THE SIGNIFICANCE OF DUPLICATIVE *VSG* GENE ACTIVATION DURING ANTIGENIC VARIATION IN AFRICAN TRYPANOSOMES

NICK ROBINSON
WELLCOME CENTRE FOR MOLECULAR PARASITOLOGY
THE ANDERSON COLLEGE
UNIVERSITY OF GLASGOW

Submitted for the degree of Doctor of Philosophy

October, 1999

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

ACKNOWLEDGEMENTS

Many thanks to my supervisor, Dave Barry, for providing continual encouragement and guidance throughout this studentship.

I would also like to acknowledge the members of the Trypanosome Research Team, especially Nils Burman and Richard McCulloch, for their advice and assistance, which has proved invaluable during my three years at WUMP.

Thanks also to Sara Melville, Vanessa Leech and Caroline Gerrard for kindly recommending the PFG conditions utilised in this thesis.

This work was supported by a studentship from the Medical Research Council, with additional funding from the Wellcome Trust.

CONTENTS

Title i

Declaration ii

Acknowledgements iii

Figures list ix

Tables list xii

Abbreviations xiii

ABSTRACT 1

CHAPTER 1

(Introduction) 4

1.1 General introduction 5

1.2 Antigenic variation 8

1.3 Structure of the VSG coat 9

1.4 The bloodstream expression sites (*BES*s) 12

1.5 The genetic mechanisms for antigenic variation 14

1.6 The metacyclic variable antigenic types (MVATs) 19

1.7 Gene expression in trypanosomes 20

1.8 Life cycle dependent stage-specific regulation of *VSG* expression 22

1.9 Regulation of a single active *BES* 24

1.10 The dynamics of trypanosome antigenic variation and hierarchical VAT (variable antigen type) expression 26

1.11 The minichromosomes of *T. brucei* 28

1.12 Differences between monomorphic and pleomorphic trypanosome lines 30

CHAPTER 2

(Materials and Methods) 32

(Reagent abbreviations) 33

2.1 Routine handling of trypanosomes 35

2.1.1 Host immunosuppression, trypanosome growth and collection, and stabilate manufacture 35

2.1.2 Trypanosome cloning 36

2.1.3 Preparation of blood smears for immunofluorescence and plasma for immune lysis 36

2.2 Serology 37

2.2.1 Immune lysis assay 37

2.2.2 Immunofluorescence 38

2.3 Reverse transcription polymerase chain reaction (RT-PCR) 38

2.3.1 RNA isolation 38

2.3.2 Reverse transcription 39

2.3.3 PCR amplification of *VSG* specific cDNA 39

2.4 Cloning of PCR amplified *VSG* specific cDNAs 40

2.4.1 Ligations 40

2.4.2 Transformations and plasmid retrieval 41

2.5 Isolation and purification of trypanosomes from blood 41

2.6 Isolation of genomic DNA 42

2.6.1 Preparation of genomic plugs from live trypanosomes 42

2.6.2 Standard genomic DNA preparations 43

2.7 Gel electrophoresis and Southern blotting 43

2.7.1 General gel electrophoresis 43

2.7.2 Genomic digestions 43

2.7.3 Pulsed field gel electrophoresis (PFGE) 44

2.7.4 Southern blotting 45

2.8 Probe manufacture and DNA hybridization 45

2.8.1 Radiolabelling 45

2.8.2 Hybridization 46

2.8.3 Stripping of nylon filters 47

2.9 Manufacture and screening of minichromosomal libraries 47
2.9.1 Isolation of minichromosomal DNA 47
2.9.2 Manufacture of the minichromosomal libraries 48
2.9.3 PCR screening of the minichromosomal libraries 49
2.10 Phenol/ chloroform extractions 49
2.11 Sequencing 50

CHAPTER 3

(The order of VSG expression in a chronic, pleomorphic, *T. brucei* infection) 51

3.1 Introduction 52
3.2 Derivation of the ILTat 1.2 pleomorphic clone 53
3.3 Initiation and progression of the chronic ILTat 1.2 infections 53
3.4 Immune lysis dissection of the first relapse peak 58
3.5 Further ILTat 1.2 infections, and additional immune lysis investigations 61
3.6 Immunofluorescent typing of the VATs 63
3.7 Reverse transcription PCR (RT-PCR) amplification of the *VSG* genes 66
3.8 Allocation of ILTat numbers to previously unidentified VATs 69
3.9 A PCR approach for analysing rapid switch products 70
3.10 Single relapse study of the high switching pleomorphic line ILTat 1.61c 74
3.11 Summary 78

CHAPTER 4

(Genetic analysis of the VSG switching mechanisms utilized by pleomorphic trypanosomes) 79

4.1 Introduction 80
4.2 Genomic analysis of the *VSG* genes activated during the first relapse peak 81

4.3 ILTat 1.2 *VSG* gene copies in expressor and non-expressor clones 82
4.4 Activation of the ILTat 1.25, 1.67 and 1.68 *VSG* genes 85
4.5 Activation of the ILTat 1.69, 1.21 and 1.64 *VSG* genes 89
4.6 RAD51 control hybridizations 93
4.7 Activation of the ILTat 1.25 *VSG* gene from the ILTat 1.61c high switching single relapse investigation 94
4.8 Pulsed field gel electrophoresis (PFGE) analysis of the *VSG* genes activated during the first relapse peak 95
4.9 Chromosomal location of the ILTat 1.2 *VSG* genes 95
4.10 PFGE analysis of the ILTat 1.25, 1.67 and 1.68 *VSG* genes 96
4.11 PFGE analysis of the ILTat 1.69, 1.21 and 1.64 *VSG* genes 100
4.12 Resolution of the 1.8 Mb cluster 101
4.13 Resolution of the minichromosomal BCs 103
4.14 Summary 104

CHAPTER 5

(Characterization of the upstream flanking region of the ILTat 1.21 minichromosomal *VSG* gene) 106

5.1 Introduction 107
5.2 Identification of restriction sites suitable for minichromosomal library manufacture 108
5.3 Construction of the minichromosomal libraries 113
5.4 Screening the minichromosomal libraries 114
5.5 Characterization of the ILTat 1.21 and 1.25 upstream flanks 121
5.6 Identification of the ILTat 1.21 upstream duplication boundary 128

CHAPTER 6

(Discussion) 136

6.1 Antigenic variation 137
6.2 Pleomorphism and monomorphism: significant differences between trypanosome lines resulting from laboratory attenuation 140
6.3 The hierarchical order of *VSG* gene expression in chronic infections 143

6.4 VAT appearance in the chronic ILTat 1.2 infection 144
6.5 Mode of *VSG* gene activation in the ILTat 1.2 and 1.61c pleomorphic infections: the predominance of duplicative transposition 146
6.6 Telomeric *VSG* genes are activated most frequently, and predominate in the early stages of infection 149
6.7 The occurrence of a dominant *BES* 153
6.8 The ILTat 1.21 minichromosomal *VSG* upstream flank 154
6.9 Future work 155

REFERENCES 161

FIGURES LIST

CHAPTER 1

1. Simplified life cycle of *T. brucei* 7
2. The arrangement of dimeric membrane-form (mf) VSG molecules attached to the lipid membrane by GPI anchors 11
3. Longitudinal section of an African trypanosome 11
4. Schematic representation of a *BES* and the chromosomal internal BC genes 13
5. The mechanisms of *VSG* activation in *T. brucei* 18
6. Typical parasitaemic profile of a trypanosome infection in a mouse host 27

CHAPTER 3

7. Derivation of the 77 switched clonal stabilates isolated from the ILTat 1.2 chronic rabbit infection 56
8. Parasitaemia of the ILTat 1.2 chronic rabbit infection over the 30 days, determined by microscopy 57
9. Immune lysis typing of the 36 trypanosome clones isolated from the first relapse peak of the ILTat 1.2 chronic rabbit infection 60
10. Summary of the onset of lytic activity in four separate rabbits against the six VATs that were isolated from the first relapse peak in rabbit A 62
11. Summary of the immunofluorescence results for the six VATs isolated from the first relapse peak 65
12. RT-PCR of the *VSG* transcripts of the six VATs isolated from the first relapse peak 67
13. Physical maps of the ILTat 1.2 and VAT A-F VSG genes 68
14. Schematic representation of the PCR method for analysing high switching VSG gene products 73
15. Derivation of the 11 switched clones generated from the ILTat 1.61c single relapse investigation 75
16. VSG gene cDNA generated by RT-PCR for the 11 switched clones isolated from the ILTat 1.61c single relapse investigation 76

CHAPTER 4

17. PFGE separation of chromosome-size DNA from 3 trypanosome stocks 81
18. ILTat 1.2 *VSG* gene copies in ILTats 1.2, 1.25, 1.67 and 1.68 83
19. ILTat 1.2 *VSG* gene copies in ILTats 1.2, 1.68, 1.21 and 1.64 84
20. Activation of the ILTat 1.25 *VSG* gene 86
21. Activation of the ILTat 1.67 *VSG* gene 87
22. Activation of the ILTat 1.68 *VSG* gene 88
23. Activation of the ILTat 1.69 *VSG* gene 90
24. Activation of the ILTat 1.21 *VSG* gene 91
25. Activation of the ILTat 1.64 *VSG* gene 92
26. *RAD51* single copy gene control probing 93
27. Activation of the ILTat 1.25 *VSG* gene from the ILTat 1.61c high switching experiment 94
28. Chromosomal location of the ILTat 1.2 *VSG* genes 96
29. ILTat 1.25 *VSG* cDNA probing of PFGE separated chromosome size DNA 97
30. ILTat 1.67 *VSG* cDNA probing of PFGE separated chromosome size DNA 98
31. ILTat 1.68 *VSG* cDNA probing of PFGE separated chromosome size DNA 99
32. Transposition of the duplicated ILTat 1.69, 1.21, and 1.64 *VSGs* 100
33. 600s PFGE resolution of the 1.8 Mb comigrating chromosomes, probed with ILTat 1.2 cDNA 101
34. 600s PFGE resolution of the 1.8 Mb comigrating chromosomes, probed with ILTat 1.2 or 1.69 cDNA 102
35. PFGE resolution of the minichromosomes 103

CHAPTER 5

36. Physical maps of the ILTat 1.21 and 1.25 *VSG* full-length cDNAs 109
37. Autoradiograph demonstrating the sparsity of restriction sites upstream of the *VSG* gene in the ILTat 1.21 minichromosomal BC 110
38. Location of the *Nsi*I sites upstream of the *VSG* gene in the ILTat 1.21 and 1.25 BCs 112
39. 120 h PFGE separation of the minichromosomes on a 1.2% LMP agarose gel stained with EtBr 113
40. PCR detection of the ILTat 1.21 upstream region from the *Sal*I/ *Nsi*I minichromosomal library 115
41. PCR detection of the ILTat 1.25 upstream region from the *Sal*I/ *Nsi*I minichromosomal library 117
42. Half-nested PCR reactions of the ILTat 1.25 and 1.21 minichromosomal PCR products 118
43. PCR amplification of the potential ILTat 1.68 and 1.69 minichromosomal products from the *Sal*I/ *Nsi*I library 119
44. Determination of the average insert size of the *Sal*I/ *Nsi*I minichromosomal library 120
45. Physical maps of the ILTat 1.25 and 1.21 minichromosomal clones 123
46. 70 bp repeat sequence detected in the ILTat 1.21 genomic clone 125
47. CLUSTAL W analysis of the 1239 bp region homologous with *ESAG3* in the ILTat 1.21 minichromosomal clone 127
48. Position of the two ORFs within the *ESAG3* homologous ILTat 1.21 sequence 128
49. Southern analysis of the ILTat 1.21 minichromosomal clone 130
50. *Age*I/ *Nsi*I and *Age*I/ *Hind*III double digests of ILTat 1.2 and 1.21 genomic DNA 131
51. Restriction analysis of the ILTat 1.21 barren region 133

CHAPTER 6

52. Construction of the full length ILTat 1.21 minichromosomal clone 158
53. The five suggested constructs for manipulation the 70 bp repeat region 160

TABLES LIST

1. Working dilutions and derivations of 37 reference antisera 64
2. Summary of the 11 VATs isolated from the rabbit A infection 69
3. *VSG* loci and activation events 105

ABBREVIATIONS

AnTat	Antwerp Trypanozoon antigen type
BAC	bacterial artificial chromosome
BC	basic copy
BES	bloodstream expression site
(B)VAT	(bloodstream) variable antigen type
EATRO	East African Trypanosomiasis Research Organisation
ELC	expression linked copy
ESAG	expression site associated gene
ETat	Edinburgh Trypanozoon antigen type
GPI	glycosylphosphatylinositol
GUTat	Glasgow University Trypanozoon antigen type
ILTat	International Laboratory for research on animal diseases, Trypanozoon antigen type
MES	metacyclic expression site
(M)VAT	(metacyclic) variable antigen type
(M)VSG	(metacyclic) variant surface glycoprotein
PARP	procyclic acidic repetitive protein
PFGE	pulsed field gel electrophoresis
UTR	untranslated region
VSG	variant surface glycoprotein

Reagent abbreviations are listed in Materials and Methods, page 33.

ABSTRACT

African trypanosomes can survive prolonged exposure to the immune responses of their mammalian hosts by constantly changing the variant surface glycoprotein (VSG) expressed in their surface coat in a process known as antigenic variation. Each parasite has a repertoire of approximately 1000 silent *VSG* genes, which are expressed differentially. The *VSG* genes are expressed exclusively at specialised transcriptional units known as the bloodstream expression sites (*BES*s). Approximately 20 of these sites exist in the trypanosome nucleus, although only one is maximally active at a time, ensuring that a single VSG is expressed in the parasite's surface coat. VSG switching is mediated by replacing the *VSG* gene in the active *BES*, or by silencing the active *BES* and initiating transcription from a new *BES*.

Investigations into trypanosome antigenic variation have described several different *VSG* switching mechanisms, although the majority of these studies were performed with laboratory attenuated trypanosome lines. Following repeated syringe passaging, these "monomorphic" lines have lost the ability to differentiate from the proliferative bloodstream form. Under normal circumstances these trypanosomes cannot be transmitted by the tsetse fly and do not develop through their life cycle. In addition, monomorphic trypanosomes display a marked reduction in their VSG switch rates, which are up to 4 or 5 orders of magnitude lower than those of non-adapted lines (Turner, C.M.R. and Barry, J.D. 1989. Parasitology **99:** 67-75; Turner, C.M.R. 1997. FEMS Microbiol. Lett. **153:** 227-231). This raises questions about the significance of the switching mechanisms observed in these lines. It has been proposed that non-adapted, or pleomorphic, trypanosomes normally have an active *VSG* switch mechanism, involving gene duplication, that is depressed, or from which a component is absent, in monomorphic lines (Barry, J.D. 1997. Parasitol. Today **13:** 212-218). The main aim of this thesis was to examine this hypothesis by analysing the switching mechanisms used during *VSG* activation during the early stages of a chronic, pleomorphic infection. Additionally, this investigation examined the chromosomal environment of the basic copy *VSG* genes, to determine whether a chromosome position effect influenced the early order of *VSG* gene expression.

The majority of the work presented in this thesis was derived from a single rabbit infection using ILTat 1.2 pleomorphic trypanosomes, which display a switch rate of 1 x 10^{-5} switches/cell/generation (Turner, C.M.J. 1997. FEMS Microbiol. Lett. **153:** 227-231). In total, 88 trypanosome clones were isolated from this infection, and it was subsequently discovered, by immune lysis assay, that they represented 11 distinct variable antigen types (VATs). To examine the reproducibility of the timing of expression, a further three ILTat- 1.2 infections were undertaken in new rabbit hosts and these infections displayed a generally reproducible pattern of VAT appearance. Genomic DNA from a representative of each of the 11 VATs was then isolated and separated by pulsed field gel electrophoresis (PFGE), or digested with *Hin*dIII, *Eco*RI or *Pst*I, and size fractionated on agarose gels. After Southern blotting, the filters were probed with VAT specific full-length cDNAs, to establish the copy number, chromosomal location and the method of activation of each of the 11 *VSG* genes. This analysis revealed that duplicative events predominated during the early stages of infection, with at least 9 (possibly 10) of the *VSG*s becoming activated by this mechanism. Only 1 VAT had activated its *VSG* gene by the transcriptional switch that accounts for 2/3 of the switching events observed in the early stages of monomorphic infections (Liu, A.Y.C. *et al.* 1985. J. Mol. Biol. **167:** 57-75). At least 8 (and probably 10) of the activated VSGs originated from telomeric loci, while there was only 1 example of duplication from a chromosome internal site, where most *VSG* genes reside, and this occurred in the second relapse peak. PFGE determined that 6 VATs possessed minichromosomal silent copies of the respectively expressed *VSG*, and it appeared that these donors were preferentially utilized in the duplicative events. The upstream flank from one of these minichromosomal donors was cloned, and it appears that the duplication boundary delimits at the 70 bp repeat region, which is located upstream of most *VSG* genes. A smaller single relapse study was also initiated using the rapidly switching pleomorphic line, ILTat 1.61c, which switches at 3 x 10^{-2} switches/ cell/generation (Turner, C.M.J. 1997. FEMS Microbiol. Lett. **153:** 227-231). This yielded only one analysable product, which was one of the same 11 genes observed in the rabbit infection, and was activated by duplication from a minichromosomal basic copy.

In summary, it appears that, during the early stages of infection, trypanosome antigenic variation is predominated by duplicative transposition of telomeric *VSG*s, many of which reside on the minichromosomes. These results contrast strongly with

the outcome of analyses of monomorphic trypanosomes, which utilize several different switching mechanisms, and most commonly display transcriptional switching during the early stages of infection. The work presented in this thesis therefore supports the hypothesis of a dedicated duplicative switching mechanism that is reduced in (or even absent from) laboratory adapted, monomorphic lines.

CHAPTER 1

INTRODUCTION

1.1 General introduction

Trypanosomes are unicellular, uniflagellate protozoans, of the order kinetoplastida, that diverged early during eukaryotic evolution, and consequently display differences in their genetics and metabolism when compared to higher organisms (Cross, 1990a). Several trypanosome species, including *Trypanosoma brucei*, *T. evansi*, *T. equiperdum*, *T. congolense* and *T. vivax* (collectively termed the African trypanosomes), are extracellular parasites that proliferate in the tissue fluids, capillary beds and vascular system of their mammalian hosts (Vickerman, 1985). Although these infections remain sub-clinical in much of the vast reservoir of African wildlife, high-level parasitaemias arise in non-indigenous domesticated cattle, resulting in decreased productivity and even death. High parasitaemias are also responsible for the severity of acute human sleeping sickness; this disease, which is caused by certain variants of *T. brucei*, is generally fatal if left untreated. *T. brucei* is now considered to comprise two morphologically indistinguishable subspecies, which are identified by their differences in host range and associated pathologies (Borst *et al.* 1981; Tait *et al.* 1985; MacLeod, 1999). One of these subgroups possesses two host range variants, *T. brucei brucei* and *T. b. rhodesiense*. The first variant, *T. b. brucei*, causes Ngana (meaning "loss of spirits" in the Zulu language) in cattle and also infects a wide range of the African wildlife, but this variant is not resistant to human serum. However, *T. b. rhodesiense*, the second variant, is resistant to human serum and is the causative agent of the acute form of human sleeping sickness. The second subspecies, *T. b. gambiense*, is also responsible for human sleeping sickness, but produces the chronic form of the disease.

In Africa, trypanosomes are transmitted between hosts mainly by an insect vector, which in the case of *T. brucei* (and indeed most of the African trypanosomes) is *Glossina*, the tsetse fly. The life cycle of *T. brucei* is illustrated in Figure 1, and is characterized by a succession of different forms adapted to the varied environments the parasite encounters (Vickerman, 1985). In the mammalian bloodstream, the slender form of the parasite divides and cannot engage in oxidative phosphorylation, since mitochondrial functions are repressed. Instead it relies on aerobic glycolysis, metabolising glucose to pyruvate, inside specialised organelles termed glycosomes.

As an infection progresses, the population becomes increasingly pleomorphic, consisting of the slender dividing form, intermediate forms, and the non-dividing short stumpy form, pre-adapted to transmission in the fly. All of these bloodstream types are covered by a coat which is made from a characteristic surface protein, the VSG (variant surface glycoprotein) (Overath *et al.* 1994), which shields the parasite from the mammalian host's immune system. Once ingested by the fly, the stumpy trypanosomes quickly differentiate into the actively dividing non-infective procyclic form, in which the VSG coat is replaced with one composed of procyclin (Roditi and Pearson, 1990). A recent study (Van Den Abbeele *et al.* 1999) has demonstrated that the number of *T. brucei* parasites in the tsetse midgut falls dramatically over the first few days following the infective bloodmeal. In this investigation, the authors estimated that about 7.5 x 10^6 trypanosomes were ingested during the bloodmeal, and demonstrated that after 3 days the parasite number of the population had dropped over 3 orders of magnitude to 1-2 x 10^3 individuals. The remaining parasite population, consisting of transformed procyclic trypanomastigotes, was shown to expand rapidly over the next 3 days, until the number of midgut parasites stabilised at 2-5 x 10^5 parasites; this population size remained constant throughout the last 18 days of the observation period.

During the population expansion within the midgut, the procyclic cells move from the posterior towards the anterior midgut, while their morphology becomes progressively elongated (Van Den Abbeele *et al.* 1999). The end-point of this midgut colonization is indicated by the presence of a long trypomastigote form, referred to as mesocyclic cells (Vickerman, 1985), which occur in the anterior midgut near the proventriculus; (Vickerman, 1985; Van Den Abbeele *et al.* 1999). This mesocyclic trypomastigote form then enters the proventriculus, changing into longer (post-mesocyclic) cells, which begin to replicate their DNA from a diploid to a tetraploid state, and migrate to the foregut and proboscis (Van Den Abbeele *et al.* 1999). Instead of entering mitosis immediately, the (tetraploid) long trypomastigote form differentiates into a long epimastigote cell (with no loss of genetic material) and it is suspected that this stage migrates to the salivary glands (Van Den Abbeele *et al.* 1999). The long epimastigote cell then divides asymmetrically generating two morphologically distinct (diploid) daughter cells, the long and short epimastigote forms (Van Den Abbeele *et al.* 1999). The short epimastigote trypanosomes then attach to the epithelial cells of the salivary glands where they multiply rapidly; the

fate of the long epimastigote stage has not been determined, although it is possible that it does not develop any further and dies (Van Den Abbeele *et al.* 1999). Development then progresses through a non-dividing intermediate stage, the nascent metacyclic (Vickerman, 1985), when the parasites begin to re-express the VSG coat, before transforming into infective metacyclic trypanomastigote cells, which are transmitted back to the mammalian host when the fly feeds.

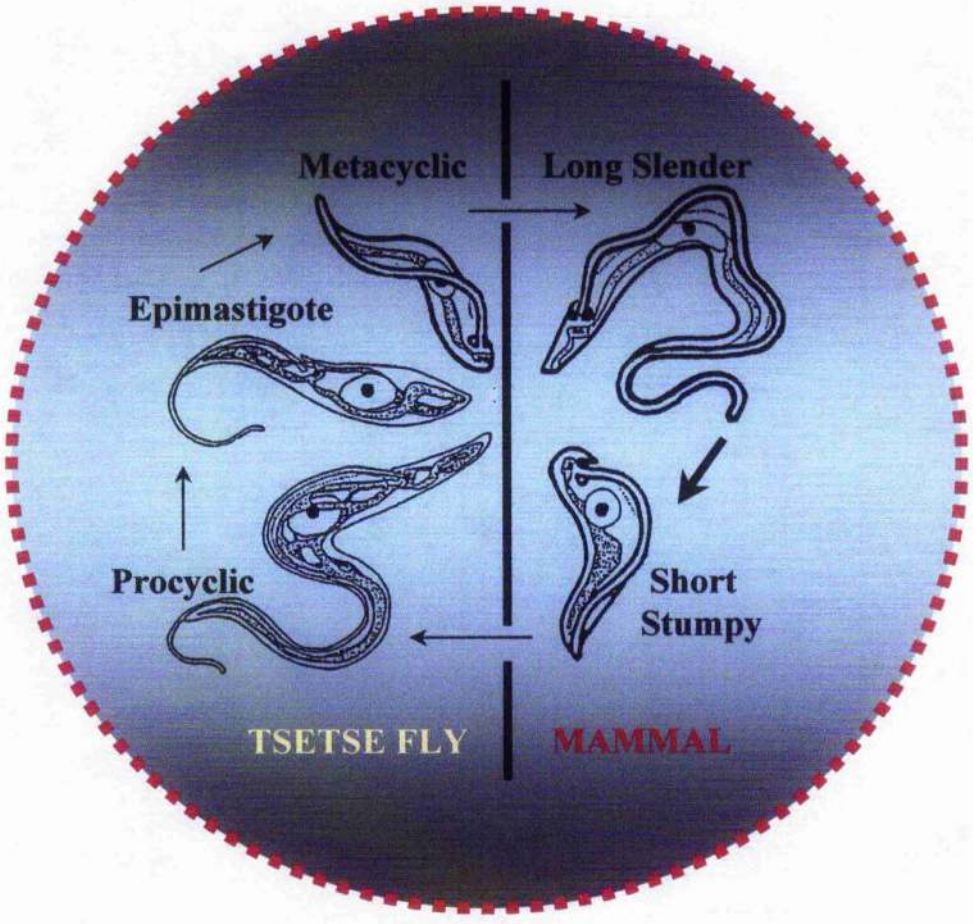


Figure 1. Simplified life cycle of *T. brucei* (adapted from Vickerman [1985]). The VSG coat is indicated in the metacyclic, long slender and short stumpy stages by a thicker outline of the trypanosome. The long slender, procyclic and epimastigote stages are proliferative, while the short stumpy and metacyclic stages are non-proliferative. Procyclic trypanosomes reside in the tsetse midgut, while the epimastigote and metacyclic stages occur in the salivary glands of the fly.

1.2 Antigenic variation

The African trypanosomes are extracellular parasites, which exposes their surface molecules to attack from the immune system of their host. As mentioned before, the bloodstream form cells are protected by a VSG coat, which masks the underlying surface antigens and inhibits the non-specific immune mechanisms of the mammal (Overath *et al.* 1994). However, the coat itself is highly immunogenic and provokes an antibody mediated response, ultimately leading to the destruction of the trypanosome. To prevent total elimination of the infecting population, the parasites has developed the ability to vary the VSG expressed in their surface coat using a mechanism known as antigenic variation (for recent reviews see (Cross, 1996; Barry, 1997a; Borst *et al.* 1998; Cross *et al.* 1998)). A single African trypanosome can potentially produce at least 100 antigenically novel surface coats (Capbern *et al.* 1977), and it seems likely that this is a conservative representation, with an estimated 1000 *VSG* genes available in the parasite's repertoire (Van der Ploeg *et al.* 1982). It therefore seems clear that, with such a vast capacity for variance, the infection can remain for prolonged periods in the bloodstream of the mammalian host. However, studies performed on cattle have revealed that chronic infection can eventually lead to the animal visibly eliminating the trypanosomes from its bloodstream (Nantulya *et al.* 1984; Nantulya *et al.* 1986; Barry, 1986). In these cases, the animals displayed no further clinical symptoms of disease and appeared resistant to subsequent infections from the same trypanosome serodeme (Nantulya *et al.* 1984; Barry, 1986). This phenomenon, known as "self-cure", is associated with the host exhausting the parasite's repertoire of VSG coats, resulting in acquired immunity to subsequent trypanosome challenges.

Antigenic variation enables the trypanosome to avoid immune-destruction and therefore prolong the infection period. The selective benefit from this strategy is obvious, since the probability of the infection being passed to a new host by the tsetse vector is directly dependent on the time that the population can survive within the bloodstream.

1.3 Structure of the VSG coat

In bloodstream form trypanosomes the dense VSG coat completely encompasses the surface of the cell membrane, associating as tightly packed dimers (Overath *et al.* 1994) (Figure 2). This molecular arrangement appears to thwart complement activation (Ferrante and Allison, 1983), while preventing other non-humoral molecules from reaching the cell membrane, and shields the invariant surface molecules (which probably do not penetrate far into the coat) from immunoglobulin (Ig). However, low molecular weight compounds are small enough to pass between the VSG molecules and reach transporters on the membrane. Larger host macromolecules, some of which are required for trypanosome metabolism, are also excluded from the membrane surface. Receptors for these molecules are contained within the flagellar pocket (see Figure 3) (Coppens *et al.* 1988; Grab *et al.* 1992; Webster and Russel, 1993), which is readily accessible to large protein complexes and Ig, but not the cellular arm of the host immune system (Borst, 1991a). It has been suggested (reviewed in Borst and Fairlamb, 1998) that humoral immune responses to these invariant proteins may not be especially effective, since the antibodies could be cleared from the flagellar pocket by endocytosis.

Each VSG monomer contains about 400-500 amino acids (Carrington *et al.* 1991) and consists of two domains separated by a "hinge" region; (Johnson and Cross, 1979; Carrington *et al.* 1991). The amino-terminal domain constitutes two-thirds of the molecule and represents the epitopes accessible to host antibodies. This region displays an enormous variation in its amino acid sequence, sharing only 13-30% sequence identity between VSGs (Miller *et al.* 1984; Carrington *et al.* 1991). However, some amino-terminus homology has been discovered between VSGs originating from the same serodeme (Olafson *et al.* 1984), and also between immunologically distinct VSGs (Barbet, 1985); it has also been demonstrated that cysteine residues are conserved within this region (Cross, 1984). Further homologies were revealed when Carrington *et al.* (1991) aligned the amino acid sequence of 19 VSGs in the context of conserved patterns of cysteine residues. The authors discovered that the amino-terminal domains of these VSGs could be divided into 3 distinct classes (types A, B and C). Despite differences in amino-terminus sequence, the individual VSG molecules appear to fold into a similar three-

dimensional structure, which has been determined for two VSGs by X-ray crystallography (Blum *et al.* 1993).

A higher degree of sequence similarity between VSGs is displayed in the carboxy-terminal domain, which consists of one or two tightly folded subdomains (Allen and Gurnett, 1983; Carrington *et al.* 1991). The greatest homology within this region occurs in the 50 amino acids closest to the carboxy-terminus (Matthyssens *et al.* 1981; Rice-Ficht *et al.* 1981), and the work of Carrington *et al.* (1991) identified four classes (types 1-4) of cysteine residue (and glycosylation site) arrangement within this domain. Additionally, different combinations of the various amino and carboxy-terminal domains were detected in this study, with the amino-terminus classes A and B associating with each of the carboxy-terminal classes 1,2 and 3.

The greatest degree of amino acid conservation between VSGs occurs in the carboxy-terminal hydrophobic tail of the VSG precursor peptide (Holder and Cross, 1981; Cross, 1990b). This pre-protein is modified in the endoplasmic reticulum, where the conserved tail is liberated from the VSG, and a glycosylphosphatidylinositol (GPI) moiety is covalently linked to the remaining protein. After the attachment of the GPI anchor, these mature molecules are referred to as membrane form (mf) VSGs. During an antigenic switch the mature mfVSGs are exocytosed to the cell surface in the flagellar pocket and then diffuse laterally to form the new coat (Overath *et al.* 1994), where the GPI anchor attaches the VSG to the phospholipid bilayer of the plasma membrane.

The VSG coat is shed when the trypanosomes differentiate to the procyclic form and is replaced by a coat of procyclin, also known as PARP (procyclic acidic repetitive protein). It appears that, during transition, the coat is mixed, containing both PARP and VSG molecules, resulting from a gradual release of the VSG, which ensures that the trypanosome is never "naked" (Ziegelbauer and Overath, 1990; Roditi and Pearson, 1990). The PARP coat is not as dense as the VSG coat and this could presumably result in the membrane-embedded transporters becoming more accessible to certain molecules in the procyclin coated insect form trypanosomes when compared to the metacyclic and bloodstream form parasites.

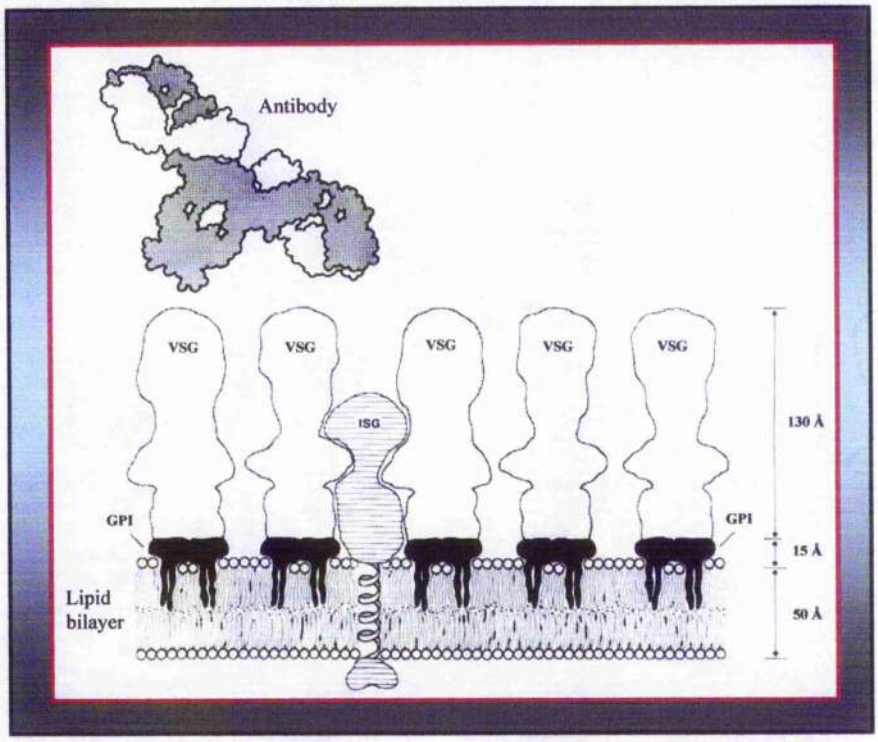


Figure 2. The arrangement of dimeric membrane-form (mf) VSG molecules attached to the lipid membrane by GPI anchors. An invariant surface glycoprotein (ISG) and a host antibody are also indicated in this schematic representation. After Overath *et al.* (1994).

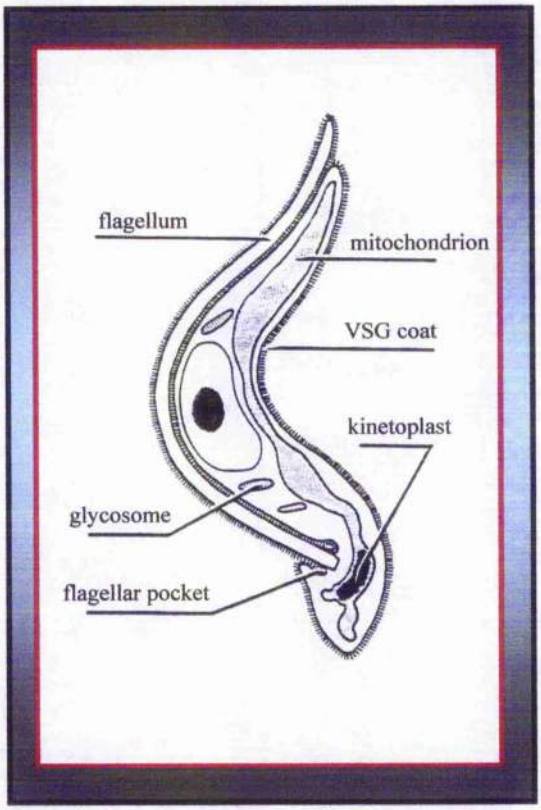


Figure 3. Longitudinal section of an African trypanosome. This schematic representation illustrates the major structural features of the cell. After Borst (1991a).

1.4 The bloodstream expression sites (*BES*s)

T. brucei has a repertoire estimated at approximately 1000 individual *VSG* genes and pseudogenes, the majority of which reside in long tandem arrays within the chromosomes; these genes are thought generally to be clustered in one or a few loci and are known as the chromosome internal, or basic copy (BC) genes (Van der Ploeg *et al.* 1982; Gibson and Borst, 1986). The remaining *VSG* genes occupy telomeric sites, predominantly on the set of about 100 minichromosomes present in the genome, which vary in size between approximately 25 and 150 kb (Van der Ploeg *et al.* 1984b; Gibson and Borst, 1986).

In the bloodstream-form trypanosomes, *VSG* gene transcription occurs exclusively at specialised telomeric regions called bloodstream expression sites (*BES*s) (Figure 4) (Johnson *et al.* 1987; Kooter *et al.* 1987; Pays *et al.*, 1989b). It is estimated that up to twenty of these units exist within each nucleus (Cully *et al.* 1985; Borst *et al.* 1990), although only one *BES* is maximally active at a time, ensuring that a single VSG is expressed in the protein coat of the parasite (Vanhamme and Pays, 1998; Chaves *et al.* 1999). The *BES* typically consists of a 45-60 kb transcriptional unit from the promoter to the end of the *VSG* gene (Johnson *et al.* 1987; Kooter *et al.* 1987; Pays *et al.*, 1989b), and is polycistronic, containing several expression site associated genes (*ESAG*s). More recently, a shorter *BES*, approximately 30 kb in length, has also been discovered (Xong *et al.* 1998). The *VSG* gene is located 5-10 kb from the chromosome end and is surrounded by two "barren" regions devoid of restriction sites. The upstream barren region extends for several kb and contains multiple imperfect 70 bp repeats, while the downstream region is composed mostly of telomeric and sub-telomeric hexanucleotide repeats, which constitute the end of the chromosome.

Homology between the individual *BES*s is high, although there are distinct differences. It has been observed that some *BES*s possess two promoters, which are arranged in tandem and separated by 13 kb of DNA (Gottesdiener *et al.* 1991; Gottesdiener *et al.* 1992). It has been estimated that this second promoter is present in half of all the *BES*s (Gottesdiener *et al.* 1992; Navarro and Cross, 1996). Much of the heterogeneity between the *BES*s appears in the *ESAG* sequences, where sequence divergence ranges from <10% (*ESAG6* and *7*) (reviewed in Borst and Fairlamb (1998)) to approximately 30% (*ESAG1*) (Cully *et al.* 1985); the *ESAG* set

encoded by different *BES*s also appears to vary significantly. The proteins encoded by the very similar *ESAG6* and *ESAG7* have been shown to form a heterodimeric transferrin receptor (Steverding *et al.* 1995), while *ESAG4* appears to be part of a large family of adenylate/guanylate cyclase genes (Alexandre *et al.* 1996), with at least nine additional members outside *BES*s (Alexandre *et al.* 1996; Borst and Fairlamb, 1998). The remaining *ESAG*s are not as well characterized (Overath *et al.* 1994), but there is some speculation that they could also be involved in membrane surface functions (reviewed by Borst and Fairlamb, 1998). However, the *ESAG8* gene is clearly not involved in membrane function, since it produces a protein that associates with the nucleus of the cell (Lips *et al.* 1996).

It has been proposed that the close association of the *ESAG* genes with the *VSG* genes in each *BES* indicates that the *ESAG* proteins are essential for the trypanosome during the bloodstream stages of its life cycle (Cully *et al.* 1985). However, more recent studies have discovered genes related to *ESAG*s (*GRESAG*s) outside the *BES*s, and have identified *ESAG* transcripts from *BES*s in procyclic cells (Pays *et al.* 1989b; Graham and Barry, 1991; Alexandre *et al.* 1996). It therefore appears that the control of expression of the *ESAG* genes is more complex than was originally suspected, and it is possible that the *ESAG* gene products are also necessary in the insect stages of the trypanosome life cycle.

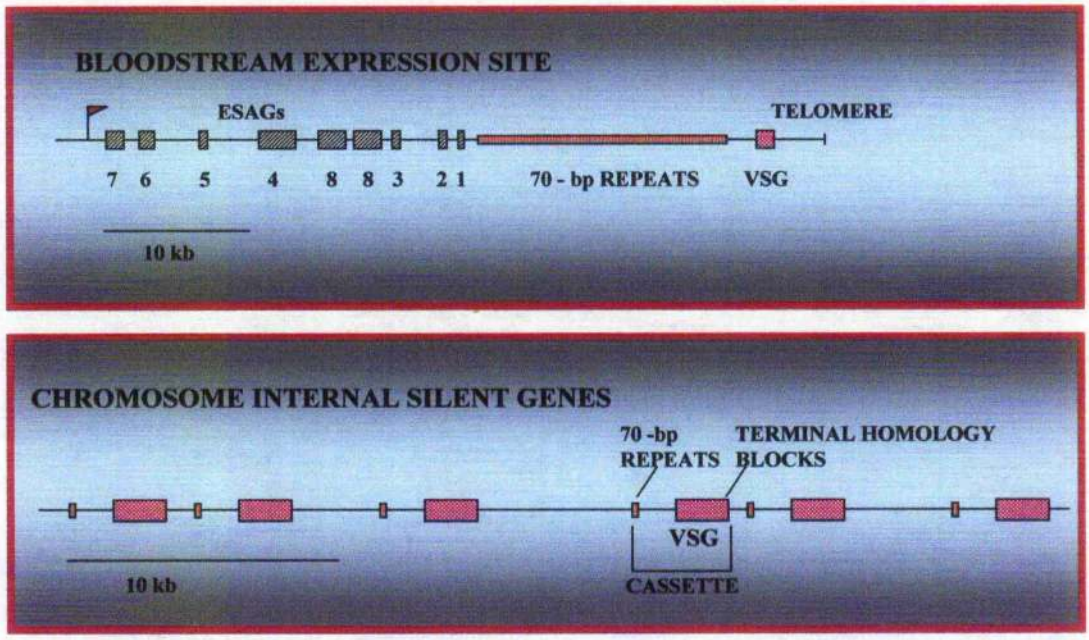


Figure 4. Schematic representation of a *BES* and the chromosomal internal BC genes. The red flag denotes the *BES* promoter. After Barry (1997a).

1.5 The genetic mechanisms for antigenic variation

A number of studies have been performed on the African trypanosomes since the early 1980s to elucidate the genetic mechanisms behind antigenic variation. These investigations have reported several different ways by which a new *VSG* gene can become expressed by the parasite. However, the majority of these studies were conducted using laboratory adapted "monomorphic" trypanosome lines (populations which do not differentiate into the stumpy bloodstream form, and are not transmissible by tsetse fly under normal circumstances), which display *VSG* switching rates several orders of magnitude lower than in non-adapted, pleomorphic lines (see section 1.12, page 25). The reduced switch rate observed in these attenuated parasites is low enough to be explained by background mutation and homologous recombination, and this has led some investigators to question the relative significance of the different switch events seen in these lines (Barry, 1997a).

The five reported types of DNA rearrangements associated with antigenic switching are summarised below and in Figure 5 (page 18):

1. Duplicative transposition of a chromosome internal gene into an active *BES*.

The BC *VSG* genes are never expressed within their own chromosome internal loci. For transcription to occur, these genes must be copied and transposed to the active *BES*, forming an expression linked copy (ELC), which replaces the previous ELC (Hoeijmakers *et al.* 1980; Pays *et al.* 1983a; Myler *et al.* 1984). This is the only route to activation for the chromosome internal genes, which constitute the majority of the *VSG* repertoire, and is thus the dominant mechanism in antigenic variation. The gene replacement occurs through recombination between regions of homology flanking the BC genes and the *BES*. In their 5' flank, the BC genes possess several of the 70 bp repeats found in the *BES* (Liu *et al.* 1983), while there is also considerable homology within the highly conserved carboxy-terminal coding region and in the 3' untranslated region (UTR) (Timmers *et al.* 1987; Lee and Van der Ploeg, 1987). The homology can extend up to 200 bp into the 3' UTR, which contains a block of 16 nucleotides, whose sequence is conserved in most *VSG* genes (Majumder *et al.* 1981). It is therefore convenient to consider the BC *VSG* genes

and their flanking regions as expression cassettes, which are copied into the *BES*, replacing the previous cassette and resulting in the expression of a new antigen type.

