGTAAAAAAATCAATCAAGA
3248    cowdria-26h04.p1c       CTTCAT*AAAATTAAGTGATGGAAT*AAAAAAATTTTTTCCAATAGTAAAAAAATCAATCAAG*

CONSENSUS                       CTTCATCAAAATTAAGTGATGGAATTAAAAAAATTTTTTCCAATAGTAAAAAAATCAATCAAGA


                                  6160       6170       6180       6190       6200       6210
9096    contig1                 AACTTTAGAAGAATTAAAGGAATTAC*TGATTACAGCAGATATTGGATATGAAAATGCATCATT
3250    WL1C4443-T7             AACTTTATAAGAATTAAAGGAATTAC*TGATTACAGCAGATATTGGATATGAAAATGCATCATT
3249    WL1H113T7               CACTTTAGAAGAATTAAGGGAATTCCGTG*TTACAGCAGATATTGGATATGAAAATGCTTCATT
3245    cowdria-81b10.p1c       AACTTTAAAAGAAT*AAAGGAATTAC*TGATTACAGCAGATATTGGATATGAAAATGCATCATT
3247    cowdria-18c05.p1c       AACTTTAGAAGAAT*AAAGGGACTAC*TGATTACAGCAGATA
3248    cowdria-26h04.p1c       AACTTTAGAAGAA*TAAAGGAATTAC*TGAT*ACAGCA

CONSENSUS                       AACTTTAGAAGAATTAAAGGAATTAC*TGATTACAGCAGATATTGGATATGAAAATGCATCATT


                                6220     6230       6240       6250       6260       6270       6280
9096    contig1                 ATTAATAAAAAAAAATTGCAGAAGCAAAATTTGATGAAGTAACTGATCATACAATAAAGCAAAAA
3250    WL1C4443-T7             ATTAATAAAAAAAAATTGCAGAAGCAAAATTTGATGAAGTAACTGA
3249    WL1H113T7               ATTAATAAAAAAAAATTGCAGA*GCAAAATT*GATGAAGTCACTTGTCTTAC*ATAGAGGCAAAT
3245    cowdria-81b10.p1c       ATTAATAAAAAAAATCGCAGAAGCAAAATTTGATGAAGTAACTGATCATACAATAAAGCAAAAA
-3252   WL1P27                                                          GATCATACAATAAAGCAAAAA

CONSENSUS                       ATTAATAAAAAAAAATTGCAGAAGCAAAATTTGATGAAGTAACTGATCATACAATAAAGCAAAAA


                                    6290       6300       6310       6320       6330       6340
9096    contig1                 CTAGCAGAAGAGATAGAAAATATTTTATTACAAGT*TGAAAAACCT*TTTTCTATA*ATAAAAA
3249    WL1H113T7               CTGGCAGAAGAGATAGAATATTTTTTTTCACAAGG*TGAAAAACCT*TTTTATATA*ATAAAAA
3245    cowdria-81b10.p1c       CTAGCAGAAGAGATAGAA
3252    WL1P27                  CTAGCAGAAGAGATAGAAAATATTTTATTACAAGT*TGAAAAACCT*TTTTCTATA*ATAAAAA
-3251   WL1H113T3                         AGATATAAAATATTTTTTTACCAGTGTAAAAAACCTGTTTTCTATATATAAAAA

CONSENSUS                       CTAGCAGAAGAGATAGAAAATATTTTATTACAAGT*TGAAAAACCT*TTTTCTATA*ATAAAAA


                                6350       6360       6370       6380       6390       6400       6410
9096    contig1                 AA*CCACATGTAATTATGATATGTGGAACTAATG*GAAACGGTA*AAACAACAACAGTAGGTAA
3249    WL1H113T7               AAACCACTTGTTATTATTATATGTGGAACTAATG*AAAACGGAA*AAACACCCCCGG-GGGAAA
-3252   WL1P27                  AA*CCACATGTAATTATGATATGTGGAACTAATG*GAAACGGTA*AAACAACAACAGTAGGTAA
-3251   WL1H113T3               AC*CCCTGTGTAATTTTG*TATGTGGAACTTATGTGAAACGGTGTAAACCACCACCGTAGGTCA
-3253   cowdria-24b10.p1c                        TATGTGGAACTAATG*GAAACGGTA*AAACAACAACAGTAGGTAA
-3254   cowdria-28c05.p1c                           TGGAACTAATG*GAAACGGTA*AAACAACAACAGTAGGTAA

CONSENSUS                       AA*CCACATGTAATTATGATATGTGGAACTAATG*GAAACGGTA*AAACAACAACAGTAGGTAA


                                      6420       6430       6440       6450       6460       6470
9096    contig1                 ATTAGCATATAAATTTAAGAATAATGGAAAAAG*TGTATTAG*TTG*CAGCGTGTGACACATTT
3249 WL1H113T7                  ATTCGCATATAATTTAAAAAATAATGAAAAAAGGTGT*TTA-TTTG*GAGGGTGGGACACATTT
-3252WL1P27                     ATTAGCATATAAATTTAAGAATAATGGAAAAAG*TGTATTAG*TTG*CAGCGTGTGACACATTT
-3251WL1H113T3                  ATTAGCATATTAATTTAAGGATTATTGAAAAAG*TGTATTAA*TTCACGGGGTGTCCCCCTTTT
-3253cowdria-24b10.p1c          ATTAGCATATAAATTTAAGAATAATGGAAAAAG*TGTATTAG*TTG*CAGCGTGTGACACATTT
-3254cowdria-28c05.p1c          ATTAGCATATAAATTTAAGAATAATGGAAAAAG*TGTATTAG*TTG*CAGCGTGTGACACATTT

CONSENSUS                       ATTAGCATATAAATTTAAGAATAATGGAAAAAG*TGTATTAG*TTG*CAGCGTGTGACACATTT
```

```
                                   6480       6490       6500       6510       6520       6530
9096    contig1                    CGTGCAGCA*GCTACAGAACAACTTACAGTATGGT*CACAAAAAG*TAGATTTTCCCATTG*TA
3249    WL1H113T7                  TGGGCAG*A*GTTCCAAAAAATTTTACAGTTTGGT*CCCCAAAAGGTAATTTTTC
-3252   WL1P27                     CGTGCAGCA*GCTACAGAACAACTTACAGTATGGT*CACAAAAAG*TAGATTTTCCCATTG*TA
-3251   WL1H113T3                  *GTTCAGCAAGCCACA*AACCACTTACAGTATGGTTCCCAAAAAG*AAG*TTTTCCCCTTGGTA
-3253   cowdria-24b10.p1c          CGTGCAGCA*GCTACAGAACAACTTACAGTATGGT*CACAAAAAG*TAGATTTTCCCATTG*TA
-3254   cowdria-28c05.p1c          CGTGCAGCA*GCTACAGAACAACTTACAGTATGGT*CACAAAAAG*TAGATTTTCCCATTG*TA

CONSENSUS                          CGTGCAGCA*GCTACAGAACAACTTACAGTATGGT*CACAAAAAG*TAGATTTTCCCATTG*TA

                                   6540     6550       6560       6570       6580       6590       6600
9096    contig1                    ACAAGTACACAAGGATCAGATGCAGCAAGTGTAGCATATCAAGCCATGC*AGCAAGCTTTAAAA
-3252   WL1P27                     ACAAGTACACAAGGATCAGATGCAGCAAGTGTAGCATATCAAGCCATGC*AGCAAGCTTTAAAA
-3251   WL1H113T3                  CCAGGCACCCA*GGATCCGATGCCGCCAGTGTAGCATATCCAACCATTCCACCAGGTTTTAAAA
-3253   cowdria-24b10.p1c          ACAAGTACACAAGGATCAGATGCAGCAAGTGTAGCATATCAAGCCATGC*AGCAAGCTTTAAAA
-3254   cowdria-28c05.p1c          ACAAGTACACAAGGATCAGATGCAGCAAGTGTAGCATATCAAGCCATGC*AGCAAGCTTTAAAA

CONSENSUS                          ACAAGTACACAAGGATCAGATGCAGCAAGTGTAGCATATCAAGCCATGC*AGCAAGCTTTAAAA

                                       6610       6620       6630       6640       6650       6660
9096    contig1                    AATGAAATAGATATTTTACTTGTTGATACTGCAGGAAGATTACACAACTATAAAAATCTAATGG
-3252   WL1P27                     AATGAAATAGATATTTTACTTGTTGATACTGCAGGAAGATTACACAACTATAAAAATCTAATGG
-3251   WL1H113T3                  ATTAAAATAAA*ATTTTACTTGGTGAAGCCGCAGGAAGGTTTCCCCACTATTAAAATCTAATGG
-3253   cowdria-24b10.p1c          AATGAAATAGATATTTTACTTGTTGATACTGCAGGAAGATTACACAACTATAAAAATCTAATGG
-3254   cowdria-28c05.p1c          AATGAAATAGATATTTTACTTGTTGATACTGCAGGAAGATTACACAACTATAAAAATCTAATGG

CONSENSUS                          AATGAAATAGATATTTTACTTGTTGATACTGCAGGAAGATTACACAACTATAAAAATCTAATGG

                                   6670       6680       6690       6700       6710       6720       6730
9096    contig1                    AAGAATTAGCAAAAATTAAACGAATTATAGGTAAACACGATAATGAAGCACCACATGATGTAGT
-3252   WL1P27                     AAGAATTAGCAAAAATTAAACGAATTATAGGTAAACACGATAATGAAGCACCACATGATGTAGT
-3251   WL1H113T3                  AAGAATTAGCAAAAATTAAACGAATTATAGGTAAACACGATAATGAAGCACCACATGATGTAGG
-3253   cowdria-24b10.p1c          AAGAATTAGCAAAAATTAAACGAATTATAGGTAAACACGATAATGAAGCACCACATGATGTAGT
-3254   cowdria-28c05.p1c          AAGAATTAGCAAAAATTAAACGAATTATAGGTAAACACGATAATGAAGCACCACATGATGTAGT

CONSENSUS                          AAGAATTAGCAAAAATTAAACGAATTATAGGTAAACACGATAATGAAGCACCACATGATGTAGT

                                       6740       6750       6760       6770       6780       6790
9096    contig1                    ATTAATACTAGACGCAACAACTGGGCAAAACGCTCTTAATCAAGTCGAAGTTTTTTTACAATTT
-3252   WL1P27                     ATTAATACTAGACGCAACAACTGGGCAAAACGCTCTTAATCAAGTCGAAGTTTTTTTACAATTT
-3251   WL1H113T3                  ATTAAATCTTAACGCCACCACTGGGCCAAACGCTTTTTATCAAGTTCAAGGTTTTTTTCCATTT
-3253   cowdria-24b10.p1c          ATTAATACTAGACGCAACAACTGGGCAAAACGCTCTTAATCAAGTCGAAGTTTTTTTACAATTT
-3254   cowdria-28c05.p1c          ATTAATACTAGACGCAACAACTGGGCAAAACGCTCTTAATCAAGTCGAAGTTTTTTTACAATTT

CONSENSUS                          ATTAATACTAGACGCAACAACTGGGCAAAACGCTCTTAATCAAGTCGAAGTTTTTTTACAATTT

                                   6800       6810       6820       6830       6840       6850
9096    contig1                    GTAAACATTAGTG*GACTTG*TTATCACAAAACTGGATGGT*ACAGCAAAAG*GTGGAGTAGTA
-3252   WL1P27                     GTAAACATTAGTG*GACTTG*TTATCACAAAACTGGATGGT*ACAGCAAAAG*GTGGAGTAGTA
-3251   WL1H113T3                  GGAAACCTTAATGTGACCTGGTTCTC*CAAAACTGGATGGGCACGCCAAAAGTGTG*AGTAGTG
-3253   cowdria-24b10.p1c          GTAAACATTAGTG*GACTTG*TTATCACAAAACTGGATGGT*ACAGCAAAAG*GTGGAGTAGTA
-3254   cowdria-28c05.p1c          GTAAACATTAGTG*GACTTG*TTATCACAAAACTGGATGGT*ACAGCAAAAG*GTGGAGTAGTA

CONSENSUS                          GTAAACATTAGTG*GACTTG*TTATCACAAAACTGGATGGT*ACAGCAAAAG*GTGGAGTAGTA

                                   6860     6870       6880       6890       6900       6910       6920
9096    contig1                    ATAAGAATAGCACAAAAATATAAATTGAATATTCATGCAATAGGAATAGGAGAACAGGTAGAAG
-3252   WL1P27                     ATAAGAATAGCACAAAAATATAAATTGAATATTCATGCAATAGGAATAGGAGAACAGGTAGAAG
-3251   WL1H113T3                  ATATAAATAGCGCACAAATATATATTGTATATTCACGCACTATAAATAGAAGAACAGGTATAAG
-3253   cowdria-24b10.p1c          ATAAGAATAGCACAAAAATATAAATTGAATATTCATGCAATAGGAATAGGAGAACAG*TAGA
-3254   cowdria-28c05.p1c          ATAAGAATAGCACAAAAATATAAATTGAATATTCATGCAATAGGAATAGGAGAACAGGTAG