Studies have revealed that the upstream array of 70 bp repeats is frequently involved in *VSG* gene recombination (Florent *et al.* 1987; Lee and Van der Ploeg, 1987), although it appears that this region is not always utilised in monomorphic trypanosomes (Donelson *et al.* 1983; Michiels *et al.* 1983; Pays *et al.* 1983c; Lee and Van der Ploeg, 1987). Indeed, the deletion or inversion of the 70 bp repeat region of the active *BES* in monomorphic trypanosomes had no effect on the incidence of *VSG* duplications into that *BES* (McCulloch *et al.* 1997). The downstream conserved region of the *VSG* coding sequence and associated environment have also been shown to act as recognition sequences for transposition (Pays *et al.* 1983c; Timmers *et al.* 1987). Homology within the *VSG* coding region itself can lead to gene partial conversion events, although this is thought to occur mainly in the later stages of infection (Bernards *et al.* 1981; Barbet and Kamper, 1993).

Although most of the experimental evidence comes from monomorphic lines, duplicative transposition has also been demonstrated in high-switching pleomorphic lines, and the duplication boundary delimits to the 70 bp repeats in every observed activation with sufficient mapping (Delauw *et al.* 1987; Shah *et al.* 1987; Matthews *et al.* 1990). It therefore appears that this repeat sequence is normally associated with recombination in pleomorphic lines, although it is clear that further experimentation is necessary, since relatively few switch events have been analysed in these trypanosome lines.

2. Duplicative transposition of a silent telomeric *VSG* gene into an active *BES*.

This mechanism is also referred to as telomere conversion, and involves the same principles as the internal gene transposition. The main difference is that the duplication can proceed beyond the boundary of the expression cassette, resulting in the replacement of large segments of sequence, and possibly the whole telomere (de Lange *et al.* 1983). Although the majority of duplicative transposition involves the chromosome internal genes, telomere conversion remains an important switching mechanism, probably providing the main route to activation for the minichromosomal genes (Barry, 1997a).

3. Reciprocal telomere recombination

Reciprocal telomere recombination occurs via a classical homologous recombination reaction and involves the simple exchange of one telomeric region (including its *VSG* gene) for another (Pays *et al.* 1985b). This leads to the expression of a new variable antigen type (VAT) if a novel *VSG* gene is incorporated into an active *BES*. It is suspected that telomeric reciprocal recombination is just a typical background cross-over event, since there are few reported instances, all of which were reported in monomorphic trypanosomes. However, one *in vitro* investigation has suggested that earlier studies may have underestimated the frequency of these reactions, confusing changes far upstream of the *VSG* gene for transcriptional switches (Rudenko *et al.* 1996). Using integrated markers, and improved pulsed field gel electrophoresis (PFGE) techniques, this study concluded that telomere exchange could be an important, rather than rare, event in monomorphic trypanosomes. Reciprocal telomere recombination, however, has not been reported in pleomorphic trypanosomes.

4. Mosaic gene formation

This process is also referred to as segmental gene conversion and involves the formation of a mosaic ELC from the combination of two or more *VSG* coding regions, producing an antigenically novel surface coat, which protects the parasite from existing antibodies in the bloodstream (Pays *et al.* 1985a; Longacre and Eisen, 1986; Roth *et al.* 1989; Thon *et al.* 1989; Kamper and Barbet, 1992). These conversions are thought to be generally associated with antigen types appearing in chronic infections when the host has acquired immunity to most of the non-mosaic genes (Roth *et al.* 1989; Thon *et al.* 1990; Barbet and Kamper, 1993). It has been suggested that the additional source of antigenic variation provided by this relatively infrequent mode of switching could be important in prolonging the infection period (Kamper and Barbet, 1992). The simplest combinations arise when a new *VSG* gene does not completely replace the previous sequence in the *BES*, producing a chimera consisting mostly of the new gene, with the downstream region derived from the *VSG* originally in the *BES* (Pays *et al.* 1983d). In such cases, the two *VSG*s

involved have substantial homology with each other. More complicated segmental conversions between *VSG* gene family members result in the formation of mosaic (or composite) genes (Pays *et al.* 1985a). Often, mosaic genes are constructed from various *VSG* members that are incomplete pseudogenes, and therefore cannot encode a VSG if expressed singly (Thon *et al.* 1989).

5. *"in situ"* switching: the transcriptional activation of a new *BES* and silencing of the previous *VSG* gene

The *in situ* switch is the second major method of antigenic variation after gene conversion. Early experiments indicated that although these two processes would often occur simultaneously, they were in fact independent of each other (Myler *et al.* 1984). *in situ* switching involves no exchange of genetic material, but requires the activation of a previously silent *BES* and the inactivation of the transcribed *BES*. This results in the formerly expressed *VSG* gene being silenced, while the *VSG* gene in the newly activated *BES* becomes transcribed. The regulatory mechanisms that ensure that only one *BES* is transcribed, whilst the other sites are maintained in an inactive state, have not yet been determined. Moreover, the potential mechanisms involved in executing a transcriptional switch between an active and inactive *BES* are only now being examined. Failure to activate the new *BES* would result in the loss of the surface coat and destruction of the parasite, while continued expression of the old *BES* would cause the production of a mixed VSG coat (Muñoz-Jordán *et al.* 1996), rendering the trypanosome vulnerable to antibodies raised against the last VSG. It had previously been hypothesised that a mobile promoter element, which could transpose from one *BES* to the next, would ensure that the two stages of *in situ* switching remained synchronous. However, *in vitro* studies on certain trypanosome variants demonstrated that more than one *BES* could be activated concurrently (Cornelissen *et al.* 1985a), suggesting a stochastic model for the control of *BES* transcription (each *BES* is activated and inactivated independently of other *BES*s). Although this model was accepted for some time, a recent study has suggested that a more complicated system must operate, in which the activation and inactivation of the individual *BES*s are not independent of each other (further details in section 1.9, page 18) (Chaves *et al.* 1999).

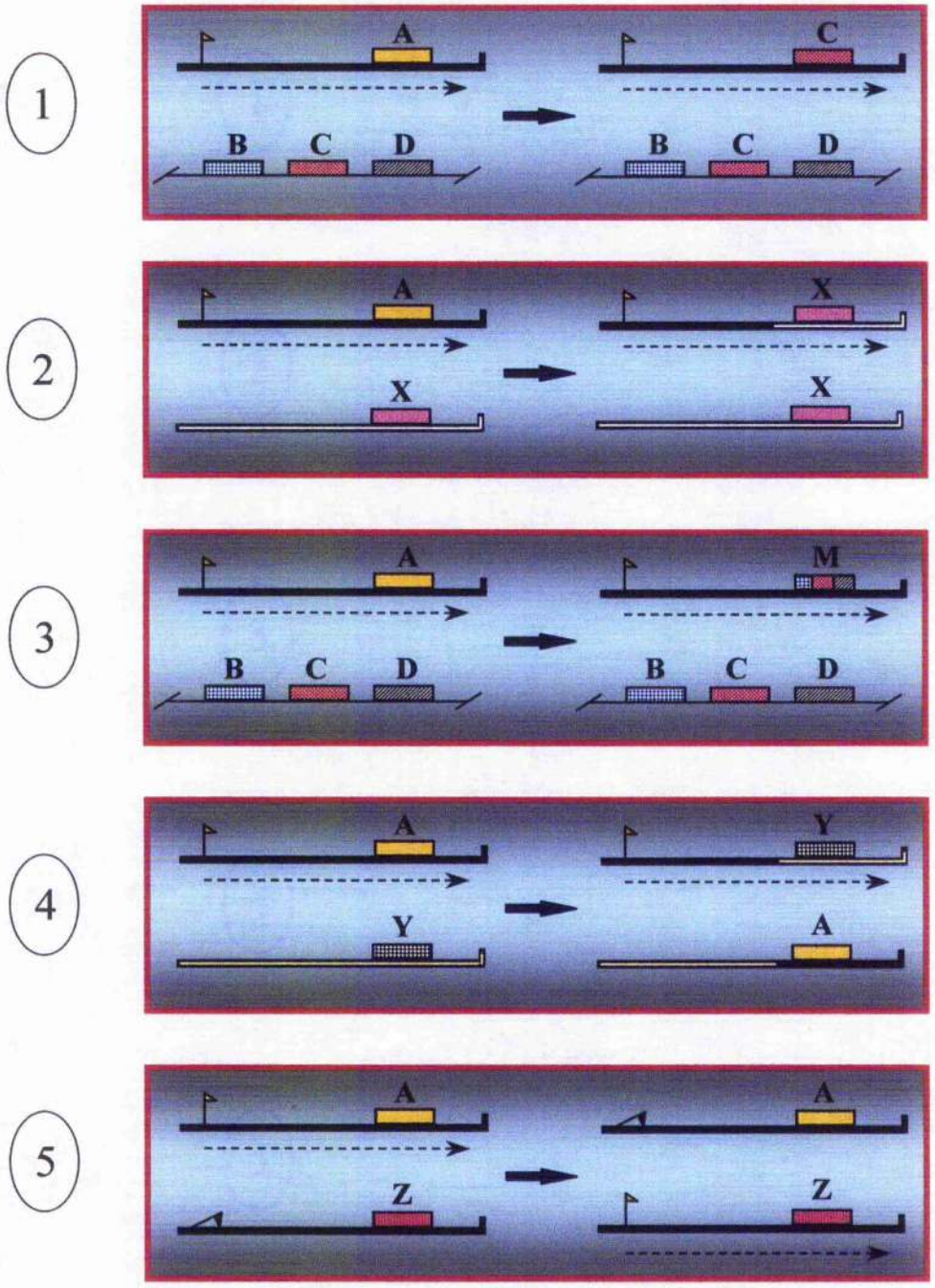


Figure 5. The mechanisms of *VSG* activation in *T. brucei* (adapted from (Borst, 1991b)): (1) duplicative transposition of a chromosome internal BC gene; (2) telomere conversion of a silent gene; (3) segmental gene conversion; (4) reciprocal telomere exchange; (5) *"in situ"* transcriptional switching. Region A depicts a *VSG* gene at the active *BES*; B, C and D represent chromosome internal BC genes; X and Y denote telomeric genes originating in silent loci; M depicts a mosaic gene; Z represents a *VSG* present at a different *BES*. Expression site promoters are denoted by flags, which are raised or lowered to represent activation / silencing respectively; transcription is indicated by an arrow under the *BES*. The vertical bars downstream of the *VSG*s represent the telomere ends.

In addition to these five mechanisms, it was thought that antigenic variation occurred during duplicative transposition by the generation of point mutations within the ELC (Lu *et al.* 1994). A more recent study, however, has concluded that point mutations are very rarely generated during the gene conversion process (Graham and Barry, 1996) and this is consistent with the earlier results of Kamper and Barbet (1992). It has also been suggested (Graham and Barry, 1996) that the high point mutation frequencies observed in earlier investigations are probably artefactual, associated with the intensive selection pressure and prolonged growth during experimentation. Indeed, Baltz *et al.* (1991) were able to generate point mutations by deliberately growing trypanosomes under strong selective pressure for prolonged periods. Graham & Barry (1996) also implied that point mutagenesis seems an unlikely mechanism for antigenic variation, since an effective switch requires that the new VSG possesses a completely novel set of exposed epitopes, and it is improbable that this could be achieved through single base changes.

1.6 The metacyclic variable antigen types (MVATs)

In the metacyclic stage of development a small, specific, subset of *VSG*s is activated (1-2% of the total *VSG* repertoire). It is at this stage of the life cycle that the parasites pass from the tsetse salivary glands into the dermal connective tissue of a new mammalian host, where a chancre develops, resulting from a local inflammatory reaction (Barry and Emery, 1984). Metacyclic VSG (*MVSG*) gene expression occurs at telomeric expression sites (*MES*s) that are separate from those used by bloodstream forms (Cornelissen *et al.* 1985b; Lenardo *et al.* 1986), and are expressed independently of the bloodstream *VSG* repertoire (Esser *et al.* 1982; Crowe *et al.* 1983; Turner *et al.* 1986). An infecting population of metacyclic parasites is polyclonal (Tetley *et al.* 1987), with the various VATs arising in fixed proportions in a particular serodeme (Hajduk *et al.* 1981), although each individual cell expresses only a single VSG in its coat. When the metacyclic parasites enter a new mammalian host, they soon differentiate to the long slender, rapidly dividing bloodstream form, but still express the *MVSG*s *in situ* from the *MES*s. These parasites, known as metacyclic-derived trypanosomes, drain from the chancre into

the lymphatic system and eventually reach the bloodstream, where they persist for approximately the first five days of infection (Esser and Schoenbechler, 1985), after which the *MES*s are silenced and the *BES*s activated at around the sixth day of infection. It is thought that the antigenically heterogeneous metacyclic population has evolved as a highly effective means of initiating and establishing an infection in a host that may have acquired immunity to various VATs in a previous infection (Barry *et al.* 1990).

Although the *MVSG* coding regions themselves are not obviously different from their bloodstream counterparts and can in fact be transposed to the *BES* during bloodstream infection, their upstream environment differs from that of other telomeric *VSG* genes. It has been demonstrated that this region is either completely devoid of the 70 bp repeats (Lenardo *et al.* 1986), or possesses just 1-2 copies of these repeats (Matthews *et al.* 1990), which presumably reduces the probability of recombination with the bloodstream *VSG* genes, thus accounting for the relative stability of the metacyclic genes. However, there is still movement of *VSG* genes through these telomeres, ensuring that the MVAT repertoire is constantly evolving (Barry *et al.* 1983); (Lenardo *et al.* 1986).

1.7 Gene expression in trypanosomes

The trypanosome genome is thought to be polycistronic and intronless, appearing almost eubacterial in its organisation (reviewed in Clayton, 1992). However, the work of Marchetti *et al.* (1998) is indicative of a more complicated genomic arrangement. This investigation demonstrated that several RNA polymerase III-transcribed small nuclear RNA genes appeared to be scattered throughout the genome, and interspersed with the closely juxtaposed RNA polymerase II transcription units; this arrangement is characteristic of eukaryotic genomes. Interestingly, the RNA polymerase III genes constituted the 3' boundary of the RNA polymerase II transcription units in the loci examined in this analysis.

A recent study has revealed that chromosome I of *Leishmania major* displays a remarkable polycistronic gene arrangement, in which the first 29 genes are all encoded on one strand, while the remaining 50 genes are encoded on the opposite strand (Myler *et al.* 1999). The only known exceptions to the polycistronic

organisation in *T. brucei* are the *MVSG* genes, which are expressed *in situ* from monocistronic, telomeric loci (Graham and Barry, 1995). Genes within a polycistron are generally packed in dense clusters and often contain tandemly arranged copies of related genes, as observed for the *T. brucei* glucose transporters (Bringaud and Baltz, 1993). However, unrelated genes can also occur within a polycistron, and this is the case for the *VSG BES* (Pays *et al.* 1989b). Although the primary transcription within a polycistron is under common control, the individual genes within the unit can display markedly different expression levels (Gibson *et al.* 1988; Bringaud and Baltz, 1993; Revelard *et al.* 1993). This indicates that gene expression is controlled primarily at the posttranscriptional level in trypanosomes.

The polycistronic primary transcripts are segmented into their constituent messengers by the RNA processing machinery (Ullu *et al.* 1993), individual mature mRNAs result resulting from the addition of a cap at the 5' end and a poly[A]$^+$ tail at the 3' end. Capping is achieved by *trans* splicing, a process first described in *T. brucei* (reviewed in Nilsen, 1995), which is mechanistically related to *cis* splicing of introns in yeast and higher eukaryotes. During *trans* splicing, the 5' ends of the nascent RNA molecules are processed by the addition of a 39 nucleotide pre-capped spliced leader, derived from a reaction with a 140 nucleotide mini-exon donor RNA (Murphy *et al.* 1986). The 3' end of mature mRNAs is generated by cleavage of the pre-mRNA at a polyadenylation site, or by trimming from the downstream splice site, followed by the addition of a poly[A]$^+$ tail, using a mechanism similar to that operating in higher eukaryotes. The *trans* splicing and polyadenylation processes appear to be coupled, and both use a polypyrimidine tract as a recognition sequence (Matthews and Gull, 1994; Vassella *et al.* 1994).

T. brucei possesses the three classical RNA polymerases observed in the other eukaryotes (Lee and Van der Ploeg, 1997). The polymerase involved in the transcription at the loci containing the *VSG* and procyclin genes appears to be resistant to very high concentrations of the drug α-amanitin, while the transcription of most other trypanosome protein-coding genes is inhibited by this drug (Kooter and Borst, 1984; Kooter *et al.* 1987). The α-amanitin insensitivity of the *VSG* and procyclin polymerase is indicative of RNA polymerase I, an enzyme complex that also transcribes ribosomal RNA genes.

1.8 Life cycle dependent stage-specific regulation of *VSG* expression

Contrary to the early assumptions that *VSG* expression could be prevented entirely by *BES* promoter repression in the procyclic cells, it has been demonstrated that the *BES* promoters remain active throughout the bloodstream and procyclic stages of the trypanosome life cycle (Zomerdijk *et al.* 1990; Rudenko *et al.* 1994). The work of Rudenko *et al.* (1994) has revealed that during the procyclic stage, a low level of transcription could be detected from many (if not all) of the *VSG* expression sites. Although this study clearly demonstrated transcripts from more than one *BES*, it did not determine whether these products were representative of transcription from all of the *BES*s, transcribed at a low level, or from just a few *BES*s transcribed at a higher level (Ansorge *et al.* 1999). Similarly, this analysis could not establish whether all of the transcriptional products originated from individual cells, or whether the heterogeneity of these transcripts resulted from a phenotypically diverse population, in which each cell produced only a single (or few) transcript(s) (Rudenko *et al.* 1994; Horn and Cross, 1997). Although this work was not entirely conclusive, it seems plausible that all the *BES* promoters display a similarly low level of transcription in the procyclic cells, and this model is widely accepted (Alarcon *et al.* 1999). Further evidence for this model was provided by an investigation by Navarro and Cross (1998). The authors placed a reporter gene 1 kb downstream of a *BES* promoter and discovered that, in the bloodstream stage, transcription from this locus differed by 1000-fold, depending on whether the *BES* was active or inactive. However, in procyclic cells the amount of transcription was independent of the previous activation state of the *BES*, and proceeded at a level 100-fold below that observed at the active *BES* in bloodstream form cells. It therefore appears that the mechanism of gene silencing operating in procyclic cells is different from that seen in the bloodstream stages, when only a single *BES* is maximally active.

The transcriptional down-regulation of the *BES*s associated with the procyclic stage occurs not only at the level of transcription initiation (Rudenko *et al.* 1994), but also by transcription attenuation (Pays *et al.* 1989a; Zomerdijk *et al.* 1990; Rudenko *et al.* 1994). This attenuation results in the transcription aborting at approximately 500-1000 bp downstream of the promoter (Pays *et al.* 1989a; Zomerdijk *et al.* 1990; Rudenko *et al.* 1994), thus preventing the expression of any gene within that polycistron. The down-regulation of *VSG* expression observed in procyclic cells

seems to be both promoter and position dependent. A *BES* promoter becomes derepressed when it is placed in the ribosomal spacer region (Rudenko *et al.* 1994), while a ribosomal promoter demonstrates no change in activity when it replaces the endogenous *BES* promoter (Rudenko *et al.* 1994; Horn and Cross, 1997).

A recent study, by Navarro *et al.* (1999), has suggested that the repression of the *BES* promoters in the procyclic form is dependent on life-stage dependent changes in chromatin structure. In this study, a *BES* promoter was substituted with a bacteriophage T7 RNA polymerase (T7RNAP) promoter in trypanosomes that had been genetically manipulated to express T7RNAP. It was discovered that T7RNAP dependent transcription was unaffected in the bloodstream form, but was repressed in procyclic cells, where the effect was demonstrable along the entire length of the *BES*. Since it is unlikely that a bacteriophage-derived polymerase would be affected by inherent transcriptional regulation, it seems probable that this repression is due to conformational changes in chromatin structure.

The control of *VSG* expression from the metacyclic repertoire, in contrast to the *BES* regulation observed in bloodstream forms, displays a strict stage specific regulation under exclusively transcriptional regulation during the parasite life cycle (Graham and Barry, 1995). More recently, it has been revealed that *MVSG* gene promoter activity is also dependent upon chromosomal location (Graham *et al.* 1998). This study analysed the effect of introducing a metacyclic promoter, linked to a reporter gene, either back into its endogenous telomere, or into the non-transcribed spacer region of ribosomal DNA. It was demonstrated that at the bloodstream stage of development the promoter was inactive at the telomere, but highly active at the chromosome-internal position, while no activity was seen at either locus in procyclic trypanosomes. The regulation of *MVSG* repertoire is the only example of "true life cycle stage-specific control of gene expression by transcription initiation" discovered thus far in the Kinetoplastida (Graham and Barry, 1995).

It was previously believed that transcription only occurred from the active *BES* in bloodstream form trypanosomes. However, it has recently been demonstrated, by analysis of nuclear RNA following UV irradiation of the cells, that an extremely low level of minor transcripts can also be detected from some of the "silent" *BES*s and *MES*s (Alarcon *et al.* 1999). An immunofluorescence assay demonstrated that the number of contaminating (non-clonal) parasites was <1 in 5000 suggesting that these

minor transcripts were not artefactual. It was also noted that, following UV irradiation, the proportion of the minor transcripts increased 2-3 fold, while the major RNA showed no corresponding increase. This discrepancy also implied that the rare transcripts were not derived from other active *BES*s in rogue cells. Evidence for this "leaky" transcription has also been observed in independent studies. For example, Ansorge *et al.* (1999) have demonstrated that, in a bloodstream form trypanosome clone expressing the 222 *VSG BES*, 20% of *ESAG6 mRNA* (which is transcribed exclusively in the *BES*s (Pays *et al.* 1989b)) originated from independent, supposedly inactive, *BES*s. In addition, an earlier study had revealed that a low level of drug-resistance was conferred when *neo* or *ble* genes were inserted 1 kb downstream from a "silent" *BES* (Navarro and Cross, 1996).

By collating the analyses of several laboratories, Alarcon *et al.* (1999) have proposed a model for the regulation of the *BES*s and *MES*s throughout the trypanosome life-cycle. They suggest an epigenetic regulatory mechanism for the control of both *BES*s and *MES*s rather than an independent, life-cycle stage specific, system for the control of the metacyclic repertoire. This hypothesis can only be substantiated by further investigations in metacyclic trypanosomes, using single-cell reporter techniques (Barry *et al.* 1998; Alarcon *et al.* 1999) and by repeating these experiments in pleomorphic bloodstream trypanosomes.

1.9 Regulation of a single active *BES*

In bloodstream form trypanosomes, only one out of the twenty *BES*s is maximally active at any particular time, while the remaining sites are "silenced", ensuring that a single VSG is expressed in the surface coat. The fundamental mechanisms behind *BES* silencing and *in situ* switching have yet to be determined, despite the growing amount of interest and research in this area. It has been observed that switching between *BES*s occurs by regulation of transcription initiation, and this is not dependent on promoter sequence specificity (Rudenko *et al.* 1995). However, the method of promoter repression remains unclear. Until very recently it was considered that *BES* repression could be controlled by telomeric silencing associated

with changes in chromatin conformation, and this was consistent with the observation that actively transcribed telomeres grow slightly faster than inactive telomeres (Pays *et al.* 1983b; Myler *et al.* 1988b). In addition, it has been shown that the active and silent *BES*s display marked differences in their sensitivity to DNAaseI (Pays *et al.* 1983b) and single-strand specific endonucleases (Greaves and Borst, 1987). However, the latest and most comprehensive analysis of *in situ* switching has suggested that the regulation of *BES* transcription cannot be controlled by telomere silencing alone, and could even be influenced by an entirely different mechanism (Chaves *et al.* 1999). This study utilized trypanosomes in which two of the *BES*s were tagged with different selectable marker genes, enabling the investigators to attempt to produce double expressors (with transcription occurring at both *BES*s, therefore resulting in a mixed VSG coat) by drug selection *in vitro*. It was found that double resistant clones arose at a low frequency, and appeared to be achieving this phenotype by rapidly switching between the two *BES*s. The authors concluded that the *BES*s were not activated / inactivated independently and suggested that the double expressors were not a stable intermediate during *in situ* switching. They also proposed that *BES*s can exist in a "pre-active" silent state, from which they can be readily activated (explaining the rapid switching between the two *BES*s during drug selection). These results indicate that the mechanism of *BES* regulation cannot be explained by telomeric silencing alone and it appears that a non-stochastic system must operate.

It has also been suggested that the novel modified nucleotide β-D-glucosyl-hydroxymethyluracil (base J) could be involved in telomere inactivation (Gommers-Ampt *et al.* 1993). This modification was originally detected following the observation that some telomeric DNA was partially resistant to digestion by the restriction digestion endonucleases *Pst*I and *Pvu*II (Bernards *et al.* 1984; Pays *et al.* 1984), and this was subsequently shown to result from presence of a previously unidentified base (Gommers-Ampt *et al.* 1991). Further analysis demonstrated that this modification correlated with base J, which substitutes approximately 0.5-1.0% of thymidine in the DNA of African trypanosomes (Gommers-Ampt *et al.* 1993). Base J appears to accumulate around non-transcribed repetitive sequences including the telomeric repeats (van Leeuwen *et al.* 1996), the 50 bp repeats that lie upstream of the *BES* promoters and the 177 bp repeats that constitute the majority of

minichromosomal sequence (van Leeuwen *et al.* 1997). This modification is also seen in the *VSG* coding sequence and associated 70 bp repeats in the silent *BES*s, but is absent from these sites at the active telomere (van Leeuwen *et al.* 1997); the changes at the silent sites are lost reversibly on transcriptional reactivation. Although base J is found in non-transcribed *VSG* genes near telomeres in bloodstream from trypanosomes, it does not occur in the chromosome-internal arrays and is completely undetectable in procyclic cells.

The association of base J with telomeric repeat sequences has led to the suggestion that a modifying enzyme could recognise and bind to these regions and operate for a distance along the neighbouring DNA (Bernards *et al.* 1984). This hypothesis is consistent with the observation that the modification becomes progressively more pronounced towards the telomere end (Bernards *et al.* 1984). Alternatively, it is possible that putative changes in chromatin structure, associated with the repetitive DNA that surrounds the *VSG* genes in the *BES*s, could act as a target for such a modifying enzyme (van Leeuwen *et al.* 1997). The fact that the active *BES* is devoid of J in the *VSG* coding sequence and flanking 70 bp repeats is suggestive of competition between transcription and DNA modification in the *BES*s. However, it has not yet been established whether base J modification is the direct cause of *BES* inactivation, or whether it is a consequence of telomeric silencing that could facilitate the repression of the inactive state (van Leeuwen *et al.* 1997).

1.10 The dynamics of trypanosome antigenic variation and hierarchical VAT (variable antigen type) expression

T. brucei is transmitted to a new mammal when metacyclic parasites are introduced into the host dermal connective tissue by the bite of an infected tsetse fly. As mentioned previously, these metacyclic trypanosomes rapidly differentiate to the long slender bloodstream form, but continue to express the metacyclic *VSG* genes over the first few days of infection. At about day 5-7 it appears that the *MES*s are silenced and transcription commences from the *BES*s, resulting in the expression of the bloodstream VAT repertoire (Esser and Schoenbechler, 1985; Barry, J. D., pers.

comm.). Often, one of the first BVATs to become activated is the major VAT ingested by the tsetse fly when it took its infective feed (Hajduk *et al.* 1981; Delauw *et al.* 1985). As the infection progresses, the parasitaemia displays a continually relapsing growth pattern consisting of recurring infective peaks, each embodying multiple VAT sub-populations, and troughs of sub-patency. The parasitaemic profile of a trypanosome infection can contrast greatly between populations in different hosts, with differences arising in the timing of emergence, number and complexity, and overall size of the relapse peaks (Barry, 1986). A typical mouse parasitaemic profile (representative of a less complicated infection) is presented in Figure 6, and displays a classical relapse pattern.

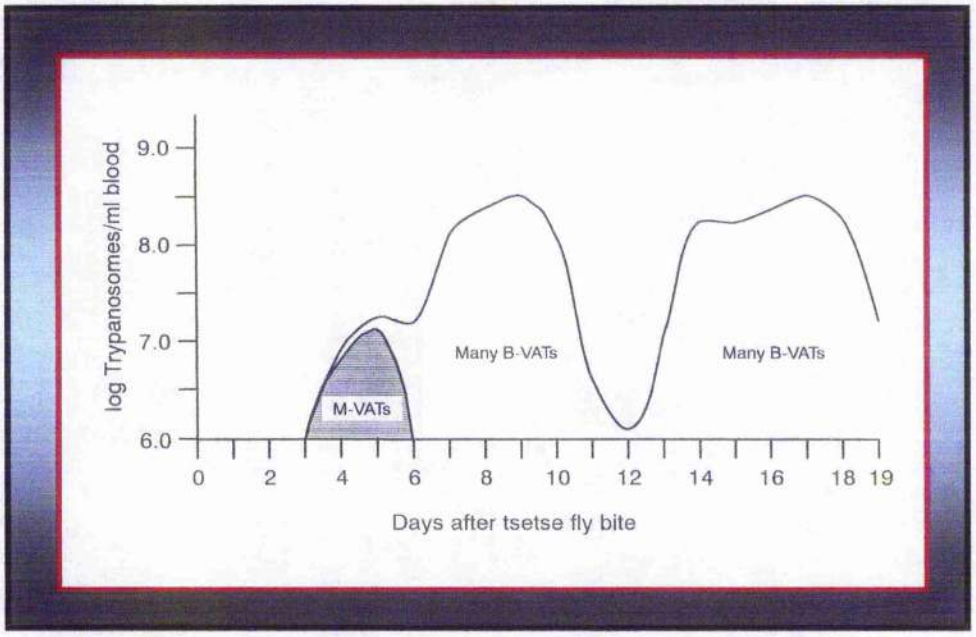


Figure 6. Typical parasitaemic profile of a trypanosome infection in a mouse host (After Barry, 1997b; data from Hajduk *et al.* 1981). The shaded sub-peak represents the expression of the metacyclic repertoire. The irregular shape of the two main parasitaemia peaks is due at least partly to the different VAT sub-populations embodied within the population, since each VAT generates a distinct sub-peak. As the infection progresses, a series of complex relapse peaks will be generated.

Investigations into the progression of antigenic variation in trypanosomes have revealed that some VATs are expressed more frequently than others, clearly demonstrating a "semi-predictable" order of antigen type appearance (Gray, 1965);

(Capbern *et al.* 1977; Miller and Turner, 1981; Timmers *et al.* 1987). It is thought that this hierarchical expression of VATs enables the parasites to maximize the infection period by ensuring efficient use of the potentially exhaustive *VSG* gene repertoire. Monomorphic studies have revealed that the early switches appear to involve predominantly telomeric BC genes (Pays *et al.* 1983a; Myler *et al.* 1984; Liu *et al.* 1985), while the chromosome-internal genes, which are presumably activated at a lower frequency, must be expressed after the initial stages of infection, since they represent the majority of the *VSG* gene repertoire (Lee and Van der Ploeg, 1987; Timmers *et al.* 1987). Towards the end of an infection, the products of relatively rare switching events, such as segmental gene conversion, become apparent (Thon *et al.* 1989; Barbet and Kamper, 1993). During an infection it is inevitable that more frequently activated *VSG* genes will become re-expressed by some trypanosomes. However, any parasite presenting a VAT that has been previously encountered by the host immune system will be eliminated, and this provides some explanation of how the less frequently activated *VSG*s arise in the later stages of infection. Although the immune response plays a significant role in controlling the population dynamics of the trypanosome infection, it is important to recognise that the actual timing of antigenic variation appears to be independent of this influence (Myler *et al.* 1985).

1.11 The minichromosomes of *T. brucei*

T. brucei possesses approximately 100 minichromosomes which vary in size between about 25 and 150 kb (Van der Ploeg *et al.* 1984b; Gibson and Borst, 1986). These chromosomes are linear and contain a tandemly repeated 177 bp element (Sloof *et al.* 1983), which constitutes approximately 80-90 % of the entire sequence. This repetitive region appears to be indigestible with most restriction enzymes (Williams *et al.* 1982; Weiden *et al.* 1991), including some that cut frequently, such as *Hae*III (Weiden *et al.* 1991). The 177 bp repeats flank a unique stretch of DNA, located in the centre of the chromosome, which extends for about 5 kb and is termed the restriction island, after the discovery of several unique restriction sites within this domain (Gull *et al.* 1998). The sub-telomeric region immediately proximal to the

177 bp repeats appear to consist of minichromosomal specific DNA (Weiden *et al.* 1991), while GC-rich repeat sequence, separated by AT-rich spacers, have been detected between the specific sequence and the telomeric repeats (Weiden *et al.* 1991). The telomeres themselves consist of a conserved $(CCCTTA)_n$ repeat that is also represented in the telomeres of the larger chromosomes (Van der Ploeg *et al.* 1984a). *VSG* genes are the only protein encoding genes that have been discovered thus far on minichromosomes (Van der Ploeg *et al.* 1984b; Weiden *et al.* 1991), and these provide the trypanosome with a large pool of silent telomeric BCs.

There has been no evidence of the *VSG* genes becoming transcribed from minichromosomal loci, and no *VSG* promoters have been discovered thus far. However, Zomerdijk *et al.* (1992) demonstrated ribosomal promoters on two minichromosomes in 427 monomorphic trypanosomes. One of these promoters was located adjacent and internal to the telomeric repeats, which has led to speculation that this region is a product of telomere healing, possibly by *de novo* telomere addition. The authors targeted a *neo* gene behind this promoter, and subsequently demonstrated that it could efficiently mediate *neo*-mRNA synthesis.

1.12 Differences between monomorphic and pleomorphic trypanosome lines

Trypanosomes rapidly become attenuated in the laboratory as a result of repeated syringe passaging. These monomorphic lines no longer differentiate to the stumpy bloodstream form and under normal circumstances cannot be transmitted by tsetse fly bite. Additionally, monomorphic trypanosomes display a marked drop in the overall rate of antigenic variation, which, at $1 \times 10^{-6} - 1 \times 10^{-7}$ switches/ cell/generation (Lamont *et al.* 1986), is up to five orders of magnitude lower then in the non-adapted, pleomorphic lines (Turner and Barry, 1989; Turner, 1997). Although this reduced switch rate has facilitated studies into antigenic variation, allowing cloned populations to remain predominantly of one VAT during expansion, it is important to consider the possibility that the selection to monomorphism is associated with alterations in the cell's recombinational or transcriptional machinery.

The diminished rate of antigenic variation in monomorphic trypanosomes is low enough to be explained by background mutation and mitotic homologous recombination, and this raises questions about the significance of the *VSG* gene activation events observed in these lines. Barry (1997a) proposed that the difference in *VSG* gene switching between monomorphic and pleomorphic trypanosomes is indicative of a specific gene switching mechanism that might include a dedicated enzyme activity catalysing specific recombination, which is reduced in, or even absent from, monomorphic lines (further details are given in sections 3.1, page 52 and section 6.5, page 146).

In monomorphic trypanosome lines, *in situ* switching predominates during the early stages of infection (Liu *et al.* 1985), while duplicative events become more common only as the infection progresses (Michels *et al.* 1983; Lee and Van der Ploeg, 1987; Timmers *et al.* 1987). Both *in situ* and duplicative events have been observed in the relatively few analyses performed using pleomorphic trypanosomes (Delauw *et al.* 1987; Shah *et al.* 1987; Matthews *et al.* 1990), but the relative frequencies of these two activation mechanisms have not been determined at any stage of the expression hierarchy. If duplicative events are controlled by a dedicated mechanism in pleomorphic lines, then it is logical to expect a higher proportion of duplicative switching, relative to *in situ* transcriptional switches, throughout the entire

expression hierarchy, when comparing pleomorphic with monomorphic populations. The main objective of the work presented in this thesis was to examine this hypothesis, by asking three key questions regarding trypanosome antigenic variation:

1. How are *VSG* genes activated during early infection, and is there a predominant switching mechanism?
2. Where are the *VSG* gene donors located, and does their chromosomal environment influence their timing of appearance in the *VSG* gene expression hierarchy?
3. Which *BES*s are involved in the early switching events, and what are the boundaries of duplication during gene conversion?

CHAPTER 2

MATERIALS AND METHODS

Reagent Abbreviations

AIX selective plates for blue/ white colony screening:
L-agar supplemented with 0.27 M ampicillin (Sigma, Ltd.), 0.5 mM IPTG (Boehringer Manheim), 0.2 M X-Gal (Boehringer Manheim)

CBSS Carter's Balanced Salt Solution (1 x):
0.023 M HEPES, 0.12 M NaCl, 5.41 mM KCl, 0.55 mM $CaCl_2$, 0.4 mM $MgSO_4$, 5.6 mM Na_2HPO_4, 0.035 M glucose, 0.05 mM phenol red, pH to 7.4

DEPC diethyl pyrocarbonate:
Used at 0.1% to remove RNAase

DMSO dimethyl sulphoxide

EtBr ethidium bromide

IPTG isopropyl-β-D-thiogalactopyranoside

NDS solution used for the manufacture of genomic plugs (1 x):
0.5 M EDTA, 0.5 M TRIS base, 0.5 M NaOH, 17 mM lauroyl sarcosine, pH adjusted to 8 or 9 with NaOH

PBS phosphate buffered saline (Sigma, Ltd.)

PSG phosphate/ sodium chloride/ glucose buffer (1 x):
0.06 M Na_2HPO_4, 3.6 mM NaH_2PO_4, 46 mM NaCl, 55 mM glucose, pH to 8

SDS sodium dodecyl sulphate

SOC SOB broth + 20 mM glucose

SOB Bacterial media (per litre):

20 g bacto-tryptone, 5 g bacto-yeast extract, 0.5 g NaCl

SSC sodium chloride/ sodium citrate solution (1 x):

0.15 M NaCl, 0.015 M $Na_3C_6H_5O_7.2H_2O$

TAE TRIS/ acetate/ EDTA buffer (1x):

0.04 M TRIS base, 0.04 M glacial acetic acid, 1 mM EDTA

TE 10 mM Tris.Cl, 1 mM EDTA

TBE TRIS/ borate/ EDTA buffer (1x):

0.089 M TRIS base, 0.089 M ortho-boric acid, 2 mM EDTA

X-Gal 5-bromo-4-chloro-3-indolyl-β-D-galactoside

2.1 Routine handling of trypanosomes

The two trypanosome clones utilized in this study were derived from a *Trypanosoma brucei* EATRO (East African Trypanosomiasis Research Organization) 795 line that is tsetse fly transmissible. The ILTat (I.L.R.A.D. – International Laboratory for Research in Animal Diseases – *Trypanozoan* antigen type) 1.2 expressor clone switches VSG at about 1 x 10^{-5} switches/cell/generation and the ILTat 1.61c clone, derived from a single metacyclic trypanosome, switches at about 3 x 10^{-2} switches/cell/generation (Turner, 1997). The ILTat 1.2 infections were undertaken in lop-eared or New Zealand White rabbits (Bantin and Kingman, Hull, U.K.), while the ILTat 1.61c investigation involved single relapse infections in ICR mice (Bantin and Kingman, Hull, U.K.). The rabbit infections were established by the intravenous injection of approximately 1 x 10^{6} clonal trypanosomes (previously grown in an immunosuppressed mouse infected with the ILTat 1.2 stabilate), while the single relapse infections were initiated by an intraperitoneal injection of the ILTat 1.61c stabilate into an immunosuppressed mouse; subsequent cloning steps were performed in non-immunosuppressed mice. Full experimental details and clone histories are given in sections 3.2-3.3 and 3.10 (pages 53 and 74 respectively). Routine procedures are described below.

2.1.1 Host immunosuppression, trypanosome growth and collection, and stabilate manufacture

Trypanosomes were grown from stabilate in ICR mice that had been immunosuppressed by cyclophosphamide treatment (250 mg.kg^{-1} body weight, Sigma Ltd.) 24 h previously. Unless otherwise stated, exsanguination was performed at the initial parasitaemic peak (typically 3-5 days post infection) by cardiac puncture into 5% sodium citrate anticoagulant in CBSS (Fairlamb *et al.* 1992) (0.1 ml of 5% citrate used per 0.9 ml collected blood).

Stabilates were prepared immediately by mixing the blood 2:1 with 22.5% DMSO (in CBSS), and freezing the sample gradually to -70°C over a 24 h period; the stabilates were then transferred to liquid nitrogen for storage. When a stabilate was required for infection a tube of it was thawed rapidly at 37°C, mixed with 0.1 ml 5%

foetal calf serum (in CBSS) and injected immediately into the immunosuppressed animal.

If a large number of parasites were required (*e.g.* for preparing genomic DNA), the infection was established in an immunosupressed ICR mouse and the blood collected at the initial peak as before, but then divided equally and injected intraperitoneally into two Wistar rats (Bantin & Kingman, Hull, U.K.). The trypanosomes were then harvested at the initial peak, typically 24-48 h later.

2.1.2 Trypanosome cloning

Trypanosome clones were isolated immediately from the blood by micromanipulation. A dilution series of the trypanosomes was prepared in 5% foetal bovine serum (GibcoBRL, Life Technologies) in CBSS, in the wells of a Terasaki dish (Greiner labortechnik). These dilutions were examined using a binocular microscope to establish the number of trypanosomes in each well. When a well containing a single trypanosome was found, 5 µl 5% foetal bovine serum (in CBSS) was carefully pipetted into that well. The entire contents were transferred by micropipette into 100 µl 5% foetal bovine (in CBSS) in a 1.5 ml eppendorf tube. This liquid was then drawn into a syringe containing 0.4 ml of air, and was injected immediately, intraperitoneally, into an ICR mouse, using the air to push the last drop of liquid from the syringe. The Terasaki plate was kept moist by lining the edges with wet tissue, and the procedure was undertaken in an atmospherically controlled environment (14°C, 80-95% humidity).

2.1.3 Preparation of blood smears for immunofluorescence and plasma for immune lysis

Thin-film blood smears were prepared by placing a drop of blood (mixed 1:1 with CBSS) on a glass microscope slide. The blood was spread along the length of the slide using a slide edge and left to air-dry. The smears where then fixed in acetone for 5 min, air-dried, and stored at 4°C sealed in polythene bags containing silica gel (to remove any moisture). If trypanosomes were not being collected for other

purposes, the small amount of blood needed for these smears was taken from the tail of the mouse using a surgical lancet.

Plasma was prepared directly from the rabbit blood. On each day of infection, in all four rabbits, 0.4 ml of blood was isolated and this was spun at maximum speed in a microcentrifuge for 10 min, to pellet the trypanosomes and the larger blood components. The supernatant was then collected and sodium azide was added to a final concentration of 0.2% (to prevent microbial growth during storage). Plasma (containing specific antibodies) was also prepared from the clonal mice in the ILTat 1.2 study (see section 3.3, page 53). At the initial peak these mice were drug-cured with cymelarsan (5 $mg.kg^{-1}$ body weight; Rhône Mérieux) and the blood was collected three days later (allowing sufficient time for antibodies to be raised); the plasma was then prepared as above. All the plasma samples were stored in 1.5 ml screw-cap eppendorf tubes at 4°C.

2.2 Serology

2.2.1 Immune lysis assay

Trypanosomes were suspended in guinea-pig complement to 5×10^6 $cells.ml^{-1}$. The plasma isolated from the clone mice (containing specific antibodies) was diluted 1:20 in guinea-pig complement (Harlan) and mixed with the trypanosome dilution in the bottom of a V-welled microtitre dish (Greiner labortechnik). Incubation was undertaken at room temperature for 1 h after which the contents of the well were mixed again. A drop (5µl) of the reaction was then placed on a microscope slide and sealed with a glass coverslip. Percentage lysis was then determined by phase-contrast light microscopy, with destroyed cells appearing as ruptured "ghosts". Each trypanosome clone was also incubated for an hour in the presence of guinea-pig complement alone, as a negative control, to ensure that all the lysis was due to the action of antiserum. The assay was performed within an hour of blood collection and the trypanosomes were stored on ice while the dilutions were prepared.