CONSENSUS                          ATAAGAATAGCACAAAAATATAAATTGAATATTCATGCAATAGGAATAGGAGAACAGGTAGAAG
```

```
                                    6930       6940       6950       6960       6970       6980
9096    contig1                     ATCTTAAAGATTTCTCCGCTAAAGAGTTCACTGCTGGACTTTTAAATATGGATAACATATTGTA
779     cowdria-61b06.p1c            TCTTAAAGATTTCTCCGCTAAA*AGTTCACTGCTGGACTTTTAAATATGGATAACATATTGTA
780     cowdria-44g05.p1c            TCTTAAAGA*TTCTCCGCTAAAGAGTTCACTGCTGGACTTTTAAATATGGATAACATATTGTA
781     cowdria-47e11.p1c                     TTCTCCGCTAAAGA*TTCACTGCTGGACTTTTAAATATGGATAACATATTGTA

CONSENSUS                           ATCTTAAAGATTTCTCCGCTAAAGAGTTCACTGCTGGACTTTTAAATATGGATAACATATTGTA


                                    6990       7000       7010       7020       7030       7040       7050
9096    contig1                     ATTTAATAGCAACTTTGCTTGTAAAGCAGTATTCTTCTTGCATTTACTGTTTAGTTTAATATTA
779     cowdria-61b06.p1c           ATTTAATAGCAACTTTGCTTGTAAAGCAGTATTCTTCTTGCATTTACTGTTTAGTTTAATATTA
780     cowdria-44g05.p1c           ATTTAATAGCAACTTTGCTTGTAAAGCAGTATTCTTCTTGCATTTACTGTTTAGTTTAATATTA
781     cowdria-47e11.p1c           ATTTAATAGCAACTTTGCTTGTAAAGCAGTATTCTTCTTGCATTTACTGTTTAGTTTAATATTA

CONSENSUS                           ATTTAATAGCAACTTTGCTTGTAAAGCAGTATTCTTCTTGCATTTACTGTTTAGTTTAATATTA


                                          7060       7070       7080       7090       7100       7110
9096    contig1                     ACTCATTGCTATTAGACATTTATTTTTTCCTAATTCAATTGATGTAATATAAAAATTTCTCGTT
779     cowdria-61b06.p1c           ACTCATTGCTATTAGACATTTATTTTTTCCTAATTCAATTGATGTAATATAAAAATTTCTCGTT
780     cowdria-44g05.p1c           ACTCATTGCTATTAGACATTTATTTTTTCCTAATTCAATTGATGTAATATAAAAATTTCTCGTT
781     cowdria-47e11.p1c           ACTCATTGCTATTAGACATTTATTTTTTCCTAATTCAATTGATGTAATATAAAAATTTCTCGTT

CONSENSUS                           ACTCATTGCTATTAGACATTTATTTTTTCCTAATTCAATTGATGTAATATAAAAATTTCTCGTT


                                      7120       7130       7140       7150       7160       7170
9096    contig1                     GCTACTAATTTTTAGTTCTCTTGTAACAAAAATTCTACTGAGTTGATACAATCATAATATAACA
779     cowdria-61b06.p1c           GCTACTAATTTTTAGTTCTCTTGTAACAAAAATTCTACTGAGTTGATACAATCATAATATAACA
780     cowdria-44g05.p1c           GCTACTAATTTTTAGTTCTCTTGTAACAAAAATTCTACTGAGTTGATACAATCATAATATAACA
781     cowdria-47e11.p1c           GCTACTAATTTTTAGTTCTCTTGTAACAAAAATTCTACTGAGTTGATACAATCATAATATAACA

CONSENSUS                           GCTACTAATTTTTAGTTCTCTTGTAACAAAAATTCTACTGAGTTGATACAATCATAATATAACA


                                    7180       7190       7200       7210       7220       7230       7240
9096    contig1                     ATAAGTTTCTATTCATACAGAAAAAATTTTTTTCTGTACACTACTTAAAAATAAAAAAAAGTAC
779     cowdria-61b06.p1c           ATAAGTTTCTATTCATACAGAAAAAATTTTTTTCTGTACACTACTTAAAAATAAAAAAAAGTAC
780     cowdria-44g05.p1c           ATAAGTTTCTATTCATACAGAAAAAATTTTTTTCTGTACACTACTTAAAAATAAAAAAAAGTAC
781     cowdria-47e11.p1c           ATAAGTTTCTATTCATACAGAAAAAATTTTTTTCTGTACACTACTTAAAAATAAAAAAAAGTAC

CONSENSUS                           ATAAGTTTCTATTCATACAGAAAAAATTTTTTTCTGTACACTACTTAAAAATAAAAAAAAGTAC


                                       7250       7260       7270       7280       7290       7300
9096    contig1                     TATACAAGCAATACTAAAAACACATCAGTTTATATCATAGTAAGTAAATAAGATAGAAACTTTG
779     cowdria-61b06.p1c           TATACAAGCAATACTAAAAACACATCAGTTTATATCATAGTAAGTAAATAAGATAGAAACTTTG
780     cowdria-44g05.p1c           TATACAAGCAATACTAAAAACACATCAGTTTATATCATAGTAAGTAAATAAGATAGAAACTTTG
781     cowdria-47e11.p1c           TATACAAGCAATACTAAAAACACATCAGTTTATATCATAGTAAGTAAATAAGATAGAAACTTTG

CONSENSUS                           TATACAAGCAATACTAAAAACACATCAGTTTATATCATAGTAAGTAAATAAGATAGAAACTTTG


                                    7310       7320       7330       7340       7350       7360       7370
9096    contig1                     TAATATTTGCAATAAAATCAATAATATATTATGTTAAACTAACATACAGGAAATATTACGTTTT
779     cowdria-61b06.p1c           TAATATTTGCAATAAAATCAATAATATATTATGTTAAACTAACATACAGGAAATATTACGTTTT
780     cowdria-44g05.p1c           TAATATTTGCAATAAAATCAATAATATATTATGTTAAACTAACATACAGGAAATATTACGTTTT
781     cowdria-47e11.p1c           TAATATTTGCAATAAAATCAATAATATATTATGTTAAACTAACATACAGGAAATATTACGTTTT

CONSENSUS                           TAATATTTGCAATAAAATCAATAATATATTATGTTAAACTAACATACAGGAAATATTACGTTTT


                                          7380       7390       7400       7410       7420       7430
9096    contig1                     CACTCTTTTAAAAAAGAAATTCATATTCATTTATACTA*TAAGTAGTTTAGAGTATTTACATGA
779     cowdria-61b06.p1c           CACTCTTTTAAAAAAGAAATTCATATTCATTTATACTA*TAAGTAGTTTAGAGTATTTACATGA
780     cowdria-44g05.p1c           CACTCTTTTAAAAAAGAAATTCATATTCATTTATACTA*TAAGTAGTTTAGAGTATTTACATGA
781     cowdria-47e11.p1c           CACTCTTTTTAAAAAGAAAT*CATATTCATTT*TACTA-TAAG*AGTT*AGAGT*TTTACATGA

CONSENSUS                           CACTCTTTTAAAAAAGAAATTCATATTCATTTATACTA*TAAGTAGTTTAGAGTATTTACATGA
```

```
                                   7440       7450       7460       7470       7480       7490     7500
9096    contig1                    AAGATCATGATATAATAGAGTATATTATCAAGTAATCTATTAAATCTGGTTGAATATATAATAA
779     cowdria-61b06.p1c          AAGATCATGATATAATAGAGTATATTATCAAGTAATCTATTAAATCTGGTTGAATATATAATAA
780     cowdria-44g05.p1c          AAGATCATGATATAATAGAGTATATTATCAAGTAATCTATTAAATCTGGTTGAATATATAATAA
781     cowdria-47e11.p1c          AAGATCATG*TATA*TAGAGT*TATT*TCAAGGAATCCATTAA*TCTGGT*GAATATATAAT*A

CONSENSUS                          AAGATCATGATATAATAGAGTATATTATCAAGTAATCTATTAAATCTGGTTGAATATATAATAA


                                          7510       7520       7530       7540       7550       7560
9096    contig1                    TATTTTCAGAATATTAGATACCTACACTATTAGTTTATCATTTAACATCCTAACAAAATTAATT
779     cowdria-61b06.p1c          TATTTTCAGAATATTAGATACCTACACTATTAGTTTATCATTTAACATCCTAACAAAATTAATT
780     cowdria-44g05.p1c          TATTTTCAGAATATTAGATACCT
781     cowdria-47e11.p1c          TATTTTAA*AA*ATTAGATAC

CONSENSUS                          TATTTTCAGAATATTAGATACCTACACTATTAGTTTATCATTTAACATCCTAACAAAATTAATT


                                      7570       7580       7590       7600       7610       7620
9096    contig1                    TTAACATTGAATCTATGATATTTTTACATATTAAGTCTTAATTTACACAACATGTAAAACACAT
779     cowdria-61b06.p1c          TTAACATTGAA

CONSENSUS                          TTAACATTGAATCTATGATATTTTTACATATTAAGTCTTAATTTACACAACATGTAAAACACAT


                                   7630       7640       7650       7660       7670       7680     7690
9096    contig1                    TTTGTATAAATTAAGAAAGTACATAAATTTAGGTAACTATATTAAAATCTCATTATAACAAACA
CONSENSUS                          TTTGTATAAATTAAGAAAGTACATAAATTTAGGTAACTATATTAAAATCTCATTATAACAAACA


                                         7700       7710       7720       7730       7740       7750
9096    contig1                    ATAACCTAGCGATTATTATATATACCTACACACAAAATAGGTAACACCAGTTTTTTAGTAGATA
CONSENSUS                          ATAACCTAGCGATTATTATATATACCTACACACAAAATAGGTAACACCAGTTTTTTAGTAGATA


                                      7760       7770       7780       7790       7800       7810
9096    contig1                    TGGGAAATCTACTTTATAAATTCAGAATTTAAGTAAAACTGTTATATTTAAGTAACATGTAGTA
CONSENSUS                          TGGGAAATCTACTTTATAAATTCAGAATTTAAGTAAAACTGTTATATTTAAGTAACATGTAGTA


                                   7820    7830       7840       7850       7860       7870     7880
9096    contig1                    AACTAATAAACATAATCACATTAAAATACAGTTATAATATATCATCTATAAATATTTGATACCA
CONSENSUS                          AACTAATAAACATAATCACATTAAAATACAGTTATAATATATCATCTATAAATATTTGATACCA


                                      7890       7900       7910       7920       7930       7940
9096    contig1                    ATAAGTATTGCTAACTATTATTTGAAAAATATCAGGCATTTTTGTAAATACAATAATCCATAAA
CONSENSUS                          ATAAGTATTGCTAACTATTATTTGAAAAATATCAGGCATTTTTGTAAATACAATAATCCATAAA


                                   7950       7960       7970       7980       7990       8000     8010
9096    contig1                    TTAAATCCAATTTAAATAAACAACAAACTTATACGAAATTAAAAATAATTATATTTGTAATCTA
CONSENSUS                          TTAAATCCAATTTAAATAAACAACAAACTTATACGAAATTAAAAATAATTATATTTGTAATCTA


                                         8020       8030       8040       8050       8060       8070
9096    contig1                    TAAACACTCACATACAAATCAGTAACACAATCCTATTTATCTTGTTTACATAAGATATAAAGCT
CONSENSUS                          TAAACACTCACATACAAATCAGTAACACAATCCTATTTATCTTGTTTACATAAGATATAAAGCT


                                      8080       8090       8100       8110       8120       8130
9096    contig1                    TGTTTTGTAAGTTAAAAAAACTTATAAGTGAAGGAGTAGTTATATTAATCAAGTATTCCATTAC
CONSENSUS                          TGTTTTGTAAGTTAAAAAAACTTATAAGTGAAGGAGTAGTTATATTAATCAAGTATTCCATTAC


                                   8140    8150       8160       8170       8180       8190     8200
9096    contig1                    TCATTTATTCATAATGACTCTAGAAAATCTCCTACATTACAAGCAAGATATATAGAAAAATATA
CONSENSUS                          TCATTTATTCATAATGACTCTAGAAAATCTCCTACATTACAAGCAAGATATATAGAAAAATATA


                                      8210       8220       8230       8240       8250       8260
9096    contig1                    AACACCCAGCTAAATTTTTTTCTTATTCACAATGAAAACCTTTAACTTATTTCAAAATTAAAAA
CONSENSUS                          AACACCCAGCTAAATTTTTTTCTTATTCACAATGAAAACCTTTAACTTATTTCAAAATTAAAAA
```

```
                                8270      8280      8290      8300      8310      8320      8330
9096    contig1                 CTATTAAAGTATACTCCTACAATAATGTAGCACTAATCTCTTATCCATGAAAATTCTATCACCA
CONSENSUS                       CTATTAAAGTATACTCCTACAATAATGTAGCACTAATCTCTTATCCATGAAAATTCTATCACCA