2.2.2 Immunofluorescence

Indirect immunofluorescence was performed on the acetone-fixed thin blood smears that were prepared for each of the clone mice. Reference antisera, which originally had been derived in either mouse or rabbit hosts, were used separately as the primary antibody (see section 3.6, page 63). Prior to the experiment, the slides were marked with a hydrophobic paint pen (Mark-Tex corp., BDH) to produce wells for antibody containment. The smears were then rehydrated for 5 min in PBS (Sigma, Ltd.), drained and dried between the wells, and transferred to a humid chamber. A drop (5µl) of the appropriately diluted primary antibody was then added to the wells and incubated at room temperature for 30 min (details of primary antibodies and dilutions are indicated in Table 3, section 3.6, page 64). The slide was then drained and washed twice with PBS, and dried between the wells. Afterwards, 5µl anti-rabbit or anti-mouse IgG fluorescein isothiocyanate (FITC) labelled conjugates (Promega) (diluted 1:100 or 1:50 respectively) were introduced into the wells and incubated at room temperature for 15 min. The slide was again washed twice in PBS and a few drops of PBS:glycerol (1:1) were added, after which the wells were sealed with a glass coverslip. Fluorescence was then determined by UV light microscopy using an arbitrary scoring system: - no fluorescence, (+) faint fluorescence, + slight fluorescence, ++ clear fluorescence and +++ strong fluorescence. Positive controls were also performed using the VAT specific mouse plasma, generated from the clonal mice, as the primary antibody. The plasma in these controls was diluted in 1 x PBS to 1:80, previously determined by an optimisation series.

2.3 Reverse transcription polymerase chain reaction (RT-PCR)

2.3.1 RNA isolation

RNA was isolated directly from infected blood (after the parasitaemia achieved 1 x $10^{7.8}$ cells.ml^{-1} or greater (Herbert and Lumsden, 1976)) with TRIzol (Gibco-BRL, Life-Technologies). The blood was mixed with TRIzol (0.1ml blood: 1 ml TRIzol) and incubated at room temperature for 5 min. Following incubation, 200 µl of

chloroform was added and the tube was shaken vigorously for 15 s before incubation at room temperature for 2-3 min. The solution was then spun at maximum speed in a microcentrifuge (cooled to 4°C) for 15 min and the colourless aqueous phase was eluted. Afterwards, 1 µl of (10 $mg.ml^{-1}$) glycogen and 500 µl of isopropanol were mixed with this eluate, and the reagents were then incubated at room temperature for 10 min. The RNA was then pelleted by centrifugation for 10 min at 4°C, after which the supernatant was removed. The pellet was washed with 1 ml 75% ethanol, air-dried and dissolved in 11 µl DEPC-treated water.

2.3.2 Reverse transcription

cDNA was generated by reverse transcription using Superscript II reverse transcriptase (GibcoBRL, Life Technologies). 1 µl of 0.5 $\mu g.\mu l^{-1}$ oligo[dT] primer (GibcoBRL, Life Technologies) was added to the 11 µl containing RNA (see Section 2.3.1, page 38) and heated to 70°C for 10 min. After cooling the solution on ice, 2 µl 10 x PCR buffer (Advanced Biotechnologies), 2 µl 25 mM $MgCl_2$, 1 µl 10 mM dNTPs (Pharmacia Biotech) and 2 µl 0.1 M DTT were added to the reaction, which was mixed and warmed to 42°C for 5 min. Next, 1 µl of Superscript II (200 $U.\mu l^{-1}$) reverse transcriptase was added and the tube incubated at 42°C for 50 min. The enzyme was then heat inactivated at 70°C for 15 min after which the solution was cooled on ice. Finally, 1 µl of RNAase H (3.8 $U.\mu l^{-1}$) was added and the reaction incubated at 37°C for 20 min to remove the single stranded RNA.

2.3.3 PCR amplification of *VSG* specific cDNA

DNA representing the expressed *VSG* gene was amplified from the cDNA using two short oligonucleotides (manufactured by Cruachem) as described by Carrington *et al.* (1991). The first primer (tbsl: GTTTCTGTACTATATTG) was a 17mer specific to the mRNA spliced leader, while the second primer (tb3ut: GTGTTAAAATATATCA) was an antisense 16mer specific to a region in the downstream untranslated region (UTR), highly conserved in all *VSG* genes. The 50 µl PCR reactions consisted of the following components: 5 µl 10 x PCR buffer (Advanced Biotechnologies), 3 µl 25mM $MgCl_2$, 1 µl 10 mM dNTPs (Pharmacia

Biotech), 2 µl 5 µM tbsl, 2 µl of both 5 µM primers, 0.5 µl (5 $U.\mu l^{-1}$) *Taq* polymerase (Advanced Biotechnologics), 1 µl cDNA from the RT step and 35.5 µl sterile, distilled water. This reaction mix was contained within a thin-walled microeppendorf tube (Perkin-Elmer corporation), and was overlaid with a drop of mineral oil (Sigma Ltd.). Amplification was performed in a Stratagene Robocycler96 for 30 cycles of 1 min at 96°C, 1 min at 42°C and 2 min at 70°C; the reaction began with a "hot start" at 96°C for 5 min and ended with a 5 min extension period at 70°C. An aliquot (usually 10 µl) of the PCR product was run on a 0.7% agarose gel at 100 V for analysis.

2.4 Cloning of PCR amplified *VSG* specific cDNAs

2.4.1 Ligations

Following fractionation on a 0.7% agarose gel, the *VSG* specific PCR products were gel purified using the QIAgen gel extraction kit (following the manufacturer's protocol) and were cloned with the "T-easy" vector (Promega) or PCR Script Amp SK(+) plasmids (Stratagene) (following the manufacturer's protocol); the ligations were performed either overnight at 4°C ("T-easy"), or at room temperature for 1 h (PCR Script). At each cloning attempt several different insert:vector ratios were prepared (*e.g.* 1, 2, 3 and 5 µl of purified PCR product) to maximize the probability of a successful ligation. The "T-easy" vector possesses an open multiple cloning site with a 3' terminal thymidine at either end. This enables *Taq* polymerase generated PCR products, which contain 3' terminal adenines, to be ligated directly into the site. The PCR Script system works by adding an infrequently cleaving restriction enzyme, which cuts at a site within the plasmid polylinker (producing blunt overhangs) into the ligation mix. PCR products generated by *Taq* polymerase must be "polished" (to remove the 3' A overhangs) using *Pfu* polymerase before the ligation step.

2.4.2 Transformations and plasmid retrieval

Transformations were performed by heat shocking competent *E. coli* DH5α cells (GibcoBRL, Life Technologies). Approximately one-third (usually 3 μl) of the ligation product was added to a 100 μl aliquot of competent cells (previously thawed on ice), mixed gently, and left on ice for 30 min. The cells were then heat shocked at 42°C for 60 s, and immediately transferred to ice for 1-2 min. Afterwards, 0.9 ml of SOC was then added to the cells, which were subsequently placed at 37°C for 1 h. After this incubation period, the cells were spun down, resuspended in 100 μl SOC, and spread on AIX selective L-agar plates, allowing blue/white colony screening (AIX plates: 0.27 M ampicillin (Sigma, Ltd.), 0.5 mM IPTG (Boehringer Manheim), 0.2 M X-Gal (Boehringer Manheim)). Transformed colonies were re-streaked on fresh AIX plates, after which a single colony was selected to inoculate 5 ml L-broth (supplemented with ampicillin to a final concentration of 10 $\mu g.ml^{-1}$). The plasmids were then retrieved from 3 ml of this overnight culture using the QIAGEN Mini kit (following the manufacturer's protocol), and the insert size was determined by restriction digestion with *Not*I ("T-easy" vector) or *Not*I/ *Eco*RI (PCR Script) of 300 ng plasmid (the enzymes were obtained from New England Biolabs). Resultant products were visualised on a 0.7% agarose gel.

2.5 Isolation and purification of trypanosomes from blood

During some procedures, such as preparing genomic DNA, it was necessary to separate the trypanosomes from blood components, and this was achieved on a Percoll gradient (Grab and Bwayo, 1982). A Percoll stock solution was prepared by adding 8.55 g of sucrose and 2.00 g of glucose to 100 ml of 100% percoll (Sigma, Ltd.), and the pH was adjusted to 7.4 by adding solid HEPES. As soon as the blood was collected, an equal volume of Percoll stock solution was added and the solution was mixed. Another 4 volumes of Percoll stock solution and 2 volumes of CBSS were then added and mixed, and this solution was then transferred to 50 ml centrifuge tubes (Nalgene, BDH) (filled to at least 80% of the total tube volume). The tubes were then centrifuged in a JA20 (Beckman) rotor (with a fixed angle of

34°) at 17500g for 15 min at 4°C. This caused the trypanosomes to separate from the blood cells, and formed a discrete band near the top of the gradient. The parasites were then collected with a disposable plastic pipette. Following pelleting (by centrifugation in a swing out rotor at 1500g for 5 min), the trypanosomes were washed in CBSS to remove the Percoll; this washing step was repeated at least once to ensure that no traces of Percoll remained.

2.6 Isolation of genomic DNA

2.6.1 Preparation of genomic plugs from live trypanosomes

It was necessary to produce minimally sheared genomic DNA, encased in agarose plugs, for the pulsed field gel electrophoresis (PFGE) analysis. Following their isolation on a Percoll gradient, trypanosomes were spun down gently (1500g) at room temperature and carefully resuspended in PSG buffer. This wash was repeated once more, and the concentration of the cells was determined using a haemocytometer under a phase contrast microscope. The trypanosomes were then spun down again and resuspended in PSG to a concentration of 1 x 10^9 cells.ml^{-1} (twice the final concentration of the plug) and warmed to 37°C for 1 min. Equal volumes of the trypanosome suspension and a 1.4% low-melting point agarose (InCert agarose, FMC Bioproducts) solution (which had been melted and cooled to 37°C previously) were mixed together by swirling, resulting in a final concentration of 5 x 10^8 cells.ml^{-1} in 0.7 % agarose. Next, 100 µl of this suspension was pipetted into each well of a plug mould (BIO-RAD) and allowed to set. The plugs were then placed in NDS at pH 9.0, supplemented with 1 mg.ml^{-1} proteinase K (Sigma, Ltd.), at 50°C for 24 h. After this period, the plugs were transferred to NDS at pH 8.0, and supplemented with 1 mg.ml^{-1} proteinase K, at 55°C for 24 h. The plugs were then stored in fresh NDS at pH 8.0 at 4°C.

2.6.2 Standard genomic DNA preparations

Trypanosomes were separated on a Percoll gradient, as described previously, and then washed in ice cold digestion buffer (10 mM Tris-HCl at pH 7.5, 1 mM EDTA and 100 mM NaCl). The cells were then centrifuged at 1500*g* and resuspended in 5 ml digestion buffer, supplemented with SDS to 1% and proteinase K to 1 $mg.ml^{-1}$ and incubated for at least 2 h at 55°C; the solution was swirled gently every 15 min. RNAase was then added to a final concentration of 100 $\mu g.ml^{-1}$ and incubated at 50°C for at least an hour. The genomic DNA was then phenol/chloroform extracted 2-3 times and chloroform extracted once; the phases were separated by centrifuging at 10 000*g* for 10 min. Wide-bore disposable plastic pipettes were used to aspire the aqueous phase, to ensure minimal shearing. Afterwards, 1/20 volume of 3 M sodium acetate was added, and the DNA was then ethanol precipitated by layering 2 volumes of 99% ethanol onto the solution and swirling the tube gently; the DNA eventually contracted into a tight pellet. The pellet was then spooled using a sealed Pasteur pipette hook and placed in 70% ethanol to remove salts. After repeated ethanol washes, the pellet was drained and air-dried, after which it was dissolved in TE to a final concentration of 1 $\mu g.\mu l^{-1}$.

2.7 Gel electrophoresis and Southern blotting

2.7.1 General gel electrophoresis

Standard DNA separations were performed on 0.7% agarose gels (GibcoBRL, Life Technologies) run at 100 V in 1 x TAE buffer, using a commercial 1 kb ladder as a size marker (GibcoBRL, Life Technologies).

2.7.2 Genomic digestions

Genomic DNA (usually 1 µg) was digested overnight with the appropriate restriction enzymes, following the manufacturer's protocol (New England Biolabs). The products were then fractionated on a 0.7 % agarose gel run at 30 V (overnight in

1 x TAE buffer) to ensure a high resolution of the bands, after which the DNA was transferred to a nylon membrane by Southern blotting (see below).

2.7.3 Pulsed field gel electrophoresis (PFGE)

Chromosome sized DNA was resolved by PFGE using the CHEF-DR III system (BIO-RAD). This method of electrophoresis is highly sensitive to changes in buffer composition and it is therefore important to ensure that the gel is made from (and the genomic plugs dialysed in) the buffer that is circulating in the tank (and not from a fresh buffer stock). At least 2 litres of the appropriate buffer were circulated in the electrophoresis tank for 10 min, after which a small sample was taken for dialysing the genomic plugs. The plugs were dialysed, at room temperature, in 1 ml of buffer, which was changed every hour for four hours. After the final change of buffer, the plugs were placed at 4°C and left overnight. The following day, 105 ml of buffer were taken from the tank and the agarose gel was prepared at the appropriate concentration. Following dialysis, the genomic plugs were placed on the comb of the gel former (one-half a genomic plug per lane) and 100 ml of the molten agarose was then poured, and allowed to set. After 30 min, the comb was removed (leaving the plugs embedded in the gel) and the wells were filled with the remaining liquid agarose. The three PFGE conditions utilised in this thesis were as follows: (i) 6 –day general separation, 1.2% agarose gel at 15°C in 0.089 M Tris-borate, 0.1 mM EDTA (1 x TB(0.1)E) (85 V, 1400-700 s pulse time, 144 h); (ii) 1.8 Mb resolving run, identical to the general run except the pulse time was fixed at 600 s; (iii) minichromosomal separation, 1.0 % agarose gel at 14°C in 0.045 M Tris-borate, 0.5 mM EDTA (0.5 x TBE) (200 V, 20 s pulse time, 16 h). The DNA was then transferred to a nylon membrane by Southern blotting.

2.7.4 Southern blotting

Prior to blotting, the agarose gels were stained with ethidium bromide and viewed and photographed on a UV transilluminator. The gel was then placed in 200 ml 0.25 M HCl for 15 min (to nick the DNA), rinsed with distilled water, and immersed in 200 ml denaturation solution (0.5 M NaOH, 1.5 M NaCl). After 30 min of denaturing, the gel was rinsed briefly with distilled water and then placed in 200 ml neutralizing solution (1 M Tris-HCl pH 8.0, 1.5 M NaCl) for 30 min. The DNA was then transferred to a nylon membrane (Micron Separations, Inc. or Hybond-N) by wet blotting using 20 x SSC as the transfer buffer (Sambrook *et al.* 1989); standard blots were left for 24 h, while PFGE blots were left for 48 h. After transfer, the DNA was crosslinked to the membrane using a UV spectrolinker.

2.8 Probe manufacture and DNA hybridization

2.8.1 Radiolabelling

The majority of the probes used in this study were manufactured from the full-length *VSG* specific cDNAs, which were excised from their plasmids by restriction digestion and gel purified using the QIAGEN gel extraction kit (following the maufacturer's protocol); radiolabelling was performed using the Prime-It II kit (Stratagene). Initially, 50 ng of purified template DNA were mixed with 10 µl of random oligonucleotides and sterile, distilled water, in a total reaction volume of 37 µl. This mixture was then heated to 95-100°C for 5 min, cooled, and centrifuged briefly. Afterwards, 10 µl 5 x primer buffer, 2 µl α-^{32}P labelled dCTP and 1 µl Klenow (5 U.µl^{-1}) were added in order, mixed carefully, and incubated at 37°C for 5 min. The resultant probes were then purified from the unincorporated nucleotides by passing them through NucTrap columns (following the manufacturer's protocol, Stratagene). Once purified, the probes were denatured at 95°C for 5 min before use.

Three probes corresponding to the amino-terminus of the ILTat 1.2, 1.21 and 1.25 VSGs were also produced. Template DNA was generated by PCR amplification from the plasmids using the tbsl (spliced leader 17mer) primer (Cruachem) and a reverse and complementary *VSG* specific internal primer (Genosys Biotechnologies Ltd.). The three internal 20mer primers were:
NILT1.2, ACATCTGACGCCACCGCT (1 kb into the ILTat 1.2 coding sequence);
ASALI, TTTTGGTGTATTAGCGCCGC (1 kb into the 1.25 coding sequence);
EAGEI, CGTCCCTTGGTGTCGCCGCC (0.4 kb into the 1.21 coding sequence).
PCR amplification was performed as in section 2.3.3, but the annealing temperature was performed at 50°C and the elongation phase of the cycle was shortened to 1 min. The PCR products were then gel purified using the QIAGEN gel extraction kit (following the manufacturer's protocol) and subsequently radiolabelled.

2.8.2 Hybridization

The nylon filters were wetted with distilled water, placed between two sheets of hybridization mesh (Amersham, Life Technologies) and transferred to a glass hybridization tube. Approximately 20 ml of Church-Gilbert solution (0.342 M Na_2HPO_4, 0.158 M $NaH_2PO_4.2H_2O$, 0.257 M SDS and 1 mM EDTA per litre) was added and the filters were prehybridized for a minimum of 1 h at 65°C in a rotating hybridization oven. The purified, denatured probe was then added, and the hybridization was left overnight at 65°C. After this hybridization step, the filters were washed at 65°C in the rotating oven with the following series of solutions: 5 x SSC, 0.1% SDS (twice); 2 x SSC, 0.1% SDS; 1 x SSC, 0.1% SDS; 0.1 x SSC, 0.1% SDS (50 ml solution used per 15 min wash). The filters were then rinsed at room temperature in 0.1 x SSC (without SDS), heat-sealed in plastic, and placed next to medical photographic film (Konica Medical Corporation) in an autoradiography cassette at -80°C for 4-168 h, depending on the expected strength of the signal.

2.8.3 Stripping of nylon filters

The nylon filters were stripped with boiling 0.1% SDS. The solution was poured onto the filters in a heat resistant container and allowed to cool to room temperature. The procedure was then repeated again, after which the filter was rinsed in 2 x SSC and was ready for reuse. The PFGE filters required a stronger stripping solution to remove the probe (0.4 M NaOH and 0.1% SDS). Following stripping the filters were placed next to medical photographic film in an autoradiography cassette at -80°C for 24 h, to ensure that the procedure had been successful.

2.9 Manufacture and screening of the minichromosomal libraries

2.9.1 Isolation of minichromosomal DNA

Minichromosomal DNA was separated on a 1.2% low-melting point (LMP) agarose gel (NuSieve GTG agarose, Flowgen) by PFGE. This gel was run at 15°C in TB(0.1)E (0.089 M Tris-borate, 0.1 mM EDTA) (85 V, 1400-600 s pulse time, 120 h) using all fourteen of the gel wells (one-half of a genomic plug per lane). The minichromosomes were then excised from the gel and dialysed overnight in 10 ml of TE at 4°C and then for 3 h at room temperature in 10 ml of TE at pH 6.5. The agarose (2.2g isolated in this experiment) was then washed twice for 30 min on ice with 5 ml of β-agarase I buffer (New England Biolabs), after which the buffer was removed. The gel slices were then melted by incubation at 65° C, cooled to 40°C, and incubated with 2μl of β-agarase I (10 $U.\mu l^{-1}$; New England Biolabs) at 40°C for 1 h. After the agarase digestion, the DNA was cleaned by phenol/chloroform and chloroform extractions, and then it was ethanol precipitated and resuspended in 50 μl TE. The amount of DNA produced from this preparation was 0.9 μg, determined by UV spectrophotometry.

2.9.2 Manufacture of the minichromosomal libraries

The DNA was divided into two aliquots and then digested overnight at 37°C with either *SalI*/*Nsi*I or *SalI*/*Bgl*II (New England Biolabs) using 1.5 µl of each enzyme (*Bgl*II and *Nsi*I at 10 $U.\mu l^{-1}$; *Sal*I at 20 $U.\mu l^{-1}$) in *Nsi*I or *Sal*I unique buffers (New England Biolabs) respectively, in a total reaction volume of 30 µl. After digestion, the DNA was phenol/chloroform extracted and then chloroform extracted. The DNA was then ethanol precipitated, and dissolved in 15 µl sterile, distilled water. The *SalI*/ *Nsi*I minichromosomal library was then prepared by ligating 30, 60 or 120 ng of the *SalI*/ *Nsi*I-digested minichromosomes into 50 ng of pBluescript K/S (+/-) (Stratagene), which had previously been digested with *SalI*/ *Pst*I (*Pst*I produces the same 4 base restriction overhang as *Nsi*I); 1 µl T4 DNA ligase (Promega, 3U. μl^{-1}) was used per ligation, in a total reaction volume of 10 µl (incubated overnight at 16°C). A negative control ligation was also prepared (with no insert) to ensure that the plasmid did not self-ligate. Next, 3 µl of each ligation was used to transform 62.5 µl XL1-BLUE MRF' *E. coli* supercompetent cells (Stratagene) by heat shocking (see section 2.4.2, page 40) and subsequently spread on AIX plates and incubated overnight at 37°C. The colonies were then washed from the 3 plates by adding 8 ml L-broth (supplemented with ampicillin to a final concentration of 10 $\mu g.ml^{-1}$) per plate. A 1 ml sample of this liquid library was then taken as a working stock and stored at 4°C. Glycerol was added to the remaining stock (15% final concentration), which was then divided into 2 ml samples and stored at -80°C. The *SalI*/ *Bgl*II library was prepared in a similar manner but ligated into pBluescript that had previously been digested with *SalI*/ *Bam*HI (*Bam*HI produces the same 4 base restriction overhang as *Bgl*II). The cloning efficiency of this library was lower than that of the *SalI*/ *Nsi*I library, and consequently two transformations were prepared from each ligation to increase the number of colonies.

2.9.3 PCR screening of the minichromosomal libraries

This method is described in section 5.4 (page 114). The sequences of the primers used in this screening method were as follows:

T7, GTAATACGACTCACTATAGGGC; T3, AATTAACCCTCACTAAAGGG; E200, CAGCTCCCTGTCTTCACTCC; E890, TAACGCAGTTGCAAGCATTG; A180, AATACAGGTTGTTGTTGCGA; A550, GTCTCTGCAAGAAATGCTAA; C190, TGCGCTAGAGTGTTTGTAGC; C390, CTGTCGCGATGCCATACAAA; D210, GTTAGCTTGTTGGTTGCGTA; D370, TGCCAGGTTGTGATACTTCT.

The primers were manufactured by Genosys Biotechnologies Ltd., and were diluted to 5 μM. The PCR reactions were set up as described in section 2.3.3 (page 39), and the amplification was performed with a hot start for 5 min at 96°C, 30 cycles of 1 min at 96°C, 1 min at 55°C and 3 min at 70°C, and a final extension at 70°C for 10 min; identical conditions were used when the ILTat 1.21 product was generated using Long-*Taq* DNA polymerase (Stratagene). However, the following conditions were used when the ILTat 1.25 product was generated with *Pfu* polymerase: 37 μl distilled water, 5 μl 10 x *Pfu* buffer (Stratagene), 2 μl 25 mM $MgCl_2$, 1 μl 10 mM dNTPs (Pharmacia Biotech), 2 μl 5 μM T7 primer, 5 μM E200 primer, 1 μl *Pfu* polymerase (2.5 $U.\mu l^{-1}$). The amplification was performed as above, except that an elongation time of 4 min was adopted. The PCR products were fractionated on a 0.7% agarose gel, purified using the QIAGEN gel extraction kit (following the manufacturer's protocol) and cloned into the PCRScript vector (Stratagene) (following the manufacturer's protocol).

2.10 Phenol/ chloroform extraction

The volume of the sample was adjusted minimally to 200 μl by the addition of TE buffer. An equal volume of phenol/ chloroform (1:1 mixture) was then added and mixed thoroughly by inversion. The two phases were separated by centrifugation in a microcentrifuge at maximum speed for 10 min, after which the aqueous layer was eluted and transferred to a new eppendorf tube. An equal volume of chloroform was

then added and the tube contents were mixed by inversion. After centrifugation at maximum speed for 5 min., the aqueous layer was eluted and added to 1/10 the original volume of 3 M sodium acetate and 2 µl of glycogen (Boehringer Manheim); 2 volumes of 99% ethanol were then added and mixed thoroughly. The tube was then transferred to -20°C for at least 20 min, after which the DNA was pelleted by centrifugation at maximum speed for 15 min. The pellet was then washed in 1 ml 70% ethanol, air-dried, and resuspended in an appropriate volume of buffer (usually TE).

2.11 Sequencing

An appropriate amount of plasmid (500 ng for sequencing from plasmids; 200 ng if purified PCR product was sequenced directly) was mixed in a thin-walled microeppendorf tube with 5 pmoles of primer and 8 µl ABI PRISM dye terminator "ready reaction" solution (Perkin-Elmer corporation) in a final volume of 20 µl. The thermal cycling was then performed with a hot start at 96°C for 4 min, 25 cycles at 96°C for 10 s, 50°C for 5 s and 60°C for 4 min on a GeneAmp PCR system 2400 (Perkin-Elmer corporation). The solution was then transferred to a 0.5 ml eppendorf tube, after which 2 µl of 3 M sodium acetate and 50 µl of 95% ethanol were added and mixed. The tube was then placed on ice for 10 min, after which the DNA was pelleted by centrifugation for 20 min. The supernatant fluid was then removed and the pellet was washed with 250 µl 70% ethanol and air-dried. Automated sequencing was then performed by the Molecular Biology Support Unit (MBSU) of the University of Glasgow.

CHAPTER 3

THE ORDER OF VSG EXPRESSION IN A CHRONIC, PLEOMORPHIC, *T. brucei* INFECTION

3.1 Introduction

Extensive investigations into antigenic variation using monomorphic trypanosomes have enabled us to observe, and interpret, the various VSG switching mechanisms these organisms have adopted to evade the immune response of their hosts. The reduced switch rate seen in these laboratory adapted lines has facilitated such studies, allowing cloned populations to remain predominantly of one VAT during expansion. However, it is clear that these parasites differ significantly from non-passaged field isolates, and it is therefore important to consider the possibility that the selection to monomorphism is associated with alterations in the cell's recombinational or transcriptional machinery. In addition, a switching rate of 1 x 10^{-6} – 1 x 10^{-7} switches/cell/generation (Lamont *et al.* 1986) is low enough to be explained by background mutation and homologous recombination, which therefore raises questions about the significance of *VSG* gene activation events observed in monomorphic parasites.

It has been proposed (Barry, 1997a) that the marked change in switch rates between monomorphic and pleomorphic trypanosomes is indicative of a specific gene switching mechanism that might include a dedicated enzyme activity catalysing site-specific recombination, which is reduced in, or even absent from, monomorphic lines. This mechanism is likely to involve a DNA repair pathway, and has been shown to be delimited upstream by the 70 bp repeat region in both monomorphic lines (Michels *et al.* 1983) and normal, fly-transmitted lines (Delauw *et al.* 1987; Matthews *et al.* 1990). However, in monomorphic lines this repeat region is not always used, perhaps due to the absence of the specific mechanism. Indeed, the deletion or inversion of the 70 bp repeat region of the active BES in monomorphic trypanosomes had no effect on the incidence of *VSG* duplications into that *BES* (McCulloch *et al.* 1997).

The aim of this investigation is to examine the order of VSG appearance in a chronic, pleomorphic infection, and determine the activation mechanisms by which these *VSG* genes become expressed. A study of this kind would be extremely complicated if using the highest switching trypanosome lines, since the populations grown from a single cell will not be phenotypically pure after the expansion period required for molecular analysis. Therefore the ILTat 1.2 (I.L.R.A.D. – International Laboratory for Research in Animal Diseases - *Trypanozoon* antigen type) line, which

switches at the lower end of the natural range of switch rates found in fly-transmitted pleomorphic trypanosomes, was utilised in these experiments (derivation and switch rate given in section 3.2 (Turner, 1997)). A single relapse investigation (section 3.10, page 74) was also undertaken using the ILTat 1.61c pleomorphic line, which is the most highly switching clone known (at 3 x 10^{-2} switches/cell/generation); a PCR approach for analysing the rapidly switching products is outlined in section 3.9 (page 70).

3.2 Derivation of the ILTat 1.2 pleomorphic clone

The pleomorphic line utilised in this study was derived from the *T. brucei* EATRO 795 stock (East African Trypanosomiasis Research Organization), originally a field isolate from bovine blood (Uhembo, Central Nyanza Province, Kenya; 1964). After an unknown number of syringe passages over a period of several years (Onyango *et al.* 1966; Miller and Turner, 1981), trypanosomes expressing ILTat 1.2 were cloned, and a stabilate SUSB 48 (State University of New York, Stony Brook) was produced. The ILTat 1.2 expressor clone displays a VSG switching rate of about 1 x 10^{-5} switches/trypanosome/generation, and upon fly transmission, emerges from the insect with an overall rate of around 1 x 10^{-2} switches/cell/generation; individual clones within this population display a range of rates between 1 x 10^{-2} and 1 x 10^{-5} switches/ cell/generation (Turner and Barry, 1989; Turner, 1997). It is therefore evident that this ILTat 1.2 clone retains pleomorphism, despite displaying switching at a rate of only one to two orders of magnitude above that of monomorphic cell lines. In addition, this line will not proliferate in liquid culture (whereas monomorphic lines do), but does grow on a semi-solid agarose plate culture (Vassella and Boshart, 1996), clearly differentiating to stumpy form cells.

3.3 Initiation and progression of the chronic ILTat 1.2 infections

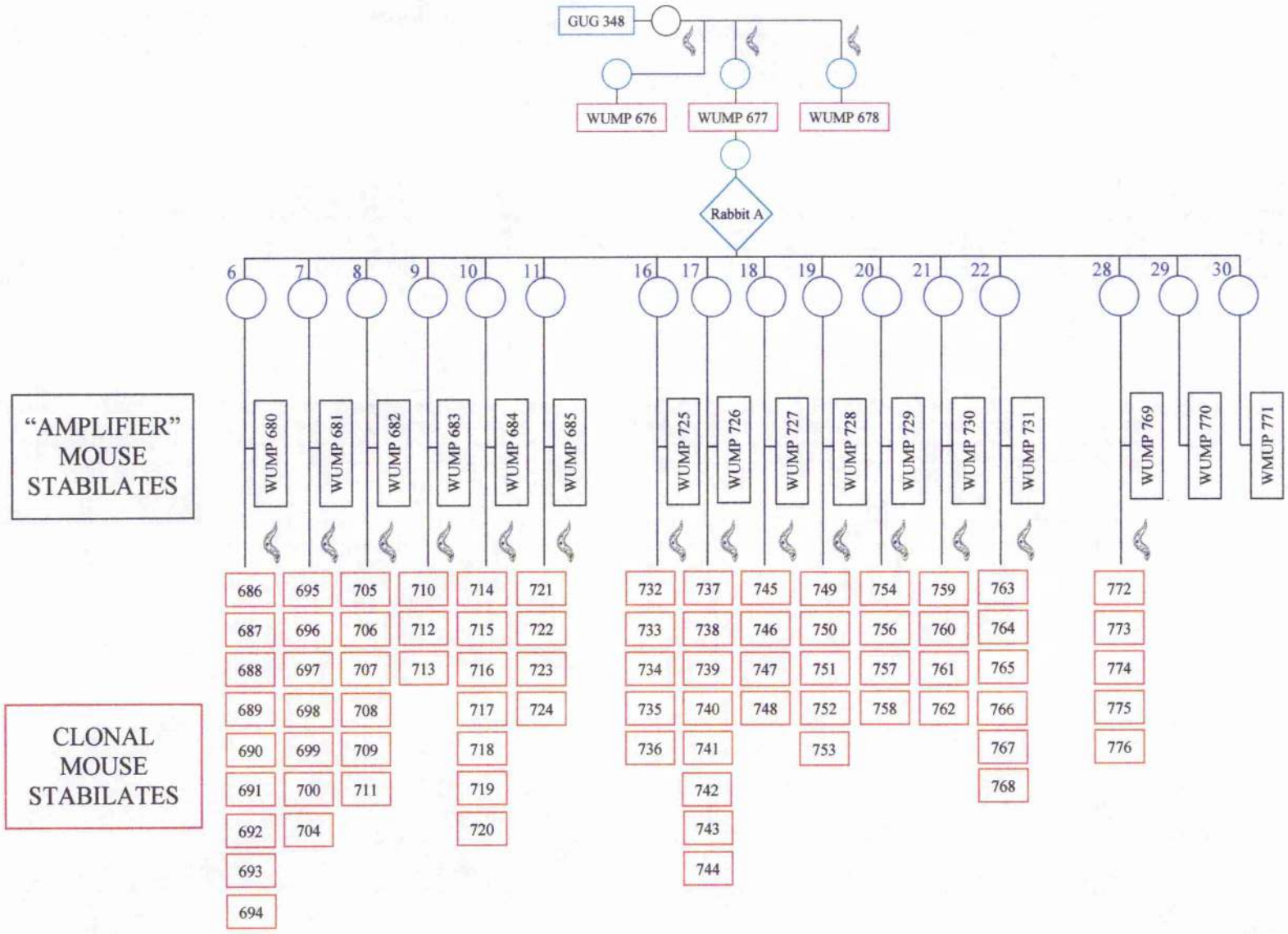
Trypanosomes were grown from the GUG (Glasgow University Genetics) 348 ILTat 1.2 stabilate in a CFLP mouse that had been immunosuppressed by cyclophosphamide treatment (Sigma Ltd, 25 mg.kg^{-1} body weight) 24 h previously. Exsanguination was performed at the initial parasitaemic peak by cardiac puncture

into sodium citrate anticoagulant in Carter's Balanced Salt Solution (CBSS), after which the trypanosomes were cloned into three immunosuppressed ICR mice to ensure the homogeneity of the ILTat 1.2 population. These mice were bled, and stabilates (WUMP 676-678) prepared when the parasitaemia achieved $10^{7.8}$ trypanosomes.ml^{-1}; a new ILTat 1.2 population was then grown from the WUMP 677 stabilate in a new immunosuppressed ICR mouse. After harvesting at the initial peak, approximately 1 x 10^6 trypanosomes were injected intravenously into a lop-eared rabbit (Bantin & Kingman, Hull, UK) from which pre-infection plasma had been collected as a control for antibodies.

On each day of infection (from days 6-30 post infection), 1 ml of blood was taken from the rabbit and 0.4 ml used for the preparation of serum; the remaining 0.6 ml was injected immediately, intraperitoneally, into an immunosuppressed ICR mouse. This step enabled the trypanosome titre to be elevated from the characteristically low level seen in rabbits prior to cloning. These "amplifier" mice were bled when the parasitaemia reached above $10^{7.8}$ trypanosomes.ml^{-1} (Herbert and Lumsden, 1976) and trypanosome clones were isolated immediately by micromanipulation. The isolation of 10 clones was attempted from each "amplifier" population, although the number of clones that arose was variable, and thus 77 clones were isolated (Figure 7); no cloning was attempted from the day 29 and 30 "amplifier" mice. All clones were stabilated in liquid nitrogen after the parasitaemia reached patency (approximately $10^{7.8}$ trypanosomes.ml^{-1}, achieved typically 3-6 days post infection). This approach to cloning trypanosomes and preparing specific antisera was based on methods developed for rapidly switching *T. vivax* that minimize growth period, and therefore the extent of antigenic variation. This also permits rapid raising of specific antisera by cymelarsan (5 mg.kg^{-1} body weight) curing of the mice in which the clones had been derived (Barry, 1986).

The rabbit parasitaemia was also measured daily, and over the 30 days of infection was typically low (Figure 8), but nevertheless yielded two visible relapse peaks, on days 7-9 and 11, and days 18-22. The parasite growth within the "amplifier" mice revealed that the rabbit was infective on days 6-11, 16-22, and 28-30, corresponding to three relapse peaks. Of the 77 clones isolated from the "amplifier" mice, 36 were isolated from the first relapse peak, 36 from the second, and 5 from the third. Subsequent immune lysis investigation of the second relapse peak revealed that only

Figure 7. Derivation of the 77 switched clonal stabilates isolated from the ILTat 1.2 chronic rabbit infection. Stabilates are enclosed in rectangles, while circles denote mouse hosts; the trypanosome symbol represents a cloning step. The GUG (Glasgow University Genetics) 348 stabilate was injected into an immunosuppressed CFLP mouse, and trypanosomes were cloned into ICR mice to produce the stabilates WUMP 676-678. The WUMP 677 stabilate was then expanded in an immunosuppressed ICR mouse before injection into the rabbit. The stabilates enclosed in black rectangles are generated from mixed population "amplifier" mice (represented by the blue circles), which were injected directly, intraperitoneally, with 0.6 ml of rabbit blood. The stabilates enclosed in red rectangles (WUMP numbers) on the lower row of the table are the switched clonal stabilates, isolated directly from the "amplifier" mouse blood; cloning mice are not indicated in this lineage. Although 10 clones were taken from each "amplifier" population, the number that survived the cloning procedure was variable, and thus 77 clones were isolated. No "amplifier" populations grew from the day 11-16 and 22-28 samples, when no trypanosomes were detectable in the rabbit blood by microscopy (see Fig. 8). Cloning was not attempted from day 29 and 30 "amplifier" blood.


GUG 348
WUMP 676
WUMP 677
WUMP 678
Rabbit A
6
7
8
9
10
11
16
17
18
19
20
21
22
28
29
30
"AMPLIFIER" MOUSE STABILATES
WUMP 680
WUMP 681
WUMP 682
WUMP 683
WUMP 684
WUMP 685
WUMP 725
WUMP 726
WUMP 727
WUMP 728
WUMP 729
WUMP 730
WUMP 731
WUMP 769
WUMP 770
WMUP 771
CLONAL MOUSE STABILATES
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
704
705
706
707
708
709
711
710
712
713
714
715
716
717
718
719
720
721
722
723
724
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
756
757
758
759
760
761
762
763
764
765
766
767
768
772
773
774
775
776

2 VATs were represented by the 36 clones, and consequently trypanosomes grown in mice from the stabilates of the day 17 and day 21 amplifier mice were treated with antiserum against those 2 VATs and cloned. This led to the isolation of another 6 trypanosome clones from the day 17 sample and 8 clones from the day 21 sample, all of which represented 2 new VATs. The day 19 "amplifier" mouse trypanosomes were then treated in the same way with antibodies against the 4 VATs from the second peak, yielding 5 more clones, all of the same VAT. In total, 96 trypanosome clones were derived form the three peaks.

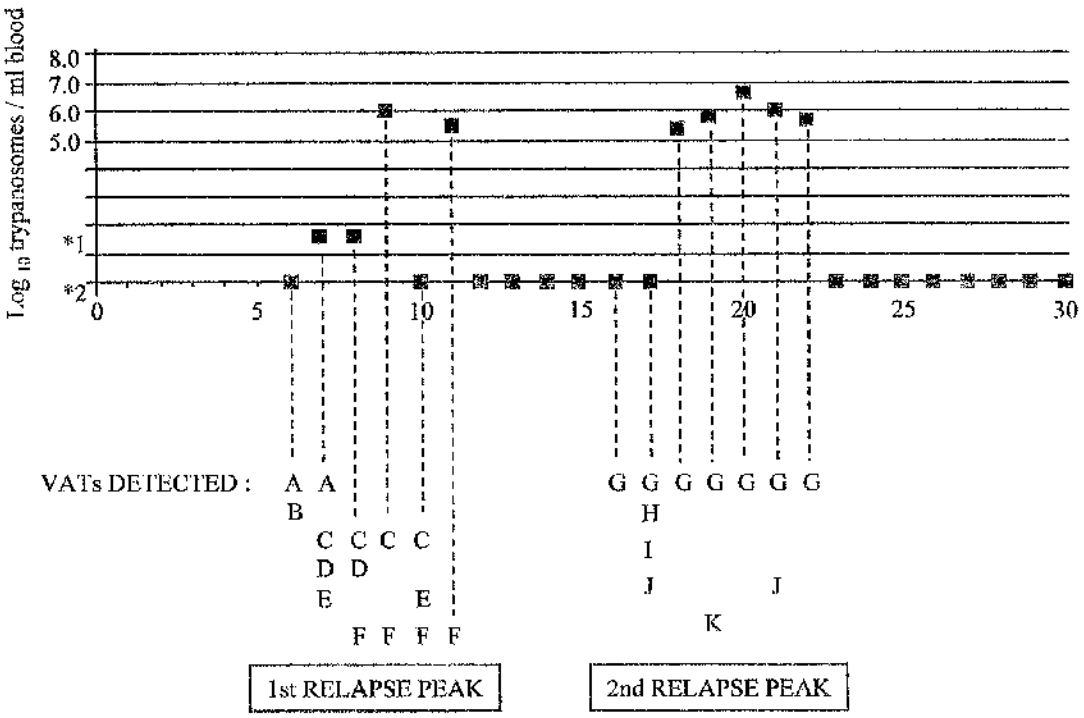


Figure 8. Parasitaemia of the ILTat 1.2 chronic rabbit infection over the 30 days of infection determined by microscopy (Herbert and Lumsden, 1976). Trypanosomes were detectable on days 7, 8, 9, 11, and days 18, 19, 20, 21, and 22. The rabbit was infective between days 6-11 (the 1st relapse peak) and days 18-22 (the 2nd relapse peak). The graph also demonstrates the days on which each of the 11 distinct VATs were present in the host blood. The individual VATs were determined by immune lysis typing of the 96 individual clones (see 3.4), and were given the temporary codes A-K. These VATs were subsequently allocated ILTat numbers: ILTat 1.25 (VAT A), ILTat 1.67 (VAT B), ILTat 1.68 (VAT C), ILTat 1.69 (VAT D), ILTat 1.21 (VAT E), ILTat 1.64 (VAT F), ILTat 1.23 (VAT G), ILTat 1.70 (VAT H), ILTat 1.71 (VAT I), ILTat 1.22 (VAT J) and ILTat 1.72 (VAT K). *1 denotes trypanosomes visible in the rabbit blood at a low unquantifiable level (arbitrary values assigned). *2 indicates the days on which trypanosomes were undetectable by microscopy.