                                      8340      8350      8360      8370      8380      8390
9096    contig1                 TTTTACAAACTTAAACACTTACACCAATGCCACACAACACTAAATATTAATCAAATATTCTTCA
CONSENSUS                       TTTTACAAACTTAAACACTTACACCAATGCCACACAACACTAAATATTAATCAAATATTCTTCA

                                  8400      8410      8420      8430      8440      8450
9096    contig1                 ATTAACTTATTCATAATGACTCTAGAAAATCTCCTACATTACAAGCAAGATATATAGAAAAATA
CONSENSUS                       ATTAACTTATTCATAATGACTCTAGAAAATCTCCTACATTACAAGCAAGATATATAGAAAAATA

                                8460    8470      8480      8490      8500      8510      8520
9096    contig1                 TCAACACCCAGCTAAATTTTTTTTCTTATTCACAATGAAAACTTTTAACTTATTTCAAAATTAA
CONSENSUS                       TCAACACCCAGCTAAATTTTTTTTCTTATTCACAATGAAAACTTTTAACTTATTTCAAAATTAA

                                    8530      8540      8550      8560      8570      8580
9096    contig1                 AAACTATTAATTCTTATATAATATAAAAACTACTCTTTCTACTAATATCCTATACAAACACTTT
CONSENSUS                       AAACTATTAATTCTTATATAATATAAAAACTACTCTTTCTACTAATATCCTATACAAACACTTT

                                8590      8600      8610      8620      8630      8640      8650
9096    contig1                 AGTATTTTATAGTACTACGTAAAAAATCCAGGTTGATCAACATCAGGCTACTTTATCTTTACTA
CONSENSUS                       AGTATTTTATAGTACTACGTAAAAAATCCAGGTTGATCAACATCAGGCTACTTTATCTTTACTA

                                      8660      8670      8680      8690      8700      8710
9096    contig1                 TACAACGTGTACTAATGATAAATAGCAACAAACACAATGTTAGATTATATTTATCAGTAACATA
CONSENSUS                       TACAACGTGTACTAATGATAAATAGCAACAAACACAATGTTAGATTATATTTATCAGTAACATA

                                  8720      8730      8740      8750      8760
9096    contig1                 GCACTTCTTATAAAAAATCTCTTAAAATACATAAGCATACCATATC
CONSENSUS                       GCACTTCTTATAAAAAATCTCTTAAAATACATAAGCATACCATATC
```

Addendum 4. DNA sequence of contig two screened against the sequences obtained from the random sequencing of the *C. ruminantium* λZAPII library. Column 1: Position where clone sequence begins. Column 2: Clone number and direction of primer used (F = forward, R = reverse)

```
                                        10        20        30        40        50        60
-5315   cowdria-50a05.p1c       TCACAAATCATTTAGCAAATTATTTCATCCAACACTT*GAAAGCAGTACAAACACTGACATTAT
5316    cowdria-30b11.p1c        CACAAATCATTTAGCAAATTATTTCATCCAACACTT-GAAAGCAGTACAAACACTGACATTAT
-5314   cowdria-1h3.p1c           ACAAATCATTTAGCGAATTATTTCATCCAACACTT*AGAAGCAGTGCAAACACTGACATTAT
-5317   cowdria-30d11.p1c                                            AGCAGTACAAACACTGACATTAT

CONSENSUS                       TCACAAATCATTTAGCAAATTATTTCATCCAACACTT*GAAAGCAGTACAAACACTGACATTAT
                                     70        80        90       100       110       120
-5315   cowdria-50a05.p1c       TGCATATGGGCTACCTGCATCTCCAGGAGCAGCAACCGGTCAAGTAGTTTTTACACCCAGTGAT
5316    cowdria-30b11.p1c       TGCATATGGGCTACCTGCATCTCCAGGAGCAGCAACCGGTCAAGTAGTTTTTACACCCAGTGAT
-5314   cowdria-1h3.p1c         TGCATATGGGCTACCTGCATCTCCAGGAGCAGCAACCGGTCAAGTAGTTTTTACACCCAGTGAT
-5317   cowdria-30d11.p1c       TGCATATGGGCTACCTGCATCTCCAGGAGCAGCAACCGGTCAAGTAGTTTTTACACCCAGTGAT

CONSENSUS                       TGCATATGGGCTACCTGCATCTCCAGGAGCAGCAACCGGTCAAGTAGTTTTTACACCCAGTGAT


                                130       140       150       160       170       180       190
-5315   cowdria-50a05.p1c       *GCAGAACAACTTAAAAAAGAAGGAAAAAATGTAATATTAATACGTCAAGAAACCAACCCAGAA
5316    cowdria-30b11.p1c       *GCAGAACAACTTAAAAAAGAAGGAAAAAATGTAATATTAATACGTCAAGAAACCAACCCAGAA
-5314   cowdria-1h3.p1c         AGCAGAACAACTTAAAAAAAAAGGAAAAAATGTAATATTAATACGTCAAGAAACCAACCCAGAA
-5317   cowdria-30d11.p1c       *GCAGAACAACTTAAAAAAGAAGGAAAAAATGTAATATTAATACGTCAAGAAACCAACCCAGAA
-5318   cowdria-30b03.p1c                                       ATACGTCAAGAAACCAACCCAGAAGATATAAA
-5319   cowdria-36d11.p1c                                       ATACGTCAAGAAACCAACCCAGAAGATATAAA

CONSENSUS                       *GCAGAACAACTTAAAAAAGAAGGAAAAAATGTAATATTAATACGTCAAGAAACCAACCCAGAA


                                     200       210       220       230       240       250
-5315   cowdria-50a05.p1c       GATATAAATGGGATGAATAGTGCTGTTGGTATTATCACCTTAAGAGGAGGAATGACTTCCCATG
5316    cowdria-30b11.p1c       GATATAAATGGGATGAATAGTGCTGTTGGTATTATCACCTTAAGAGGAGGAATGACTTCCCATG
-5314   cowdria-1h3.p1c         GATATAAATGGGATGAATAGTGCTGTTGGTATTATCACCTTAAGAGGAGGAATGACTTCCCATG
-5317   cowdria-30d11.p1c       GATATAAATGGGATGAATAGTGCTGTTGGTATTATCACCTTAAGAGGAGGAATGACTTCCCATG
-5318   cowdria-30b03.p1c               TGGGATGAATAGTGCTGTTGGTATTATCACCTTAAGAGGAGGAATGACTTCCCATG
-5319   cowdria-36d11.p1c               TGGGATGAATAGTGCTGTTGGTATTATCACCTTAAGAGGAGGAATGACTTCCCATG

CONSENSUS                       GATATAAATGGGATGAATAGTGCTGTTGGTATTATCACCTTAAGAGGAGGAATGACTTCCCATG


                                  260       270       280       290       300       310
-5315   cowdria-50a05.p1c       CAGCAGTAGTTGCTAGAGGCATGGGTAAGCCATGCATTTGTAGTGTTAATAACATTTTTATAGA
5316    cowdria-30b11.p1c       CAGCAGTAGTTGCTAGAGGCATGGGTAAGCCATGCATTTGTAGTGTTAATAACATTTTTATAGA
-5314   cowdria-1h3.p1c         CAGCAGTAGTTGCTAGAGGCATGGGTAAGCCATGCATTTGTAGTGTTAATAACATTTTTATAGA
-5317   cowdria-30d11.p1c       CAGCAGTAGTTGCTAGAGGCATGGGTAAGCCATGCATTTGTAGTGTTAATAACATTTTTATAGA
-5318   cowdria-30b03.p1c       CAGCAGTAGTTGCTAGAGGCATGGGTAAGCCATGCATTTGTAGTGTTAATAACATTTTTATAGA
-5319   cowdria-36d11.p1c       CAGCAGTAGTTGCTAGAGGCATGGGTAAGCCATGCATTTGTAGTGTTAATAACATTTTTATAGA
-5313   cowdria-1h4.p1c                                 GTAAACCACGCATCTGGTGTGTTAATCACCATATTATATA

CONSENSUS                       CAGCAGTAGTTGCTAGAGGCATGGGTAAGCCATGCATTTGTAGTGTTAATAACATTTTTATAGA


                                320       330       340       350       360       370       380
-5315   cowdria-50a05.p1c       TAAAAGC*GAACAATTCTTTTATACAAATACAGGAATAAAAGTATATAAAGGTGATAATATCAC
5316    cowdria-30b11.p1c       TAAAAGC*GAACAATTCTTTTATACAAATACAGGAATAAAAGTATATAAAGGTGATAATATCAC
-5314   cowdria-1h3.p1c         TAAAAGC*GAACAATTCTTTTATACAAATACAGGAATAAAAGTATATAAAGGTGATAATATCAC
-5317   cowdria-30d11.p1c       TAAAAGC*GAACAATTCTTTTATACAAATACAGGAATAAAAGTATATAAAGGTGATAATATCAC
-5318   cowdria-30b03.p1c       TAAAAGC*GAACAATTCTTTTATACAAATACAGGAATAAAAGTATATAAAGGTGATAATATCAC
-5319   cowdria-36d11.p1c       TAAAAGC-GAACAATTCTTTTATACAAATACAGGAATAAAAGTATATAAAGGTGATAATATCAC
-5313   cowdria-1h4.p1c         TAAAAGC*GAGCAATTCTTTTATACAAATACAGGAATACAAGTAAATAAAGGTGATAATATCAC

CONSENSUS                       TAAAAGC*GAACAATTCTTTTATACAAATACAGGAATAAAAGTATATAAAGGTGATAATATCAC
```

```
                                    390       400       410       420       430       440
-5315   cowdria-50a05.p1c    AATAAA*TGGATGCAATGGAGAA*GTTATTTTAGGTGTTATTAAGACAACACTACCAAAATTAG
5316    cowdria-30b11.p1c    AATAAA*TGGATGCAATGGAGAA*GTTATTTTAGGTGTTATTAAGACAACACTACCAAAATTAG
-5314   cowdria-1h3.p1c      AATAAA*TGGATGCAATGGAGAA*GTTATTTTAGGTGTTATTAAGACAACACTACCAGAATTAG
-5317   cowdria-30d11.p1c    AATAAA*TGGATGCAATGGAGAA*GTTATTTTAGGTGTTATTAAGACAACACTACCAAAATTAG
-5318   cowdria-30b03.p1c    AATAAA*TGGATGCAATGGAGAA*GTTATTTTAGGTGTTATTAAGACAACACTACCAAAATTAG
-5319   cowdria-36d11.p1c    AATAAA*TGGATGCAATGGAGAA*GTTATTTTAGGTGTTATTAAGACAACACTACCAAAATTAG
-5313   cowdria-1h4.p1c      AATAGAGTGGATGCAATGGAGAAAGTT

CONSENSUS                   AATAAA*TGGATGCAATGGAGAA*GTTATTTTAGGTGTTATTAAGACAACACTACCAAAATTAG


                                    450       460       470       480       490       500       510
-5315   cowdria-50a05.p1c    ATGAAAGTTTTTATGATC
5316    cowdria-30b11.p1c    ATGAAAGTTTTTATGATCTAATGGAATGGGTAGACGAAATAAGAACATTAAAAGTTATGGCTAA
-5314   cowdria-1h3.p1c      ATGAAAGTTTTTATGATC
-5317   cowdria-30d11.p1c    ATGAAAGTTTTTATGATCTAATGGAATGGGTAGACGAAATAAGAACATTAAAAGTTATGGCTAA
-5318   cowdria-30b03.p1c    ATGAAAGTTTTTATGATCTAATGGAATGGGTAGACGAAATAAGAACATTAAAAGTTATGGCTAA
-5319   cowdria-36d11.p1c    ATGAAAGTTTTTATGATCTAATGGAATGGGTAGACGAAATAAGAACATTAAAAGTTATGGCTAA
-5312   WL1.C61.T7                      GATCTAATGGAATGGGTAGACGAAATAAGAACATTAAAAGTTATGGCTAA
5320    WL1.C61.WL1F                    GATCTAATGGAATGGGTAGACGAAATAAGAACATTAAAAGTTATGGCTAA

CONSENSUS                   ATGAAAGTTTTTATGATCTAATGGAATGGGTAGACGAAATAAGAACATTAAAAGTTATGGCTAA


                                         520       530       540       550       560       570
5316    cowdria-30b11.p1c    TGCTGATACACCAGAAGATGCAGAAATATCAATGAATTTCAAAGCAGATGGTATAGGTTTATGC
-5317   cowdria-30d11.p1c    TGCTGATACACCAGAAGATGCAGAAATATCAATGAATTTCAAAGCAGATGGTATAGGTTTATGC
-5318   cowdria-30b03.p1c    TGCTGATACACCAGAAGATGCAGAAATATCAATGAATTTCAAAGCAGATGGTATAGGTTTATGC
-5319   cowdria-36d11.p1c    TGCTGATACACCAGAAGATGCAGAAATATCAATGAATTTCAAAGCAGATGGTATAGGTTTATGC
-5312   WL1.C61.T7           TGCTGATACACCAGAAGATGCAGAAATATCAATGAATTTCAAAGCAGATGGTATAGGTTTATGC
5320    WL1.C61.WL1F         TGCTGATACACCAGAAGATGCAGAAATATCAATGAATTTCAAAGCAGATGGTATAGGTTTATGC

CONSENSUS                   TGCTGATACACCAGAAGATGCAGAAATATCAATGAATTTCAAAGCAGATGGTATAGGTTTATGC


                                      580       590       600       610       620       630       640
5316    cowdria-30b11.p1c    AGAACAGA
-5317   cowdria-30d11.p1c    AGAACAGAACATATGTTTTTTTCAGACAAGAGAATAAGCATAGTACAAGAAAT
-5318   cowdria-30b03.p1c    AGAACAGAACATATGTTTTTTTCAGACAAGAGAATAAGCATAGTACAAGAAATGATCGTATCAG
-5319   cowdria-36d11.p1c    AGAACAGAACATATGTTTTTTTCAGACAAGAGAATAAGCATAGTACAAGAAATGATCGTATCAG
-5312   WL1.C61.T7           AGAACAGAACATATGTTTTTTTCAGACAAGAGAATAAGCATAGTACAAGAAATGATCGTATCAG
5320    WL1.C61.WL1F         AGAACAGAACATATGTTTTTTTCAGACAAGAGAATAAGCATAGTACAAGAAATGATCGTATCAG
5321    WL1J39WF                                                                GATCGTATCAG

CONSENSUS                   AGAACAGAACATATGTTTTTTTCAGACAAGAGAATAAGCATAGTACAAGAAATGATCGTATCAG


                                    650       660       670       680       690       700       710
-5318   cowdria-30b03.p1c    ATAAGAAAGAAGAAAGAGCAGTAGCATTAGAAAAACTAGAAATTATGCAAAAAGAAGACTTCAA
-5319   cowdria-36d11.p1c    ATAAGAAAGAAGAAAGAGCAGTAGCATTAGAAAAACTAGAAATTATGCTAAAAGAAGACTTCAA
-5312   WL1.C61.T7           ATAAGAAAGAAGAAAGAGCAGTAGCATTAGAAAAACTAGAAATTATGCAAAAAGAAGACTTCAA
5320    WL1.C61.WL1F         ATAAGAAAGAAGAAAGAGCAGTAGCATTAGAAAAACTAGAAATTATGCAAAAAGAAGACTTCAA
5321    WL1J39WF             ATAAGAAAGAAGAAAGAGCAGTAGCATTAGAAAAACTAGAAATTATGCAAAAAGAAGACTTCAA

CONSENSUS                   ATAAGAAAGAAGAAAGAGCAGTAGCATTAGAAAAACTAGAAATTATGCAAAAAGAAGACTTCAA


                                         720       730       740       750       760       770
-5318   cowdria-30b03.p1c    GAAAATATTTACACATACATTAGACAAACAGGTAACTATTAGGTTACT
-5319   cowdria-36d11.p1c    GAAAATATTTACACATGCATTAGACAAACAGGTAACT*TTAGGTT
-5312   WL1.C61.T7           GAAAATATTTACACATACATTAGACAAACAGGTAACTATTAGGTTACTCGA
5320    WL1.C61.WL1F         GAAAATATTTACACATACATTAGACAAACAGGTAACTATTAGGTTACTCGA
5321    WL1J39WF             GAAAATATTTACACATACATTAGACAAACAGGTAACTATTAGGTTACTCGATCCTCCACTACAC
5323    cowdria-13c12.p1c                                                CTCGATCCT*CACTA*AC
5322    cowdria-4c10.p1c                                                  TCGATCCTCCACTACGC

CONSENSUS                   GAAAATATTTACACATACATTAGACAAACAGGTAACTATTAGGTTACTCGATCCTCCACTACAC
```

```
                                        780       790       800       810       820       830
5321    WL1J39WF                GAGTTCTTACCAGATAATGATGATGCTATACAAGAAATATCATCCAGAACAGG*AAAATCCTTA
5323    cowdria-13c12.p1c       GAGTTCTTACACGATAATGATGATGCTATA*AAGAAATATCATCCAGAACAGG*AAAATCCTTA
5322    cowdria-4c10.p1c        GAGTTCTTACCAGATAATGATGATGCTATACAAGAAATATCATCCAAAACAGGCAAAATCCTTA

CONSENSUS                       GAGTTCTTACCAGATAATGATGATGCTATACAAGAAATATCATCCAGAACAGG*AAAATCCTTA


                                    840       850       860       870       880       890
5321    WL1J39WF                GAAAAACTAAAAAATAGGATTCTTTATTTATTAGAAAAAAATCCTATGTTAGGACATCGTGGAT
5323    cowdria-13c12.p1c       GAAAAACTAAAAAATAGGATTCTTTATTTATTAGAAAAAAATCCTATGTTAGGACATCGTGGAT
5322    cowdria-4c10.p1c        GAAAAACTAAAAAATAGGATTCTTTATTTATTAGAAAAAAATCCTATGTTAGGACATCGCGGAT

CONSENSUS                       GAAAAACTAAAAAATAGGATTCTTTATTTATTAGAAAAAAATCCTATGTTAGGACATCGTGGAT


                                900       910       920       930       940       950       960
5321    WL1J39WF                GCAGACTTGCAATATCATATCCTGAAATATATGAGATGCAAGTTAAAGCAATATTTCTTGCTAT
5323    cowdria-13c12.p1c       GCAGACTTGCAATATCATATCCTGAAATATATGAGATGCAAGTTAAAGCAATATTTCTTGCTAT
5322    cowdria-4c10.p1c        GCAGACTTGCAATATCATATCCTGAAATATATGAGATGCAAGTTAAAGCAATATTTCTTGCTAT

CONSENSUS                       GCAGACTTGCAATATCATATCCTGAAATATATGAGATGCAAGTTAAAGCAATATTTCTTGCTAT


                                    970       980       990      1000      1010      1020
5321    WL1J39WF                TAAAGAATTGCAAGAAGAAACAGGTATTAAAGAAATCATTCCTGAAATTATGATACCTTTAATT
5323    cowdria-13c12.p1c       TAAAGAATTGCAAGAAGAAACCGGTATTAAAGAAATCATTCCTGAAATTATGATACCTTTAATT
5322    cowdria-4c10.p1c        TAAAGAATTGCAAGAAGAAACAGGTATTAAAGAAATCATTCCTGAAATTATGATACCTTTAATT

CONSENSUS                       TAAAGAATTGCAAGAAGAAACAGGTATTAAAGAAATCATTCCTGAAATTATGATACCTTTAATT


                              1030      1040      1050      1060      1070      1080
5321    WL1J39WF                ATGTCTGAAGAAGAGATTATTGT*AA*TCAAAG*AACTAG*TAAAT*AACATT*GCAAAACAAT
5323    cowdria-13c12.p1c       ATGTCTGAAGAAGAGATTATTGT*AA*TCAAAG*AACTAG*TAAAT*AACATT*GCAAAACAAT
5322    cowdria-4c10.p1c        ATGTCTGAAGAAGAGATTATTGTGAAGTCAAAGGAACTAGCTAAATCAACATTCGCAA

CONSENSUS                       ATGTCTGAAGAAGAGATTATTGT*AA*TCAAAG*AACTAG*TAAAT*AACATT*GCAAAACAAT


                              1090      1100      1110      1120      1130      1140      1150
5321    WL1J39WF                TTGATAATCCTAAATACTTACTTGGAACAATGATAGAACTACCAAAAGCAGCATTAATTGCAGA
5323    cowdria-13c12.p1c       TTGATAATCCTAAATACTTACTTGGAACAATGATAGAACTACCAAAAGCAGCATTAATTGCAGA

CONSENSUS                       TTGATAATCCTAAATACTTACTTGGAACAATGATAGAACTACCAAAAGCAGCATTAATTGCAGA


                                  1160      1170      1180      1190      1200      1210
5321    WL1J39WF                TAAAATTGC*AAACATGTACAATTCTTCAGTTTTG*TACAAATGA
5323    cowdria-13c12.p1c       TAAAATTGCAAAACATGTACAATTCTTCAGTTTTGGTACAAATGATCTTACTCAAACCACTTTA
5324    WL1H221-WF                                                ATGATCTTACTCAAACCACTTTA
9095    contig2                                                   ATGATCTTACTCAAACCACTTTA

CONSENSUS                       TAAAATTGCAAAACATGTACAATTCTTCAGTTTTGGTACAAATGATCTTACTCAAACCACTTTA


                               1220      1230      1240      1250      1260      1270      1280
5323    cowdria-13c12.p1c       GGAGTTTCTCGAGATGATTCTGCAAGTTTTATAGGTACATATCGAGATTTAGGAATTATAAAAC
5324    WL1H221-WF              GGAGTTTCTCGAGATGATTCTGCAAGTTTTATAGGTACATATCGAGATTTAGGAATTATAAAAC
9095    contig2                 GGAGTTTCTCGAGATGATTCTGCAAGTTTTATAGGTACATATCGAGATTTAGGAATTATAAAAC
5329    cowdria-56f11.p1c                                                                       C

CONSENSUS                       GGAGTTTCTCGAGATGATTCTGCAAGTTTTATAGGTACATATCGAGATTTAGGAATTATAAAAC


                                     1290      1300      1310      1320      1330      1340
5323    cowdria-13c12.p1c       ACGATCCTTTTG*AAACCTT*AG*ATATTGATGG*AGTAGG*AAAATTAATATCAATCGCAGTA
5324    WL1H221-WF              ACGA
9095    contig2                 ACGATCCTTTTG*AAACCTT*AG*ATATTGATGG*AGTAGG*AAAATTAATATCAATCGCAGTA
5329    cowdria-56f11.p1c       TCGATCCTTTTG*AAACCTT*AG*ATATTGATGG*AGTAGG*AAAATTAATATCAATCGCAGTA
5325    WL1T43-T7                CGATCCTTTTG*AAACCTT*AG*ATATTGATGG*AGTAGG*AAAATTAATATCAATCGCAGTA
5326    WL1A142WF                 GATCCTTTTG*AAACCTT*AG*ATATTGATGG*AGTAGG*AAAATTAATATCAATCGCAGTA
5327    WL1M717WF                 GATCCTTTTG*AAACCTT*AG*ATATTGATGG*AGTAGG*AAAATTAATATCAATCGCAGTA
5328    cowdria-49c03.p1c            CCTTTTGGAAACCTT-A--ATATTGATGG-AGTAGGCAAAATTAATATCAATCGCAGTA

CONSENSUS                       ACGATCCTTTTG*AAACCTT*AG*ATATTGATGG*AGTAGG*AAAATTAATATCAATCGCAGTA
```

```
                                    1350       1360       1370       1380       1390       1400
5323    cowdria-13c12.p1c           AGTT
9095    contig2                     AGTTTAGGTAAAAAAGAATCTCCAGATATCAAAATTGGAGTATGTGGAGAACATGGAGGAAACT
5329    cowdria-56f11.p1c           AGTTTAGGTAAAAAAGAATCTCCAGATATCAAAATTGGAGTATGTGGAGAACATGGAGGAAACT
5325    WL1T43-T7                   AGTTTAGGTAAAAAAGAATCTCCAGATATCAAAATTGGAGTATGTGGAGAACATGGAGGAAACT
5326    WL1A142WF                   AGTTTAGGTAAAAAAGAATCTCCAGATATCAAAATTGGAGTATGTGGAGAACATGGAGGAAACT
5327    WL1M717WF                   AGTTTAGGTAAAAAAGAATCTCCAGATATCAAAATTGGAGTATGTGGAGAACATGGAGGAAACT
5328    cowdria-49c03.p1c           AGTTTAGGTAAAAAAGAATCTCCAGATATCAAAATTGGAGTATGTGGAGAACATGGAGGAAACT

CONSENSUS                           AGTTTAGGTAAAAAAGAATCTCCAGATATCAAAATTGGAGTATGTGGAGAACATGGAGGAAACT


                                    1410       1420       1430       1440       1450       1460       1470
9095    contig2                     TTGAATCTATACAATTCTTTTCAAGGCTAAACATAAACTATATTTCCTGTTCTCCATACAGAAT
5329    cowdria-56f11.p1c           TTGAATCTATACAATTCTTTTCAAGGCTAAACATAAACTATATTTCCTGTTCTCCATACAGAAT
5325    WL1T43-T7                   TTGAATCTATACAATTCTTTTCAAGGCTAAACATAAACTATATTTCCTGTTCTCCATACAGAAT
5326    WL1A142WF                   TTGAATCTATACAATTCTTTTCAAGGCTAAACATAAACTATATTTCCTGTTCTCCATACAGAAT
5327    WL1M717WF                   TTGAATCTATACAATTCTTTTCAAGGCTAAACATAAACTATATTTCCTGTTCTCCATACAGAAT
5328    cowdria-49c03.p1c           TTGAATCTATACAATTCTTTTCAAGGCTAAACATAAACTATATTTCCTGTTCTCCATACAGAAT