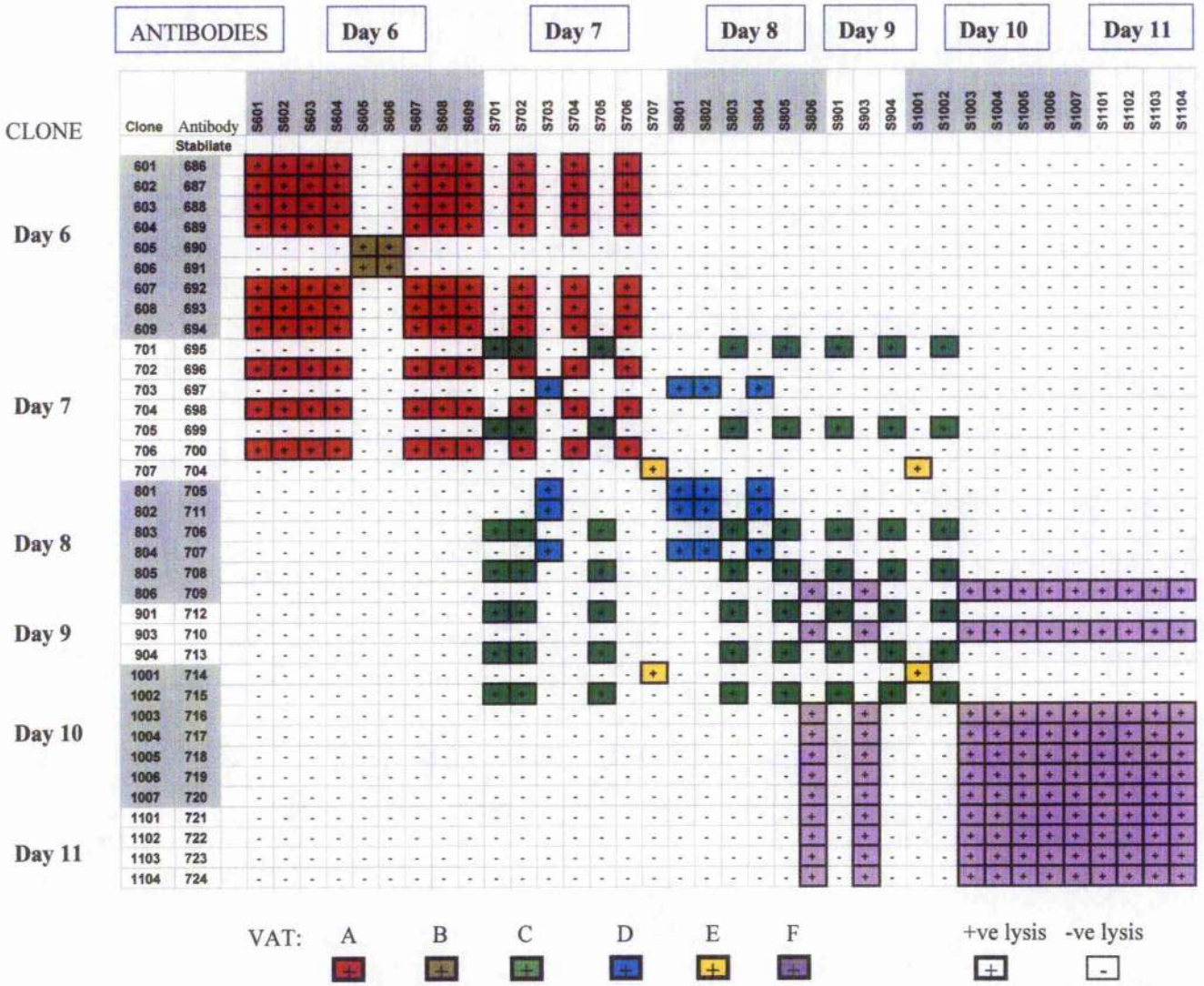
3.4 Immune lysis dissection of the first relapse peak

The number of distinct VATs represented by the 36 trypanosome clones from the first relapse peak was determined by immune lysis typing. Plasma, containing specific antibodies, was produced from each clone mouse following the elimination of infection by cymelarsan treatment. Each trypanosome clone was then grown from stabilate until the parasitaemia reached approximately $10^{7.8}$ cells.ml^{-1}, and cross-tested against all 36 antisera. The assay was performed with trypanosomes suspended in guinea-pig complement to 5 x 10^6 cells.ml^{-1}, and incubated in 1:20 diluted antiserum for 1 h at room temperature. Percentage lysis was then quantified by light microscopy, counting at least 100 parasites; every positive displayed >99% lysis. Each trypanosome clone was also incubated for an hour in the presence of guinea-pig complement alone, as a negative control, to ensure that all the lysis was due to the action of the antiserum.

The 1296 individual immune lysis results (summarised in Figure 9) revealed that 6 VATs (allocated the temporary working codes A-F) were represented in the first relapse peak. Each antiserum displayed a distinct lysis pattern, reacting exclusively with trypanosome clones expressing the same antigen coat (with the exception of antiserum S702, which cross-reacted with a second VAT). Figure 9 also illustrates the change in VAT composition over the 5 days of the first relapse peak, progressing from the top left to the bottom right of the chart.

Identical immune lysis typing was undertaken for the second relapse peak and revealed a further 5 unique VATs. Figure 8 demonstrates the days on which each VAT was present in the rabbit infection.

Figure 9. Immune lysis typing of the 36 trypanosome clones isolated from the first relapse peak of the ILTat 1.2 chronic rabbit infection. Immune lysis cross-reaction was performed between each of the 36 clones (WUMP stabilate numbers indicated) against each of their 36 corresponding antisera. Trypanosomes were incubated with the relevant antibodies in the presence of guinea-pig complement at room temperature for 1 hr in the bottom of a 96 well V-bottomed microtitre plate. A minus sign indicates no trypanosome lysis, whilst a plus sign denotes complete lysis (>99%). Clones of the same VAT (*e.g.* 690 and 691) display an identical cross-reactive pattern and have been allocated a unique colour and a temporary working code (*e.g.* B). The VATs were subsequently allocated ILTat numbers : ILTat 1.25 (VAT A), ILTat 1.67 (VAT B), ILTat 1.68 (VAT C), ILTat 1.69 (VAT D), ILTat 1.21 (VAT E), and ILTat 1.64 (VAT F).


ANTIBODIES
Day 6
Day 7
Day 8
Day 9
Day 10
Day 11
CLONE
Clone
Antibody
Stabilate
S601 S602 S603 S604 S605 S606 S607 S608 S609 S701 S702 S703 S704 S705 S706 S707 S801 S802 S803 S804 S805 S806 S901 S903 S904 S1001 S1002 S1003 S1004 S1005 S1006 S1007 S1101 S1102 S1103 S1104
601 686
602 687
603 688
604 689
605 690
606 691
607 692
608 693
609 694
701 695
702 696
703 697
704 698
705 699
706 700
707 704
801 705
802 711
803 706
804 707
805 708
806 709
901 712
903 710
904 713
1001 714
1002 715
1003 716
1004 717
1005 718
1006 719
1007 720
1101 721
1102 722
1103 723
1104 724
Day 6
Day 7
Day 8
Day 9
Day 10
Day 11
VAT: A B C D E F
+ve lysis
-ve lysis

3.5 Further ILTat 1.2 infections, and additional immune lysis investigations

To eliminate the possibility of the isolated VATs having become activated in the "amplifier" mouse, rather than in the rabbit, six trypanosome clones representing each VAT (see Figure 10 for details) were grown from stabilate and tested in immune lysis against the daily plasma samples from rabbit A. Specific antibodies arose only after the corresponding VAT was detectable in the rabbit, typically 3-6 days later (Figure 10, Table 2). Identical analysis of the VATs from the second relapse peak revealed a similar trend, confirming that all 11 VATs were generated during the original infection, rather than during expansion in the "amplifier" mice.

In order to analyse the reproducibility of appearance of the VATs in different infections, the same ILTat 1.2 infecting population was grown from stabilate (WUMP 677) in an ICR mouse and then treated with antisera against all 11 VATs isolated from rabbit A, to remove existing switch products. After washing, the surviving trypanosomes were injected into another two lop-eared rabbits (B and C), and also into a New Zealand White rabbit (D) (Bantin & Kingman, Hull, UK). Plasma samples were harvested daily (or on alternate days for rabbit D), and tested in immune lysis against each of the VATs present in the initial infection (results for the first relapse peak are summarised in Figure 10).

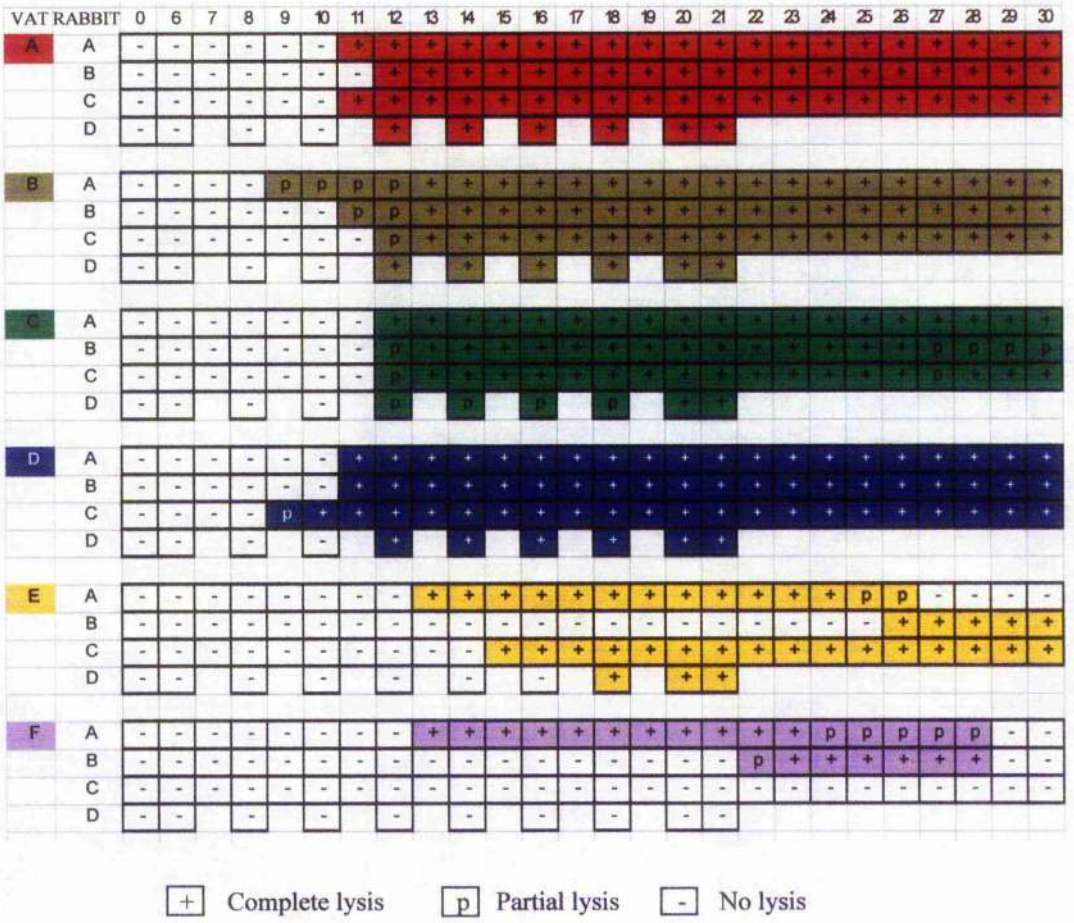


Figure 10. Summary of the onset of lytic activity in four separate rabbits against the six VATs that were isolated from the first relapse peak in rabbit A. Daily plasma samples were prepared from each of the lop-eared rabbits (A-C) over the 30 days of infection. Plasma was sampled on alternate days from the New Zealand White rabbit (D) until day 21 when the infection was terminated. One clone was grown from stabilate for each of the VATs : clone 601 (VAT A, WUMP 686), clone 605 (VAT B, WUMP 690), clone 803 (VAT C, WUMP 706), clone 804 (VAT D, WUMP 707), clone 1001 (VAT E, WUMP 714), and clone 1101 (VAT F, WUMP 721). Trypanosomes were incubated with the relevant plasma in the presence of guinea-pig complement at room temperature for 1 h in the bottom of a 96 well V-bottomed microtitre plate. A plus sign indicates complete lysis (>99%), while a minus sign denotes no lysis; partial lysis (≈ 10-60%) is represented by the letter p. The VATs were subsequently allocated ILTat numbers: ILTat 1.25 (VAT A), ILTat 1.67 (VAT B), ILTat 1.68 (VAT C), ILTat 1.69 (VAT D), ILTat 1.21 (VAT E), and ILTat 1.64 (VAT F).

The immune lysis investigation revealed that VATs A, B, C, and D (later allocated the ILTat numbers 1.25, 1.67, 1.68, and 1.69 respectively) appeared early in infection, with consistent timing, in all four hosts (antibodies detected between days 9 and 12). VAT E (ILTat 1.21) was also present in all four infections, but displayed a greater spread in its time of emergence (activity detected between days 13 and 26). VAT F (ILTat 1.64), however, elicited antibodies in rabbits A (day 13) and B (day 22), but did not appear in rabbits C and D. There was more variation in the timing of antibody onset for the second relapse peak (data not shown), with only VAT H (ILTat 1.70) appearing in every host.

It is clear from these studies that the VATs emerging early in infection have a higher probability of expression, and are consequently represented more frequently, than those appearing at a later time. As a result of this, the order of VAT expression in the second relapse peak is less predictable than in the first relapse peak. Indeed, this variation can be seen even within the first relapse peak, where the VATs activated towards the end of the peak (E and F) display a more irregular expression pattern than the VATs presented at the very start of the infection (A to D). This phenomenon, in which VATs are expressed in a "semi-predictable" order of antigen type appearance, has been observed experimentally elsewhere (Miller and Turner, 1981), and was originally demonstrated by Gray (1965), van Meirvenne *et al.* (1975), and Capbern *et al.* (1977).

3.6 Immunofluorescence typing of the VATs

Indirect immunofluorescence was performed as previously described (Van Meirvenne *et al.* 1975), on acetone-fixed thin blood smears using 37 existing reference antisera as the primary antibody. The antisera, which were originally derived from either mouse or rabbit hosts, were diluted in 1 x PBS prior to use (see Table 1). Anti-rabbit (or anti-mouse) Ig fluorescein isothiocyanate (FITC) labelled conjugates, were used as the secondary antibody.

The immunofluorescence study yielded 3 clear serotypes from the first relapse peak (Figure 11): ETat 1.7 (VAT D), ETat 1.9 (VAT E), and GUTat 7.13 (VAT F); in the second relapse peak VAT J was determined as GUTat 7.1. The equivalent ILTat numbers are given in Table 2 (section 3.8, page 69) and the significance of these findings is discussed in the final chapter of this thesis.

Antiserum	Dilution *	Host	Antiserum	Dilution *	Host
ILTat 1.3	200	rabbit	ETat 1.1	300	rabbit
ILTat 1.4	100	rabbit	ETat 1.3	100	rabbit
ILTat 1.61	300	mouse	ETat 1.4	200	rabbit
ILTat 1.62	100	rabbit	ETat 1.5	100	rabbit
			ETat 1.6	100	rabbit
GUTat 7.1	500	mouse	ETat 1.7	200	rabbit
GUTat 7.2	500	mouse	ETat 1.9	100	rabbit
GUTat 7.3	400	rabbit	ETat 1.10	100	rabbit
GUTat 7.4	200	rabbit	ETat 1.11	100	rabbit
GUTat 7.5	300	rabbit	ETat 1.12	100	rabbit
GUTat 7.6	400	rabbit	ETat 1.13	300	rabbit
GUTat 7.8	100	rabbit	ETat 1.14	100	rabbit
GUTat 7.9	200	rabbit	ETat 1.15	100	rabbit
GUTat 7.10	100	rabbit	ETat 1.16	75	rabbit
GUTat 7.11	100	rabbit	ETat 1.17	100	rabbit
GUTat 7.12	200	rabbit	ETat 1.18	100	rabbit
GUTat 7.13	200	mouse	ETat 1.19	50	rabbit
GUTat 10.1	200	mouse	ETat 1.20	75	rabbit
GUTat 11.7	200	mouse	ETat 1.21	100	rabbit

Table 1. Working dilutions and derivations of the 37 reference antisera.

ILTat – ILRAD *Trypanozoon* antigen type

GUTat – Glasgow University *Trypanozoon* antigen type

ETat – Edinburgh *Trypanozoon* antigen type

* reciprocal (1 in X dilution)

TRYPANOSOME VARIABLE ANTIGEN TYPE (VAT)												
	VAT A		VAT B		VAT C		VAT D		VAT E		VAT F	
Antibodies	Result	control*	Result	control*	Result	control *	Result	control*	Result	control*	Result	control*
GUTat 7.1	-	n	-	n	-	n	-	n	-	n	-	n
GUTat 7.13	+	+++	(+)	+++	(+)	++	-	n	+	+++	+++	+++
GUTat 7.2	-	n	-	n	-	n	-	n	-	n	-	n
GUTat 7.3	-	n	-	n	-	n	-	n	-	n	-	n
ILTat 1.3	-	n	-	n	-	n	-	n	-	n	-	n
ILTat 1.4	(+)	+++	(+)	+++	+	++	-	n	+	+++	+	+++
ILTat 1.61	-	n	-	n	-	n	-	n	-	n	-	n
ILTat 1.62	-	n	-	n	-	n	-	n	-	n	-	n
GUTat 10.1	-	(+)	-	(+)	-	(+)	-	(+)	-	(+)	-	(+)
GUTat 11.7	-	(+)	-	(+)	-	(+)	-	(+)	-	(+)	-	(+)
GUTat 7.12	-	(+)	-	(+)	-	(+)	-	(+)	-	(+)	-	(+)
GUTat 7.4	-	(+)	-	(+)	-	(+)	-	(+)	-	(+)	-	(+)
GUTat 7.5	-	(+)	-	(+)	-	(+)	-	(+)	-	(+)	-	(+)
GUTat 7.6	(+)	++	(+)	(+)	-	+	-	+	-	++	-	+
GUTat 7.8	(+)	++	(+)	(+)	-	+	-	+	-	++	(+)	+
GUTat 7.9	-	++	-	(+)	-	+	-	+	-	++	-	+
GUTat 7.10	(+)	++	-	(+)	-	+	-	+	-	++	(+)	+
GUTat 7.11	-	++	-	(+)	(+)	+	-	+	-	++	(+)	+
ETat 1.1	(+)	+++	(+)	+++	-	++	-	+	(+)	+++	(+)	+++
ETat 1.3	(+)	+++	(+)	+++	(+)	++	-	+	(+)	+++	(+)	+++
ETat 1.4	(+)	+++	(+)	+++	(+)	++	-	+	(+)	+++	(+)	+++
ETat 1.5	(+)	+++	(+)	+++	-	++	(+)	+	(+)	+++	(+)	+++
ETat 1.6	(+)	+++	(+)	+++	(+)	++	-	+	(+)	+++	(+)	+++
ETat 1.7	(+)	+++	(+)	+++	(+)	++	++	++	(+)	+++	+	+++
ETat 1.9	(+)	+++	-	+++	-	++	(+)	++	+++	+++	+	+++
ETat 1.10	+	+++	(+)	+++	(+)	++	(+)	++	(+)	+++	+	+++
ETat 1.11	+	+++	(+)	+++	(+)	++	(+)	++	+	+++	+	+++
ETat 1.12	+	+++	-	+++	-	++	(+)	++	(+)	+++	(+)	+++
ETat 1.13	(+)	+++	(+)	+++	-	+++	(+)	+++	+	+++	(+)	+++
ETat 1.14	+	+++	(+)	+++	(+)	+	(+)	+++	+	+++	(+)	+++
ETat 1.15	+	+++	(+)	+++	(+)	+	(+)	+++	+	+++	(+)	+++
ETat 1.16	(+)	+++	-	+++	-	+	(+)	+++	(+)	+++	(+)	+++
ETat 1.17	++	+++	+	+++	(+)	+	+	+++	+	+++	+	+++
ETat 1.18	(+)	++	+	+++	+	+	(+)	+++	+	+++	(+)	++
ETat 1.19	+	++	(+)	+++	(+)	+	(+)	+++	+	+++	+	++
ETat 1.20	(+)	++	+	+++	(+)	+	(+)	+++	+	+++	(+)	++
ETat 1.21	(+)	++	(+)	+++	+	+	(+)	+++	+	+++	(+)	++

Figure 11. Summary of the immunofluorescence results for the six VATs isolated from the first relapse peak. Indirect immunofluorescence was performed on acetone-fixed thin blood smears using 37 existing reference antibodies (either rabbit or mouse) as the primary antibody. Trypanosomes were visualized with fluorescein isothiocyanate (FITC) labelled conjugates. Each reaction was given an arbitrary score according to the extent of the fluorescence: - no fluorescence, (+) faint fluorescence, + slight fluorescence, ++ clear fluorescence, +++ strong fluorescence. * Positive controls were performed on each slide using mouse antisera specific to the individual VATs as the primary antibody; (n) indicates no positive control.

3.7 Reverse Transcription PCR (RT-PCR) amplification of the *VSG* genes

Clones representing each VAT were grown from stabilate in immunosuppressed ICR mice until the parasitaemia approached $10^{7.8}$ trypanosomes.ml^{-1}. The stabilates used were: WUMP 686 (VAT A), WUMP 690 (VAT B), WUMP 706 (VAT C), WUMP 707 (VAT D), WUMP 704 (VAT E), and WUMP 720 (VAT F). Approximately 1-5 μg mRNA was isolated directly from 100 μl blood, using 1 ml TRIzol (Gibco-BRL), and reverse transcribed to single strand cDNA using oligo [dT] primer (Gibco-BRL). The expressed VSG genes were then amplified by PCR using two short specific oligonucleotides as described by Carrington *et al.* (1991). The first primer (tbsl: GTTTCTGTACTATATTG) was a 17mer specific to the mRNA spliced leader, while the second primer (tb3ut: GTGTTAAAATATATCA) was a 16mer specific to a region in the downstream UTR, highly conserved in all *VSG* genes. Amplification was performed for 30 cycles of 1 min at 96° C, 1 min at 42° C and 2 min at 70° C, in a final reaction volume of 50 μl (1 μl of the 20 μl reverse transcribed product was used as template).

Figure 12A demonstrates the major PCR products generated from the poly[A]$^+$ RNAs of the six VATs isolated from the first relapse peak, which varied in size from approximately 1.5 to 1.8 kb. In the ILTat 1.2 and VAT B PCR reactions a second minor PCR product of around 0.8 kb was also seen (Figure 12B), although this band was much fainter and appeared only occasionally in the VAT A, C, D, E and F reactions.

Following fractionation on a 0.7% agarose gel, the products were cloned using the 'T-easy' vector, or PCRScript Amp SK (+). Three clones were produced from ILTat 1.2, and VAT A-E DNA; only 1 clone was generated for VAT F. As trypanosomes are variable in expression of VSGs, this approach can yield cDNA clones from minor VATs in the trypanosome clones. Partial sequence (400 bases from the amino- and carboxy- termini) was therefore obtained from each of the clones, and also directly from the purified PCR products. The majority of the cloned products proved to derive from the predominant PCR product, although there were some discrepancies. All three VAT B clones failed to match the VAT B PCR product, and were found to have identity with the ILTat 1.2 clones. Consequently, three new

clones were produced following RT-PCR of new mRNA isolated from trypanosomes grown from a different stabilate (WUMP 691); these clones matched the VAT B PCR sequence. The third VAT A clone also seemed to have derived from a different expressor, since its amino-terminus sequence did not match that of other two clones or of the PCR product. Interestingly, this clone did display complete homology with the other two clones at its carboxy-terminus, and it therefore seems possible that this minor expressor was activated by a recombinational event, in which part of the previous *VSG* gene remained in the expression site.

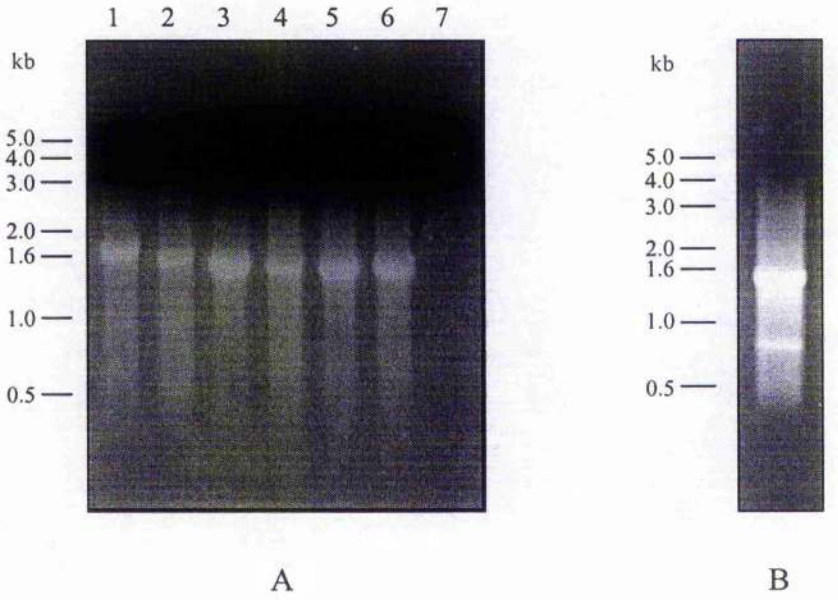


Figure 12. RT-PCR of the *VSG* transcripts of the six VATs isolated from the first relapse peak. Panel A. EtBr stained 0.7% agarose gel of the VSG gene cDNA products generated by RT-PCR of mRNA from the six VATs identified in the first relapse peak. PCR products originating from VAT A to F (WUMP 686, 690, 706, 707, 704, and 720 respectively) poly[A]$^+$ RNA were loaded in lanes 1-6 respectively (10 µl of the 50 µl reaction loaded per lane); a negative (no template) control was also loaded in lane 7. Panel B displays the PCR product generated from VAT B (WUMP 691), demonstrating the second minor PCR product, at 0.8 kb, that was often seen. Molecular markers are indicated to the left of each gel.

BLAST searching against the NCBI database revealed a typical VSG sequence pattern: general conservation at the downstream end of the sequence, reflecting peptide sequence conservation, and considerable variation at the amino-terminal coding end, which contains variable epitopes of the VSG. VAT A was identified from the database as ILTat 1.25, and VAT E as ILTat 1.21 (consistent with the ETat 1.9 immunofluorescence result; J.D. Barry, pers. comm.), with both clones perfectly matching the existing sequence. The ILTat 1.2 clones displayed exact homology with the database sequence at the carboxy-terminus; the amino-terminus (approximately 500 bp) was missing from the database. Two clones derived from the second relapse peak were also identified from the database: VAT G (ILTat 1.23) and VAT J (ILTat 1.22, consistent with the GUTat 7.1 immunofluorescence result).

Full length cDNA sequence was also determined from the first clone of each of the VATs A-F, and also from ILTat 1.2, allowing internal restriction sites to be determined (Figure 13).

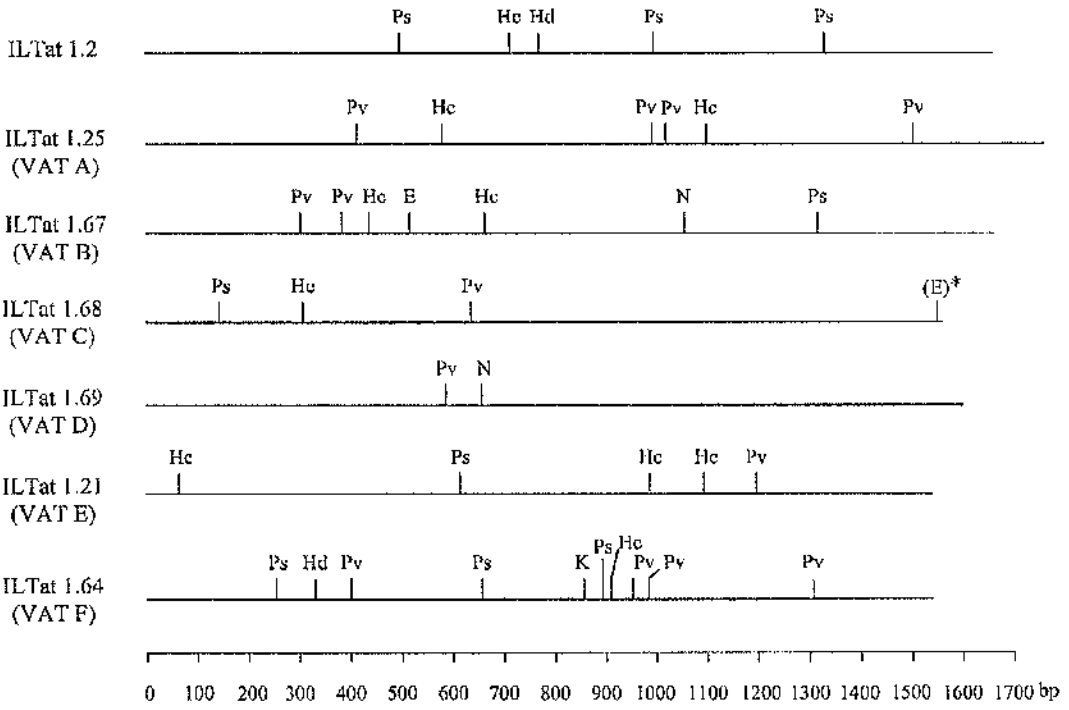


Figure 13. Physical maps of the ILTat 1.2 and VAT A-F VSG genes (starting with the last seventeen nucleotides of the spliced leader). Restriction sites were determined from the full length cDNA sequence of each VAT. Abbreviations: E, *Eco*RI; Hc, *Hinc*II; Hd, *Hind*III; K, *Kpn*I; N, *Not*I; Ps, *Pst* I; Pv, *Pvu*II. * This *Eco*RI site was not seen in the other two VAT C clones, and probably represents a PCR or cloning artefact.

3.8 Allocation of ILTat numbers to previously unidentified VATs

Immunofluorescence and DNA sequencing had identified six of the eleven VATs isolated from the rabbit A infection. Five of these VATs corresponded to previously allocated ILTat numbers: ILTat 1.25 (VAT A), ILTat 1.21 (VAT E), ILTat 1.64 (VAT F), ILTat 1.23 (VAT G), and ILTat 1.22 (VAT J). The remaining six VATs were allocated the new ILTat numbers 1.67-1.72 according to their order of emergence from the rabbit infection. The ILTat numbers are summarised in Table 2, which also demonstrates the equivalent ETat and GUTat numbers; the Table also indicates the days on which each of the VATs was present in the rabbit A infection, and the timing of antibody onset in rabbits A-D. These VATs will now be referred to in the text by their ILTat codes, rather than their working letters.

VAT				Time isolated from rabbit A	Number of clones isolated	Day of antibody onset in rabbits			
CODE	ILTat	ETat	GUTat	days		A	B	C	D
A	1.25*			6,7	10	11	12	11	12
B	1.67			6	2	9	11	12	12
C	1.68			7-10	7	12	12	12	12
D	1.69	1.7 i		7,8	4	11	11	9	12
E	1.21*	1.9 i		7,10	2	13	26	15	18
F	1.64		7.13 i	8-11	11	13	22	-	-
G	1.23*			16-22	35	26	-	-	-
H	1.70			17	1	22	12	22	14
I	1.71			17	3	23	23	27	-
J	1.22*	1.2	7.1 i	17, 21	8	23	-	-	18
K	1.72			19	5	23	-	14	16

Table 2. Summary of the 11 VATs isolated from the rabbit A infection. * indicates VAT identified by DNA sequencing; i denotes immunofluorescence result.

3.9 A PCR approach for analysing rapid switch products

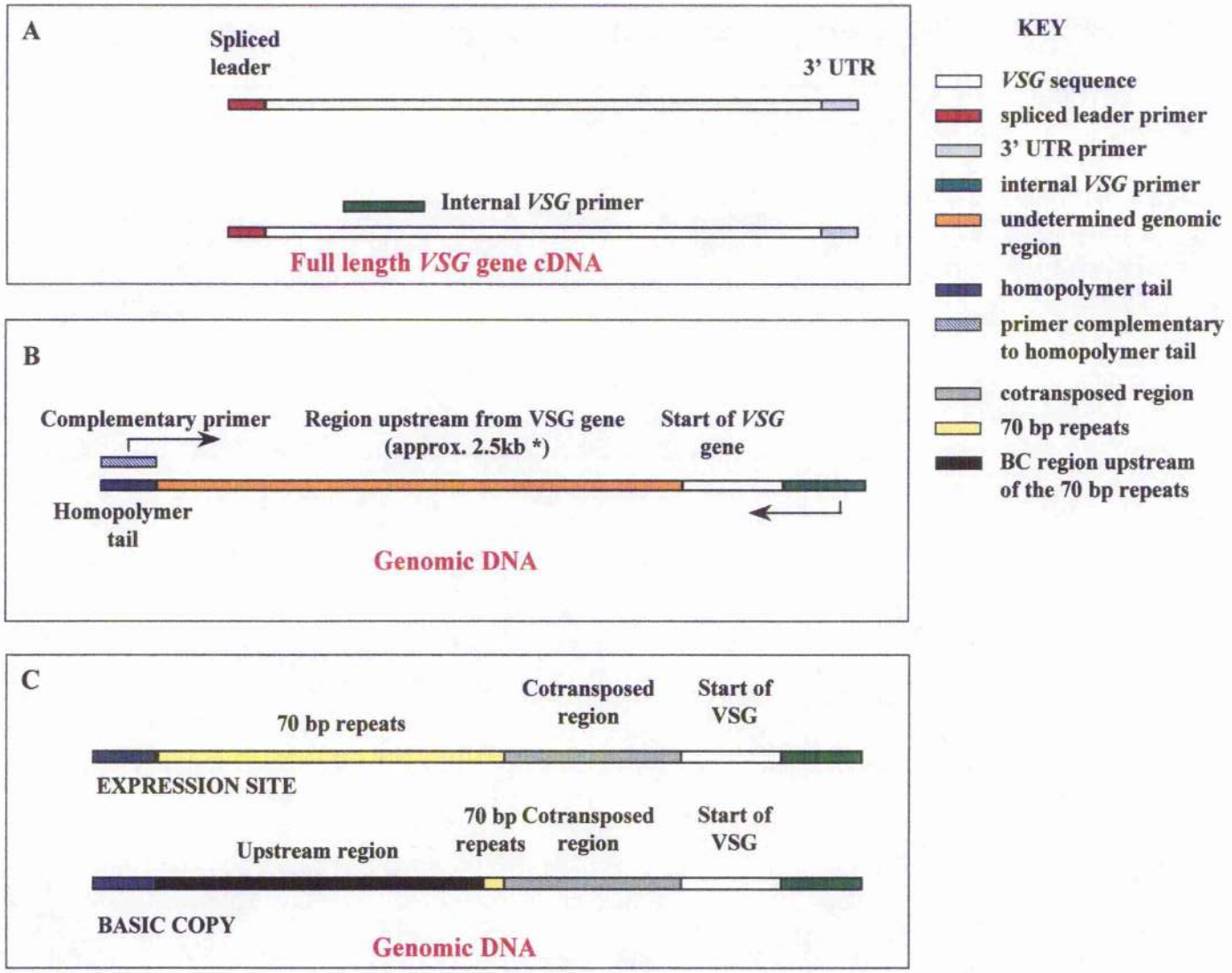
Although the ILTat 1.2 line retains pleomorphism, its switch rate is a few orders of magnitude lower than the overall rates seen in lines that have not undergone such extensive passaging. However, as discussed earlier, it would be impossible to analyse the myriad of switch products that could be generated in a chronic infection, if initiated with one of these high switching lines. Single relapse investigations, such as those performed by Miller and Turner (1981), would reduce the complexity of the study, but the antigenic variation rate of the most rapidly switching trypanosomes (approximately 3 x 10^{-2} switches/cell/generation) could complicate genomic Southern analysis. In order to circumvent this problem, an experimental procedure was designed in which high switching products from a single relapse study could be scrutinized by PCR analysis (summarised in Figure 14).

This method determines whether a switch event has occurred by a duplicative or non-duplicative mechanism by analysing the region upstream of the *VSG* genes before and after the gene activation. Expression site sequence should be quite distinct from the BC sequence(s) displaying a longer run of the 70 bp repeats (as discussed in section 1.4, page 9), allowing the *VSG* gene copy number, and therefore the switch mechanism, to be resolved. In the case of a duplicative event both ELC and BC sequence should be detectable, but during a transcriptional activation only the BC(s) should be apparent. It is necessary to examine at least 2.5 kb upstream of the *VSG* sequence to ensure that the area upstream of the cotransposed region and 70 bp repeats is examined, since this is where the differences between BC and ELC will be observed (beyond the duplication boundary). If the BC(s) contain a long array of 70 bp repeats (*e.g.* ILTat 1.21; see section 5.6, page 128) then the analysis must be continued further upstream. This method requires the upstream fragments, containing the 70 bp repeats, to be cloned and therefore a recombination deficient *E. coli* strain should be utilised. The methodology for this approach is described below.

A clonally-initiated infection with a trypanosome from a high switching line (*e.g.* ILTat 1.61c) would be allowed to progress to the first relapse peak, when the mixed population of parasites would be re-cloned into new mice. These subclones would then be harvested once the parasitaemia achieved patency, and genomic DNA and mRNA would be isolated. Subsequent cloning of the cDNA (produced by RT-PCR as in section 3.7) would allow specific amino-terminus *VSG* gene primers to be

manufactured for each of the switched products. This primer could then be used to amplify single strand product, upstream of the *VSG* gene, from both the expressor and initiator genomic DNA. After tailing this product with a homopolymer tail, double stranded DNA could then be generated by amplifying with the *VSG* gene specific primer and an oligonucleotide complementary to the tail. It should then be possible to determine whether the VSG has switched via a duplicative or transcriptional mechanism by the cloning and sequencing of these PCR products. In the case of a duplicative activation, two (or more) products should be observed from the expressor DNA (representing the BC(s) and ELC), while only the BC(s) should be detectable from the initiator DNA. During a transcriptional activation, however, only the BC(s) will be seen from either DNA source since no ELC will be formed.

Figure 14. Schematic representation of the PCR method for analysing high switching VSG gene products. Panel A depicts the generation of the cDNA by RT-PCR (using spliced leader and 3' UTR primers as in section 3.7) from the clones isolated from the first relapse peak. Following the cloning of the PCR product, an internal amino-terminus VSG gene specific primer could be used to generate single strand genomic DNA upstream of the VSG gene. Panel B shows this product being tailed, and then amplified to double strand DNA using the internal VSG primer and an oligonucleotide complementary to the homopolymer tail. Panel C indicates the sequence differences, upstream of the VSG gene, between genes in an expression site and those in internal arrays. In the case of a duplicative activation, both the ELC and BC sequence(s) should be produced from the expressor genomic DNA, whilst only the BC sequence(s) should be detected from the initiator genomic DNA. * approximately 2.5 kb upstream of the *VSG* gene should be sequenced to ensure that the boundary of duplication (probably within the 70 bp repeats) is passed.


A
Spliced leader
3' UTR
Internal VSG primer
Full length VSG gene cDNA
B
Complementary primer
Region upstream from VSG gene (approx. 2.5kb *)
Start of VSG gene
Homopolymer tail
Genomic DNA
C
70 bp repeats
Cotransposed region
Start of VSG
EXPRESSION SITE
Upstream region
70 bp repeats
Cotransposed region
Start of VSG
BASIC COPY
Genomic DNA
KEY
VSG sequence
spliced leader primer
3' UTR primer
internal VSG primer
undetermined genomic region
homopolymer tail
primer complementary to homopolymer tail
cotransposed region
70 bp repeats
BC region upstream of the 70 bp repeats

3.10 Single relapse study of the high switching pleomorphic line ILTat 1.61c

A single relapse study was undertaken using the high switching line ILTat 1.61c. This line was also derived from the SUSB 48 stabilate, but initiated from a single metacyclic trypanosome produced by a tsetse fly. The cloned trypanosome was allowed to expand in a mouse for five days before a stabilate, GUP 2812 (Glasgow University Protozoology), was produced. This population was previously shown to be 98% homologous to ILTat 1.61 by immunofluorescence (J.D. Barry, pers. comm.). The switching rate of this line had also been determined (Turner, 1997), which, at about 3 x 10^{-2} switches/cell/generation, is the highest rate observed in trypanosomes.

The GUP 2812 stabilate was injected into an immunosuppressed ICR mouse, and the parasites were grown until the parasitaemia achieved $10^{7.2}$ trypanosomes.ml^{-1} blood. Ten single trypanosomes were then isolated immediately by micromanipulation, and injected intraperitoneally into new, immunosuppressed, ICR mice. The 7 resulting clones were grown for 5-7 days, when 250 µl of blood was isolated from each mouse by tail-bleeding into heparinised capillary tubes. Genomic DNA was then isolated from 200 µl whole blood, and stabilates were prepared from the remaining 50 µl. This genomic DNA was prepared as a control to ensure that the VSG switch occurred at the relapse, rather than during the growth phase at the initial peak. The infection was then meant to progress to the first relapse peak, but the ICR mice appeared to be very sensitive to the infection, resulting in only one of the clone mice surviving beyond the initial peak. On the 10^{th} day of infection the remaining clone mouse achieved a parasitaemia of $10^{7.8}$ trypanosomes.ml^{-1} blood. Parasites were harvested by tail-bleeding, and 10 trypanosome clones were isolated, and injected intraperitoneally into new, immunosuppressed, mice. The parasitaemia had risen to $10^{8.7}$ trypanosomes.ml^{-1} blood by day 11, and a further 10 clones were injected into new mice; stabilates were made from both days 10 and 11. These clones were then expanded until the parasitaemia reached approximately $10^{7.8}$ trypanosomes.ml^{-1} blood, when the mice were exsanguinated; stabilates were prepared, and genomic DNA was isolated, following the separation of the trypanosomes from the blood by a Percoll gradient. mRNA was also prepared from 100 µl blood, using 1 ml TRIzol.

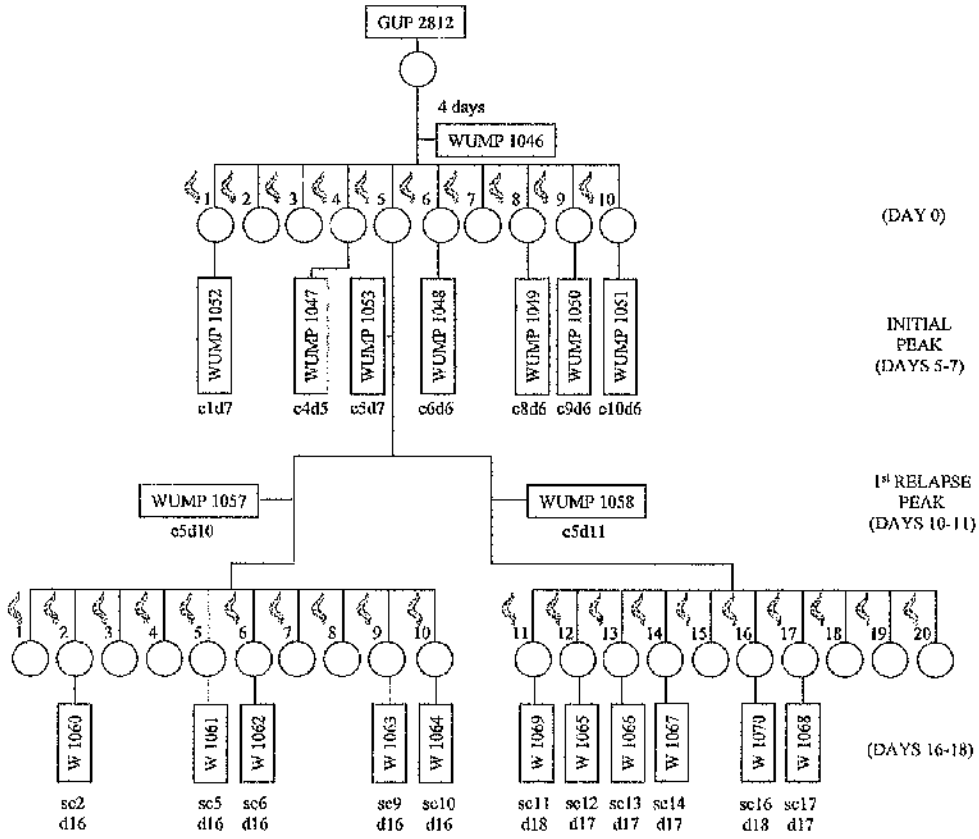


Figure 15. Derivation of the 11 switched clones generated from the ILTat 1.61c single relapse investigation. The trypanosome population was established in an ICR mouse from the GUP 2812 stabilate. 10 clones were then isolated and injected into new mice. Stabilates were prepared from 50 µl blood at the initial peak, and infections were continued until the first relapse peak; only clone mouse 5 survived the infection. 10 clones were sampled from mouse 5 on day 10, and a further 10 clones were isolated on day 11. The trypanosomes were then harvested when the parasitaemia achieved approximately $10^{7.8}$ cells.ml^{-1}, when stabilates were prepared. Stabilates are enclosed in rectangles, while clone mice are represented as circled numbers; the working clone numbers are indicated under the stabilate numbers (c1d7–clone 1, isolated day 7; sc2d16-switched clone 2, isolated day 16). Three of the initial cloning attempts, and nine of the 1st relapse cloning attempts, were unsuccessful. The cloning steps are indicated by the trypanosome symbol.

In total, 11 switched clones were isolated from the first relapse peak (Figure 15). Full length cDNAs were then generated from the mRNA by RT-PCR (as described in section 3.7, page 66). These PCR products were fractionated on a 0.7% agarose gel (see Figure 16) and gel purified. A single PCR product of approximately 1.7 kb was generated from 9 of the 11 switched clones. Two distinct products were amplified from the sc6d16 template (switched clone 6, isolated day 16; Figure 16, lane 3), suggesting that either a switch had occurred during the expansion of the clone, or that more than one trypanosome had been injected into the mouse at the cloning step. No product was generated for sc5d16 (Figure 16, lane 2).