CONSENSUS                           TTGAATCTATACAATTCTTTTCAAGGCTAAACATAAACTATATTTCCTGTTCTCCATACAGAAT


                                         1480       1490       1500       1510       1520       1530
9095    contig2                     TCCAATAGCAAGGCTAATAGCTGCACAATGCACAATATTCAATAAGGATATAAATTAATTCTTT
5329    cowdria-56f11.p1c           TCCAATAGCAAGGCTAATAGCTGCACAATGCACAATATTCAATAAGGATATAAATTAATTCTTT
5325    WL1T43-T7                   TCCAATAGCAAGGCTAATAGCTGCACAATGCACAATATTCAATAAGGATATAAATTAATTCTTT
5326    WL1A142WF                   TCCAATAGCAAGGCTAATAGCTGCACAATGCACAATATTCAATAAGGATATAAATTAATTCTTT
5327    WL1M717WF                   TCCAATAGCAAGGCTAATAGCTGCACAATGCACAATATTCAATAAGGATATAAATTAATTCTTT
5328    cowdria-49c03.p1c           TCCAATAGCAAGGCTAATAGCTGCACAATGCACAATATTCAATAAGGATATAAATTAATTCTTT

CONSENSUS                           TCCAATAGCAAGGCTAATAGCTGCACAATGCACAATATTCAATAAGGATATAAATTAATTCTTT


                                    1540       1550       1560       1570       1580       1590       1600
9095    contig2                     TTTAAGAGTAAAACACATTTATAATGTAGTTATTATAACTACTAGATAATAATAGCCATATAAT
5329    cowdria-56f11.p1c           TTTAAGAGTAAAACACATTTATAATGTAGTTATTATAACTACTAGATAATAATAGCCATATAAT
5325    WL1T43-T7                   TTTAAGAGTAAAACACATTTATAATGTAGTTATTATAACTACTAGATAATAATAGCCATATAAT
5326    WL1A142WF                   TTTAAGAGTAAAACACATTTATAATGTAGTTATTATAACTACTAGATAATAATAGCCATATAAT
5327    WL1M717WF                   TTTAAGAGTAAAACACATTTATAATGTAGTTATTATAACTACTAGATAATAATAGCCATATAAT
5328    cowdria-49c03.p1c           TTTAAGAGTAAAACACATTTATAATGTAGTTATTATAACTACTAGATAATAATAGCCATATAAT

CONSENSUS                           TTTAAGAGTAAAACACATTTATAATGTAGTTATTATAACTACTAGATAATAATAGCCATATAAT


                                    1610       1620       1630       1640       1650       1660       1670
9095    contig2                     CTATTAATATCATAATTAACAGTTTATAATAATTTAAACATACAATAAACACATTATGTTATTT
5329    cowdria-56f11.p1c           CTATTAATATCATAATTAACAGTTTATAATAATTTAAACATACAATAAACACATTATGTTATTT
5325    WL1T43-T7                   CTATTAATATCATAATTAACAGTTTATAATAATTTAAACATACAATAAACACATTATGTTATTT
5326    WL1A142WF                   CTATTAATATCATAATTAACAGTTTATAATAATTTAAACATACAATAAACACATTATGTTATTT
5327    WL1M717WF                   CTATTAATATCATAATTAACAGTTTATAATAATTTAAACATACAATAAACACATTATGTTATTT
5328    cowdria-49c03.p1c           CTATTAATATCATAATTAACAGTTATTAATAATTTAAACATACAATAAACACATTATGTTATTT
-5330   WL1A133_-T7                                                                          GTTATTT

CONSENSUS                           CTATTAATATCATAATTAACAGTTTATAATAATTTAAACATACAATAAACACATTATGTTATTT


                                         1680       1690       1700       1710       1720       1730
9095    contig2                     *CTACAC*TACATA*TCTACA*TATATTATCC*AACAGTTT*ACCCTTGGCTTT*GAGTA*TTT
5329    cowdria-56f11.p1c           *CTACAC*TACATA*TCTACA*TATATTATCC*AACAGTTT*ACCCTTGGCTTT*GAGTA*TTT
5325    WL1T43-T7                   *CTACAC*TACATA*TCTACA*TATATTATCC*AACAGTTT*ACCCTTGGCTTT*GAGTA*TTT
5326    WL1A142WF                   *CTACAC*TACATA*TCTACA*TATATTATCC*AACAGTTT*ACCCTTGGCTTT*GAGTA*TTT
5327    WL1M717WF                   *CTACAC*TACATA*TCTACA*TATATTATCC*AACAGTTT*ACCCTTGGCTTT*GAGTA*TTT
5328    cowdria-49c03.p1c           *CTACAC*TACATA*TCTACA*TATATTATCC*AACAGTTT*ACCCTTGGCTTT*GAGTA*TTT
-5330   WL1A133_-T7                 TCAACACCTACATAATCTTCAATATATTATCCCAACAGTTTTACCCTTGGCTTTTGAGTATTTT

CONSENSUS                           *CTACAC*TACATA*TCTACA*TATATTATCC*AACAGTTT*ACCCTTGGCTTT*GAGTA*TTT
```

```
                                 1740      1750      1760      1770      1780      1790      1800
9095    contig2                  TATGATTGA*GACAGTAACTTTACATTCCACTTCTTAGATCGATAATACATTGGAATATTTTAT
5329    cowdria-56f11.p1c        TATGATTGA*GACAGTAACTTTACATTCCACTTCTTAGATCGATAATACATTGGAATATTTTAT
5325    WL1T43-T7                TATGATTGA*GACAGTAACTTTACATTCCACTTCTTAGA
5326    WL1A142WF                TATGATTGA*GACAGTAACTTTACATTCCACTTCTTAGATCGATAATACATTGGAATATTTTAT
5327    WL1M717WF                TATGATTGA*GACAGTAACTTTACATTCCACTTCTTAGA
5328    cowdria-49c03.p1c        TATGATTGA*GACAGTAACTTTACATTCCACTTCTTAGATCGATAATACATTGGAATATTTTAT
-5330   WL1A133_-T7              TATGATTGAAGACAGTAACTTTACATTCCACTTTTTAGATCGATAATACATTGGAATATTTTAT
5331    cowdria-52f12.p1c                                          GATCGATA*TACATTGGAATATTTTAT

CONSENSUS                        TATGATTGA*GACAGTAACTTTACATTCCACTTCTTAGATCGATAATACATTGGAATATTTTAT


                                      1810      1820      1830      1840      1850      1860
9095    contig2                  TACTCTACAATGCATATTCATAATAAATACTAGAAATCATAATTTGACAATGTATTACTATCAG
5329    cowdria-56f11.p1c        TACTCTACAATGCATATTCATAATAAATACTAGAAATCATAATTTGACAATGTATTACTATCAG
5326    WL1A142WF                TACTCTACAATGCATATTCATAATAAATACTAGAAATCATAATTTGACAATGTATTACTATCAG
5328    cowdria-49c03.p1c        TACTCTACAATGCATATTCATAATAAATACTAGAAATCATAATTTGACAAT
-5330   WL1A133_-T7              TATTCTACAATGCATATTCATAATAAATACTAGAAATCATAATTTGACAATGTATTACTATCAG
5331    cowdria-52f12.p1c        TACTCTACAATGCATATTCATAATAAATACTAGAAATCATAATTTGACAATGTATTACTATCAG

CONSENSUS                        TACTCTACAATGCATATTCATAATAAATACTAGAAATCATAATTTGACAATGTATTACTATCAG


                                 1870      1880      1890      1900      1910      1920      1930
9095    contig2                  TAGTGTAGTACTCCTTCTAATAAACACATTCATGCATACAATATATTACATAATTTTATTTATG
5329    cowdria-56f11.p1c        TAGTGTAGTACTCCTTCTAATAAACACATTCATGCATACAATATATTACATAATT-TATTTATG
5326    WL1A142WF                TAGTGTAGTACTCCTTCTAATAAACACATTCATGCATACAATATATTACATAATT-TATTTATG
-5330   WL1A133_-T7              TAGTGTAGTACTCCTTCTAATAAACACATTCATGCATACAATATATTACATAATTTTATTTATG
5331    cowdria-52f12.p1c        TAGTGTAGTACTCCTTCTAATAAACACATTCATGCATACAATATATTACATAATTTTATTTATG

CONSENSUS                        TAGTGTAGTACTCCTTCTAATAAACACATTCATGCATACAATATATTACATAATTTTATTTATG


                                      1940      1950      1960      1970      1980      1990
9095    contig2_                 TAGTATATATCCTGTATAAATATACAAACATATATTTATATTGCTAAAGTTAGATATTTTATTA
5329    cowdria-56f11.p1c        TAGTATATATCCTGTATAAATATACAAACATATATTTATATTGCTAAAGTTAGATATTTTATTA
5326    WL1A142WF                TAGTATATATCCTGT-TAAATATACAAACATATATTTATATTGCTAAAGTTAGATATT-TATTA
-5330   WL1A133_-T7              TAGTATATATCCTGTATAAATATACAAACATATATTTATATTGCTAAAGTTAGATATTTTATTA
5331    cowdria-52f12.p1c        TAGTATATATCCTGTATAAATATACAAACATATATTTATATTGCTAAAGTTAGATATTTTATTA

CONSENSUS                        TAGTATATATCCTGTATAAATATACAAACATATATTTATATTGCTAAAGTTAGATATTTTATTA


                                 2000      2010      2020      2030      2040      2050      2060
9095    contig2                  TAAA*TCTTTTGAATACTTTTAATCTATACAACTATA*ACCTTATGTTCTATTCCTCATGTTTT
5329    cowdria-56f11.p1c        TAAAATCTTTTGAATACCTTTAA
5326    WL1A142WF                TAAA*TCTTTTGAATACTTTTAATCTATACAAC--TATACCT*ATGTTCTATTCCTCATGTTTT
-5330   WL1A133_-T7              TAAA*TCTTTTGAATACTTTTAATCTATACAACTATA*ACCTTATGTTCTATTCCTCATGTTTT
5331    cowdria-52f12.p1c        TAAA*TCTTTTGAATACTTTTAATCTATACAACTATA*ACCTTATGTTCTATTCCTCATGTTTT

CONSENSUS                        TAAA*TCTTTTGAATACTTTTAATCTATACAACTATA*ACCTTATGTTCTATTCCTCATGTTTT


                                      2070      2080      2090      2100      2110      2120
9095    contig2                  CATGATCAACATCAGGCTACTTTATCTTTACTATACAACGTGTACTAATGATAAATAGCAACAA
5326    WL1A142WF                CATGATCAACATCAGGCTACTTTATCTTTACTATACAACGTGTACTAATGATAAATAGCAACAA
-5330   WL1A133_-T7              CATGATCAACATCAGGCTACTTTATCTTTACTATACAACGTGTACTAATGATAAATAGCAACAA
5331    cowdria-52f12.p1c        CATGATCAACATCAGGCTACTTTATCTTTACTATACAACGTGTACTAATGATAAATAGCAACAA
5332    cowdria-80e03.p1c                     TACTTTATCTTTACTATACAACGGGTGCTAATGATAAATAGCAACAA

CONSENSUS                        CATGATCAACATCAGGCTACTTTATCTTTACTATACAACGTGTACTAATGATAAATAGCAACAA


                                 2130      2140      2150      2160      2170      2180      2190
9095    contig2                  ACACAATGTTAGATTATATTTATCAGTAACATAGCACTTCTTATAAAAAATCTCTTAAAATACA
5326    WL1A142WF                ACACCATGTTAGATTATA*TTATCAGTAACATAGCACTTCTT-ATAAAAATCTTTTAAAATACA
-5330   WL1A133_-T7              ACACAATGTTAGATTATATTTATCAGTAACATAGCACTTCTTATAAAAAATCTCTTAAAATACA
5331    cowdria-52f12.p1c        ACACAATGTTAGATTATATTTATCAGTAACATAGCACTTCTTATAAAAAATCTCTTAAAATACA
5332    cowdria-80e03.p1c        ACACAATGTTAGATTATATTTATCAGTAACATAGCACTTCTTATAAAAAATCTCTTAAAATACA

CONSENSUS                        ACACAATGTTAGATTATATTTATCAGTAACATAGCACTTCTTATAAAAAATCTCTTAAAATACA
```

```
                                        2200       2210       2220       2230       2240       2250
9095    contig2                 TAAGCATACCATATCATTATATCCTATAATACTAGTAAGACATGTATAAACACATATATATCAG
5326    WL1A142WF               TA-CC*TACCATATCATTTTATCC*ATAATACTAGGGAGACA
-5330   WL1A133_-T7             TAAGCATACCATATCATTATATCCTATAATACTAGTAAGACATGTATAAACACATATATATCAG
5331    cowdria-52f12.p1c       TAAGCATACCATATCATTATATCCTATAATACTAGTAAGACATGTATAAACACATATATATCAG
5332    cowdria-80e03.p1c       TAAGCATACCATATCATTATATCCTATAATACTAGTAAGACATGTATAAACACATATATATCAG

CONSENSUS                       TAAGCATACCATATCATTATATCCTATAATACTAGTAAGACATGTATAAACACATATATATCAG


                                   2260       2270       2280       2290       2300       2310
9095    contig2                 TATGACAAAGTCCAATAACTATATTATTCATGACACAACATTAAGTACTACGAATAATGTTTAC
-5330   WL1A133_-T7             TATGACAAAGTCCAATAACTATATTATTCATGACACAACATTAAGTACTACGAATAATGTTTAC
5331    cowdria-52f12.p1c       TATGACAAAGTCCAATAACTATATTATTCATGACACAACATTAAGTACTACGAATAATGTTTAC
5332    cowdria-80e03.p1c       TATGACAAAGTCCAATAACTATATTATTCATGACACAACATTAAGTACTACGAATAATGTTTAC

CONSENSUS                       TATGACAAAGTCCAATAACTATATTATTCATGACACAACATTAAGTACTACGAATAATGTTTAC


                                2320       2330       2340       2350       2360       2370       2380
9095    contig2                 ACAATTAAATATTGATTAGCTACTTAAATAATAAAAAACAATTTATCATTACAATTATATATGT
-5330   WL1A133_-T7             ACAATTAAATATTGATTAGCTACTTAAATAATAAAAAACAATTTATCATTACAATTATATATGT
5331    cowdria-52f12.p1c       ACAATTAAATATTGATTAGCTACTTAAATAATAAAAAACAATTTATCATTACAATTATATATGT
5332    cowdria-80e03.p1c       ACAATTAAATATTGATTAGCTACTTAAATAATAAAAAACAATTTATCATTACAATTATATATGT

CONSENSUS                       ACAATTAAATATTGATTAGCTACTTAAATAATAAAAAACAATTTATCATTACAATTATATATGT


                                     2390       2400       2410       2420       2430       2440
9095    contig2                 ATATAAAAAAACAGTACCGTCTTTATACATCTAAATAATATATATTGTAACTACTCATCAGATC
-5330   WL1A133_-T7             ATATAAAAAAACAGTACCGTCTTTATACATCTAAATAATATATATTGTAACTACTCATCAGA
5331    cowdria-52f12.p1c       ATATAAAAAAACAGTACCGTCTTTATACATCTAAA*AATATATATTGTAACTACTCATCA
5332    cowdria-80e03.p1c       ATATAAAAAAACAGTACCGTCTTTATACATCTAAATAATATATATTGGAACTACTCATCAGATC
5333    cowdria-38c04.p1c                                                                    C

CONSENSUS                       ATATAAAAAAACAGTACCGTCTTTATACATCTAAATAATATATATTGTAACTACTCATCAGATC


                                 2450       2460       2470       2480       2490       2500       2510
9095    contig2                 ATTAAATTATTTAACTATACAACAAGTAATAATAATCTATAAATCAAGACTCTTTACTTTAAAA
5332    cowdria-80e03.p1c       ATTAAATTATTTAACTATACAACAAGTAATAATAATCTATAAATCAAGACTCTTTACTTTAAAA
5333    cowdria-38c04.p1c       ATTAAATTATTTAACTATACAACAAGTAATAATAATCTATAAATCAAGACTCTTTACTTTAAAA

CONSENSUS                       ATTAAATTATTTAACTATACAACAAGTAATAATAATCTATAAATCAAGACTCTTTACTTTAAAA


                                      2520       2530       2540       2550       2560       2570
9095    contig2                 AACTTTATTGATATCTATATATACTATATGTTAGGGTGGCGGTAAAAAGGCTTATGCACTTTTT
5332    cowdria-80e03.p1c       AACTTTATTGATATCTATATATACTATATGTTAGGGTGGCGGTAAAAAGGCTTATGCACTTTTT
5333    cowdria-38c04.p1c       AACTTTATTGATATCTATATATACTATATGTTAGGGTGGCGGTAAAAAGGCTTATGCACTTTTT

CONSENSUS                       AACTTTATTGATATCTATATATACTATATGTTAGGGTGGCGGTAAAAAGGCTTATGCACTTTTT


                                   2580       2590       2600       2610       2620       2630
9095    contig2                 CAAAAATTCATTATCAGATTTTTCTATAATCTGGTTAGTTTTCCTACCATATATATCTTCAGTA
5332    cowdria-80e03.p1c       CAAAAATTCATTATCAGATTTTTCTATAATCTGGTTAGTTTTCCTACCATATATATCTTCAGTA
5333    cowdria-38c04.p1c       CAAAAATTCATTATCAGATTTTTCTATAATCTGGTTAGTTTTCCTACCATATATATCTTCAGTA

CONSENSUS                       CAAAAATTCATTATCAGATTTTTCTATAATCTGGTTAGTTTTCCTACCATATATATCTTCAGTA


                                2640       2650       2660       2670       2680       2690       2700
9095    contig2                 TTGATATTATTATTAAAAAAAGTTAGATACTCACAAGAAACAATAAGTTTTATTGCTTCATGCA
5332    cowdria-80e03.p1c       TTGATATTATTATTAAAAAAAGTTAGATACTCACAAGAAACAATAAGTT-TATTGCTTCATGCA
5333    cowdria-38c04.p1c       TTGATATTATTATTAAAAAAAGTTAGATACTCACAAGAAACAATAAGTTTTATTGCTTCATGCA

CONSENSUS                       TTGATATTATTATTAAAAAAAGTTAGATACTCACAAGAAACAATAAGTTTTATTGCTTCATGCA


                                     2710       2720       2730       2740       2750       2760
9095    contig2                 TGCTATTTTACGCAGCATTTAATATTTACACCCAAACAACCGATAACATATCAGCCTTACATGT
5332    cowdria-80e03.p1c       TGCTATTTTACGCAGCATTTAATATTTACACCCAAA
5333    cowdria-38c04.p1c       TGCTATTTTACGCAGCATTTAATATTTACACCCAAACAACCGATAACATATCAGCCTTACATGT

CONSENSUS                       TGCTATTTTACGCAGCATTTAATATTTACACCCAAACAACCGATAACATATCAGCCTTACATGT
```

```
                                2770       2780       2790       2800       2810       2820
9095    contig2                 TGTATTATGTTTTACTCCTAAAATACAGCTAACATTAAAAACAGAGAGATATATAGGCATAATA
5333    cowdria-38c04.p1c       TGTATTATGTTTTACTCCTAAAATACAGCTAACATTAAAAACAGAGAGATATATAGGCATAATA

CONSENSUS                       TGTATTATGTTTTACTCCTAAAATACAGCTAACATTAAAAACAGAGAGATATATAGGCATAATA


                                2830     2840       2850       2860       2870       2880       2890
9095    contig2                 TTCAGCATACTAGCTTCATTTTTGTGGATAGTAACTACTATTTATACTATTAGTTATATGAGAC
5333    cowdria-38c04.p1c       TTCAGCATACTAGCTTCATTTTTGTGGATAGTAACTACTATTTATACTATTAGTTATATGAGAC

CONSENSUS                       TTCAGCATACTAGCTTCATTTTTGTGGATAGTAACTACTATTTATACTATTAGTTATATGAGAC


                                    2900       2910       2920       2930       2940       2950
9095    contig2                 ACAATGATAAAAATAACAAAAAGCAATCTATATTT*TATGCGTGTTTTGCAGCAAGTATTGGAT
5333    cowdria-38c04.p1c       ACAATGATAAAAATAACAAAAAGCAATCTATATTT-TATGCGTGTTTTGCAGCAAGTATTGGAT

CONSENSUS                       ACAATGATAAAAATAACAAAAAGCAATCTATATTT*TATGCGTGTTTTGCAGCAAGTATTGGAT


                                2960       2970       2980       2990       3000       3010       3020
9095    contig2                 GTACAATGGGAATTGCTTTTTCAGGAAATCTATTAACATTATTCATATTTTATGAGTTGCTAAC
5333    cowdria-38c04.p1c       GTACAATGGGA

CONSENSUS                       GTACAATGGGAATTGCTTTTTCAGGAAATCTATTAACATTATTCATATTTTATGAGTTGCTAAC


                                      3030       3040       3050       3060       3070       3080
9095    contig2                 TATTAGCACTTACCCTCTTGTTACATATTATGCAAATCACGAATCACAAATTTCTGGTAGGTAT
CONSENSUS                       TATTAGCACTTACCCTCTTGTTACATATTATGCAAATCACGAATCACAAATTTCTGGTAGGTAT


                                  3090       3100       3110       3120       3130       3140
9095    contig2                 TATATGGGAATACTGCTGGGAACTTCAATGTTATTATTTTTACCAGCAATCATCATTACTTATA
CONSENSUS                       TATATGGGAATACTGCTGGGAACTTCAATGTTATTATTTTTACCAGCAATCATCATTACTTATA


                                3150     3160       3170       3180       3190       3200       3210
9095    contig2                 ACATTAGCGGCACCTTAGATTTTACAAAAGGTGGCATATTACCATCAAGCATCTCTAGCGTTTT
CONSENSUS                       ACATTAGCGGCACCTTAGATTTTACAAAAGGTGGCATATTACCATCAAGCATCTCTAGCGTTTT


                                    3220       3230       3240       3250       3260       3270
9095    contig2                 TTTAATGAGCTTATTATTCCTATTCATTTACAGTATAGGAAAAACTGCATTAATGCCTATACAT
CONSENSUS                       TTTAATGAGCTTATTATTCCTATTCATTTACAGTATAGGAAAAACTGCATTAATGCCTATACAT


                                3280       3290       3300       3310       3320       3330       3340
9095    contig2                 TCATGGTTACCAAGAGCTATGGTTGCCCCTACTCCAGTAAGTGCACTATTACATGCAGTAGCAG
CONSENSUS                       TCATGGTTACCAAGAGCTATGGTTGCCCCTACTCCAGTAAGTGCACTATTACATGCAGTAGCAG


                                      3350       3360       3370       3380       3390       3400
9095    contig2                 TTGTCAAATCTGGAGTTTTTGCTTTAATAAAGGTTTTATTGTATATAATCGGGTTAGAAAGACT
CONSENSUS                       TTGTCAAATCTGGAGTTTTTGCTTTAATAAAGGTTTTATTGTATATAATCGGGTTAGAAAGACT


                                  3410       3420       3430       3440       3450       3460
9095    contig2                 ACAAGCGTTAGTTCAAACACATAACAATATATTGATGTACGCTTCAGCAATTACAATTCTCTTA
CONSENSUS                       ACAAGCGTTAGTTCAAACACATAACAATATATTGATGTACGCTTCAGCAATTACAATTCTCTTA


                                3470       3480       3490       3500       3510       3520       3530
9095    contig2                 GCATCCTTTATAGCAATAAAACAAACAAACTTAAAGAAGTTATTAGCATATTCAACAATTTCAC
CONSENSUS                       GCATCCTTTATAGCAATAAAACAAACAAACTTAAAGAAGTTATTAGCATATTCAACAATTTCAC


                                      3540       3550       3560       3570       3580       3590
9095    contig2                 AGCTTTCTTATATAACAATAGCAGTATCATTGTATACAGAACGTGCCGTTGATATTTCTATATT
CONSENSUS                       AGCTTTCTTATATAACAATAGCAGTATCATTGTATACAGAACGTGCCGTTGATATTTCTATATT


                                  3600       3610       3620       3630       3640       3650
9095    contig2                 TCAAATGATATCGCATGCATTTGCAAAAATAACATTATTCTTTACTGCAGGAGCAATATATACC
CONSENSUS                       TCAAATGATATCGCATGCATTTGCAAAAATAACATTATTCTTTACTGCAGGAGCAATATATACC
```

```
                                3660     3670      3680      3690      3700      3710      3720
9095    contig2                 AAAACAGGAAAAAAATACTTAAATGAACTTCAAGGTATTGGTAAATCCATGCCAATAACAATGA
CONSENSUS                       AAAACAGGAAAAAAATACTTAAATGAACTTCAAGGTATTGGTAAATCCATGCCAATAACAATGA

                                    3730      3740      3750      3760      3770      3780
9095    contig2                 CAGCGTTCTCTATAGGAGCTGCTGCAATGATAGGCATTCCTCCTGCTGTAACATTTTGGGGAAA
CONSENSUS                       CAGCGTTCTCTATAGGAGCTGCTGCAATGATAGGCATTCCTCCTGCTGTAACATTTTGGGGAAA

                                3790      3800      3810      3820      3830      3840      3850
9095    contig2                 ATTTTTTATTATATCAGAATCATTAAATCAAAATATTATGTCTGTAGTACTAGTATTAATAGCA
-453    cowdria-37c04.p1c                                               GTCTGTAGTACTAGTATTAATAGCA
-454    cowdria-66e08.p1c                                               GTCTGTAGTACTAGTATTAATAGCA

CONSENSUS                       ATTTTTTATTATATCAGAATCATTAAATCAAAATATTATGTCTGTAGTACTAGTATTAATAGCA

                                     3860      3870      3880      3890      3900      3910
9095    contig2                 AGCACCATACTAAATACAATATATTTTATCCCTATTATATACAATGCGTTTTATGTCCCATGTA
-453    cowdria-37c04.p1c       AGCACCATACTAAATACAATATATTTTATCCCTATTATATACAATGCGTTTTATGTCCCATGTA
-454    cowdria-66e08.p1c       AGCTCCATACTAAATACAATATATTTTATCCCTATTATATACAATGCGTTTTATGTCCCATGTA