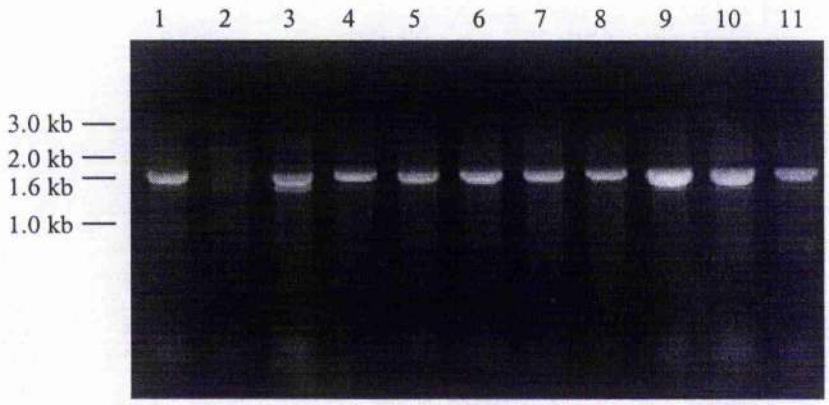


Figure 16. VSG gene cDNA generated by RT-PCR for the 11 switched clones isolated from the ILTat 1.61c single relapse investigation. The PCR products were run on a 0.7% agarose gel and stained with EtBr. Lanes 1-11: sc2d16 (switched clone 2, isolated day 16), sc5d16, sc6d16, sc9d16, sc10d16, sc12d17, sc13d17, sc14d17, sc17d17, sc11d18, and sc16d18 (10 µl of the 50 µl reaction was loaded per lane). All the products are approximately 1.7 kb, with the exception of sc6d16 (lane 3), in which two products are visible; the PCR failed to amplify the sc5d16 template (lane 2).

The 9 clones, from which the single PCR product had been generated, were grown from stabilate in immunosuppressed ICR mice, and immune lysis was performed (as described in section 3.4) using antisera specific to the six VATs isolated from the

first relapse peak of the rabbit A infection. All the clones displayed >99% lysis with ILTat 1.25 (VAT A) antiserum, and did not cross react with the other five antisera. This homogeneity was unexpected, since the first relapse populations of these high switching trypanosomes should consist of multiple VATs. However, although the overall switching rate of an ILTat 1.61c population is about 3 x 10^{-2} switches/cell/generation, individual clones within these populations demonstrate a variety of rates down to 1 x 10^{-4} switches/cell/generation (Turner, 1997). It therefore seems probable that the clone (c5d7) from which the switched clones were derived was one of the variants at the lower end of this switching scale, or that it gave rise to more stable variants. This is consistent with the percentage homotype (>99% ILTat 1.25) determined by immune lysis; if the trypanosomes had been switching at the high end of the scale the homotype percentage would have fallen substantially during the growth of the clone population, and subsequent expansion from the stabilate.

The purified PCR product from the first switched clone, sc2d16, was then cloned into the 'T-easy' vector. 400 bp sequence was obtained from either terminus of the cDNA, and this, in conjunction with *Hin*cII, and *Pvu*II restriction mapping, confirmed that product was identical to ILTat 1.25 cDNA. One by-product of the more stable switch products of ILTat 1.61c was that the activation mechanism could be studied by Southern analysis (section 4.7, page 94), rather than by the PCR method outlined in section 3.9 (page 70).

3.11 Summary

The chronic ILTat 1.2 infection produced a typically low level infection in the rabbit host, yielding three relapse populations over the 30 day period. Subsequent immune lysis investigation revealed that eleven distinct VATs were represented by the 88 clones (analysed by immune lysis) from the first and second relapse peaks. The number of clones isolated per VAT in the first relapse peak was varied, with no single antigen type predominating. However, in the second relapse peak ILTat 1.23 was by far the major VAT, present on every day of the relapse.

Three further ILTat 1.2 infections initiated from the same stabilate, and pre-treated with antisera against the eleven VATs, allowed the reproducibility of the initial infection to be examined. Four of the VATs (ILTats 1.25, 1.67, 1.68, 1.69) from the first relapse peak elicited antibodies at a similar time in all four hosts, while one VAT (ILTat 1.21) showed more variation in the timing of its onset, and the final VAT (ILTat 1.64) only appeared in two of the four rabbits. The second relapse peak demonstrated much greater variation with only one of the five VATs appearing in all of the infections (ILTat 1.70). This pattern is consistent with the semi-predictable hierarchy of infection described by Miller & Turner (1981) and was originally demonstrated by Gray (1965), van Meirvenne *et al.* (1975), and Capbern *et al.* (1977).

Due to the sensitivity of the ICR mice to the parasites, only 11 switched clones were isolated from the first relapse peak of the ILTat 1.61c infection. Although this line displays an overall switching rate of about 3×10^{-2} switches/cell/generation, individual trypanosomes demonstrate a variety of rates down to 1×10^{-4} switches/cell/generation (Turner, 1997). It seems probable that switched clones were generated from a cell at the lower end of this range, since at least 9 of the 11 expanded populations were determined to be >99% ILTat 1.25 by immunolysis.

CHAPTER 4

GENETIC ANALYSIS OF THE VSG SWITCHING MECHANISMS UTILIZED BY PLEOMORPHIC TRYPANOSOMES

4.1 Introduction

Following the identification and characterization of the six *VSG* genes isolated from the first relapse peak of the chronic ILTat 1.2 infection, the next phase in the investigation was to examine the mode of activation utilised in each of the switch events. The aim of this study was to test whether pleomorphic trypanosomes have a preference for duplicative transposition, as opposed to the predominance of "*in situ*" switching events seen in monomorphic lines (Liu *et al.* 1985).

By probing a blot of digested genomic DNA with radiolabelled *VSG* gene specific cDNA it was possible to determine the number of BCs of each gene, and establish whether the switch had occurred by a duplicative or non-duplicative event. In the case of a duplicative transposition an additional band, representing the ELC, was detected in the expressor genomic DNA, but was absent from non-expressor DNA.

Pulsed-field gel electrophoresis (PFGE) was also employed to determine the chromosomal localization of the *VSG* genes, using separation conditions kindly provided by S.E. Melville (Department of Pathology, Cambridge University). Previously EATRO 795 DNA had been resolved under the same conditions at Cambridge and the individual chromosomes distinguished by hybridization with chromosome specific probes (Figure 17). Since the DNA in this investigation was also derived from EATRO 795, it was possible to infer the chromosomal location of the *VSG* genes by comparison with this standard.

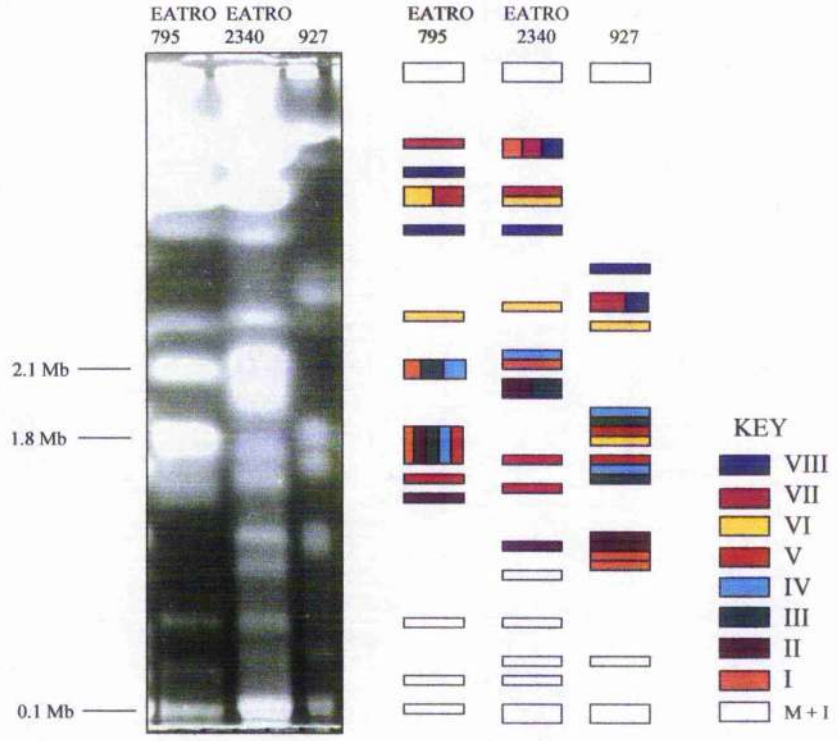


Figure 17. PFGE separation of chromosome-size DNA from 3 trypanosome stocks performed by S.E. Melville, Department of Pathology, Cambridge University. DNA from EATRO 795, EATRO 2340 and TREU 927 (Trypanosomiasis Research Edinburgh University) was separated on a 1.2% agarose gel at 85V, with a ramped pulse frequency of 1400-700s for 144h, and transferred to a nylon membrane by Southern blotting. The individual chromosomes were identified by hybridization with specific probes. Chromosomes I to VIII were resolved under these conditions, and their localisation is indicated by the colour coded boxes (representing the gel bands) to the right of the EtBr stained gel. The uncoloured boxes represent intermediate and minichromosomes (M + I). Approximate chromosomal size is indicated to the left of the gel picture.

4.2 Genomic analysis of the *VSG* genes activated during the first relapse peak

Clones representing each switched VAT, and ILTat 1.2, were grown from stabilate in immunosuppressed ICR mice, until the parasitaemia achieved approximately $10^{7.8}$ trypansosomes.ml^{-1}, when the parasites were collected and injected into two rats.

The stabilates used were: WUMP 677 (ILTat 1.2), WUMP 686 (ILTat 1.25), WUMP 690 (ILTat 1.67), WUMP 706 (ILTat 1.68), WUMP 707 (ILTat 1.69), WUMP 714 (ILTat 1.21), and WUMP 721 (ILTat 1.64) (see Figure 9). The trypanosomes were harvested when the rat parasitaemias achieved approximately $10^{8.1}$ cells.ml^{-1}, and were separated from blood components on a Percoll gradient, prior to the preparation of genomic DNA (see Materials and Methods, sections 2.5 and 2.6.2, pages 41 and 43).

1 μg of the genomic DNA was digested with *Eco*RI, *Hin*dIII, or *Pst*I, fractionated on a 0.7% agarose gel, Southern blotted onto nylon membrane and bound by UV radiation. These blots were then probed with VSG-specific cDNA, which had been excised from the plasmid, gel purified (QIAGEN gel extraction kit), and α^{32}P radiolabelled using the 'Prime-It II' kit (Stratagene), and washed to a final stringency of 0.1 x SSC, 0.1% SDS at 65°C. Full length radiolabelled cDNAs (approximately 1.6 kb) were used as probes in every case. However, probing with ILTat 1.67 cDNA produced a strong cross hybridization with other *VSG* gene sequences, and therefore a 400 base amino-terminus fragment was PCR amplified from the plasmid. This fragment was radiolabelled, and used as a more specific probe.

4.3 ILTat 1.2 *VSG* gene copies in expressor and non-expressor clones

The cDNA for the infecting VSG, ILTat 1.2, hybridized to four fragments in the *Hin*dIII digest of the 1.2 DNA (Figure 18, lane 1, and see also Figure 19, lane 1). This represents two genes, since there is an internal *Hin*dIII site within the coding sequence (physical maps of *VSG* genes in section 3.7) and is consistent with the ILTat 1.2 hybridization pattern previously reported (Williams *et al.* 1980). Two fragments comigrated at 1.1 kb, while two larger bands were observed close together at 6.2-6.5 kb. A similar pattern was seen in the 1.67 genomic DNA (lane 3), which also displayed four hybridizing fragments, although there was variation in the sizes of the larger fragments, allowing these two bands to be resolved. Only two fragments were detected in the 1.25 and 1.68 lanes, indicating that one of the 1.2 copies had been lost in these clones; the larger fragments also demonstrated size

differences. This alteration in fragment length between trypanosome clones indicates that the 1.2 genes are telomeric, the changes probably occurring in the telomere tract lying immediately downstream of the *VSG* gene.

To ensure that none of the bands was a product of partial digestion, the filter was reprobed with a PCR product encompassing the coding sequence of the trypanosome *RAD51* single copy gene (McCulloch and Barry, 1999) (Figure 26, page 93), which displayed a single band per track, indicating complete digestion. This control hybridization also demonstrated that the intense bands in the ILTat 1.2 track at 1.1 kb and 6.2-6.5 kb were doublets, rather than differences in DNA loading (although more DNA was present in the 1.2 and 1.67 lanes).

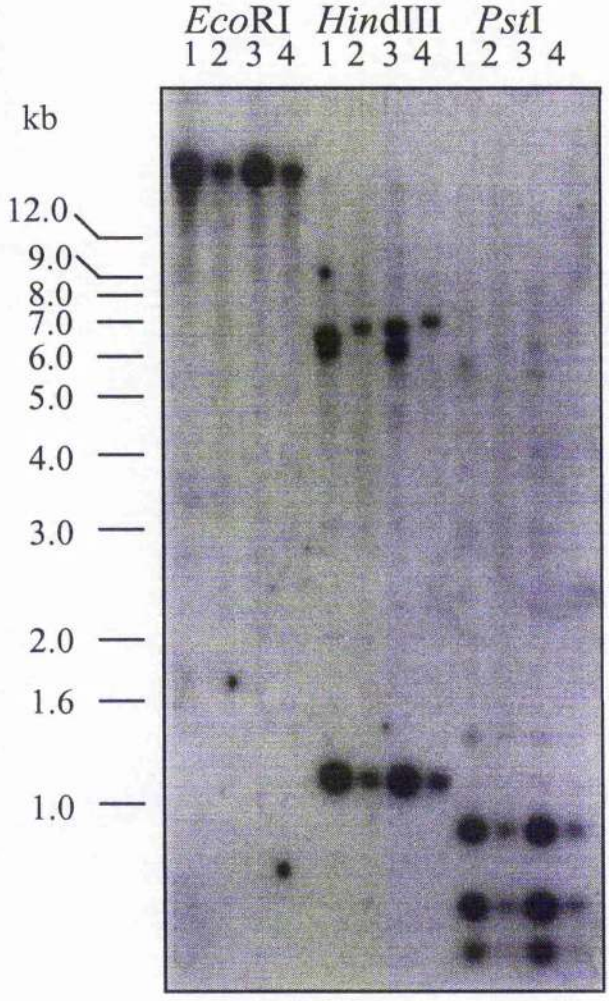


Figure 18. ILTat 1.2 *VSG* gene copies in ILTats 1.2, 1.25, 1.67 and 1.68. Genomic DNA from ILTat 1.2 (lane 1), ILTat 1.25 (lane 2), ILTat 1.67 (lane 3) and ILTat 1.68 (lane 4) was digested with *Eco*RI, *Hin*dIII or *Pst*I, separated on a 0.7% agarose gel and Southern blotted. The filter is probed with ILTat 1.2 VSG cDNA, and washed to 0.1 x SSC, at 65°C. Molecular markers (kb) are indicated to the left of the panel.

The *Pst*I digests were less informative than the *Hin*dIII digests as there are three internal sites within the coding sequence, therefore producing fragments of equal sizes from both telomeric copies. However, three clear hybridization products were

detectable at 0.9 kb (from the first internal site to another upstream of the *VSG* gene), 0.5 kb and 0.4 kb (two internal products). Again, the hybridization was more intense in the 1.2 and 1.67 lanes, confirming the double copy of the 1.2 gene in these clones. Very faint hybridization was also seen between 5 kb and 6 kb in all four lanes. These bands represent the probe binding to the last 300 bp of the *VSG* gene before the telomere, and the hybridization pattern is identical (but 1 kb smaller), to that observed for the telomeric fragments of the *Hin*dIII digests.

Digestion with *Eco*RI produced large fragments beyond the scale of the size markers (>12 kb), indicating that a considerable stretch upstream of the *VSG* gene is devoid of *Eco*RI restriction sites. High intensity hybridization in the ILTat 1.2 and 1.68 lanes again confirmed that two gene copies were present in these clones.

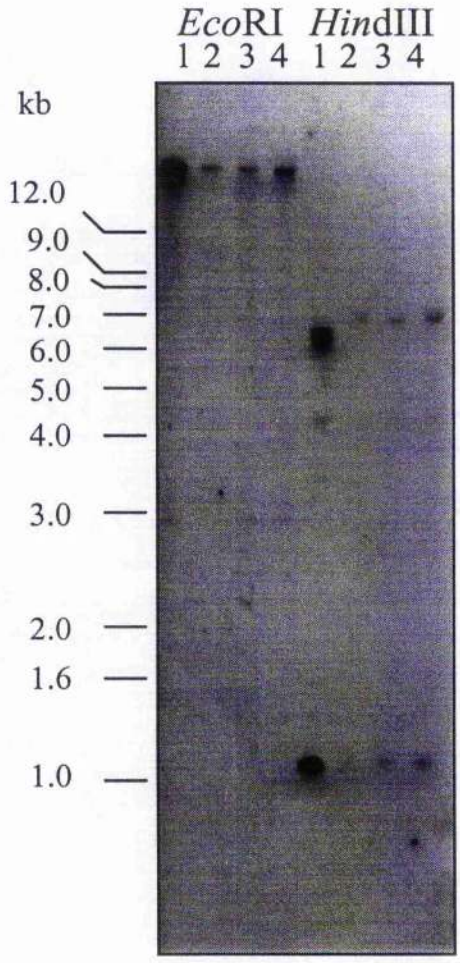


Figure 19. ILTat 1.2 *VSG* gene copies in ILTats 1.2, 1.68, 1.21 and 1.64. Genomic DNA from ILTat 1.2 (lane 1), ILTat 1.68 (lane 2), ILTat 1.21 (lane 3) and ILTat 1.64 (lane 4) was digested with *Eco*RI or *Hin*dIII, separated on a 0.7% agarose gel and Southern blotted. The filter was probed with ILTat 1.2 VSG cDNA, and washed to 0.1 x SSC, at 65°C. The *Pst*I tracks were too faint for graphical reproduction. The minor band at about 4.2 kb in lane 1 of the *Hin*dIII digest was observed in the other three lanes, but this region was also too faint to be represented graphically. Molecular markers (kb) are indicated to the left of the panel.

One of the two ILTat 1.2 *VSG* gene copies was also lost during each of the ILTat 1.69, 1.21 and 1.64 switches (Figure 19, lanes 2-4), and produced the same hybridization pattern that was observed with the 1.25 and 1.68 genomic DNA. Again, the two single bands were detectable at 1.1 kb and approximately 7 kb in the *Hin*dIII digests, with the larger fragment varying in size between clones probably due to changes in telomere length. The *Pst*I (data not shown), and *Eco*RI digests also yielded the same sized fragments as seen in the 1.25 and 1.68 hybridizations.

It was therefore apparent that five of the six VSG switches from the first relapse peak involved the loss of one of the ILTat 1.2 *VSG* gene copies, while both copies remained unaltered in the remaining activation (1.67).

4.4 Activation of the ILTat 1.25, 1.67 and 1.68 *VSG* genes

Probing with ILTat 1.25 *VSG* cDNA produced the most complicated hybridization pattern of all the six switched clones (Figure 20). Due to the presence of a large number of bands, the *Eco*RI and *Pst*I digests were difficult to interpret, but it was possible to explain the *Hin*dIII lanes. In the non-expressor lanes of this digest (1,3, and 4) there were two telomeric fragments at approximately 5-6 kb and 9 kb; three weaker bands (and an extremely faint fourth fragment) were also detectable. These results suggest that this trypanosome genome possesses two telomeric BCs of the ILTat 1.25 *VSG* gene, and also contains several highly homologous sequences (probably from closely related *VSG*s) which produce the fainter cross-reaction. Reprobing of the filter with an amino-terminus encoding fragment of the 1.25 cDNA did not change the hybridization pattern, demonstrating that the similarity between these cross-reacting sequences and the 1.25 *VSG* was occurring at the less conserved end of the gene (data not shown). This is indicative that the cross-reacting sequences are members of a *VSG* gene family.

In the trypanosome clone expressing the ILTat 1.25 gene, there was an extra fragment observed around 9.0 kb (lane 2), which corresponds to a classical ELC of the gene: the product of duplicative transposition. Again, the single copy *RAD51* (Figure 26) reprobing of this filter revealed that this additional product was not due

to partial digestion. In conclusion, it seems that there are two silent, telomeric gene copies of 1.25, one of which was activated by duplication into a *BES*. It also seems probable that the duplicated *VSG* gene was transposed to the *BES* previously occupied by the 1.2 gene, since one of 1.2 gene copies were lost during this activation (section 4.3, page 82).

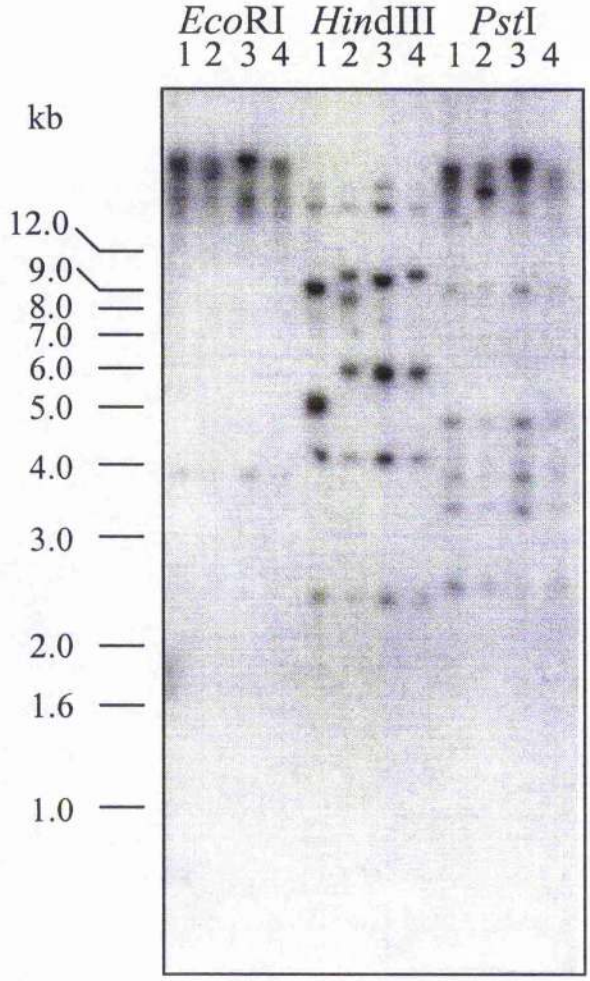


Figure 20. Activation of the ILTat 1.25 *VSG* gene. Genomic DNA from ILTat 1.2 (lane 1), 1.25 (lane 2), 1.67 (lane 3) and 1.68 (lane 4) was digested with *Eco*RI, *Hin*dIII or *Pst*I, separated on a 0.7% agarose gel and Southern blotted. This is the same filter as in Fig. 18, here probed with the 1.25 *VSG* cDNA, and washed to 0.1 x SSC, at 65°C. Molecular markers (kb) are indicated to the left of the panel.

The initial hybridization of ILTat 1.67, using the full length cDNA as a probe, produced an ambiguous result, with multiple bands apparent in every lane. Therefore, a 400 base amino-terminus fragment was PCR amplified from the plasmid, radiolabelled, and used to probe the filter; this hybridization proved to be more specific. Although *Eco*RI and *Pst*I sites are present within the coding sequence of the 1.67 *VSG* gene, they do not fall within the region encompassed by the 400 bp probe, and therefore single bands were observed (Figure 21).

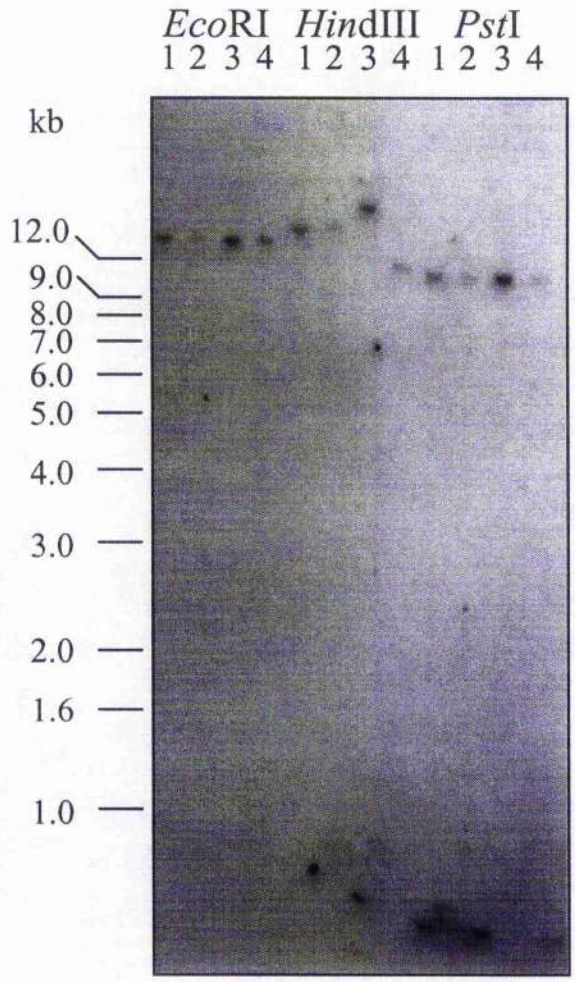


Figure 21. Activation of the ILTat 1.67 *VSG* gene. Genomic DNA from ILTat 1.2 (lane 1), 1.25 (lane 2), 1.67 (lane 3) and 1.68 (lane 4) was digested with *Eco*RI, *Hin*dIII or *Pst*I, separated on a 0.7% agarose gel and Southern blotted. This is the same filter as in Figs. 18 and 20, here probed with the first 400 bases of the 1.67 *VSG* cDNA (starting from the spliced leader), and washed to 0.1 x SSC, at 65°C. Molecular markers (kb) are indicated to the left of the panel.

This activation seems to have occurred by an *in situ* switch involving a single copy gene, since no additional band was generated in the expressor lane 3. This is consistent with the observation (in section 4.3) that the second ILTat 1.2 gene copy was retained after this event. The *BES*s are located on telomeres, and the occupancy of a *BES* by the 1.67 gene explains the variation in fragment sizes seen in the *Hin*dIII digests. This telomeric variation is not seen in the *Eco*RI or *Pst*I digests due to the internal restriction sites contained within the *VSG* coding sequence, resulting in the telomere being cleaved from the amino-terminus encoding *VSG* fragment. The ILTat 1.67 probe was specific to the first 400 bp of the *VSG* coding sequence and therefore only hybridized with the non-variable fragments produced by *Eco*RI or *Pst*I digestion.

Initially, the probing with ILTat 1.68 cDNA appeared to produce a pattern consistent with an *in situ* switch (Figure 22). Two gene copies were apparent in every lane; the more intense band was large and telomeric (>12 kb). The other was smaller (at about 5.9 kb in the *Hin*dIII digest) and possibly internal since the fragments were equal in size across all four clones (although this could also be explained by internal restriction sites downstream of a telomeric *VSG* gene). No additional ELC band was detected in the expressor lane (3). However, this activation is more complicated than it appears.

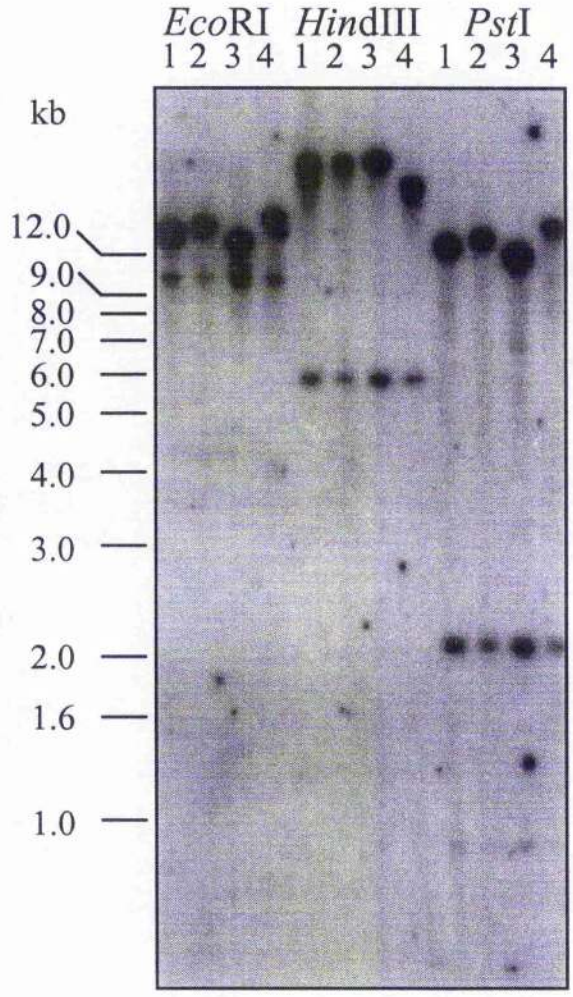


Figure 22. Activation of the ILTat 1.68 *VSG* gene. Genomic DNA from ILTat 1.2 (lane 1), 1.25 (lane 2), 1.67 (lane 3) and 1.68 (lane 4) was digested with *Eco*RI, *Hin*dIII or *Pst*I, separated on a 0.7% agarose gel and Southern blotted. This is the same filter as in Figs. 18, 20 and 21 here probed with the 1.68 *VSG* cDNA, and washed to 0.1 x SSC, at 65°C. Molecular markers (kb) are indicated to the left of the panel.

Close examination of the hybridization of the larger, telomeric fragments revealed that, although the size variation was reproducible between the *Eco*RI and *Pst*I digests, the *Hin*dIII generated fragments displayed a more variable pattern. For

example, in the *Eco*RI and *Pst*I digests, the ILTat 1.68 telomeric fragment was larger those of the 1.2, 1.25 or 1.67 clones, but in the *Hin*dIII digest it appeared smaller than its counterparts. This variation, which was not reproducible between the digests, could not be explained by changes in telomere length alone and suggested that recombination was occurring during the activation of the 1.68 gene. Further evidence for a recombinational event was seen in section 4.3 (page 82), when the second 1.2 gene copy was lost in the 1.68 clone. Later chromosomal analysis demonstrated that an ELC is produced during this activation, and is accompanied by an undetermined, additional rearrangement event (results shown in section 4.10, page 96).

There was a single *Pst*I restriction site in the ILTat 1.68 coding sequence, but this did not result in any additional bands since it occurred at the start of the amino-terminus encoding region. An *Eco*RI site was also present at the very end of the coding sequence of the cDNA clone used to generate the probe, although this was probably a PCR or cloning artefact since it was not detected in the other two clones. Telomeric variation would not have been observed if this site was present in the genomic DNA.

4.5 Activation of the ILTat 1.69, 1.21 and 1.64 *VSG* genes

The clearest hybridization patterns were produced when probing with ILTat 1.69, 1.21 and 1.64 *VSG* gene specific cDNAs, with all three activations occurring by ELC formation (Figures 23-25).

*Hin*dIII digestion revealed that two telomeric basic copies of the ILTat 1.69 *VSG* gene were present in this trypanosome line, represented by the fragments at 4.4-5.1 kb and 6.2-7.0 kb (Figure 23). A third band at about 3.4 kb was apparent in the expressor lane 2, indicating that this activation had occurred by ELC formation. The other two digests produced large fragments (>12 kb), although in both examples the smaller ELC could be detected in the 1.69 lane.

Following restriction digestion, it is more common to find that the fragment derived from the ELC is larger than the BC band. This is often due to the long 70 bp "barren region" upstream of the *VSG* gene; the stretch of sequence between the *VSG* gene in the *BES* and the telomere can also result in a larger fragment being liberated.

However, in the case of 1.69 it is possible that the telomeric fragment of the BC is actually longer than that of the *BES*, since the smallest band in the expressor lane is more intense than the other bands in all digests, suggesting that this represents the ELC.

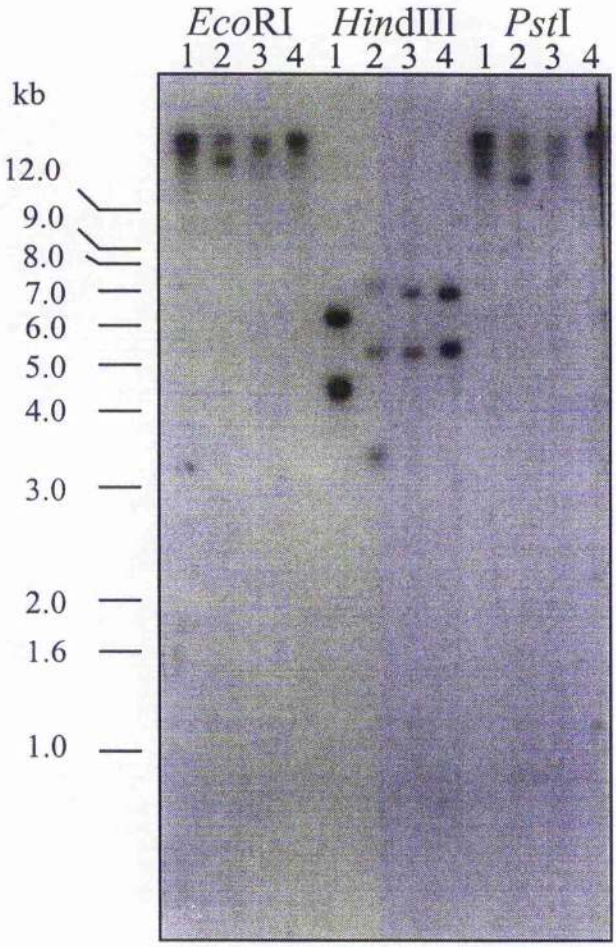


Figure 23. Activation of the ILTat 1.69 *VSG* gene. Genomic DNA from ILTat 1.2 (lane 1), 1.69 (lane 2), 1.21 (lane 3) and 1.64 (lane 4) was digested with *Eco*RI, *Hin*dIII or *Pst*I, separated on a 0.7% agarose gel and Southern blotted. This is the same filter as in Fig. 19, here probed with the 1.69 *VSG* cDNA, and washed to 0.1 x SSC, at 65°C. Molecular markers (kb) are indicated to the left of the panel.

A single telomeric fragment, probably representing a single copy gene, was observed between 5.5 and 6.3 kb in the *Hin*dIII digests of the ILTat 1.21 *VSG* cDNA probing (Figure 24). The larger band at about 8.5 kb in the expressor lane indicated that this *VSG* had also become activated by ELC formation. The *Eco*RI digests produced large fragments (>12 kb) that could not be resolved, although the hybridization in the 1.21 lane was strong enough to indicate a doublet (compare with the *RAD51* probing; section 4.6, Fig. 26, page 93). A *Pst*I restriction site in the coding sequence of 1.21 explained the hybridization pattern seen in the tracks digested with this enzyme. The band at 0.7 kb was the internal fragment spanning to a site upstream of the *VSG* gene, while the larger band (varying from 3.3 to 4.0 kb between clones) was the telomeric fragment running downstream from the internal site.

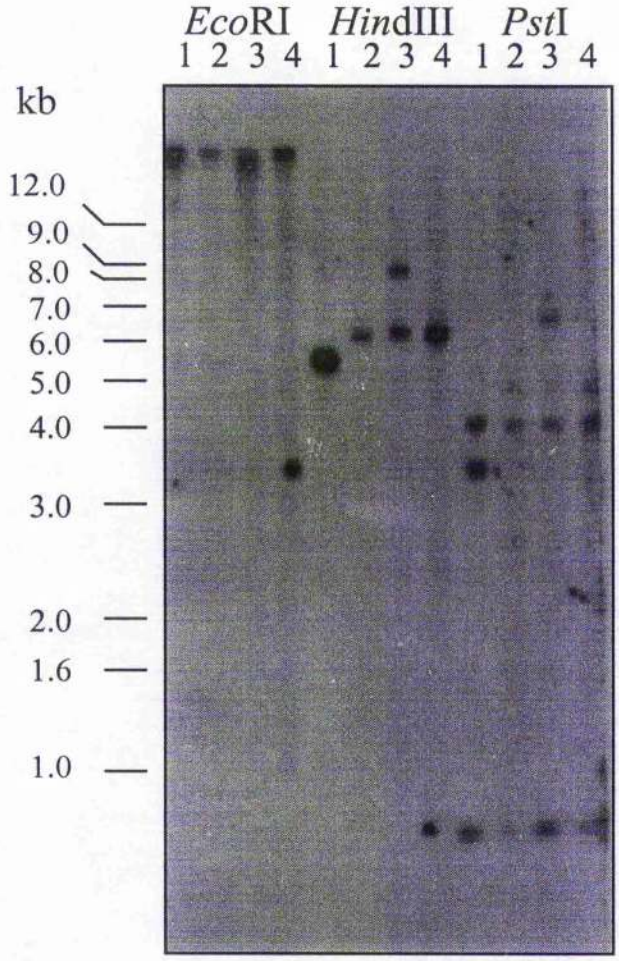


Figure 24. Activation of the ILTat 1.21 *VSG* gene. Genomic DNA from ILTat 1.2 (lane 1), 1.69 (lane 2), 1.21 (lane 3) and 1.64 (lane 4) was digested with *Eco*RI, *Hin*dIII, or *Pst*I, separated on a 0.7% agarose gel and Southern blotted. This is the same filter as in Figs. 19 and 23, here probed with the 1.21 *VSG* cDNA, and washed to 0.1 x SSC, at 65°C. Molecular markers (kb) are indicated to the left of the panel.

An additional band between approximately 4.0 and 4.7 kb was also seen in the *Pst*I digests of the four clones. This fragment was most likely a partial product generated by base J modification (see Introduction, section 1.9, page 25) of the *Pst*I site internal to the coding sequence, which renders the site indigestible in some DNA molecules. The band was 0.7 kb (the size of the internal fragment) larger than the major telomeric product, confirming that base J modification was a plausible explanation. When the filter was subsequently reprobed with a 400 base amino-terminus coding 1.21 fragment, both the downstream telomeric and ELC bands (which were previously a product of the full length cDNA probe hybridizing to the telomeric fragments downstream of the internal *Pst*I site) were no longer detectable (data not shown). However, the faint telomeric product was still present (along with the expected 0.7 kb internal fragment), indicative of a partial restriction fragment, stretching from the telomere to the *Pst*I site upstream of the *VSG* gene. The fact that no partial product was generated from the largest (6.7 kb) *Pst*I band demonstrates

that this represents the ELC fragment, since base J modification does not occur at the active telomere.

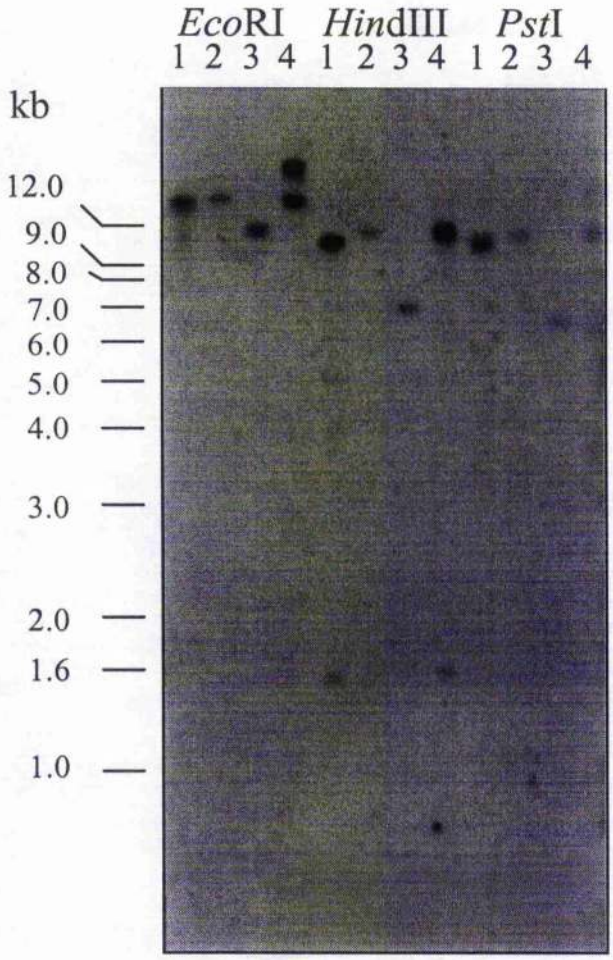


Figure 25. Activation of the ILTat 1.64 *VSG* gene. Genomic DNA from ILTat 1.2 (lane 1), 1.69 (lane 2), 1.21 (lane 3) and 1.64 (lane 4) was digested with *Eco*RI, *Hin*dIII or *Pst*I, separated on a 0.7% agarose gel and Southern blotted. This is the same filter as in Figs. 19, 23 and 24 here probed with the 1.64 *VSG* cDNA, and washed to 0.1 x SSC, at 65°C. Molecular markers (kb) are indicated to the left of the panel.

The ILTat 1.64 *VSG* activation also occurred by duplicative transposition (Figure 25). In the *Eco*RI digests a long, single telomeric fragment was present, with an additional, larger ELC band in the expressor track. The single telomeric fragment was also seen in the *Hin*dIII digests between approximately 6.5 and 10.5 kb; the duplicated copy appeared to be the same size as the BC, producing strong hybridization in the 1.64 track. A faint band at 1.6 kb was also detected, corresponding to an upstream fragment stretching from the internal site at 333 bp from the start of the *VSG* gene coding sequence. The *Pst*I hybridizations were very faint, although the telomeric variation could just be made out at around 6-10 kb. Another two (or three if there was a *Pst*I site upstream of the *VSG* gene) bands should have been observed in these tracks, although the exposure time was probably too short for these to be seen by autoradiography.

4.6 RAD51 control hybridizations

By probing the filters with a known single copy gene it was possible to determine whether the enzyme digestions had been complete. A PCR product encompassing the coding sequence of the trypanosome *RAD51* gene (a homologue of the bacterial *RecA* gene) was generated, radiolabelled, and used to reprobe the filters used for the study of the *VSG* activations. Single products were observed in all cases at approximately 9.0 kb in the *Eco*RI and *Hin*dIII digests, and 2.4 kb in the *Pst*I lanes, demonstrating that digestion was complete. The signal intensity of this control hybridization also revealed the relative amounts of genomic DNA that had been loaded in each lane. This enabled an accurate interpretation of the various hybridization effects observed during the *VSG* probings, such as the doublets observed in the ILTat 1.2 hybridizations (Figures 18 & 19).

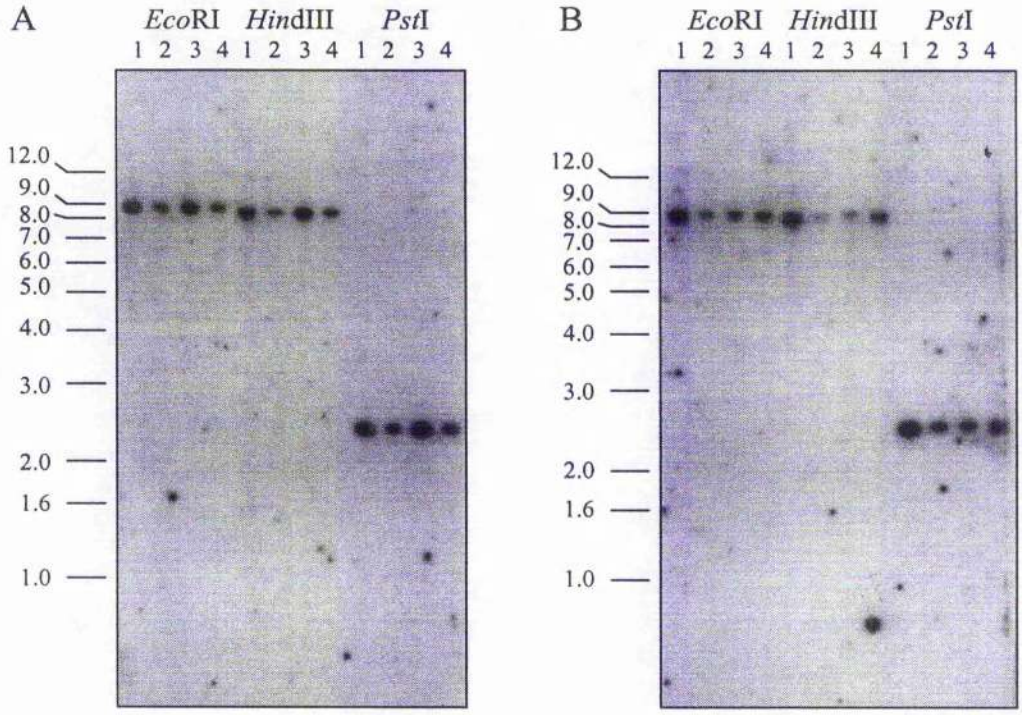


Figure 26. *RAD51* single copy gene control probing. Genomic DNA was digested with *Eco*RI, *Hin*dIII, or *Pst*I, separated on a 0.7% agarose gel and Southern blotted, and probed with a PCR product encompassing the coding sequence of the trypanosome *RAD51* gene. Panel A: ILTat 1.2 (lane 1), 1.25 (lane 2), 1.67 (lane 3) and 1.68 (lane 4). This is the same filter as in Figs. 18 and 20-22. Panel B: ILTat 1.2 (lane 1), 1.69 (lane 2), 1.21 (lane 3) and 1.64 (lane 4). This is the same filter as in Figs. 19 and 23-25. The filters were washed to 0.1 x SSC, at 65°C. Molecular markers (kb) are indicated to the left of each panel.