CONSENSUS                       AGCACCATACTAAATACAATATATTTTATCCCTATTATATACAATGCGTTTTATGTCCCATGTA

                                  3920      3930      3940      3950      3960      3970
9095    contig2                 ACACTAATAAATATGCTGAAGCTCCTATACCCATGCTAATTGCCATTTCTATAACAACAATATG
-453    cowdria-37c04.p1c       ACACTAATAAATATGCTGAAGCTCCTATACCCATGCTAATTGCCATTTCTATAACAACAAAATG
-454    cowdria-66e08.p1c       ACACTAATAAATATGATGAAGCTCCTATACCCATGCTAATTGCCATTTCTATAACACCAATATG

CONSENSUS                       ACACTAATAAATATGCTGAAGCTCCTATACCCATGCTAATTGCCATTTCTATAACAACAATATG

                                3980     3990      4000      4010      4020      4030      4040
9095    contig2                 TACAATATTATTATTTTTATATCCTGACGTAATATTCAATATAATAAATCACCTCAAATAAGAT
-453    cowdria-37c04.p1c       TACAATATTATTATTTTTATATCCTGACGTAATATTCAATATAATAAATCACCTCAAATAAGAT
-454    cowdria-66e08.p1c       TGCAATATTATTATTTTTATATCTTGACGTAATATTCAATATAATAAATCACTTCAAATAAGAT

CONSENSUS                       TACAATATTATTATTTTTATATCCTGACGTAATATTCAATATAATAAATCACCTCAAATAAGAT

                                    4050      4060      4070      4080      4090      4100
9095    contig2                 AATGTATATAAATACATTGGTTAATACCATATAAAATATTCTATTTGGAGTACAAAATAAGAAT
-453    cowdria-37c04.p1c       AATGTATATAAATACATTGGTTAATACCATATAAAATATTCTATTTGGAGTACAAAATAAGAAT
-454    cowdria-66e08.p1c       AATGTATATAAATACATTGGTTAATACCATATAAAATATTCTATTTGGAGTACAAAATAAGAAT

CONSENSUS                       AATGTATATAAATACATTGGTTAATACCATATAAAATATTCTATTTGGAGTACAAAATAAGAAT

                                4110      4120      4130      4140      4150      4160      4170
9095    contig2                 TAGTTAAATATTATTTTGAATTGATAACGTATTATTTACACTATATATTAATCAAGTCATTAAT
-453    cowdria-37c04.p1c       TAGTTAAATATTATTTTGAATTGATAACGTATTATTTACACTATATATTAATCAAGTCATTAAT
-454    cowdria-66e08.p1c       TAGTTAAATATTATTTTGAATTGATAACGTATTATTTACTCTATATATTAATCAAGTCATTAAT
-452    WL1F2537                                        TAACGGATTAGTGACAGTATATCGTGGTCAAGCGGTTAAT

CONSENSUS                       TAGTTAAATATTATTTTGAATTGATAACGTATTATTTACACTATATATTAATCAAGTCATTAAT

                                     4180      4190      4200      4210      4220      4230
9095    contig2                 TATAAGTATAGATTCTACATTCTACATACTACACCTAAACAACTAC*AATTAAATACTTCAATT
-453    cowdria-37c04.p1c       TATAAGTATAGATTCTACATTCTACATACTACACCTAAACAACTAC*AATTAAATACTTCAATT
-454    cowdria-66e08.p1c       TATAAGTATAGATTCTACATTCTACATACTACACTTAAACAACTAC*AATTAAATACTTCAATT
-452    WL1F2537                TATAAGTATAGGTTGGACATTGTAGGTGCTACACCTAAACAACTACGAATTAAATACTTCAATT

CONSENSUS                       TATAAGTATAGATTCTACATTCTACATACTACACCTAAACAACTAC*AATTAAATACTTCAATT

                                  4240      4250      4260      4270      4280      4290
9095    contig2                 TGAAATATAAATACAATATTAAAATAGTCGTAAACTAGTAGCTTATCACTTAAATTATAATACC
-453    cowdria-37c04.p1c       TGAAATATAAATACAATATTAAAATAGTCGTAAACTAGTAGCTTATCACTTAAATTATAATACC
-454    cowdria-66e08.p1c       TGAAATATAAATACAATATTAAAATAGTCGTAAACTAGTAGCTTATCACTTAAATTATAATACC
-452    WL1F2537                TGAAATATAAATACAATATTAAAATAGTCGTAAACTAGTAGCTTATCACTTAAATTATAATACC

CONSENSUS                       TGAAATATAAATACAATATTAAAATAGTCGTAAACTAGTAGCTTATCACTTAAATTATAATACC
```

```
                                4300     4310       4320       4330       4340       4350       4360
9095    contig2                 AACCAAAATCATTAATTATTCTTTTTCA*CATATTACTTAAAGTAAAAAATAGAATATAAGATT
-453    cowdria-37c04.p1c       AACCAAAATCATTAATTATTCTTTTTCA*CATATTACTTAAAGTAAAAAATAGAATATAAGATT
-454    cowdria-66e08.p1c       AACCAAAATCATTAATTATTCTTTTTCA*CATATTACTTAAAGTAAAAAATAGAATATAAGATT
-452    WL1F2537                AACCAAAATCATTAATTATTCTTTTTCA*CATATTGCTTAAAGTAAAAAATAGAATATAAGATT

CONSENSUS                       AACCAAAATCATTAATTATTCTTTTTCA*CATATTACTTAAAGTAAAAAATAGAATATAAGATT


                                    4370       4380       4390       4400       4410       4420
9095    contig2                 AAAATATAACCCTAAGTTAGGAGTACTTTATCTACACTAGATCGACTCGAGAACCAGAGA*CCC
-453    cowdria-37c04.p1c       AAAATATAACCCTAAGTTAGGAGTACTT*ATCTACA
-454    cowdria-66e08.p1c       AAAATAGAACCTTAAGTTAGGAGTACTTTATCTACACTAGATCGACTCTAGAACTAGTGACCCC
-452    WL1F2537                AAAATATAACCGTAAGTTAGGAG

CONSENSUS                       AAAATATAACCCTAAGTTAGGAGTACTTTATCTACACTAGATCGACTCGAGAACCAGAGACCCC


                                4430
9095    contig2                 AAACGTATTAG
-454    cowdria-66e08.p1c       AAGG

CONSENSUS                       AAACGTATTAG
```

Addendum 5. DNA sequence obtained by using primers WL2TP_gapF and WL2TP_gapR.

Column 1: DNA sequenced

Primer WL2TP1_gapF: 5'- CAC TTA CAC CAA TGC CAC AC -3'

Primer WL2TP1_gapR: 5'- TTA CCG CCA CCC TAA CAT ATA G -3'

```
                    10         20         30         40         50         60         70
cowdria     CCATGAAAATTCTATCACCATTTTACAAACTTAAACACTTACACCAATGCCACACAACACTAAATATTAATCAAATATT
phage       CCATGAAAATTCTATCACCATTTTACAAACTTAAACACTTACACCAATGCCACACAACACTAAATATTAATCAAATATT
clone41                                                  CACTAAATATTAATCAAATATT
clone125                                                 CACTAAATATTAATCAAATATT

CONSENSUS   CCATGAAAATTCTATCACCATTTTACAAACTTAAACACTTACACCAATGCCACACAACACTAAATATTAATCAAATATT
```

Primer WL2TP1_gapF: 5'- CAC TTA CAC CAA TGC CAC AC -3'

```
                   80         90        100        110        120        130        140
cowdria     CTTCAATTAACTTATTCATAATGACTCTAGAAAATCTCCTACATTACAAGCAAGATATATAGAAAAATATCAACACCCA
phage       CTTCAATTAACTTATTCATAATGACTCTAGAAAATCTCCTACATTACAAGCAAGATATATAGAAAAATATCAACACCCA
clone41     CTTCAATTAACTTATTCATAATGACTCTAGAAAATCTCCTACATTACAAGCAAGATATATAGAAAAATATCAACACCCA
clone125    CTTCAATTAACTTATTCATAATGACTCTAGAAAATCTCCTACATTACAAGCAAGATATATAGAAAAATATCAACACCCA

CONSENSUS   CTTCAATTAACTTATTCATAATGACTCTAGAAAATCTCCTACATTACAAGCAAGATATATAGAAAAATATCAACACCCA

                  150        160        170        180        190        200        210
cowdria     GCTAAATTTTTTTTCTTATTCACAATGAAAACTTTTAACTTATTTCAAAATTAAAAACTATTAATTCTTATATAATATA
phage       GCTAAATTTTTTTTCTTATTCACAATGAAAACTTTTAACTTATTTCAAAATTAAAAACTATTAATTCTTATATAATATA
clone41     GCTAAATTTTTTTTCTTATTCACAATGAAAACTTTTAACTTATTTCAAAATTAAAAACTATTAATTCTTATATAATATA
clone125    GCTAAATTTTTTTTCTTATTCACAATGAAAACTTTTAACTTATTTCAAAATTAAAAACTATTAATTCTTATATAATATA

CONSENSUS   GCTAAATTTTTTTTCTTATTCACAATGAAAACTTTTAACTTATTTCAAAATTAAAAACTATTAATTCTTATATAATATA

                  220        230        240        250        260        270        280        290
cowdria     AAAACTACTCTTTCTACTAATATCCTATACAAACACTTTAGTATTTTATAGTACTACGTAAAAAATCCAGGTTGATCAA
phage       AAAACTACTCTTTCTACTAATATCCTATACAAACACTTTAGTATTTTATAGTACTACGTAAAAAATCCAGGTTGATCAA
clone41     AAAACTACTCTTTCTACTAATATCCTATACAAACACTTTAGTATTTTATAGTACTACGTAAAAAATCCAGGTTGATCAA
clone125    AAAACTACTCTTTCTACTAATATCCTATACAAACACTTTAGTATTTTATAGTACTACGTAAAAAATCCAGGTTGATCAA

CONSENSUS   AAAACTACTCTTTCTACTAATATCCTATACAAACACTTTAGTATTTTATAGTACTACGTAAAAAATCCAGGTTGATCAA

                  300        310        320        330        340        350        360
cowdria     CATCAGGCTACTTTATCTTTACTATACAACGTGTACTAATGATAAATAGCAACAAACACAATGTTAGATTATATTTATC
phage       CATCAGGCTACTTTATCTTTACTATACAACGTGTACTAATGATAAATAGCAACAAACACAATGTTAGATTATATTTATC
clone41     CATCAGGCTACTTTATCTTTACTATACAACGTGTACTAATGATAAATAGCAACAAACACAATGTTAGATTATATTTATC
clone125    CATCAGGCTACTTTATCTTTACTATACAACGTGTACTAATGATAAATAGCAACAAACACAATGTTAGATTATATTTATC

CONSENSUS   CATCAGGCTACTTTATCTTTACTATACAACGTGTACTAATGATAAATAGCAACAAACACAATGTTAGATTATATTTATC

                  370        380        390        400        410        420        430
cowdria     AGTAACATAGCACTTCTTATAAAAAATCTCTTAAAATACATAAGCATACCATATCATTATATCCTATAATACTAGTAAG
phage       AGTAACATAGCACTTCTTATAAAAAATCTCTTAAAATACATAAGCATACCATATCATTATATCCTATAATACTAGTAAG
clone41     AGTAACATAGCACTTCTTATAAAAAATCTCTTAAAATACATAAGCATACCATATCATTATATCCTATAATACTAGTAAG
clone125    AGTAACATAGCACTTCTTATAAAAAATCTCTTAAAATACATAAGCATACCATATCATTATATCCTATAATACTAGTAAG

CONSENSUS   AGTAACATAGCACTTCTTATAAAAAATCTCTTAAAATACATAAGCATACCATATCATTATATCCTATAATACTAGTAAG
```

```
                440        450        460        470        480        490        500       510
cowdria         ACATGTATAAACACATATATATCAGTATGACAAAGTCCAATAACTATGTTATTCATGACACAACATTAAGTACTACGAA
phage           ACATGTATAAACACATATATATCAGTATGACAAAGTCCAATAACTATATTATTCATGACACAACATTAAGTACTACGAA
clone41         ACATGTATAAACACATATATATCAGTATGACAAAGTCCAATAACTATATTATTCATGACACAACATTAAGTACTACGAA
clone125        ACATGTATAAACACATATATATCAGTATGACAAAGTCCAATAACTATGTTATTCATGACACAACATTAAGTACTACGAA

CONSENSUS       ACATGTATAAACACATATATATCAGTATGACAAAGTCCAATAACTATATTATTCATGACACAACATTAAGTACTACGAA

                          520        530        540        550        560        570        580
cowdria         TAATGTTTACACAATTAAATATTGATTAGCTACTTAAATAATAAAAAACAATTTATCATTACAATTATATATGTATAT
phage           TAATGTTTACACAATTAAATATTGATTAGCTACTTAAATAATAAAAAACAATTTATCATTACAATTATATATGTATAT
clone41         TAATGTTTACACAATTAAATATTGATTAGCTACTTAAATAATAAAAAACAATTTATCATTACAATTATATATGTATAT
clone125        TAATGTTTACACAATTAAATATTGATTAGCTACTTAAATAATAAAAAACAATTTATCATTACAATTATATATGTATAT