4.7 Activation of the ILTat 1.25 *VSG* gene from the 1.61c high switching single relapse investigation

The activation mechanism of the ILTat 1.25 clone isolated from the ILTat 1.61c rapidly switching single relapse experiment was also determined by Southern analysis. Genomic DNA was prepared from trypanosomes (as in section 4.2, page 81) for ILTat 1.61c (GUP 2812) and the ILTat 1.25 expressor c2d16 (WUMP 1060). Afterwards, 1 μg of this DNA was digested with *Hin*dIII, *Bam*HI, or *Eco*RI, and fractionated on a 0.7 % agarose gel. Following Southern blotting, the filter was probed with VSG-specific DNA, excised from the c2d16 plasmid.

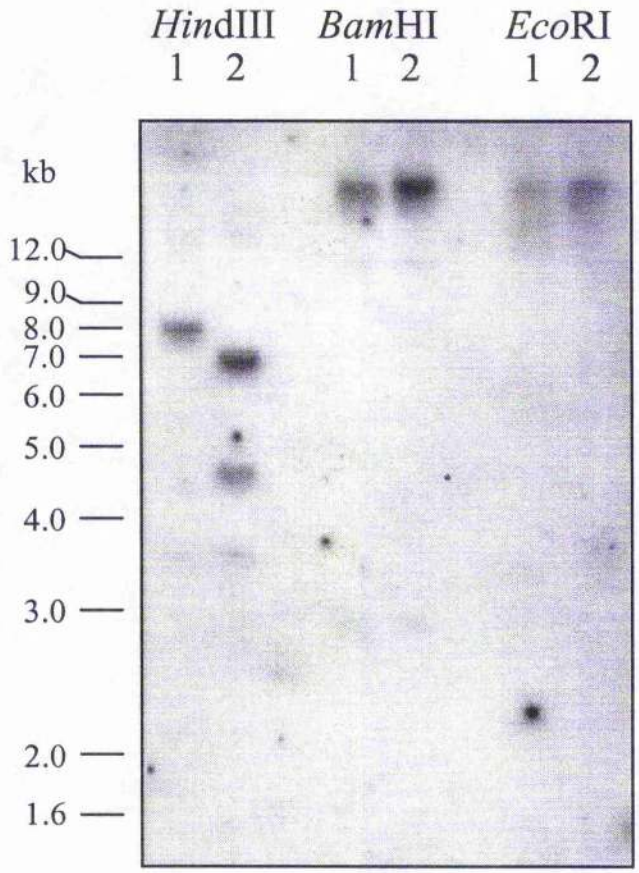


Figure 27. Activation of the ILTat 1.25 *VSG* gene from the ILTat 1.61c high switching experiment. Genomic DNA from ILTat 1.61c (lane 1) and 1.25 (lane 2) was digested with *Hin*dIII, *Bam*HI or *Eco*RI, separated on a 0.7% agarose gel and Southern blotted. The filter was probed with the 1.25 *VSG* cDNA (excised from the c2d16 plasmid) , and washed to 0.1 x SSC, at 65°C. An additional cross-reacting band at 2.1 kb was too faint for graphical reproduction. Molecular markers (kb) are indicated to the left of the panel.

The hybridization revealed that, in this study, the ILTat 1.25 *VSG* gene had become activated by ELC formation, duplicated from a single telomeric donor. In the *Hin*dIII digests the BC and ELC can be seen between approximately 4.6 and 8.0 kb. The *Bam*HI and *Eco*RI digests were less informative, producing large fragments (>12 kb).

4.8 Pulsed Field Gel electrophoresis (PFGE) analysis of the *VSG* genes activated during the first relapse peak

PFGE was utilised to determine the chromosomal location of the *VSG* gene BCs and ELCs. Trypanosomes were grown from stabilate (as in sections 4.2 and 4.7, pages 81 and 94) and separated from blood on a Percoll gradient. Genomic plugs were then prepared by immobilizing the live cells in 0.7% low melting point agarose (5×10^7 trypanosomes per 100 µl plug). Gels were run on the CHEF-DR III system using half a genomic plug per lane; 3 PFG conditions were utilised during this investigation. The 6-day general separations were run on a 1.2% agarose gel at 15°C in 0.089 M Tris-borate, 0.1 mM EDTA (85V, 1400-700s pulse time, 144 h). The 1.8 Mb chromosomal cluster was expanded under the same conditions, except that a fixed pulse time of 600s was used. The minichromosomal separations were run on a 1.0% agarose gel at 14°C in 0.045 M Tris-borate, 0.5 mM EDTA (200V, 20s pulse time, 16h).

The resultant gels were Southern blotted onto nylon membrane and bound by UV radiation. These blots were then probed with the α-^{32}P radiolabelled *VSG*-specific cDNAs (as in section 4.2, page 81) and washed to a final stringency of 0.1 x SSC, 0.1% SDS at 65°C.

4.9 Chromosomal location of the ILTat 1.2 *VSG* genes

Probing the 6-day general separation with ILTat 1.2 *VSG* cDNA revealed two clear bands in the expressor lane (Figure 28). Hybidization also occurred at the slot due to a large proportion of these two chromosomes remaining entrapped within the plug during separation. The first 1.2 copy was detected on the 2.1 Mb cluster (representing 3 comigrating chromosomes – see section 4.1, page 80), while the second occurred on the 1.8 Mb group (consisting of 5 chromosomes). The 1.8 Mb copy was retained in the 1.67 clone, but lost during the activation of clones 1.25, 1.68, 1.69, 1.21 and 1.64. It can therefore be deduced that the 1.8 Mb copy is located at the *BES*, remaining unchanged during the transcriptional switch seen in 1.67, but replaced by the new ELC in the five duplicative events.

Two ILTat 1.2 gene copies were also detected at identical locations in the 1.61c clone; both these copies were retained during the 1.25 duplicative activation event (produced from the single relapse study). This suggests that the 1.61c line is utilising a *BES* separate from the one occupied by the 1.2 *VSG* gene (on the 1.8 Mb chromosome).

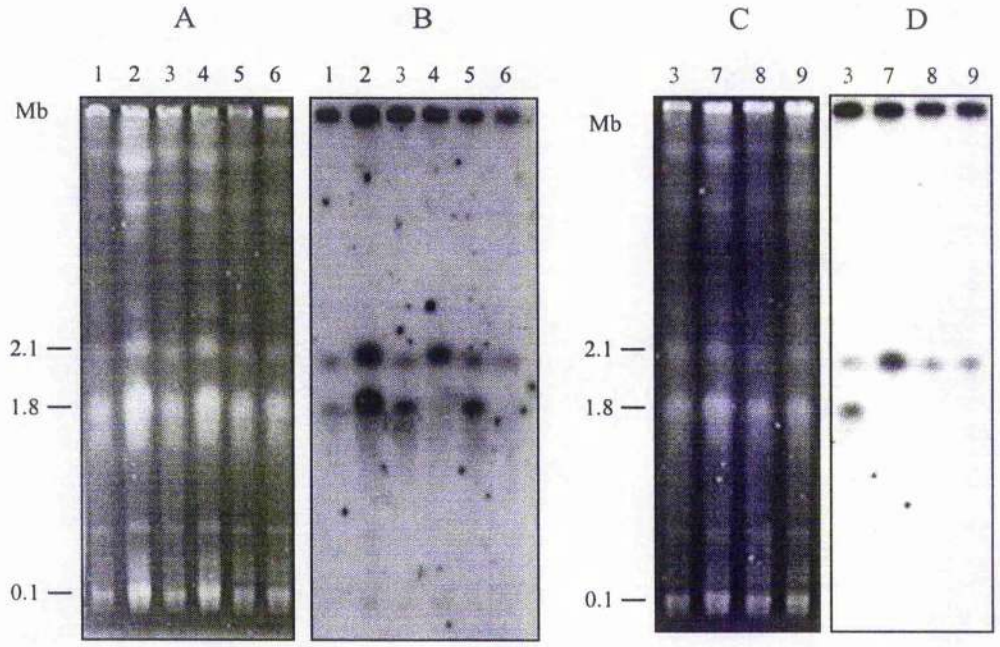


Figure 28. Chromosomal location of the ILTat 1.2 *VSG* genes. Chromosome-size DNA of ILTat 1.61c (lane 1), 1.25 (from c2d16, lane 2), 1.2 (lane 3), 1.25 (lane 4), 1.67 (lane 5), 1.68 (lane 6), 1.69 (lane 7), 1.21 (lane 8) and 1.64 (lane 9) was separated on a 1.2% agarose gel at 85V, with a ramped pulse frequency of 1400-700s for 144h. The gels (panels A and C) were then Southern blotted, and the resultant filter was probed with ILTat 1.2 *VSG* cDNA (panels B and D). The filter was washed to 0.1 x SSC, at 65°C. Approximate chromosomal size is indicated (Mb) to the left of panels A and C.

4.10 PFGE analysis of the ILTat 1.25, 1.67 and 1.68 *VSG* genes

The most ambiguous PFG results were produced when the filter was reprobed with ILTat 1.25 cDNA, because the cross-reacting bands seen in the previous genomic blots complicated the pattern. An ELC was observed on the 1.8 Mb cluster in the two expressor lanes (2 and 4). Unfortunately there was more DNA present in the

genomic plugs of both these clones, but the difference in intensity between the expressor and non-expressor tracks was more marked here than in any of the other probings (see Figures 28, 30 and 31); this confirms that these bands were real, rather than differences in loading intensity. A minichromosomal fragment at 0.1 Mb was also seen in every lane, and this accounts for one of the telomeric BCs seen in Figure 20.

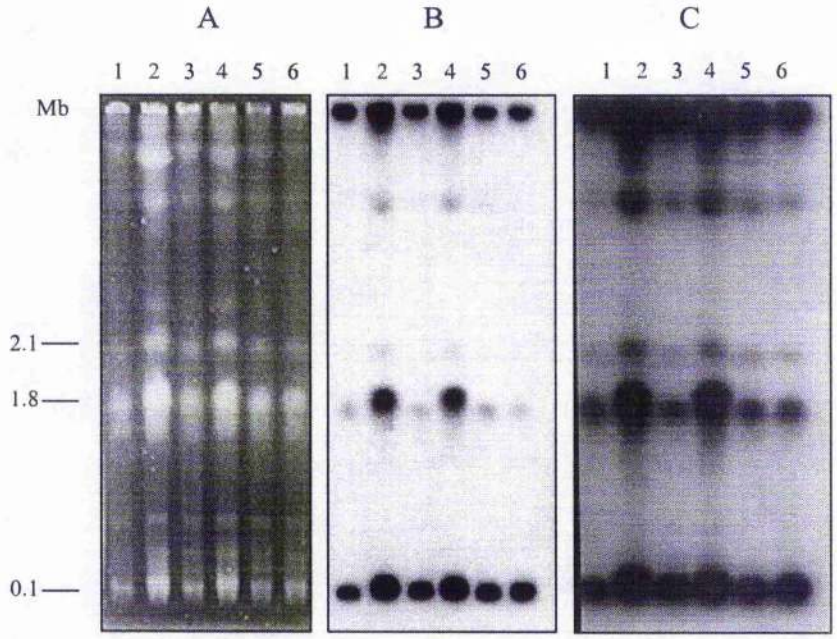


Figure 29. ILTat 1.25 *VSG* cDNA probing of PFGE separated chromosome size DNA. DNA from ILTat 1.61c (lane 1), 1.25 (from c2d16, lane 2), 1.2 (lane 3), 1.25 (lane 4), 1.67 (lane 5) and 1.68 (lane 6) was separated on a 1.2% agarose gel at 85V, with a ramped pulse frequency of 1400-700s for 144h. The gel (panel A) was then Southern blotted, and the resultant filter was probed with ILTat 1.25 *VSG* cDNA (panel B). The filter was washed to 0.1 x SSC, at 65°C. Panel C displays a longer exposure of the same blot. This is the same filter as in Fig. 28 panel B. Approximate chromosomal size is indicated (Mb) to the left of panel A.

The previous genomic digests had revealed that the ILTat 1.25 clone activated from the 1.2 rabbit infection possessed two ILTat 1.25 gene copies, while only 1 copy was observed in the clone activated from the 1.61c single relapse mouse study. It could therefore be determined that the three fainter PFGE bands, apparent between the slot and the minichromosomes, were the cross-reactions seen in Figures 20 and 27. The

second 1.25 *VSG* gene copy in the ILTat 1.2 derived clones must therefore reside on a large chromosome that has remained in the slot under these PFGE conditions. The slot also hybridized in clone 1.61c because much of the cross-reacting DNA remained in the plug during separation. DNA transfer from the slot is not quantitative, and it appears that signal strength of hybridization to the chromosomes that remain within the slot under this separation is weaker than the signal produced by the partially trapped DNA (see Figure 32, panels C and D). This phenomenon could explain why the slot hybridization of the ILTat 1.2 derived samples is no stronger than in the ILTat 1.61 derived clones.

It appeared that the ILTat 1.25 ELC (produced in the rabbit infection) replaced the 1.2 expressed copy that was previously residing in the active *BES* (Figure 28). However, in the 1.25 activation derived from the single relapse mouse experiment (from 1.61c) this 1.2 copy remained unaltered (Figure 28), demonstrating that this ELC must have been expressed from another *BES*, also present in the 1.8 Mb cluster.

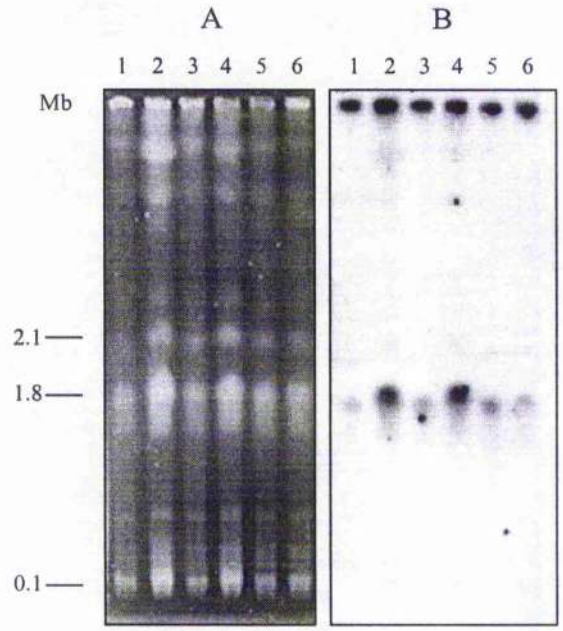


Figure 30. ILTat 1.67 *VSG* cDNA probing of PFGE separated chromosome size DNA. DNA from ILTat 1.61c (lane 1), 1.25 (from c2d16, lane 2), 1.2 (lane 3), 1.25 (lane 4), 1.67 (lane 5) and 1.68 (lane 6) was separated on a 1.2% agarose gel at 85V, with a ramped pulse frequency of 1400-700s for 144h. The gel (panel A) was then Southern blotted, and the resultant filter was probed with the 400 base amino-terminus ILTat 1.67 specific *VSG* cDNA (panel B). The filter was washed to 0.1 x SSC, at 65°C. This is the same filter as in Figs. 28 (panel B) and 29 (panels B and C). Approximate chromosomal size is indicated (Mb) to the left of panel A.

Reprobing the filter with the 400 base amino-terminus probe of ILTat 1.67 produced a single band in every lane (Figure 30), indicating that no genetic rearrangements had occurred during the activation of this gene. It also appeared that, like the ILTat 1.2 *BES*, the 1.67 *BES* activated in this *in situ* switch was localised on the same 1.8 Mb chromosomal cluster.

The unusual hybridization pattern observed from the genomic digest of ILTat 1.68 DNA (Figure 22) was elucidated from the PFGE results (Figure 31). The probing revealed that one of the 1.68 *VSG* gene BCs was minichromosomal, while the second copy probably resided on a larger, slot chromosome. Hybridization was also observed at 1.8 Mb in the expressor track, but the minichromosomal basic copy was lost during this activation. It therefore seems plausible that this gene was copied from the minichromosome and transposed to the *BES*, replacing the 1.2 copy (as seen in Figure 18), but the minichromosomal copy was then lost by another rearrangement event. An alternative explanation is that the activation occurred by reciprocal telomere recombination between the ILTat 1.68 minichromosome and the *BES* occupied by the 1.2 *VSG*, with the subsequent loss of the 1.2 copy (present on the minichromosome after the switch).

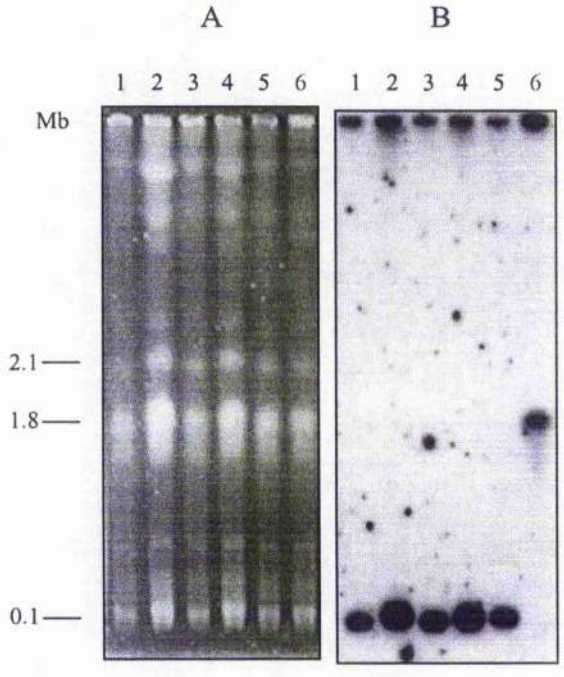


Figure 31. ILTat 1.68 *VSG* cDNA probing of PFGE separated chromosome size DNA. DNA from ILTat 1.61c (lane 1), 1.25 (from c2d16, lane 2), 1.2 (lane 3), 1.25 (lane 4), 1.67 (lane 5) and 1.68 (lane 6) was separated on a 1.2% agarose gel at 85V, with a ramped pulse frequency of 1400-700s for 144h. The gel (panel A) was then Southern blotted and the resultant filter was probed with ILTat 1.68 *VSG* cDNA (panel B). The filter was washed to 0.1 x SSC, at 65°C. This is the same filter as in Figs. 28 (panel B), 29 (panels B and C) and 30 (panel B). Approximate chromosomal size is indicated (Mb) to the left of panel A.

4.11 PFGE analysis of ILTat 1.69, 1.21 and 1.64 *VSG* genes

PFGE analysis confirmed that the ILTat 1.69, 1.21 and 1.64 *VSG* genes had become activated by duplicative transposition (Figure 32). In each case an ELC was seen in the expressor lane at the 1.8 Mb cluster (Figure 32, panels C, D and E), and this was accompanied by the loss of the 1.8 Mb copy of the 1.2 gene (Figure 32, panel B).

Panel D displays the filter probed with ILTat 1.21 *VSG* cDNA, which revealed the single, minichromosomal BC. There is no slot hybridization in the non-expressor lanes of this panel because the minichromosomal DNA is small enough to migrate unhindered from the plug. However, slot hybridization does occur in the expressor lane due to the partial entrapment of the 1.8 Mb chromosome containing the active *BES*.

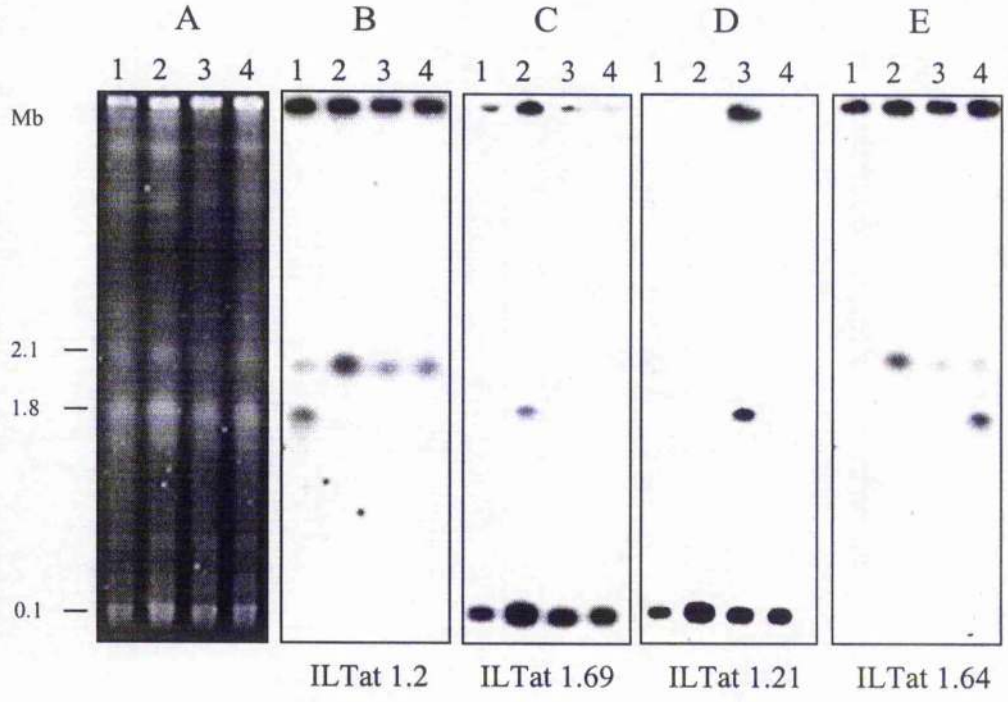


Figure 32. Transposition of the duplicated ILTat 1.69, 1.21, and 1.64 *VSG*s. Chromosome-size DNA from ILTat 1.2 (lane 1), 1.69 (lane 2), 1.21 (lane 3) and 1.64 (lane 4) was separated on a 1.2% agarose gel at 85V, with a ramped pulse frequency of 1400-700s for 144h. The gel (panel A) was then Southern blotted, and the resultant filter was initially probed with ILTat 1.2 *VSG* cDNA (panel B). The remaining three panels display the same filter probed sequentially with ILTat 1.69 (panel C), 1.21 (panel D) and 1.64 (panel E). The filters were washed to 0.1 x SSC, at 65°C. Approximate chromosomal size is indicated (Mb) to the left of panel A.

A minichromosomal copy was also detected when probing with ILTat 1.69 *VSG* cDNA (panel C), but slot hybridization was also observed in the non-expressor lanes, indicating that a second BC must reside on a large unresolved chromosome; this is consistent with the two copies seen in section 4.5 (page 89).

The ILTat 1.64 hybridization (panel E) revealed another single copy gene (seen in section 4.5, page 89) with the BC apparent on the 2.1 Mb chromosomal cluster; partial entrapment explains the slot hybridization also observed in this probing.

4.12 Resolution of the 1.8 Mb cluster

It was possible to separate the five 1.8 Mb comigrating chromosomes by fixing the pulse time of the general run at 600s (previously ramped at 1400-700s) (Figure 33).

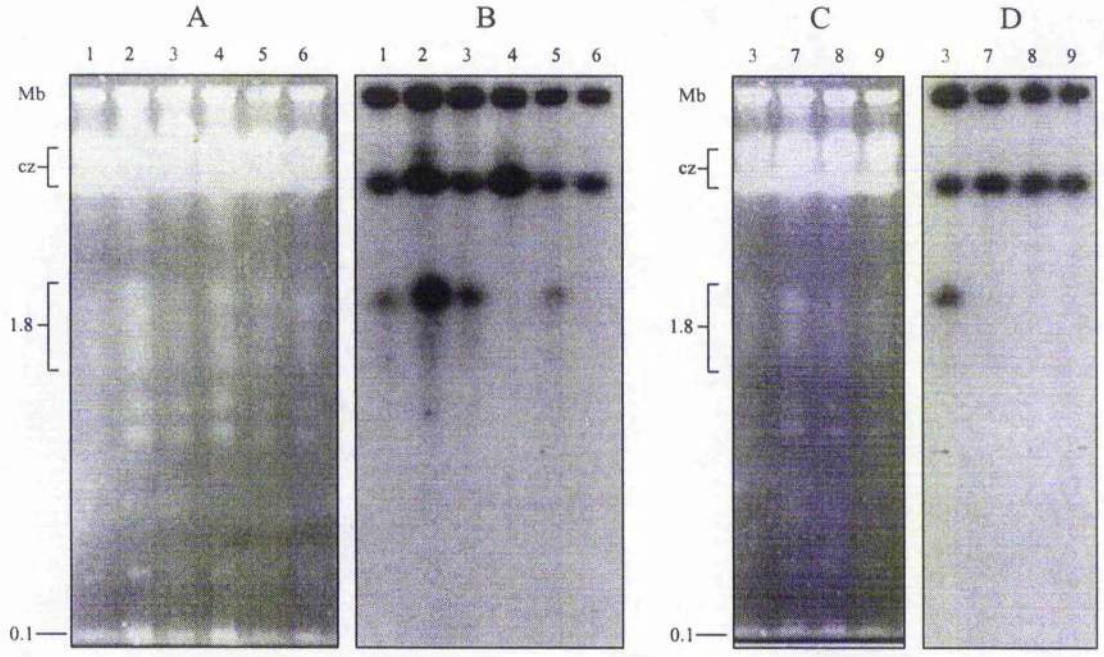


Figure 33. 600s PFGE resolution of the 1.8 Mb comigrating chromosomes, probed with ILTat 1.2 cDNA. Chromosome-size DNA from ILTat 1.61c (lane 1), 1.25 (from c2d16, lane 2), 1.2 (lane 3), 1.25 (lane 4), 1.67 (lane 5), 1.68 (lane 6), 1.69 (lane 7), 1.21 (lane 8) and 1.64 (lane 9) was separated on a 1.2% agarose gel at 85V, with a fixed pulse frequency of 600s for 144h. The gels (panels A and C) were then Southern blotted, and the resultant filters were probed with ILTat 1.2 *VSG* cDNA (panels B and D). The filters were washed to 0.1 x SSC, at 65°C. Approximate chromosomal size is indicated to the left of panels A and D (cz-compression zone).

The two ILTat 1.2 gene copies were detected when the 600s separated filter was probed with 1.2 *VSG* cDNA. The 2.1 Mb chromosome was localised to the smaller band of the compression zone, while the *BES* was apparent on the largest of the 1.8 Mb chromosomes.

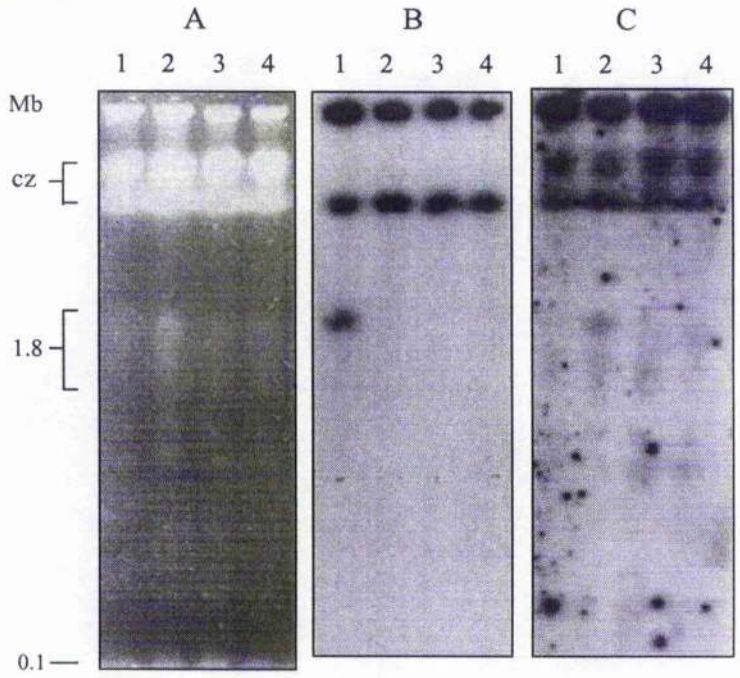


Figure 34. 600s PFGE resolution of the 1.8 Mb comigrating chromosomes, probed with ILTat 1.2 or 1.69 cDNA. Chromosome-size DNA from ILTat 1.2 (lane 1), 1.69 (lane 2), 1.21 (lane 3) and 1.64 (lane 4) was separated on a 1.2% agarose gel at 85V, with a fixed pulse frequency of 600s for 144h. The gel (panel A) was then Southern blotted, and the resultant filter was probed sequentially with 1.2 and 1.69 *VSG* cDNA (panels B and C respectively). Panels A and B are identical to panels C and D, respectively, in Fig. 33. The filters were washed to 0.1 x SSC, at 65°C. Approximate chromosomal size is indicated to the left of panel A (cz-compression zone).

Reprobing of the filter produced unconvincing results because the harsh stripping procedure (required for PFGE blots) removed too much of the non-compression zone DNA. However, in the 1.69 reprobing the ELC was observed on the largest 1.8 Mb chromosome (Figure 34). This confirms that this duplicated copy was transposed to the same *BES* that was formerly occupied by the 1.2 gene. Hybridization was also seen on both bands in the compression zone, and this is probably due to the large chromosome, previously seen in the slot of the general separation. The minichromosomes migrated to the bottom of the gel in the 600s separation, and were not transferred to the filter during the Southern blot.

4.13 Resolution of the minichromosomal BCs

The 16h PFGE separation resolved the intermediate and minichromosomes (Figure 35). Several different size classes of minichromosome were present between approximately 35 and 100 kb, and hybridization was observed at about 100 kb (ILTat 1.25), 90 kb (1.68 and 1.69), and 55 kb (1.21); again the 1.68 minichromosomal copy was absent from the expressor track. Probing with the other *VSG* cDNAs produced only compression zone hybridization (data not shown). The apparent signals seen in panel F slots occurred because the wells of these blots were marked with pencil.

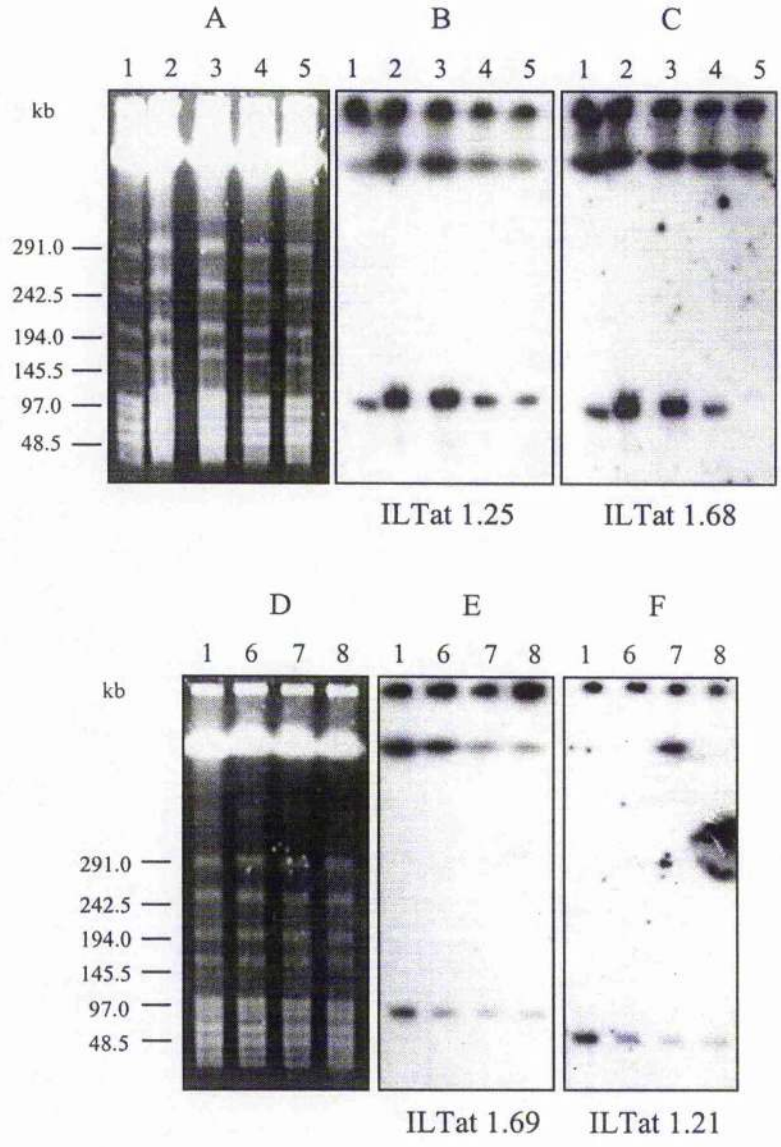


Figure 35.
PFGE resolution of the minichromosomes. Chromosome-size DNA from ILTat 1.2 (lane 1), 1.25 (from c2d16, lane 2), 1.25 (lane 3), 1.67 (lane 4), 1.68 (lane 5), 1.69 (lane 6), 1.21 (lane 7) and 1.64 (lane 8) was separated on a 1.0 % agarose gel at 200V, with a pulse frequency of 20s for 16 hours. The gels (panels A and D) were then Southern blotted, and the resultant filters were probed with 1.25 (panel B), 1.68 (panel C), 1.69 (panel E) and 1.21 (panel F) *VSG* cDNA. The filters were washed to 0.1 x SSC, at 65°C. Molecular markers (kb) are indicated to the left of panels A and D.

4.14 Summary

The genetic analysis of the six VATs seen in the first relapse peak has revealed that four became activated by duplicative transposition (ILTats 1.25, 1.69, 1.21 and 1.64), with a fifth possibly using this mechanism in conjunction with another event (1.68). ILTat 1.67 was the only *VSG* gene to be expressed following an *in situ* switch. Similar analysis of the second relapse peak demonstrated that all five *VSG* genes were activated by ELC formation (data not shown). Telomeric *VSG*s were the predominant source of ELCs, with fifteen of the eighteen silent genes observed in the first and second relapse peaks localised at the chromosome ends. All the gene study data are summarized in Table 3 (page 105).

Minichromosomal BCs apparently were donors in four of the six first relapse peak VATs (1.25, 1.68, 1.69, and 1.21), while the BCs of the other two VATs were single copy and either metacyclic (1.64), or located at a *BES* (1.67). Two of the VATs from the second relapse peak also possessed minichromosomal copies (1.23 and 1.72). The 1.21 ELC was duplicated from its single BC localised on a 35 kb minichromosome. It also seems likely that the 1.25 ELC was copied from its minichromosomal BC, because the independent 1.25 activation, observed in the single relapse mouse study (from 1.61c), used the single BC present on a 100 kb minichromosome. The activation of the 1.68 *VSG* gene also possibly involved the minichromosomal copy, since the whole 90 kb chromosome disappeared during the event. It is therefore evident that the minichromosomal *VSG* genes play an important role in antigenic variation at this early stage of infection.

In every duplicative activation detected in the first relapse peak, the ELC was accompanied by the loss of the 1.8 Mb ILTat 1.2 *VSG* copy, indicating that the same *BES* was utilised in each case. The 600s PFGE separation established that this *BES* was present on the largest of the five comigrating 1.8 Mb chromosomes.

Table 3. *VSG* loci and activation events.

Relapse peak	VSG	Activation mechanism	Silent gene		Activated gene	
			Number & location[c]	Minichromosomal genes	*BES*[d] size	Fate of 1.2 *VSG*
First	1.25[a]	duplication	1 telo	1	1.8 Mb	retained
	1.25	duplication	2 telo	1	1.8 Mb	lost
	1.67	"*in situ*"	1 telo	0	1.8 Mb	retained
	1.68	?[b]	1 telo, 1 int	1	1.8 Mb	lost
	1.69	duplication	2 telo	1	1.8 Mb	lost
	1.21	duplication	1 telo	1	1.8 Mb	lost
	1.64	duplication	1 telo	0	1.8 Mb	lost
Second	1.23	duplication	2 telo	1	ND	ND
	1.70	duplication	2 telo	0	ND	ND
	1.71	duplication	1 int	0	ND	ND
	1.22	duplication	1 telo	0	ND	ND
	1.72	duplication	1 telo, 1 int	1	ND	ND

a Switch event derived from the ILTat 1.61c single relapse mouse experiment

b Activation appeared to involve duplication and another event

c telo – telomeric; int – apparently chromosome-internal

d Five chromosomes comigrate at 1.8 Mb (Figure 17). It has been inferred from the pattern of 1.2 *VSG* loss that all the duplicative activations in the first relapse peak of the chronic rabbit infection converted the same *BES*. The 1.25 switch event produced from the 1.61c start clone (single relapse mouse study) utilised a separate *BES* on the 1.8 Mb cluster, while the 1.67 *VSG* (activated by an "*in situ*" switch occupied a third independent site also detected at 1.8 Mb

ND not determined

CHAPTER 5

CHARACTERIZATION OF THE UPSTREAM FLANKING REGION OF THE ILTat 1.21 MINICHROMOSOMAL *VSG* GENE

5.1 Introduction

The next phase of the investigation was to characterize the upstream flank of the *VSG* BCs that were activated by a recombinational event during the first relapse peak, and subsequently to determine the extent of the duplication involved in the formation of the ELC. In order to examine this upstream duplication boundary, it was necessary to isolate BC genomic clones of the relevant *VSG* genes. An ILTat 1.2 library (*Sau*3A partial digests in λ-GEM-12) had been manufactured previously by Nils Burman, and this was screened with the radiolabelled *VSG* cDNAs of ILTats 1.25, 1.68, 1.69, 1.21 and 1.64 (the five VATs activated in the first relapse peak by duplicative means). Although clones were isolated from four of these VATs, no clones could be derived for ILTat 1.21, which was a single copy minichromosomal gene. It therefore seemed possible that minichromosomal sequences were not represented in this library, in which case the ILTat 1.25, 1.68 and 1.69 clones would contain the non-minichromosomal BCs. Most telomeric sequences tend to be under-represented in libraries, while minichromosomal sequence, which consists almost entirely of either telomere or repeat sequence (Weiden *et al.* 1991), appears to be almost unclonable when manufacturing a standard size selected library (10-20 kb insert) (*e.g.* only 4 out of 18000 clones were found to contain minichromosomal repetitive sequence in the TREU927 BAC library at Cambridge University - S. Melville, pers. comm.).

An alternative approach was required for isolating clones of the minichromosomal genes, which seemed to be the preferred donors for duplicative transposition during the first relapse peak. This chapter describes the strategy that was employed for the isolation and characterization of these minichromosomal genomic clones. By preparing double restriction digests of genomic DNA, using one enzyme with a site internal in the *VSG* gene (*e.g. Sal*I) and a series of enzymes that possess 6-bp recognition sequences (and would not digest within the *VSG* coding sequence), it was possible to identify telomeric fragments using Southern analysis. In most of the digests, the minichromosomal fragment appeared as a consistently sized band, occurring above the range of the molecular markers (>12kb), and was distinguishable from the other large telomeric chromosomes. However, in some of the digests this minichromosomal fragment disappeared, indicating that the second

enzyme was cutting upstream of the *VSG* gene. It was important that the region upstream of the *VSG* gene extended at least 1.5-2.0 kb, to ensure that the fragment was large enough to encompass both the co-transposed region and the putative 70 bp repeats. However, it was necessary to limit the fragment size to minimise the amount of the minichromosomal repetitive sequence contained further upstream, in order to facilitate the subsequent cloning step.

Having identified an appropriate combination of restriction enzymes, it was possible to produce a minichromosomal library. Minichromosomal DNA was separated by PFGE, purified and digested with the two enzymes. This digested DNA was cloned into pBluescript to produce the library, which was then screened for the relevant *VSG* gene.

As each *VSG* requires a separate library based on its unique combination of enzyme sites, there was time for only two VATs to be studied using this approach. ILTat 1.21 was chosen because it possessed only one BC, which was minichromosomal. This VAT was only represented by 2 of the 36 trypanosome clones isolated from the first relapse peak, although it appears to play an important role in early infection, becoming activated in all four rabbit infections. ILTat 1.21 was also the predominant VAT seen in the single relapse studies of Miller and Turner, with 32 ILTat 1.21 activations occurring in 41 single relapse experiments in rats (Miller and Turner, 1981). ILTat 1.25, which was represented by 10 of the 36 clones from the first relapse peak of the ILTat 1.2 study, was investigated because it became activated at the start of infection in all four rabbit hosts, and was therefore a dominant early VAT. ILTat 1.25 also appeared to be a major VAT in the Miller and Turner study, occurring in 27 of 41 single relapses (Miller and Turner, 1981). Although there were 2 BCs of this *VSG* gene in the ILTat 1.2 line, it became activated from the single, minichromosomal, BC in the ILTat 1.61c study, and was therefore likely to have been duplicated from the same locus during the activation occurring in the ILTat 1.2 rabbit infection.

5.2 Identification of restriction sites suitable for minichromosomal library manufacture

It was first necessary to identify a restriction site within the *VSG* gene that could be utilised to liberate the BC upstream fragment. By using an enzyme that cut

approximately 0.5 to 1.0 kb downstream from the 5' end of the *VSG* coding sequence, it was possible to remove the 3' end and associated telomere from the BC fragment, whilst leaving a *VSG* specific tag attached which could be utilised in the screening of the subsequent library. *Sal*I was found to be a convenient enzyme for use with both ILTat 1.21 and 1.25 DNA, while *Age*I also proved useful when working with the 1.21 DNA. Both these enzymes have a 6-bp recognition sequence and their restriction sites in ILTat 1.21 and 1.25 cDNA are indicated in the physical maps below (Figure 36).

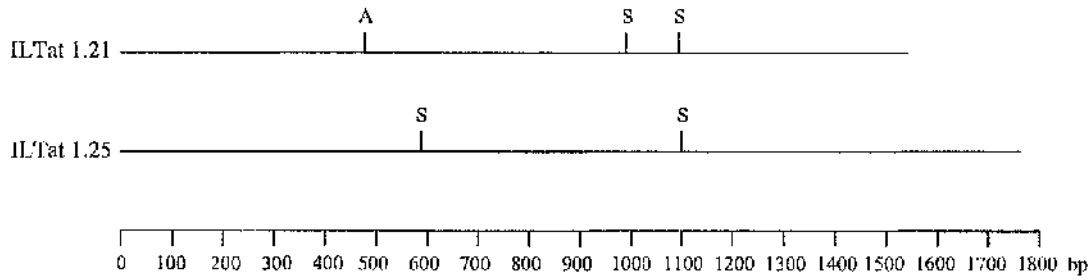


Figure 36. Physical maps of the ILTat 1.21 and 1.25 *VSG* full-length cDNAs (starting with the last seventeen nucleotides of the spliced leader – see Section 3.7, page 68). Restriction sites were determined from the full-length cDNA sequence of the VATs. Abbreviations: A, *Age*I; S, *Sal*I.

ILTat 1.2 genomic DNA was digested with *Sal*I (or *Age*I) and also double digested with *Sal*I (or *Age*I) plus a series of other enzymes that possess a 6-base recognition sequence (which did not cut within the *VSG* coding sequence). These digests were size fractionated on a 0.7% agarose gel and transferred to nylon membranes by Southern blotting. The filters where then probed with either ILTat 1.21 or 1.25 α-^{32}P radiolabelled specific cDNA. For ILTat 1.21, the probe was made from a 400 bp product (running downstream from the spliced leader) which had been PCR amplified from the cDNA clone, and was therefore specific to restriction fragments upstream of the *Age*I or *Sal*I sites. The ILTat 1.25 probe was produced from a 1 kb DNA region (again amplified from a cDNA clone, and running downstream from the start of the gene) which encompassed both the *Sal*I sites present in the cDNA. This probe therefore hybridized to the 0.5 kb internal *Sal*I fragment as well as the fragment upstream of the *VSG* beyond the first *Sal*I restriction site.

The ILTat 1.21 hybridizations revealed that very few 6-bp restriction sites are located upstream of this *VSG* gene; this result was consistent with a previous investigation of the ILTat 1.4 minichromosomal upstream domain (Williams *et al.* 1982). The hybridization typically produced a large, constantly sized, single band (see Figure 37) indicating that there was no restriction site upstream of the *VSG* for a considerable distance (and possibly no site at all) along the entire chromosome. In the ILTat 1.21 hybridization presented in section 4.5 (Figure 24) a smaller band was present in the *Hin*dIII and *Pst*I digests, revealing that sites for these enzymes did lie upstream of the *VSG* gene. However, these fragments (which include the full length of the *VSG* and its associated telomere) appeared to be too small for the cloning purposes, and the upstream sites were later shown to lie within the cotransposed region immediately upstream of the gene.

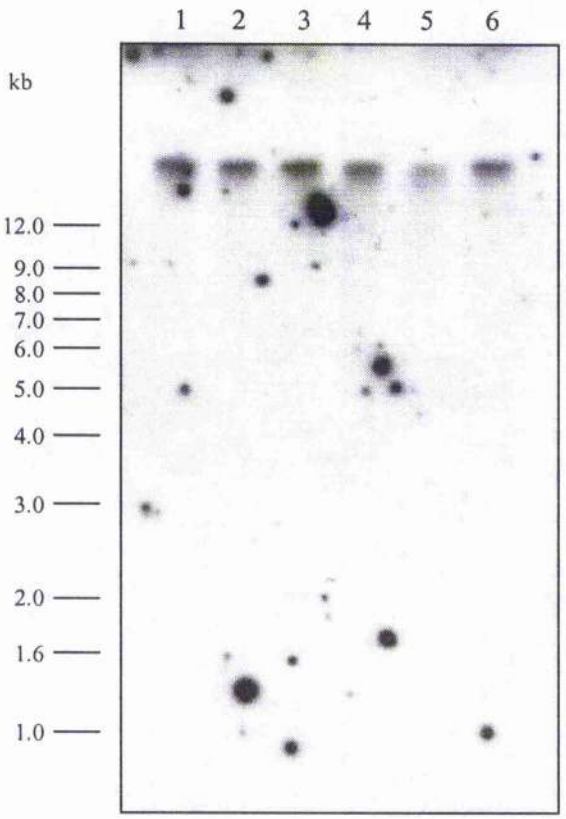


Figure 37. Autoradiograph demonstrating the sparsity of restriction sites upstream of the *VSG* gene in the ILTat 1.21 minichromosomal BC. ILTat 1.2 genomic DNA was digested with *Age*I (lane 1) or double digested with *Age*I/ *Kpn*I (lane 2), *Age*I/ *Nco*I (lane 3), *Age*I/ *Spe*I (lane 4), *Age*I/ *Xho*I (lane 5) and *Age*I/ *Xma*I (lane 6) and fractionated on a 0.7% agarose gel. The DNA was then transferred to a nylon membrane by Southern blotting and probed with a short stretch of ILTat 1.21 cDNA running 400 bases downstream from the spliced leader. The filter was washed to 0.1 x SSC, at 65°C. Molecular markers are indicated to the left of the panel.

When the ILTat 1.2 genomic DNA was double digested with *Sal*I/ *Nsi*I, the ILTat 1.21 amino-terminus probe hybridized to a 2.8 kb fragment, indicating that an *Nsi*I

site exists 1.8 kb upstream of this BC (Figure 38, panel A). This was the largest upstream fragment that could be found for the ILTat 1.21 BC, and it seemed plausible that this would span both the cotransposed and barren regions. Partial products were also seen at approximately 6 kb and 3.9 kb, due to incomplete digestion at the internal *Sal*I sites; the 6 kb band was also seen when the DNA was digested with *Nsi*I alone (data not shown). A later investigation demonstrated that these partial products probably were due to base J modification of the *Sal*I sites (section 5.5, page 121).

The ILTat 1.25 hybridizations produced a more complicated pattern; Figure 38, panel B (which is a duplicate of panel A) shows a filter probed with the 1.25 cDNA. This hybridization was harder to interpret because there are two 1.25 BCs, as well as additional cross-reacting bands (seen in section 4.4, page 85), and also partial products probably due to the base J modification of the *Sal*I sites. The 8-bp recognising *Not*I enzyme (lane 3) produced the simplest result, with a large band visible above the scale of the size markers (probably the minichromosomal copy) and a second band at about 9-10 kb representing the second BC; other fainter cross-reacting bands were apparent between these two bands.

*Nsi*I restriction sites were observed upstream of the *VSG* gene in both the ILTat 1.25 BCs, probably producing the small fragments visible at 1.3 kb and 1.8 kb; the three fainter products possibly represented the cross-reacting bands. However, these fragments only extended 0.8 kb and 1.3 kb upstream of the respective *VSG* genes, implying that these restriction sites were located within the cotransposed region, and were therefore too short to be utilised for the determination of the duplication boundary.

Digestion with *Sal*I/ *Bgl*II produced a 4.0 kb ILTat 1.25 fragment (Figure 38, panel B, lane 4), although it was impossible to determine the BC from which it was derived. It appeared that partial digestion had occurred in this digest due to the persistence of the two dominant bands (representing the intact BCs) that had been observed previously in the *Not*I digest (it was not determined if this was due to base J modification). The 4.0 kb fragment was also observed when ILTat 1.61c genomic DNA (which possesses a single minichromosomal BC of the ILTat 1.25 *VSG*) was digested with the same enzyme combination (data not shown); this suggested that the band was probably the duplicative donor in the 1.2 DNA. This *Sal*I/ *Bgl*II fragment, which extends 3.5 kb upstream of the *VSG* gene, represents an ideal size, and should

encompass both the cotransposed and 70 bp repeat regions, while remaining a suitable length for cloning.

Very faint hybridization can also be seen at about 0.5 kb which is due to the 1 kb probe binding to the 0.5 kb *Sal*I internal fragment (see Figure 36). This band should have been of a similar intensity as the other main bands in each lane, and the weakness of the signal was probably due to the partial *Sal*I digestion.

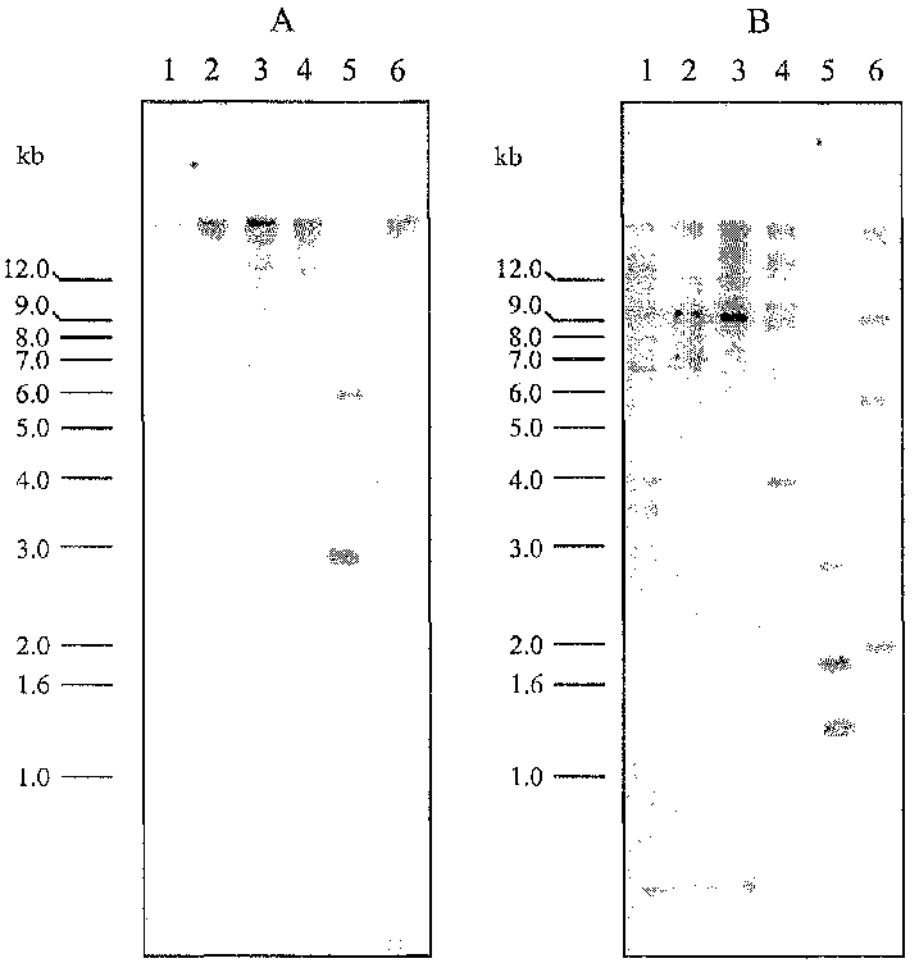


Figure 38. Location of the *Nsi*I sites upstream of the *VSG* gene in the ILTat 1.21 and 1.25 BCs. ILTat 1.2 genomic DNA was double digested with *Sal*I/ *Sac*I (lane 1), *Sal*I/ *Sac*II (lane 2), *Sal*I/ *Not*I (8 bp recognition sequence, lane 3), *Sal*I/ *Bgl*II (lane 4), *Sal*I/ *Nsi*I (lane 5) and *Sal*I/ *Kpn*I (lane 6), size fractionated on a 0.7% agarose gel and transferred to a nylon membrane by Southern blotting. The series of digests was loaded twice on the same gel allowing two identical filters to be produced, which are shown in panels A and B. Panel A is probed with the 400 base ILTat 1.21 cDNA fragment, while panel B is probed with the 1 kb ILTat 1.25 cDNA fragment (both probes start at the spliced leader). The filters were washed to 0.1 x SSC, at 65°C. Molecular markers are indicated to the left of each panel.

5.3 Construction of the minichromosomal libraries

Following the identification of restriction enzymes suitable for the library manufacture, it was necessary to prepare purified minichromosomal DNA. PFGE was used to separate the chromosome sized DNA from ILTat 1.25, 1.67, 1.69 and 1.64 genomic plugs. The separation was performed a 1.2% low-melting point agarose gel at 15°C in 0.089 M Tris-borate, 0.1 mM EDTA (85 V, 1400-600s pulse time, 120 h) using all fourteen of the gel wells (one-half of a genomic plug per lane); the PFGE conditions were modified from the "general" run described in section 4.8 (page 95). The gel was run for 120 h rather than 144 h to ensure that the minichromosomes, which migrate close to the edge of the gel under the "general" conditions, remained within the gel when it was composed of low-melting point agarose. This shorter run time resulted in a more compressed fractionation, but the minichromosomes were still well separated from the other DNA (Figure 39).

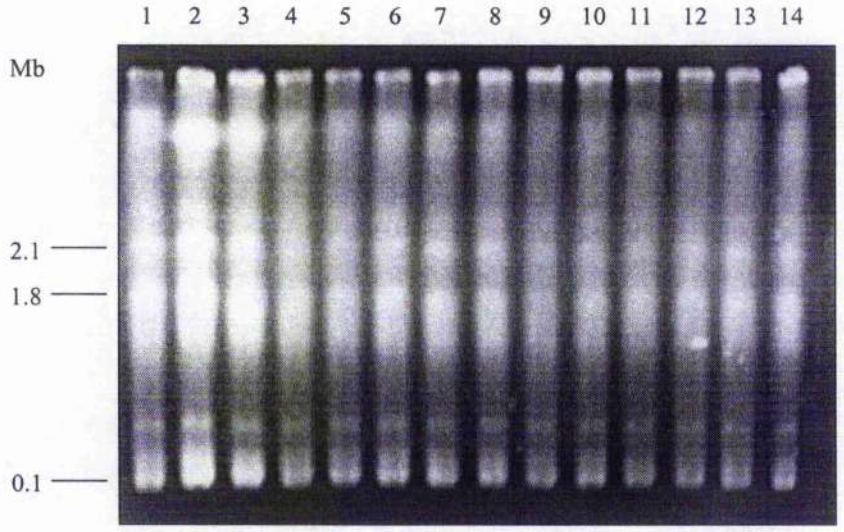


Figure 39. 120 h PFGE separation of the minichromosomes on a 1.2% LMP agarose gel stained with EtBr. Chromosome-size DNA from ILTat 1.25 (lanes 1 and 2), 1.69 (lanes 3-6), 1.64 (lanes 7 to 10) and 1.67 (lanes 11-14) was separated at 85 V, with a ramped pulse frequency of 1400-700 s for 120 h. Approximate chromosomal size is indicated to the left of the panel.

The minichromosomes were excised from the gel and dialysed overnight in 1 x T.E. at 4°C and then for 3 h at room temperature in 1 x T.E. at pH 6.5. The agarose was then removed with the enzyme agarase following the supplier's protocol, and the resultant DNA was cleaned by phenol/ chloroform extraction and ethanol precipitation.

It has been estimated that a single trypanosome nucleus contains 0.097 pg of DNA (Borst *et al.* 1982), and therefore each genomic plug, which holds 5 x 10^7 trypanosomes, should contain 4.85 μg of DNA. Since the minichromosomes represent about 10 % of the genome (Van der Ploeg *et al.*, 1984b), 0.485 μg of minichromosomal DNA should be produced per genomic plug, resulting in a total of 3.395 μg from the fourteen one-half plugs. However, the experimental yield, at 0.9 μg DNA (determined by UV spectrophotometry), was significantly lower than this theoretical maximum. This diminished yield meant it was not possible to construct a partial or size-selected library, and the DNA was therefore divided into two aliquots and then digested to completion with either *Sal*I/ *Nsi*I or *Sal*I/ *Bgl*II. The DNA was then ligated directly into pBluescript KS (+/-), which had previously been digested with *Sal*I/ *Pst*I or *Sal*I/ *Bam*HI (*Pst*I and *Bam*HI produce the same 4 base restriction overhang as *Nsi*I and *Bgl*II respectively). The ligated products were then used to transform supercompetent XL1-BLUE MRF' *E. coli* cells. The colonies were dissolved in L-broth + ampicillin to produce a liquid library (details in Materials and Methods, section 2.9.2, page 48).

5.4 Screening the minichromosomal libraries

A PCR approach was adopted for screening the libraries using a method adapted from a protocol originally used to amplify unknown regions of DNA upstream from a characterized gene (Luo and Cella, 1994). In both libraries the ILTat 1.21 and 1.25 fragments were orientated with the *VSG* sequence pointing towards the T3 promoter of the plasmid (inserted at the *Sal*I restriction site) and the upstream flank facing the T7 promoter (inserted at the *Pst*I or *Bgl*II restriction sites). It was therefore possible

to screen for the relevant gene fragment using the T7 promoter and an opposing specific primer internal to the *VSG* sequence.

Internal ILTat 1.21 *VSG* specific primers were designed to be used in conjunction with the T7 primer to amplify the 1.21 upstream fragment from the *SalI*/ *NsiI* prepared library. These *VSG* primers were 20 nucleotides in length, and reverse and complementary to the ILTat 1.21 cDNA sequence between 180-200 bp (E200), or 870-890 bp (E890) (see Materials and Methods, section 2.9.3, page 49). 1µl of the liquid library (titred at 4.85 x 10^3 colony forming units.μl^{-1}) was used as the template in the PCR reaction.

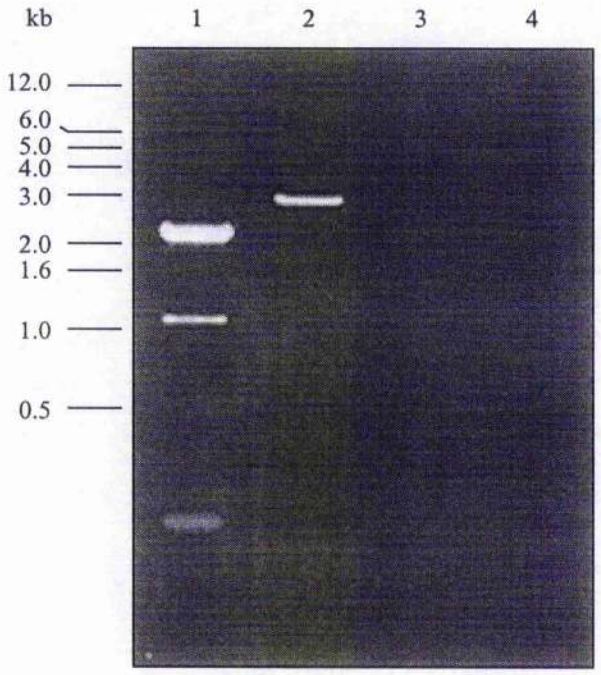


Figure 40. PCR detection of the ILTat 1.21 upstream region from the *SalI*/ *NsiI* minichromosomal library. The fragments were amplified from the *SalI*/ *NsiI* liquid library stock using the plasmid T7 promoter primer and an opposing *VSG* specific primer, either E200 or E890 (reverse and complementary to the ILTat 1.21 cDNA sequence between 180-200 bp or 870-890 bp respectively). The amplification was performed with a hot start for 5 min at 96°C, 30 cycles of 1 min at 96°C, 1 min at 55°C and 3 min at 70°C, and a final extension at 70°C for 10 min. The PCR products were then run on a 0.7% EtBr-stained agarose gel. The primers used in each reaction were: lane 1, T7+E200; lane 2, T7+E890; lane 3, E200 only; lane 4, E890 only. Molecular markers are indicated to the left of the panel.

In section 5.2 (page 108) it was shown that the *NsiI* site existed 1800 bp upstream of the ILTat 1.21 *VSG* gene. Therefore, the expected PCR product size using the E200 and E890 primers was 2080 bp and 2770 bp respectively (80 bp of this product was plasmid sequence from the T7 promoter to the *PstI* restriction site). Products approximating these sizes were seen when the PCR reaction was run on a 0.7% EtBr-stained agarose gel (Figure 40, lane 1 and 2 respectively); two other minor products were also produced when the E200 primer was used. Additionally, single primer controls were performed using only E200 or E890 (without the T7 primer) and these proved negative (Figure 40, lanes 3 and 4); a negative result was also produced from a T7 single primer control (data not shown).

The ILTat 1.25 upstream fragment (with an *NsiI* site 800 bp upstream of the *VSG* gene) was also amplified from the *SalI/ NsiI* library in the same way, using ILTat 1.25 specific primers. When used in conjunction with the T7 primer, the two *VSG* specific oligonucleotides, A180 and A550 (reverse and complementary to the ILTat 1.25 cDNA sequence between 160-180 bp and 530-550 bp respectively; see section 2.9.3, page 49), produced PCR fragments of the expected sizes at 1025 bp and 1415 bp (65 bp of this was vector sequence) (Figure 41, lanes 1 and 2). Again, no products were generated in the single primer control reactions.

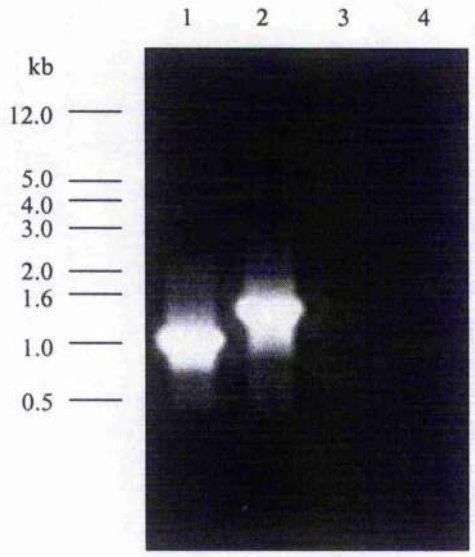


Figure 41. PCR detection of the ILTat 1.25 upstream region from the *SalI*/ *Nsi*I minichromosomal library. The fragments were amplified from the *SalI*/ *Nsi*I liquid library stock using the plasmid T7 promoter primer and an opposing *VSG* specific primer, either A180 or A550 (reverse and complementary to the ILTat 1.25 cDNA sequence between 160-180 bp or 530-550 bp respectively). The amplification was performed with a hot start for 5 min at 96°C, 30 cycles of 1 min at 96°C, 1 min at 55°C and 3 min at 70°C, and a final extension at 70°C for 10 min. The PCR products were then run on a 0.7% EtBr-stained agarose gel. The primers used in each reaction were: lane 1, T7+A180; lane 2, T7+A550; lane 3, A180 only; lane 4, A550 only. Molecular markers are indicated to the left of the panel.

Although the products appeared to be the correct size it was still possible that the bands were PCR artefacts. A half-nested PCR reaction was utilised to examine the fragments in more detail. The T7+E200 and T7+A180 PCR reactions were repeated but using the 2770 bp ILTat 1.21 and 1415 bp 1.25 PCR products as the templates instead of the library cells. This half-nested reaction produced the expected band sizes at 2080 bp (for 1.21) and 1025 bp (for 1.25) (Figure 42), suggesting that these fragments represented the correct *VSG* associated sequence.

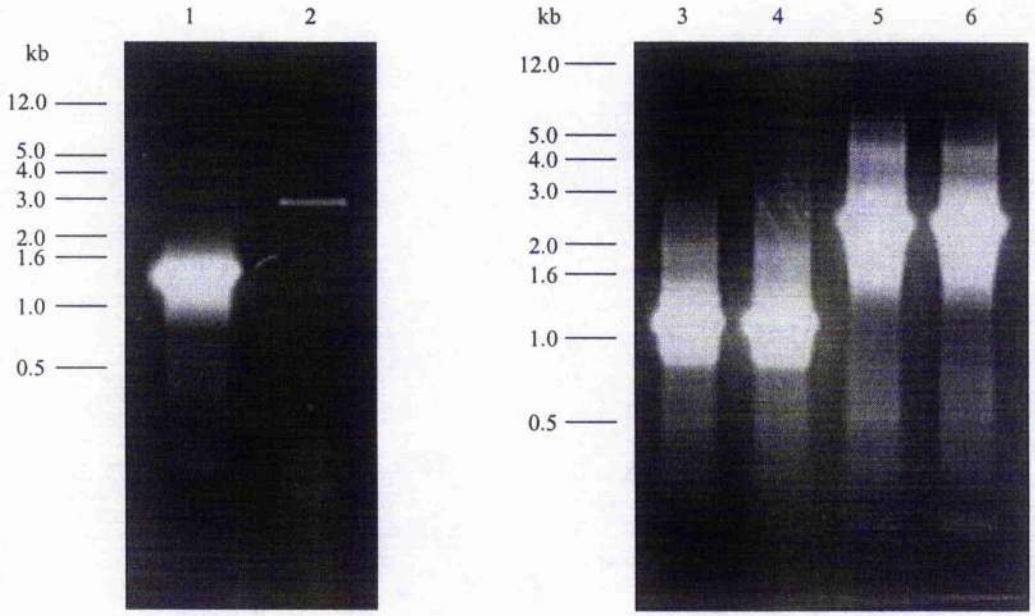


Figure 42. Half-nested PCR reactions of the ILTat 1.25 and 1.21 minichromosomal PCR products. Panel A. The ILTat 1.25 1415 bp (lane 1) and ILTat 1.21 2770 bp (lane 2) products were amplified from the liquid library stock using the T7 promoter primer and the *VSG* primers A550 and E890 respectively. The amplification was performed with a hot start for 5 min at 96°C, 30 cycles of 1 min at 96°C, 1 min at 55°C and 3 min at 70°C, and a final extension at 70°C for 10 min. The PCR products were then run on a 0.7% EtBr-stained agarose gel. Panel B. These two products were then used as the template for the half-nested PCR using the same T7 primer and the nested *VSG* primer (A180 for ILTat 1.25 and E200 for ILTat 1.21); the PCR conditions were identical to the first round of amplification. The products were then run on a 0.7% EtBr-stained agarose gel: lanes 3 and 4, ILTat 1.25 product; lanes 5 and 6, ILTat 1.21 product. Molecular markers are indicated to the left of each panel.

The *SalI*/ *Nsi*I library was also screened for ILTat 1.68 and 1.69 sequences using the same PCR approach. The *SalI*/ *Nsi*I fragments of these *VSG* genes (and their upstream flanks) had not been examined previously by restriction mapping, and it was therefore impossible to determine the orientation in which the fragments would insert into the vector. As a result of this, the PCR was performed using both the T7 and T3 pBluescript promoter primers; again two internal primers (20 nucleotides in length) were designed for each *VSG* (C190 and C390 for ILTat 1.68; D210 and D370 for ILTat 1.69; see section 2.9.3, page 49). PCR products were generated when the

T3 primer was used, suggesting that the *Nsi*I site occurred within (or downstream of) the *VSG* gene while the *Sal*I site existed in the upstream sequence. This result was consistent with the ILTat 1.69 full-length cDNA sequence (sequenced after this experiment), which possessed a single *Nsi*I restriction site (and no *Sal*I site) within the coding sequence; the 1.68 cDNA did not contain either of these two restriction sites. The T3+D210 (ILTat 1.69) reaction produced a band at approximately 1000 bp, while a product of about 1200 bp was generated from the T3+D370 amplification (Figure 43).

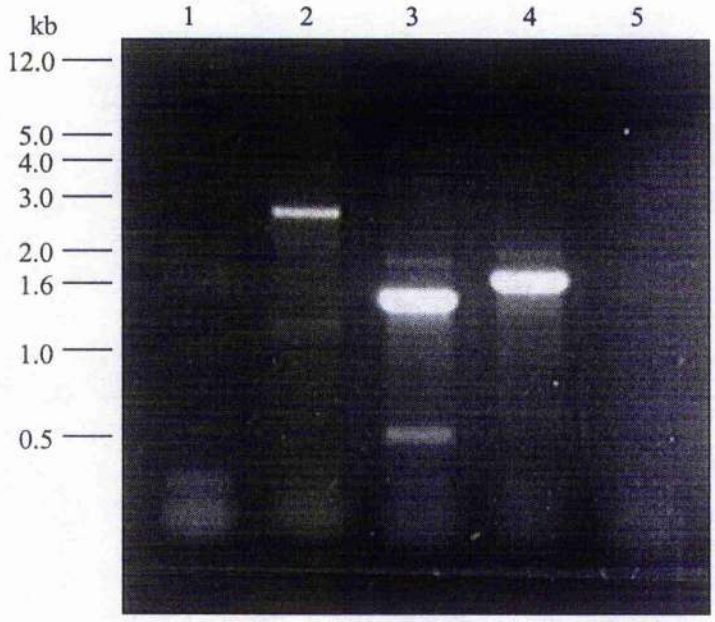


Figure 43. PCR amplification of the potential ILTat 1.68 and 1.69 minichromosomal products from the *Sal*I/ *Nsi*I library. The fragments were amplified from the *Sal*I/ *Nsi*I liquid library stock using the plasmid T3 promoter primer and an opposing *VSG* specific primer, either C190 or C390 (ILTat 1.68, lanes 1 and 2 respectively), or D210 or D370 (ILTat 1.69, lanes 3 and 4 respectively). The amplification was performed with a hot start for 5 min at 96°C, 30 cycles of 1 min at 96°C, 1 min at 55°C and 3 min at 70°C, and a final extension at 70°C for 10 min. The PCR products were then run on a 0.7% EtBr-stained agarose gel. A T3 single primer control was run in lane 5. Molecular markers are indicated to the left of the panel.

The 160 bp size difference between these two products was as expected, indicating that these bands probably represented the correct *VSG* upstream fragment. One of the ILTat 1.68 amplifications (T3+C390) produced a band at 1900 bp, but the T3+C190 reaction yielded no product (Figure 43). It is possible that the C190 primer was incompatible with the T3 primer resulting in the reaction failing. However, it is more likely that the 1900 bp band was an artefact, and the ILTat 1.68 upstream fragment was not represented in this library. Even if the ILTat 1.68 and 1.69 products represented the correct sequence they were probably too short to be used for the genomic analysis, with the hypothetical *SalI* restriction sites lying approximately 800 bp (ILTat 1.69) and 1500 bp (ILTat 1.68) upstream of the *VSG* gene.

The average insert size of the *SalI*/ *Nsi*I minichromosomal library was examined in 10 random clones. These clones were taken from a library plate and grown as an overnight culture in L-broth + ampicillin, after which the plasmids were retrieved using the QIAGEN Mini kit. The insert size was then determined by PCR amplification using the T3 and T7 promoter primers (Figure 44). This revealed that the average insert size was about 1.0 to 1.5 kb. It is clear that the ILTat 1.21 *SalI*/ *Nsi*I fragment, at 2.8 kb, must be one of the largest inserts in this library.

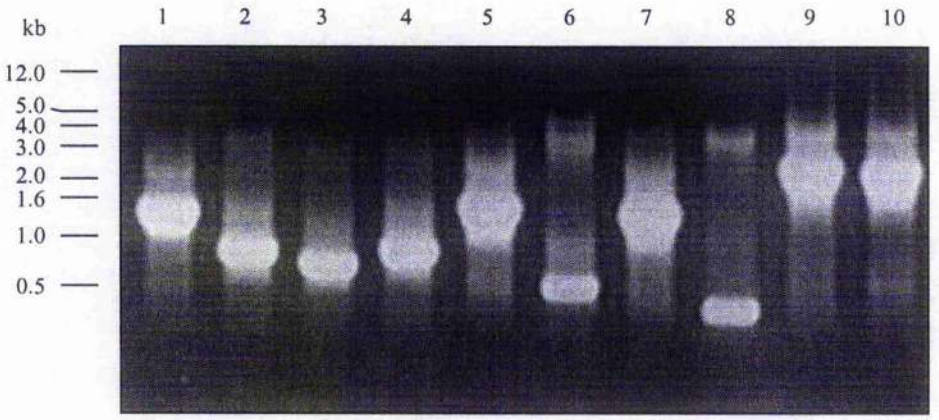


Figure 44. Determination of the average insert size of the *SalI*/ *Nsi*I minichromosomal library. Ten random clones were taken from a library plate and grown overnight in L-broth + ampicillin. The plasmids were then retrieved using the QIAgen miniprep kit. The insert size was examined by PCR amplification of the fragment using the T3 and T7 promoter primers. The amplification was performed using Long-*Taq* (Stratagene) with a hot start for 5 min at 96°C, 30 cycles of 1 min at 96°C, 1 min at 55°C and 10 min at 70°C, and a final extension at 70°C for 10 min. The PCR products were then run on a 0.7% EtBr-stained agarose gel. Molecular markers are indicated to the left of the panel.

PCR amplification was also utilised to screen the *SalI*/ *Bgl*II digested minichromosomal library. However, the A180 and A550 primers failed to produce fragments of the expected size, suggesting that the 4.0 kb ILTat 1.25 fragment was not represented in the library. It is possible that smaller restriction fragments were preferentially cloned during the library manufacture, resulting in the exclusion of this comparatively large product. Alternatively, the fragment could be inherently unclonable due to some part of the sequence; for example a long run of 70 bp or 177 bp repeats. A *SalI*/ *Bgl*II digested 3.5 to 4.5 kb size selected pBluescript library was also prepared from ILTat 1.2 genomic DNA, but screening (using conventional hybridization techniques) failed to identify the ILTat 1.25 4.0 kb fragment. Inverse PCR was also attempted on this size selected DNA, as well as with the *SalI*/ *Bgl*II cut minichromosomal DNA, but again no ILTat 1.25 specific products were found.

5.5 Characterization of the ILTat 1.21 and 1.25 upstream flanks

Pfu polymerase was used to amplify the 1025 bp ILTat 1.25 product ensuring maximum copying fidelity. However, a commercial mix of *Taq* and *Pfu* polymerases (Long-*Taq*, Stratagene) was required to generate the 2080 bp ILTat 1.21 product, which would not amplify with the *Pfu* polymerase alone. These PCR products were then ligated into the PCRScript Amp SK(+) plasmid, and three clones were isolated from each product and sequenced.

Physical maps of the ILTat 1.21 and 1.25 minichromosomal clones are presented in Figure 45. BLAST analysis of these sequences revealed 98-100% identity with the ILTat 1.21 and 1.25 *VSG* cDNA database sequences over the expected region (200 bp or 180 bp respectively). The *VSG* homology began immediately after the spliced leader sequence (ending GTTTCTGTACTATATTG) in the cDNA sequences and ended exactly at the other end of the clones, confirming that the correct fragments had been isolated from the minichromosomal library. The BLAST search did not reveal any further trypanosome sequence homology in the 800 bp upstream of the ILTat 1.25 *VSG* sequence. However, additional identities were discovered in the ILTat 1.21 upstream region.

Homology with 70 bp repeat sequence was identified over the first 250 bp at the upstream end of the ILTat 1.21 clone, producing multiple hits against the NCBI

database. Figure 46 displays the first 300 bp of sequence upstream from the *Nsi*I site, and demonstrates that three and a third of the repeats were encompassed within the clone. An alignment of each of these repeats against the 70 bp consensus sequence (Aline *et al.* 1985); (Shah *et al.* 1987) is also included in the Figure. It is apparent that the *Nsi*I site occurred within the 70 bp repeats, which is quite unusual, since these repeats are renowned for the absence of restriction sites, and consequently the 70 bp array is often referred to as a "barren" region. The only other 6 bp restriction enzyme recognition sequence that occurred within the cloned repeats was *Ssp*I, which appeared to be highly conserved in the 70 bp sequence.

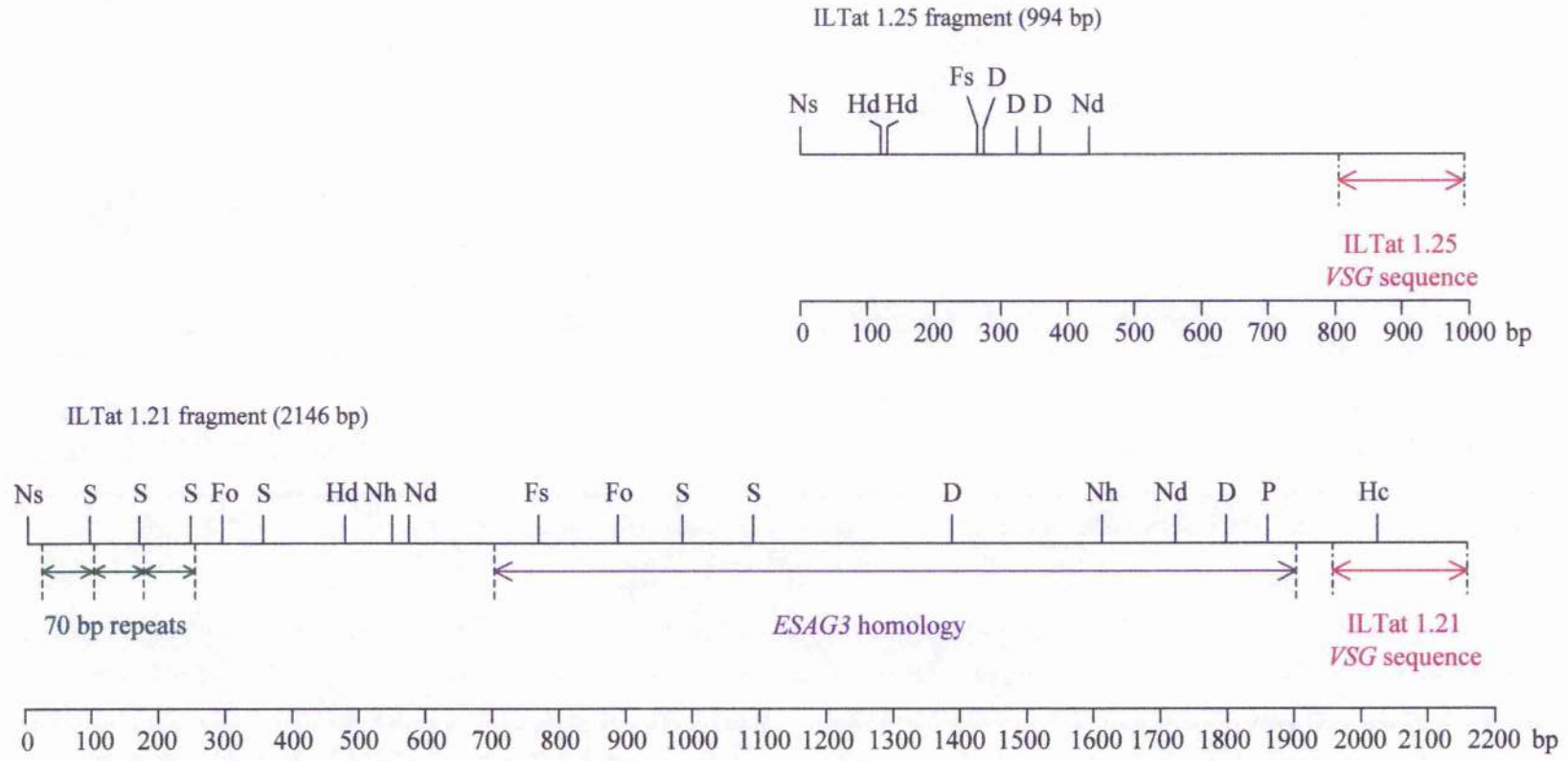


Figure 45. Physical maps of the ILTat 1.25 and 1.21 minichromosomal clones encompassing the upstream flanks, limited by the *Nsi*I restriction sites and approximately 200 bp of the *VSG* genes. Regions of homology with previously characterized trypanosome sequences were identified by BLAST and CLUSTAL W analysis. Restriction sites were determined from the full-length sequence of the clones. Abbreviations: D, *Dra*I; Fs, *Fsp*I; Fo, *Fok*I; Hc, *Hinc*II; Hd, *Hin*dIII; Nd, *Nde*I; Nh, *Nhe*I; Ns, *Nsi*I; P, *Pst*I, S, *Ssp*I.

The BLAST search also revealed that the ILTat 1.21 sequence displayed an 88% identity with *ESAG3* between 811 and 1164 bp. Subsequent CLUSTAL W analysis between the ILTat 1.21 cotransposed and *ESAG3* (taken from Alexandre *et al.*, 1988) demonstrated that the region between 707 and 1918 bp was homologous to the entire *ESAG3* gene (Figure 47), displaying 79% and 49% identity at 707-1237 bp and 1238-1770 bp respectively. Although this is an *ESAG3* pseudogene, two ORFs were identified within this *ESAG3* homologous sequence by the Vector-NTI analysis program, and occurred at 707-1219 and 1263-1818 bp (Figure 48). Both ORFs are in the same reading frame as (and correspond to segments of) the published *ESAG3* ORF.

AnTat 1.1 companion sequence (Pays *et al.* 1983d) homology was also discovered by the BLAST search, revealing a 91% identity between 1365 and 1398 bp. A CLUSTAL W analysis revealed that the ILTat 1.21 clone was 64% homologous to the first 439 bp of the companion sequence between 1221 and 1660 bp (the position relative to the *ESAG3* homology is displayed in Figure 48). This part of the companion sequence had been shown previously to be 80% homologous to part of the *ESAG3* gene (Alexandre *et al.* 1988).

A.

```
      *1                                  *2                                                                  *3
  1   CTGCATATAC GAATATAATA ATAAGAGCAG TAATAGTAAT GATAATAATG ATAATAATAG AAGAGTGTTG TGAGTGTGTG TATACGAATA TTATAATAAT
                                                                                *4
101   AGTAATGATA ATAATAATAG TAATGAAAAT AATAGAAGAG TGTTGTGAGT GTGTGTATAC GAATATTATA ATAATAGTAA TGATAATAAT AATAGTAATG

201   AAAATAATAn AAGAGTGTTG TGAGTGTGTG TGTATACnAA TATTATAATA AGAGCAGCGG TGAAAACTAT ATGGATGCGA GGACATGACG CGCTTTATCC
```

B.

70 bp alignment with consensus sequence:

```
(*1)                                                                ATATACGAATATAATAATAAGAG   27 bp (1/3 repeat)

(*2) CAGTAATAGTAATGATAATAATGATAA---TAATAGAAGAGTGTTGTGAGTGTGTGT---ATACGAATATTATAATAA---   72 bp

(*3) TAGTAATGATAATAATAATAGTAATGAAAATAATAGAAGAGTGTTGTGAGTGTGTGT---ATACGAATATTATAATAA---   75 bp

(*4) TAGTAATGATAATAATAATAGTAATGAAAATAATAnAAGAGTGTTGTGAGTGTGTGTG-TATACnAATATTATAATAAGAG   80 bp

(*c) CAGTRRTRRTRRTRRTRR---------------TAGGAGAGTGTTGTGAGTGTGT---ATATACGAATATTATAATAAGAG   63 bp

      (CONSENSUS SEQUENCE)
```

Figure 46. 70 bp repeat sequence detected in the ILTat 1.21 genomic clone. Panel A. The first 300 bp from the upstream end of the genomic clone, encompassing three and a third repeats. The repetitive sequence begins at base 5 (bases 2-6 are part of the *Nsi*I recognition sequence) and is denoted by *1. The red vertical bars indicate the start of three complete repeats, which have been allocated the symbols *2, *3 and *4 respectively. The blue vertical bar at 254 bp represents the end of the repetitive sequence. Underlined regions indicate *Ssp*I restriction sites. Panel B. Alignment of the three and a third 70 bp repeats against the consensus sequence (derived from Aline *et al.* (1985). R=Purine nucleotide. Gaps have been introduced to maximize the alignment. The lengths of the repeats are indicated after the sequence. Underlined nucleotides represent divergence from the consensus. The TRR repeats in the consensus sequence display a great variance in number between repeats (Aline *et al.* 1985).

Figure 47. CLUSTAL W analysis of the 1239 bp region homologous with *ESAG3* in the ILTat 1.21 minichromosomal clone. The *ESAG3* sequence was taken from Alexandre *et al.* 1988 and represents the entire ORF. The first 530 bp of the ILTat 1.21 sequence displayed 79% homology with *ESAG3*, while the next 680 bp exhibited 49% identity. Gaps have been inserted by the CLUSTAL W programme to maximize the alignment.