CONSENSUS       TAATGTTTACACAATTAAATATTGATTAGCTACTTAAATAATAAAAAACAATTTATCATTACAATTATATATGTATAT

                        590        600           610           620        630        640
cowdria         AAAAAAACAGTACCGTCTTTATACATCTAAATAATATATATTGTAACTACTCATCAGATCATTAAATTATTTAACTATA
phage           AAAAAAACAGTACCGTCTTTATACATCTAAATAATATATATTGTAACTACTCATCAGATCATTAAATTATTTAACTATA
clone41         AAAAAAACAGTACCGTGTTTATACATCTAAATAATATATATTGTAACTACTCATCAGATCATTAAATTATTTAACTATA
clone125        AAAAAAACAGTACCGTCTTTATACATCTAAATAATATATATTGTAACTACTCATCAGATCATTAAATTATTTAACTATA

CONSENSUS       AAAAAAACAGTACCGTCTTTATACATCTAAATAATATATATTGTAACTACTCATCAGATCATTAAATTATTTAACTATA

                650        660        670        680        690        700        710
cowdria         CAACAAGTAATAATAATCTATAAATCAAGACTCTTTACTTTAAAAAACTTTATTGATATCTATATATA
phage           CAACAAGTAATAATAATCTATAAATCAAGACTCTTTACTTTAAAAAACTTTATTGATATCTATATATA
clone41         CAACAAGTAATAATAATCTATAAATCAAGACTCTTTACTTTAAAAAACTTTATTGATATCTATATATACTATATGTTAG
clone125        CAACAAGTAATAATAATCTATAAATCAAGACTCTTTACTTTAAAAAACTTTATTGATATCTATATATACTATATGTTAG

CONSENSUS       CAACAAGTAATAATAATCTATAAATCAAGACTCTTTACTTTAAAAAACTTTATTGATATCTATATATACTATATGTTAG

                         720        730        740
clone41         GGTGGCGGTAAAAAGGCTTATGCACTTTTTCAAAAA
clone125        GGTGGCGGTAAAAAGGCTTATGCACTTTTTCAAAAA

CONSENSUS       GGTGGCGGTAAAAAGGCTTATGCACTTTTTCAAAAA
```

Primer WL2TP1_gapR: 5'- TTA CCG CCA CCC TAA CAT ATA G -3'

Addendum 6. DNA consensus sequence of the 10.9 kb contig. Column 1: Clone number and direction: (F = forward, R = reverse).

DNA sequence obtained from screening the contig against the sequences obtained from the random sequencing of the *C. ruminantium* λZAPII library

Primer WL2TP1_gapF: 5'- CAC TTA CAC CAA TGC CAC AC -3'

Primer WL2TP1_gapR: 5'- TTA CCG CCA CCC TAA CAT ATA G -3'

DNA sequence representing the gap

Last 348 bp of the WL1M636 PCR product sequence (see Figure 15, page 66)

TCATACTCAACATTTATGTATTACTTAGAACAAATACTGACCTTATTATAAAATCTATCTACTAAACAGCACAAATCTATCTTAAGTCTACACA
AAATACTATAATAAATATAGCTTATTACATACTTCATATATCTATACAGACAAAAACTGTAACTCCCAAATATTACACAGGCAATAATACATAT
TAAATTCAATTATTTTCAAGAAAGATATAGTTACACACTAAATTATTACAAATTCCTCCTGATATTTTGCTTGCAGGTACATTTCTTTATCAAT
ACATTTTCAAGATAAGTTGCAGGTAATGTTTGAAAATTTCGTTCTTTCATACCTAAATACAAACATTGCAAAACAGAAGATCTTCGGCTACCTG
AAAGAAGCGATGTTAAAACTTTTATTTTCATTTATAGCCAAAACATATTAATGCTAGTATATTTTACCGATTATATTACCTTTTTAATAAAAAT
ACAACTTCACATTTATGGAATCATTACAATAGAAAAATAAAAAACATTAAGATAACTAGTTTAACTAGTTATAAATCCTCCTATACATGTTACA
TAACTATTTCACTACTAAATATTATGAGACATTCTAGATAACCTTAATATACAGATACCTAAAATAACCATAATGAAGGACAATAACTGTCCCA
TAGTAATCCAATTAAATAAAATATAACCCACTTGAACATCTGGTTCTCTTACAAATTCAATAAAAAAACGTACTATTCCATACCATATCATAAA
AATAGAAAATAACATACCTTGATATGATTTTACCTTAGTAAAAAAAAATAATAAATTCATTACTACAAACAATAATAATCCTTCAAAAAAAGCT
TCATATAACTGACTAGGATGCCTATAAAAAAAATCTCCACTATTTTGAAAAATCATTCCAAACCGTGTATTTGTAACTTTACCATATAACTCAC
CATTAATAAAATTCGCAATACGTCCTAAAAATATTCCGACAGGTACAGCACACAAACACAAATCAATTGCTGATAAAAACTTTATTTTATACTT
TTTACAAAAAATATACATAGTACAAAAAAGCCCTAAACTTGCTCCATGAAATGACATACCACCTTTCCATAACTTAAACATTTCAATAGGAAAA
CTCATATAGAAATTTAAGTTATAAAATAATATATATCCTATCCTTCCACCAAGAATCATTCCAGTAACCCACCATGAAATTATCGACTTATAAC
TTTCTGGAGTAAAAACCTTGTACTTATCTATTTTCTGCACATACCAATAAGCAAATAAAATACCTATTATATATGCTAAAGAATACCATCGGAT
CGATAACAACCCTATCTTTAAAGCCACTGGATCTATATTCATAAAATATAATTACCTCTTTAATAATTTTAATAAATAACTATATACTTCCATA
AAAGCATTAATACAGTTTTTTCTATATTCTCAACTATAAACAATAGAGAATTTGGCATCTTCTAAAATTTAAAAACATATCTGGCAATACATAT
TCTTAAAAATAACGTCATTTCACATATCTATAAAAAGTAAAAATATCAACCAATAATAACCATGTTGATTTAGTACAATTTATTAAACTAATTC
CATTTTAACAAAATGTTTAATGTTCTAATTTAGATTCACATAAGATTATACATACAAGCTTTCATTAACTATAAATAGAATAAAAAATATAACG
TAATACCAAGTTGCGTTACACTTTATTGCACAGTCTATAATTTATCAATCATAAACAAATTAAAAAGAAAGTTTAGTATTAGATTAACAAAGCC
TGTAGTGATTCTAACATATATTCACAATACAATATATGTATAATGTATCATTAAATACAAAAATAATAAGTAAAAACTCAATTTGACACACTCA
TTATTTAATATAATTAGATATATAACATATGTTTTCACTATACAAAATTATCAGTACTATTTTTAATATTAAATTGATACATAATAATTAGTTA
TGGTATTAGTATGATAAAACATTAGGTTTTATTCTATCATAGAAACTATACAAACCTTAATATCCACAATCCTTTACTATAGAAATTTACATAT
TCTTAAAAGAAAATAGCAAGATATTATACATCTATTATCGCCAACATCAAAACTTGTCTAGCATAAAACCTTTGTACCATTATTATATAAATAA
TCAACAACAAGGTACTACCTAGTACCTACAAATTCTATTATAGATTCATCTTAAGCTATAATGTATTTATCAAGAAATATCCTATTTAATAATA
GATATTAATCATCTCACACTTATGAAATTACACTTCACTAATTTCACAATATGTAGTTTTTAAATTAATATTTAAATCTCGAAAGAAATTTAGG
TAATAAAATTTATTAGTTATATTAAACATTTTATAGTAATAAATAGTATTTTTACTATTTATTATATGTCATACAACACATTATTAAACTAACA
AGTTTTGCCAAATCCTAATATGTTCTCCAAAAAGCACCACTAGCAACACTTATTTTACTAATGATACTACCTAGCATTAATAAATTGATAATTG
TAACATCTAGAACTAACTCTCATATATCTTAAATATAAACAGTAAGACATTATCTATCACCCTAAATTCTATCATAATTACAGCATTTGTTTTA
TTTTATATAAAGACACCACTTTTTATACTTTAATAAGTCCATAAAAGTATAATTAACAGAACACTTGTTTACATATTGCTTGCACACTTTATGG
TATAAAACTTCATTAATTTTACAACCTGTAGCACCTACATTAACTCTTCAATCTTGAAAAATTAGACAACAAACTACCACCAACTATACTAAAT
GTTTTATAAACACAAATAACATTTTTTCTATTATTGGATATTACATCACATATCACTAAATTAATAAGTTTTTCCATAAACTTTAACGTGATGT
TTCTCTAGATGTATTACTAGCAGTGCCTGTTACAACTGCTTCTTCTATGTCACTATCTAGCAGTTCAGTAATTTCACAACCTGTAGCACCTACA
TTAACTCTTGAATCTTGAACAGCAGGATCAGGTAATAAAGTATCATCATCTACCCTAGGAACTGTATCAATATCCCCAGTACTTTTTTCCTCTT
GTAGATTATCTAATGCCTGCTGACCAAACACACCAGCTTTACAGGCATTATAAATAGGCATATTAACACTTTCAACAAACATAAATTCATATTT
GCCAATTTTGATATATTCAATGTCATCAAACATTTTTAGGTGATCTTTTATAGATAAGTAAGCTGTAGTACTAAGACCATCTAATTTTGATACT
ATTTTTCGATCTTTCCTACGCAGAAAATGACATATTAAAGTAATAGCTATCATTGTTCCAGTTATTCCAGCTATTACACTAGATAATACAAACT
CATGACCAAATTTAGCATCGCAACACAACATCAATACATTTGTGGCCATAAAAATACCTGCAAATCCAATAATAAACGCTGCACAGACTCTTGA
TATTTTATTCTTATACGATAGTAAATACAATCCTCCTAGAAACATTATAAAATACGATACAACAGAACAGTATAAAGCTATTCTATCAAATTTA
AAAAAATGACATGATAAAGCCACAAGTTTTAAAAGATACACAGATCCTAGTATTAGAAAAGTAGTACATCTTAACAGTAATTGCCTTTTTGTAA
GCTCATTGTTACCTGCCATAAGAAACAAATTGTGTTTTGTACTTACTTTCTTTCCTAACTCAGTAAGCTGTAAATTTCCTTTGCCATCATTGAC
CATAAAACTCTGTTTTTCAGCTTCCTGTTTTATATGTGATAAAATTCTCTCTCTTTTCGTAGAAGGAACATCATCATTTAGAAGATCCTCAGAA
ACAGTACGTATACCTTGATTATATTCTCTATACTGCCATGAAAGTAAATTCATACTACTTGCTTCTTCTTGAACACTTCTATTCCCAACATAAG
ATAATAAAAGAACAATAGGAATAACAATAGTAGGAGCAAATACTACAGTAAAACAACATCTTAATATTTTCTCACGTATAGTTAATAAAGACCA
CCTTTCAGCATATAACACAGAATCATATATTTCATGACTTAATGGTTTACTACTGACATCATATTTTTCAACTCGAAAAACTCGATCAGTACCA
GCTTGTGTTTCTAACTCATCAACAAAATCTCTTTGACTTTTTAAAATTATATCATATATATCACTTTCAGGAGTGTGCTCTTTAGCACCATCAC
ATGTAGAAATAATAAAAACAGTTTTAGGATTAGAATAAGCTTTACCTTCACGATAAGTACATTCAGAAATACCAATATCATAACTGGAATGCAC
TACTGGCAAAAACCCTAGCGAACGTATAAATGATACAGTCTTACGTATATCGTCTTTATTAGTAGAAGTATGTATTAATACCCTTTGTAAAAAA
GATTCATCTACCGCATCATCTTTATTTTGAATAACAAAAGCAGCAGGAGTAAGATTTCTTATTACCTTTCCTGAACAAAACATACCAATACTTG
GATTATAAAATATGGTATATCTTAAATTGTCAAGCCATTCCTTATAAAACTTATACTTAGCAGGATCATACATTAGTTTTTAACTTATTACCTA
TTTACAATCACAATACTACATAATTACTATTTTACATTAATTTTTTAAATATTATAAGTAAAAGTTTTACTACAATAAGGGCAAACTATTTCTT
GACCATTTCTGATAGTCAAATATATTTTCGGATGTTCAGTATAATCAACATCACTTCCTTCCCCATTACAAGATACTATTTGCTCTTCAAGTAC
ATCTATTTTTTTATCATGATTAGACATAAATATAATTTTTAATATTAATAAAACACAACTCTTATAACATCATAATAAATAACAAGCTACCTTT
CTTACCATTCAAATATATATTATTATACACTTTGTATAATCAAATTATTATATTACTAGCAAATACATGTTTACTACATTAAACTTATTTACTA
CTATCAACATAGCTTTATGAATAATTCTATAGACAAAAACTCCACTTTTTCTTTCCTTGATTAAAGAATATACCTTAGATAAGTCATCATAAAA