ILTat TAAGAAATGAAATATAGTTTAT-GTGGACTTTATTTG----GCTTTTCTGACCTAC-GTGTGTTTGTGCGCAATGTCTCTTGCTGGTGATCTAACAAATAACATAACAAATATGGCAAAC 114
ESAG3 ------ATGAAATTTGGTATCTCGAGAACTTTAATCTTACTGCTGCTGTCACATTTAGTGTGTGTGAGAGCTGTATTCCTTCAAGATGAAGAAGTAAATAACATAACAAATATGACACAC 114
ILTat AATAATACTAAGAAACCCAGAGTCTTTGTGGTTGTTGTTCAAACTCATGCAACACAAGGATGGTGTAGGCTCATATTGAGTGCGACGCTAAGCAATGTTGAGATTATTACAATTGCGATG 234
ESAG3 AATGGTACTAAGAAACCCAGAGTCTTTGTGGTTGTCGTTCAAACTCATGCAACACAAGGATGGTGTAGGCTCATATTGAGTGCGAAACTAAGCAATGTTGAAATTGTCACAATTGCGATG 234
ILTat GGAGGTTTGTATAGCCATATTAAAAGGCCTAAATGGGTTATGGAATATTTAGAAAAACAGAATGCAAGTGATGAGGATATTGTAATTCATTCAGACGGATCTGATGTTATTGTTAGTGAG 354
ESAG3 GGAGGTTTGTACAGTCATGTTAAGAGGCCTAGATGGGTTATGGAGTTTTTAGAGAAACAGAATGCAAGTGATGAGGATATTGTAATTCATTTGGACGGATCGGATATTATTGTTTCTGAC 354
ILTat CAACAGAAGTATGAAAATGCTGTGGAATATTTTATACAAAGTACTGCCGAAAATGAAAATAAGTTTAGTGAAGATGAAATAGTGAAAGAAAATCAAATATCTCCGATTCAGTTTATAGCA 474
ESAG3 GGAGAAAAATATGAAAATGCCGTCGAATATTTTATGCAGAATACTGCTGAAAATGAAGAGAAGTTTAGTGGAGAAGATATAGTAAAAGAGAAGCATATTTCCCCTATTATGTTCATGACA 474
ILTat ACTGCTGATTGTACTGCCCCTCCAGTTGGACCTGATAGTTTATTCTGATCACAA--AACACACCCAGCACGCTGCCAATGGTTTTTTGAAATGGCAAGTAAGACAGAAGGAGAGAAAGGA 592
ESAG3 ACATCAGAGTGTTATGCTCCTC-AGTTGGATTTGCTTATGAACTCAGATCACAAGGAGCATGTGCAG--AGATGTTACTGGTTTTTCAATGTCGGCTACAGAAAAGAAGAGGATAAAAAA 591
ILTat CTACCGCTTCTACTTAAATTTCCCGACACTGCCAAACCATACTTCGGCGGCGGTAATTTAATAGCAAGAGTGTGGGCATTTAAGAGGTTTAAAGCCGCATTCGATGGCGTCCTTGCACTG 712
ESAG3 TTACCCCATTTACTGAAATCACCTCCTAGTGGAAAATCTTATCTTAGTGGTGGT---TTGATAGCCAGAGTCTGCGCATTTACAAAATTTGAATATGCATTTGGAGAGTTACT--AAAGG 705
ILTat GC--TGATGAATGGTGGTCAGAACAGAGTATTTATATTCCTTTGTTGTTGTGGAGTGTGA[AACAAGAAGACAAAGTGAGCCCTAAATTCGTTTTGCCGAGAGGTCTGATAGGCTTAGATT 830
ESAG3 GAAGTGGGGAGTGGTGGTCAGAACAAAGCATTCACAAACCGTTGCTAATATGGAGTGCCA]TTCAGGAAGAAGCTGTTGAACAAAGATTTGTTTTGAAGAGAGGAATAATTGGATTGAATT 826
ILTat ACGGAGAGCGGTTTTTCTC-AGTCCCT----TACGAAAATTCAATGGAACGTGCTGCGTTCCTTCACTTCCAGGACACGCGCTAGCATGGCAGGAGAAGGCTAGAAAAATTATATCAAA 945
ESAG3 ATGAAGAGAGATTTTTTACTATTCCCAATAGTGGGGTGATT--AGAGAA---GCTCCATTCATTCATTTCCTGGGCAGCCAATTGCATGGGAAGAGAAAGCGAAATTAATTGTGAAAAA 941
ILTat CCTTACTTGGTATAAAAAGTGGCAAGCAAAGAAATTTAATACGAATCACACTCTTGGGAA[GCTGCAAGCATATGCACTATCGCAAGAAAAATTGGATTTTGTGTATGAACCAGTTTGTGG 1065
ESAG3 TCTTAGTTGGTATAAAGAGTTGAAAGGAAGAGAATTTAATTCGAAGGATA---TGGGCAAAGTAGAAACTTATGTAGTAGGTGGAGAGAATTGGGATTTTCGATATGAAAGAATTTGTGG 1058
ILTat TGACGCCGTAGACACTGGGGATTTGTTTAAAGCTAAAAAACAGTAGCAAACCAGAAATAG]CTACATCTATGAGATTATGCCGCTGCAGCGCAGCAGGACACCGCAGCAGGCTTCAACGAC 1185
ESAG3 GGACGCG----------TTCAAGGAG 1074
ILTat ATGGACATTATTAAACGCACAAATTTTAAGCTGACTCACATCGAAGCCATCTGCACCGAT 1245
ESAG3 AATGAC-TTGTTTAAGGCAAA-----GAAACTG---------- 1100

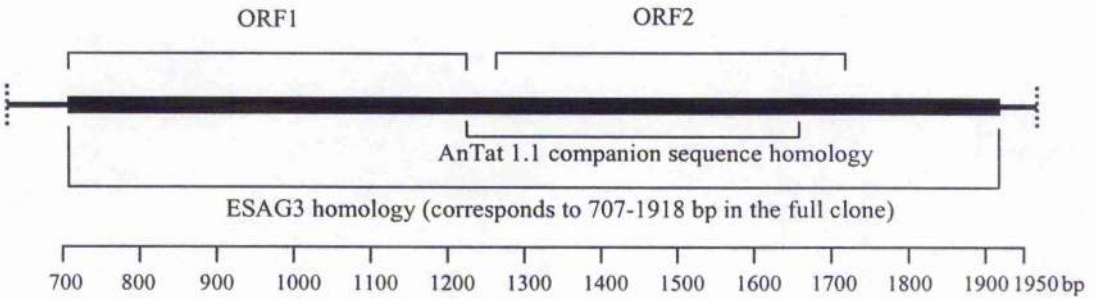


Figure 48. Position of the two ORFs within the *ESAG3* homologous ILTat 1.21 sequence. The ORFs were determined using the Vector-NTI analysis program. ORF1 and ORF2 extend from 707-1219 and 1263-1808 bp respectively, and are in the same frame as the published *ESAG3* sequence. The 439 bp region displaying 64% homology to the AnTat 1.1 companion sequence is also indicated. The homology begins at the start codon of the *ESAG3* ORF and continues for the majority of this ORF, with the ILTat 1.21 sequence appearing to be slightly longer than the *ESAG3* ORF (see Figure 49).

The presence of the *ESAG3* homologous sequence downstream of the 70 bp repeats suggests that minichromosomal telomeric sequences result from recombinational events with other telomeric regions within the genome. It also seems likely that the cotransposed sequence, which is largely comprised of the *ESAG3* sequence, is not utilised as a recognition sequence with the corresponding region of a *BES* for the ILTat 1.21 duplication event.

5.6 Identification of the ILTat 1.21 upstream duplication boundary

A Southern analysis was employed to determine whether the upstream boundary of duplication was encompassed within the ILTat 1.21 *Nsi*I/ *Sal*I BC fragment. ILTat 1.2 and 1.21 genomic DNAs were digested with *Nsi*I and *Hin*dIII, and double digested with *Sal*I/ *Nsi*I and *Sal*I/ *Hin*dIII, separated on a 0.7% agarose gel and transferred to nylon membrane by Southern blotting. The filter was then probed with the α-^{32}P radiolabelled ILTat 1.21 400 bp fragment, homologous to the amino-terminus encoding region of the *VSG* gene. The hybridization revealed an ELC,

distinct from the BC, in the expressor lanes of the *Nsi*I and *Hin*dIII single digests (Figure 49, panel A, lanes 2 and 4). However, in the *Sal*I/ *Nsi*I and *Sal*I/ *Hin*dIII double digests the ELC and BC fragments were equal in size (Figure 49, panel A, lanes 5-8), indicating that the duplication boundary must lie upstream of the *Nsi*I site. The ELC and BC fragments observed in the single digests were distinct, due to differences in telomere tract length between the two chromosomal loci.

Two fainter products were also detected in the *Sal*I/ *Nsi*I and *Sal*I/ *Hin*dIII double digests (Figure 49, panel A, lanes 5-8), probably due to the incomplete digestion of the two *Sal*I sites (within the *VSG*). Subsequent reprobing of the filter with the PCR product encompassing the *RAD51* single copy gene revealed that all the digests had proceeded to completion (Figure 49, panel B; an *Nsi*I site is present at 705 bp in the 995 bp *RAD51* fragment resulting in two hybridization products in these digests, lanes 1-2 and 5-6). This suggested that the partial digestions were due to base J modification of the *Sal*I sites. The *RAD51* probing also demonstrated that slightly more of the ILTat 1.2 DNA had been loaded per lane when compared with the 1.21 DNA. This confirmed that the increased hybridization intensity seen in the expressor lanes of the *Sal*I/ *Nsi*I and *Sal*I/ *Hin*dIII double digests during the ILTat 1.21 probing was due to the combination of both the BC and ELC signals (Figure 49, panel A, lanes 6 and 8).

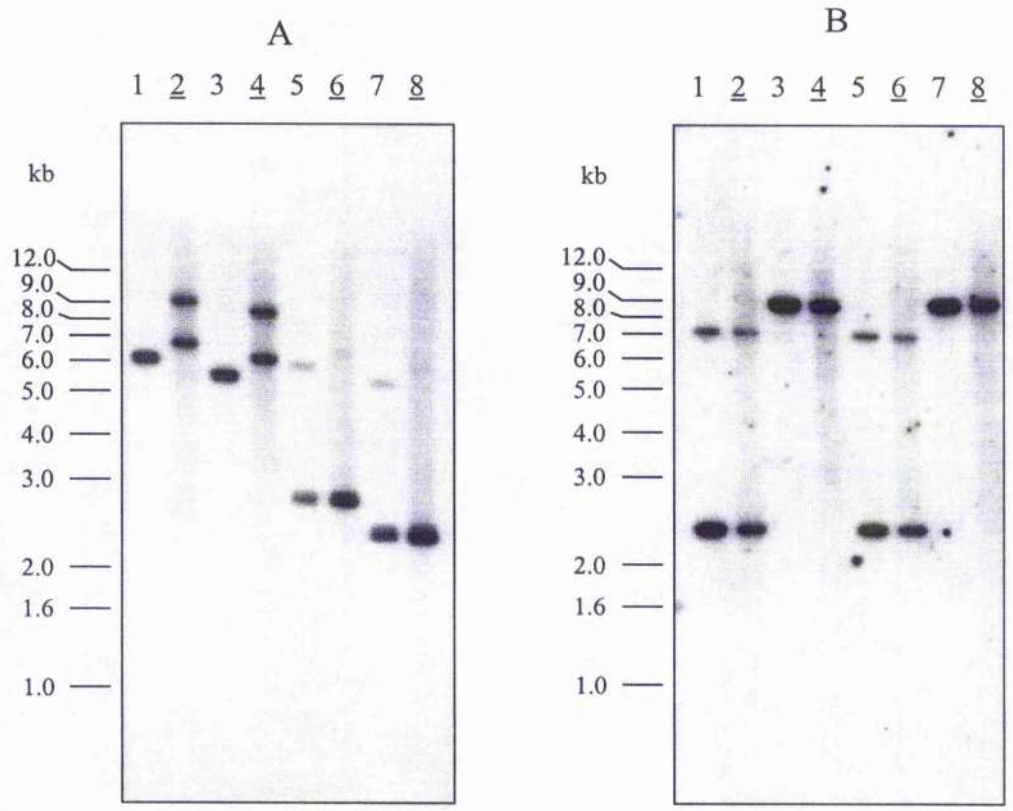


Figure 49. Southern analysis of the ILTat 1.21 minichromosomal clone. Panel A. ILTat 1.2 or 1.21 genomic DNA was digested with *Nsi*I or *Hin*dIII, and double digested with *Sal*I/ *Nsi*I or *Sal*I/ *Hin*dIII, size fractionated on a 0.7% agarose gel, and transferred to a nylon membrane by Southern blotting. The filter was then probed with the α-^{32}P radiolabelled ILTat 1.21 400 bp fragment, homologous to the amino-terminus encoding region of the VSG. Panel B displays the same filter stripped and reprobed with a PCR product encompassing the coding sequence of the trypanosome RAD51 single copy gene. The digests were loaded in the following order: *Nsi*I, lanes 1 and 2; *Hin*dIII, lanes 3 and 4; *Sal*I/ *Nsi*I, lanes 5 and 6; *Sal*I/ *Hin*dIII lanes 7 and 8. The underlined lanes indicate the digests that involved the ILTat 1.21 expressor DNA. The filters were washed to 0.1 x SSC at 65°C. Molecular markers are indicated to the left of each panel.

The double digests were repeated using *Age*I (rather than *Sal*I), and subsequent Southern analysis revealed that the minor bands seen in Figure 49A were not present in these new digests (Figure 50). This confirmed that the partial fragments must have resulted from the base J modification of the *Sal*I sites. The signal strength in the expressor lanes of Figure 50 should have been more intense than the non-expressor lanes due to the presence of the additional comigrating ELC fragment. However, examination of the EtBr-stained gel of the blot revealed that approximately twice the amount of ILTat 1.2 DNA was loaded when compared with the ILTat1.21 DNA (data not shown), and this accounts for the discrepancy from the predicted hybridization intensity.

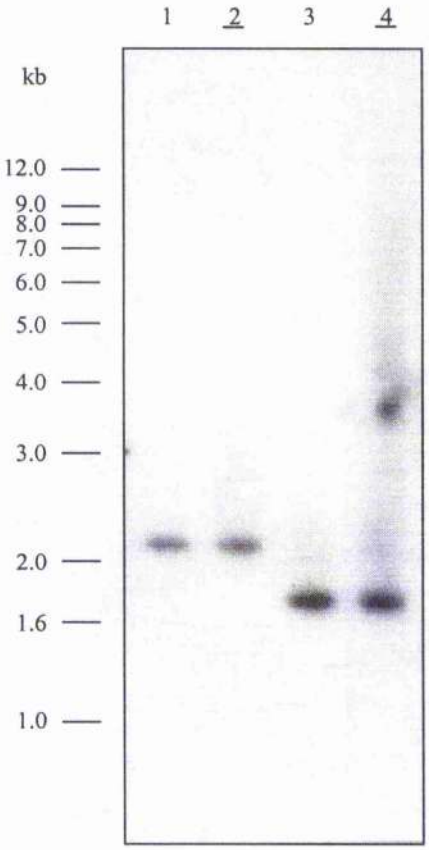


Figure 50. *Age*I/ *Nsi*I and *Age*I/ *Hin*dIII double digests of ILTat 1.2 and 1.21 genomic DNA. Digested DNA was size fractionated on a 0.7% agarose gel, and transferred to a nylon membrane by Southern blotting. The filter was then probed with the α-^{32}P radiolabelled ILTat 1.21 400 bp fragment, homologous to the amino-terminus of the VSG. The digests were loaded in the following order: *Age*I/ *Nsi*I, lanes 1 and 2; *Age*I/ *Hin*dIII, lanes 3 and 4. The underlined lanes indicate the digests that involved the ILTat 1.21 expressor DNA. The filters were washed to 0.1 x SSC at 65°C. Molecular markers are indicated to the left of the panel.

Although the duplication boundary occurred beyond the *Nsi*I site that demarcated the ILTat 1.21 BC clone, the cotransposed and associated 70 bp repeat sequences proved to be essential for the further analysis of the upstream minichromosomal *VSG* flank. An experimental approach was adopted using this cloned sequence to ascertain the length of the 70 bp repeat array. Genomic DNA from both expressor (ILTat 1.21) and non-expressor (ILTat 1.2) clones was double digested with *Hin*dIII and a series of frequently cleaving enzymes (4 bp recognition sequence), size fractionated on a 0.7% agarose gel, and transferred to nylon membrane by Southern blotting. The filter was then probed with an α-^{32}P radiolabelled cotransposed fragment, which was excised from the ILTat 1.21 minichromosomal clone using *Hin*dIII/ *Fok*I, and was homologous to the region 286-472 bp downstream of the *Nsi*I site and the 70 bp repeats (Figure 45). This fragment probably displayed unique homology with the ILTat 1.21 upstream fragment, since it did not contain any of the *ESAG3* homologous sequence. It was hoped that some of these 4 bp restriction sites, which occur in random DNA at approximately every 256 bp, would cut only after the barren region, and therefore could be utilised to liberate the entire 70 bp repeat array.

It was necessary to select restriction enzymes that would not cut within the three and a half 70 bp repeats encompassed by the genomic clone, or within the cotransposed region between the repeats and the *Hin*dIII site. It seemed plausible that the entire 70 bp array could be devoid of some of these restriction sites, since they were not represented within the first few repeats. However, it was important to examine several restriction enzymes because the 70 bp repeats are not completely conserved (Shah *et al.* 1987), and this degeneracy often results in the formation of new restriction sites.

Five frequently cleaving enzymes were selected by the criteria described above, and four of these, *Rsa*I, *Hae*III, *Mbo*I and *Sau*3AI, were used in combination with *Hin*dIII to digest the ILTat 1.2 and 1.21 genomic DNA. A single digest was also prepared using *Alu*I, which possessed a restriction site immediately downstream of the *Hin*dIII site. In addition, *Hin*dIII and *Hin*dIII/ *Dde*I (a 5 bp recognising enzyme) digests were used in the analysis. The results of this study are displayed in Figure 51.

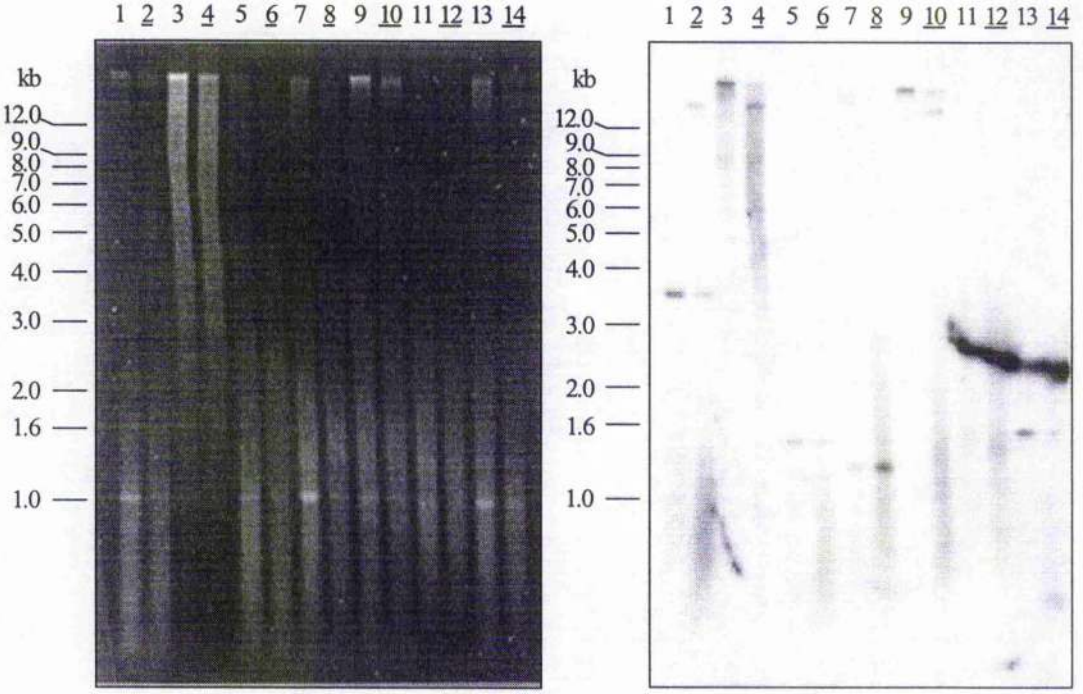


Figure 51. Restriction analysis of the ILTat 1.21 barren region. ILTat 1.2 and 1.21 genomic DNAs were digested with *Alu*I or *Hin*dIII, or double digested with *Hin*dIII/ *Dde*I, *Rsa*I, *Hae*III, *Mbo*I or *Sau*3AI, size fractionated on a 0.7% agarose gel, and transferred to a nylon membrane by Southern blotting. The filter was then probed with an α-^{32}P radiolabelled cotransposed fragment, which was excised from the ILTat 1.21 genomic clone using *Hin*dIII/ *Fok*I , and is homologous to the region 286-472 bp downstream of the *Nsi*I site (before the 70 bp repeats, Figure 45). The digests were loaded in the following order: *Alu*I (lanes 1 and 2), *Hin*dIII (lanes 3 and 4), *Hin*dIII/ *Dde*I (lanes 5 and 6), *Hin*dIII/ *Rsa*I (lanes 7 and 8), *Hin*dIII/ *Hae*III (lanes 9 and 10), *Hin*dIII/ *Mbo*I (lanes 11 and 12) and *Hin*dIII/ *Sau*3AI (lanes 13 and 14). The underlined lanes represent the digests that involved the ILTat 1.21 expressor DNA. Panel A. EtBr stained gel of the digests. Panel B autoradiograph of the subsequent hybridization. The filter was washed to 0.1 x SSC at 65C. Molecular markers are indicated to the left of each panel.

A clear autoradiograph could not be produced from this restriction analysis (Figure 51) despite preparing another two repeat filters. It appeared that degradation was occurring in the DNA stocks, especially in the ILTat 1.21 DNA, and this problem was compounded by an extremely weak hybridization signal, possibly resulting from the short length of the *Fok*I/ *Hin*dIII probe (only 182 bases long). Although no significant conclusions can be drawn from this preliminary result due to the poor quality of the hybridization, it was possible to construct a hypothetical interpretation from these data.

In the *Alu*I digest (lanes 1 and 2) a band representing the BC fragments was observed at 3.5 kb in both the ILTat 1.2 and 1.21 genomic DNA; a large additional band (>12 kb), representing the ELC fragment, was also detected in the expressor lane. This suggested that an *Alu*I site existed 3.5 kb upstream of the cotransposed *Hin*dIII site (and therefore approximately 5.0 kb from the start of the *VSG* gene), and since different sized restriction fragments were produced from the BC and ELC, it appeared to reside beyond the duplication boundary.

The *Hin*dIII digest a yielded a large fragment in the ILTat 1.2 DNA (lane 3), representing the BC, but it was harder to interpret the 1.21 digestion (lane 4). It was expected that this digest would generate two distinct fragments, derived from the BC and ELC. One of the fragments was obvious, but the second (larger) band was very faint and only just distinguishable above the background smear. It is possible that this fragment was partially degraded, resulting in the weaker signal (larger products are more vulnerable to the effects of degradation). Overall, the ILTat 1.21 DNA appeared more degraded than the 1.2 DNA, and this could explain why the large 1.2 *Hin*dIII product appeared unaffected in this way.

An unexpected result was obtained for the *Hin*dIII/ *Hae*III double digest (lanes 9 and 10). The hybridization pattern was identical to the *Hin*dIII single digest (lanes 3 and 4), indicating that no *Hae*III sites exist along the entire chromosomal region upstream of the cotransposed *Hin*dIII site. The EtBr gel demonstrates that the *Hae*III digestion was successful, with the majority of the DNA becoming reduced in size when compared with the *Hin*dIII digest. This result suggested that 177 bp minichromosomal repeats, which do not possess *Hae*III sites (Weiden *et al.* 1991), must lie extremely close to (or immediately upstream of) the 70 bp repeats.

No clear ELC was observed in the ILTat 1.21 DNA of the *Hin*dIII/ *Dde*I, *Hin*dIII/ *Rsa*I digests, but single bands were detected at 1.4 and 1.3 kb respectively in both

1.2 and 1.21 DNA (lanes 5-6 and 7-8). This indicated that both of these restriction sites, which occurred within the 70 bp repeat array, were encompassed by the duplicated copy.

It is possible that the *Alu*I site upstream of the ILTat 1.21 *VSG* gene lies beyond the 70 bp repeats, since it exists outside the duplication boundary. The apparently close proximity of the 177 and 70 bp repeats suggests that this *Alu*I site occurs in the first of the 177 bp repeats, which have been shown to possess a highly conserved *Alu*I site (Sloof *et al.* 1983). If these assumptions are correct, then the 70 bp array stretches from 2-5 kb upstream of the *VSG* gene, and the upstream duplication boundary lies within the last 2 kb of these repeats (the first kb was demonstrated to be contained within the boundary by the *Hin*dIII/ *Dde*I digest).

Although this interpretation of the autoradiograph appears logical, and is consistent with previous findings, it is clear that there are too many anomalies for an accurate conclusion to be drawn at this stage. There are three spurious results in addition to the dubious band seen in the *Hin*dIII and *Hin*dIII/ *Hae*III digests. The first of these was the faint, diffuse, hybridization seen near the top of the lane in the ILTat 1.2 *Hin*dIII/ *Rsa*I digest, which appeared reproducibly in the two further repeat filters. Secondly, the faint "band" which was observed at about 400 bp in the *Hin*dIII/ *Sau*3AI 1.21 digest appeared again as an abnormally strong "signal" (possibly background) in one repeat filter, and was absent from the second repeat filter. Finally, no convincing signal was produced from the *Hin*dIII/ *Mbo*I digests in any of the filters.

In order to confirm the hypothetical duplication boundary, this experiment will be repeated using new genomic DNA to overcome the degradation problems, and will utilise an alternative method for manufacturing the probe (such as gamma-radiolabelling) to attempt to enhance the hybridization specificity.

CHAPTER 6

DISCUSSION

6.1 Antigenic variation

African trypanosomes can survive prolonged periods within the hostile environment of the mammalian bloodstream by constantly changing the protective VSG coat expressed on their membrane surface, in a process commonly referred to as antigenic variation. This enables small, antigenically novel sub-populations of the parasite to escape immune-destruction and proliferate during the refractory period in which the host immune system must recognise and respond to the new VSGs. The population dynamics of the infection are obviously greatly influenced by the host immune system, and the majority of the trypanosomes are killed at the end of each parasitaemic wave (Barry and Turner, 1991). However, the immune response itself appears not to influence the timing of antigenic variation (Myler *et al.* 1985), but fashions the infection by eliminating the trypanosomes that have not switched, or have changed to an ineffective or previously encountered VSG coat.

Approximately 1000 *VSG* genes are available in the parasite's repertoire (Van der Ploeg *et al.* 1982), and this vast capacity for variance enables the infection to persist for extensive periods in the mammalian host. However, as discussed earlier (Introduction, section 1.2, page 8), prolonged infections in cattle can eventually lead to the eradication of the infection (Nantulya *et al.* 1984; Barry, 1986), but it is thought only after the host has acquired immunity to the entire spectrum of *VSG* genes. The genetic cost of antigenic variation is immense, with approximately 10% of the trypanosome genetic repertoire dedicated to the *VSG*s, but the resultant evasion strategy proves to be one of the most effective mechanisms employed by parasites to elude immune-destruction.

Several other pathogenic micro-organisms also utilize antigenic variation to evade the host immune system (Deitsch *et al.* 1997). Interestingly, the spirochete *Borrelia* has evolved a system that is remarkably similar to that of the African trypanosomes (Barbour, 1990; Borst, 1991b; Donelson, 1995). *B. hermsii*, the causative agent of relapsing fever in western United States and Canada, is the most comprehensively studied of the *Borrelia* species, and the mechanism of antigenic variation has been characterized extensively (Barbour, 1993). The majority of the cell's genes reside on a 1 Mb linear DNA molecule, but the genome also includes circular plasmid-like DNAs, and several linear DNA molecules of 10-200 kb, known as the linear plasmids or minichromosomes (Barbour, 1993). An outer membrane protein, termed

the variable major protein (vmp), appears to be the immunodominant antigen involved in the variation, and genes responsible for these proteins seem to occupy linear plasmids between the size of 28 and 32 kb (Plasterk *et al.* 1985; Kitten and Barbour, 1990). The majority of these *vmp* genes exist as silent copies on the "storage plasmids" (Barbour, 1993), while one *vmp* located near the telomere of a different plasmid (the "expression plasmid") is expressed in each organism (Kitten and Barbour, 1990). A single cloned organism can give rise to at least 40 antigenically distinct serotypes (Plasterk *et al.* 1985), the majority of which appear to arise following the duplicative transposition of a silent *vmp* gene into the expression plasmid, substituting the original *vmp* (Plasterk *et al.* 1985; Kitten and Barbour, 1990).

It was suspected for a number of years that *B. burgdorferi*, the organism responsible for Lyme disease, would possess a similar system for antigenic variation, as the infection displays many characteristics of relapsing fever. However, this hypothesis was confirmed only recently, following the discovery of a genetic locus called the *vmp*-like sequence (*vls*) that closely resembles the *vmp* sequence of *B. hermsii* (Zhang *et al.* 1997). This locus was identified on a 28 kb linear plasmid, which incorporates a *vls* expression site and 15 additional silent *vls* cassettes.

Both borreliae and trypanosomes rely on an invertebrate vector to transfer the infection between mammalian hosts. Antigenic variation enables the parasites to remain in the mammalian bloodstream for prolonged periods, and this increases the probability that they will be transferred to a new host by the vector. The similarity between the borrelia and trypanosome systems is a clear example of convergent evolution, where two distinct extracellular parasites have independently evolved analogous mechanisms for increasing their viability in the mammalian host.

The malaria parasite also utilizes antigenic variation to prevent immune-destruction within the mammalian bloodstream, but employs a strategy different (at both the phenotypic and genetic levels) from the one observed in the African trypanosomes (Borst *et al.* 1995; Deitsch *et al.* 1997). In *Plasmodium falciparum*, the variant antigens occur on the surface of the host erythrocyte, during an intracellular stage of the life cycle when the parasite undergoes proliferation. These surface antigens, one group of which are collectively known as PfEMP1 (*Plasmodium falciparum*-infected erythrocyte membrane protein 1) (Magowan *et al.* 1988), appear to have evolved to allow the infected host cells to bind to the vascular endothelium, resulting in their

retention at the vascular beds. This process prevents the cells from passing through the spleen, which under normal circumstances would recognise the damaged erythrocytes and remove them from the circulation. However, the proteins themselves elicit an antibody response by the humoral immune system, and this has necessitated the development of an immune-evasion strategy.

PfEMP-1 is encoded by the *var* gene family (Su *et al.* 1995), which have been detected in both chromosome internal and subtelomeric domains (Su *et al.* 1995; Fischer *et al.* 1997). In early ring stage parasites, it appears that different *var* gene transcripts are detectable at the same time from several chromosomes in an individual cell (Chen *et al.* 1998; Scherf *et al.* 1998). However, a recent Northern analysis by Taylor *et al.* (unpublished work reviewed in Newbold, 1999) discovered only one full length mRNA at this stage, although multiple transcripts could be detected at the 5' end of different *var* genes by RT-PCR. This result indicates that many *var* promoters are active in a single cell, but some regulatory mechanism (such as transcription attenuation, or selective degradation) ensures that only one mature transcript is formed. In support of this view, it has been found that only one PfEMP-1 type is detectable at the erythrocyte surface in trophozoite-infected cells (Chen *et al.* 1998). It seems that *var* gene expression does not occur from a unique expression site, since expressed *var* genes were mapped to several different chromosomal loci (Su *et al.* 1995; Fischer *et al.* 1997). Subsequent analysis has suggested that *var* switching and expression occurs *in situ*, controlled at the transcriptional level by an undetermined mechanism (Scherf *et al.* 1998).

Recently, it has been suggested that that other surface molecules, known as rifins, are also a major source of antigenic variation (Kyes *et al.* 1999; Wahlgren and Bejarano, 1999). These polypeptides are encoded by the *rif* (repetitive interspersed family) gene family, and appear to reside in sub-telomeric loci. It is estimated that 200 *rif* genes are represented in the genome, making it the largest gene family in *P. falciparum*, and potentially enabling the parasite to express a vast array of surface antigens on the erythrocyte membrane.

In addition to the African trypanosomes, and certain species of *Borrelia* and *Plasmodia*, there are several other pathogens which take advantage of phenotypic variation to prolong their survival in the host (Deitsch *et al.* 1997). These include *Neisseria*, which can vary the pilins expressed on its cell surface (Meyer *et al.* 1990; Seifert, 1996), and *Giardia*, which is able to change its variable surface proteins

(Nash, 1997). All of these systems are complex and require significant investment at both the genetic and phenotypic levels, but they are nevertheless essential for prolonged survival of the organisms.

Intricate immune-evasion mechanisms have developed as a direct result of the evolutionary conflict between parasite and host. The more effective the parasite's strategy, the more natural selection will favour the host which can evolve a corresponding defence. In turn, the development of a new host response will result in the natural selection of parasites that have acquired a novel evasion mechanism. This phenomenon is the basis of the Red Queen theory (Van Valen, 1973), a hypothesis used by evolutionary biologists to explain parasite-host, predator-prey, and competitor-competitor interaction. The theory simply states that two competing organisms are constantly developing new ways to improve their chance of monopolising a shared resource. This results in an evolutionary advantage constantly switching from one organism to the other, since the acquisition of a selective benefit in one competitor necessitates a change in the less fit individual. In this way, parasite and host co-evolution has led to development and refinement of the formidable mammalian immune system, and the highly effective immune-evasion mechanism, antigenic variation.

6.2 Pleomorphism and monomorphism: significant differences between trypanosome lines resulting from laboratory attenuation

Repeated syringe passaging in the laboratory has resulted in many trypanosome lines manifesting significant modifications in their population dynamics and life cycle when compared with non-passaged field isolates. These laboratory attenuated, or monomorphic, lines no longer differentiate to the stumpy bloodstream form and under normal circumstances cannot be transmitted by tsetse fly bite. In addition to these morphological changes, there is also a massive drop in the rate of antigenic variation, which, at $1 \times 10^{-6} - 1 \times 10^{-7}$ switches/cell/generation (Lamont *et al.* 1986), is up to five orders of magnitude lower than in the non-attenuated, pleomorphic lines (Turner, 1997). *Borrelia* also appears to become attenuated in the laboratory, displaying a loss of virulence as a result of *in vitro* passaging (Koomey, 1997; Zhang

and Williams, 1997); coincident with this attenuation is the loss of plasmids in Lyme disease (Xu *et al.* 1996), although it is not clear if this is the basis for the phenotypic changes.

The reduced switch rate associated with monomorphism has facilitated study of trypanosome antigenic variation by enabling cloned populations to remain predominantly of one VAT during expansion, and this has led to a comprehensive understanding of various genetic processes involved in VSG switching. However, the diminished rate of antigenic variation is low enough to be explained by background mutation and general mitotic homologous recombination, and this raises questions about the significance of the *VSG* gene activation events observed in monomorphic lines.

As discussed earlier, natural selection favours trypanosomes that evolve the most effective strategies for antigenic variation, and the success of these schemes is influenced by the rate and mode of *VSG* gene switching. It is therefore important to consider the possibility that the artificial selection to monomorphism is associated with changes in the cell's recombinational or transcriptional machinery. It has been proposed (Barry, 1997a) that the marked alteration in switch rates between monomorphic and pleomorphic trypanosomes is indicative of a specific gene switching mechanism, that might include a dedicated enzyme activity catalysing specific recombination, which is reduced in, or even absent from, monomorphic lines. This model suggests that duplicative transposition is initiated by a DNA double-strand break (DSB) in the *BES* 70 bp repeats (for further details see section 6.5, page 146). Under these circumstances, the monomorphic trypanosomes would have to rely on alternative minor switching mechanisms, or possibly background recombination, to sustain VSG changes.

DSB repair in vertebrates and yeast is known to be undertaken by more than one pathway (Van Dyck *et al.* 1999). Vertebrate DSBs are repaired primarily by Ku-dependent non-homologous end-joining (Kanaar *et al.* 1998), but Rad52-dependent homologous recombination also appears to play an important role in this process (Liang *et al.* 1998; Takata *et al.* 1998). *RAD52*$^{-/-}$ knockouts in mouse (Rijkers *et al.* 1998) and chicken (Yamaguchi-Iwai *et al.* 1998) cells produce an overall reduction in homologous recombination, but do not affect the cells' sensitivity to X-ray radiation. This indicates that the rate of DSB repair is unaffected by these mutations, demonstrating that the repair mechanism can occur entirely by the Ku-directed non-

homologous end-joining route in the absence of Rad52. However, in yeast cells, where the Ku-mediated process represents only a minor pathway for DSB repair (Critchlow and Jackson, 1998), *RAD52* mutants display an extreme sensitivity to ionizing radiation and exhibit severe recombinational defects (Game, 1993). In a similar manner, the reduced switching rate of the monomorphic trypanosomes could be indicative of the parasite utilising less significant modes of antigenic variation following the loss, or reduction, of the predominant gene switching mechanism.

The ILTat 1.2 study, presented in this thesis, was undertaken to elucidate the relative importance of the various *VSG* switching mechanisms during the early stages of a chronic, pleomorphic rabbit infection. A study of this kind would be impractical using the trypanosome lines with an overall switch rate of 1 x 10^{-2} switches/cell/generation, since there is a considerable risk of the clones disintegrating phenotypically during their analysis. This problem was overcome by using the ILTat 1.2 line, which switches at the lower rate of 1 x 10^{-5} switches/ cell/generation. The reduction in switch rate probably occurred during the unrecorded period of syringe passaging (Onyango *et al.* 1966; Miller and Turner, 1981) before the SUSB 48 stabilate was prepared. However, it is clear that the parasites have not become monomorphic during this period, since they still yield the stumpy bloodstream form and remain fly transmissible, emerging from the insect with a typical pleomorphic overall switch rate of 1 x 10^{-2} switches/cell/generation (Turner, 1997). In addition, this line will not proliferate in liquid culture (whereas monomorphic lines do), but does grow on a semi-solid agarose plate culture (Vassella and Boshart, 1996), clearly differentiating to stumpy form cells. It is also worth remembering that individual clones within a pleomorphic infection display a range of rates between approximately 1 x 10^{-2} and 1 x 10^{-5} switches/ cell/generation (Turner and Barry, 1989; Turner, 1997). Therefore, it appears that the ILTat 1.2 clone, which switches 1-2 orders of magnitude faster than monomorphic lines, falls within this natural range of pleomorphic switch rates.

In addition to the ILTat 1.2 chronic infection, a single relapse investigation using the ILTat 1.61c pleomorphic line is presented in this thesis. ILTat 1.61c, which is derived from a single metacyclic cell, displays the highest recorded VSG switch rate at about 3 x 10^{-2} switches/cell/generation, and a PCR approach for analysing these rapid switch products is proposed in section 3.9 (page 70).

6.3 The hierarchical order of *VSG* gene expression in chronic infections

It is now widely accepted that the hierarchical expression of VATs during antigenic variation enables the parasite to make efficient use of its *VSG* repertoire, and therefore maximize the infection period. As previously discussed, the probability of an infection being transmitted by tsetse fly to a new mammal is directly dependent on the time that the parasite can survive in the original host and the level of parasitaemia. This strong selective pressure has resulted in the refinement of the *VSG* expression hierarchy. The "semi-predictable" order of antigen type appearance (Gray, 1965; Capbern *et al.* 1977; Miller and Turner, 1981; Timmers *et al.* 1987) appears to result from a range in the activation frequencies of the various *VSG*s, and it may now be possible to propose the relative contributions of different switching mechanisms to this phenomenon. The *VSG* genes with the highest probability of activation are those residing at telomeres, and in the ILTat 1.2 pleomorphic study presented in this thesis, the genes appear to be activated predominantly by duplication, whereas in monomorphic lines they are mainly transcribed following *in situ* switches (Liu *et al.* 1985). Presumably the duplication of chromosome internal genes becomes apparent as the infection progresses, since these genes represent the bulk of the *VSG* repertoire. It is likely that, in a chronic pleomorphic infection, the telomeric genes are constantly being duplicated at a high rate, and this will inevitably lead to re-expression of some *VSG*s that have previously been encountered by the host. Obviously, all switches to such genes will be lethal, and eventually the less frequently activated chromosome internal gene products will emerge, as the host acquires immunity against the reservoir of telomeric *VSG*s. As the infection reaches the final stages, when the host has exhausted the parasite's gene repertoire, the mosaic genes resulting from segmental gene conversion ultimately appear (Thon *et al.* 1989). These gene products must occur very infrequently, and tend to arise in the terminal phase of infection (Barbet and Kamper, 1993); (Thon *et al.* 1989).

To date, the most comprehensive analysis of hierarchical expression was performed by Capbern *et al.* (1977) who studied 101 different VATs in *T. equiperdum* infections of rabbits, and discovered that the VATs formed three hierarchical groups, rather than a specific series. The early ("précoce") group consisted of VATs that could be predicted to appear within the first three weeks of the rabbit infection, the

middle ("semi-tardif") group followed thereafter, while the late ("tardif") group involved the VATs that occurred after the first month of infection. It is possible that the first two groups represent *VSG*s duplicated from the telomeric and chromosome internal sites respectively, while the late group has been demonstrated to include infrequently activated genes (Thon *et al.* 1989).

6.4 VAT appearance in the chronic ILTat 1.2 infections

The chronic ILTat 1.2 infections were terminated after 30 days in the lop-eared rabbits and after 21 days in the New Zealand White rabbit and during this time the trypanosomes achieved first and second relapse waves of parasitaemia. It is therefore evident that the eleven VATs observed in these infections are representative of the early switch products, generated at the start of a *VSG* hierarchy. As mentioned earlier, duplicative transposition appeared to be the major method for gene activation during the first and second relapse peaks, and telomeric *VSG*s were the predominant source of ELCs, with fourteen of the seventeen silent genes localised at the chromosome ends (Table 3, section 4.14, page 105).

Although these ILTat 1.2 infections are only representative of the first ("précoce") group of the antigenic hierarchy, it appears that there is also a degree of order even within this period. For example, the number of clones isolated per VAT in the first relapse peak was varied, with no single antigen type predominating, but in the second relapse peak ILTat 1.23 was by far the major VAT, present on every day of the relapse. In addition, four of the VATs observed in the first relapse peak (ILTats 1.25, 1.67, 1.68 and 1.69) elicited antibodies at a similar time in all four hosts, while one VAT (ILTat 1.21) showed more variation in the timing of its appearance, and the final VAT (ILTat 1.64) only emerged in two of the four rabbits. A much greater variation was seen in the second relapse peak, however, with only one of the five VATs appearing in all of the infections (ILTat 1.70). One of the minor VATs identified from this peak (ILTat 1.71) appeared to be duplicated from a chromosome internal BC, and it is logical to assume that this could be one of the most frequently activated *VSG*s from the internal arrays.

It seems likely that the sub-hierarchical variation observed in the early group will also occur within the middle and late groups. The probability of VAT appearance

will depend on several factors that could include the chromosomal location of the *VSG*, the gene orientation relative to the chromosome (Van der Werf *et al.* 1990) and the homology of the flanking sequences between the BC and the *BES*. It is apparent that other variables also influence, and complicate, the hierarchy and sub-hierarchy; for example, it has been observed that some VATs have a much greater probability of appearance after an intermediate *VSG* gene has been expressed (Miller and Turner, 1981).

Several of the VATs seen in the rabbit infections had been observed previously in other studies. Miller and Turner (1981) demonstrated that ILTats 1.21, 1.22, 1.23 and 1.25 appeared as primary switch products in a single relapse analysis of seven different VATs. This study revealed that ILTat 1.21 (underlining indicates those VATs identified in the ILTat 1.2 infection presented in this thesis) was the most frequently activated VAT (arising in 32 of 41 first relapses in rats), followed by 1.4 (33 of 47), 1.25 (27 of 41), 1.26 (14 of 35), 1.24 (15 of 41), 1.22 (6 of 41) and 1.23 (2 of 36). These results enlarge the dataset produced from the ILTat 1.2 rabbit infection, since the Miller and Turner study was initiated using trypanosomes also derived from EATRO 795, and closely related to the SUSB 48 stabilate. It is therefore evident that the VATs isolated from the ILTat 1.2 rabbit infection represent some of the most frequently activated VSGs.

Fly transmission of EATRO 795 into goats had also identified ILTat 1.21 and 1.69 switched products in the early stages of every infection (Barry and Emery, 1984). In addition, the ILTat 1.69 (ETat 1.7), 1.21 (ETat 1.9) and 1.22 (GUTat 7.1, ETat 1.2) VATs have been detected in related trypanosome stocks (field isolates), and their activation is typically observed early in infection (McNeillage *et al.* 1969; Barry *et al.* 1985). Two VATs, ILTats 1.22 and 1.64 (GUTat 7.13), have also been observed in the metacyclic population of the tsetse fly (Barry *et al.* 1983, Cornelissen *et al.* 1985b).

6.5 Mode of *VSG* gene activation in the ILTat 1.2 and 1.61c pleomorphic infections: the predominance of duplicative transposition

Duplicative transposition dominated in the ILTat 1.2 study, with nine, and possibly ten, of the eleven switching events occurring by this mode of activation. At least eight, and probably ten, of the donor genes resided at telomeric loci, while only one of the duplications (ILTat 1.71) definitely involved an internal BC (Table 3, section 4.14, page 105). In contrast, there was only one example of an "*in situ*" activation (ILTat 1.67). The single analysable product generated from the ILTat 1.61c single relapse investigation was also activated by the duplication of a telomeric BC. These results are remarkably different from observations in monomorphic trypanosomes, where single relapse analysis has shown that about two-thirds of switch events occur by the *in situ* mechanism, and about one-third occur by duplicative transposition (Liu *et al.* 1985). Duplicative events appear to dominate only later in these monomorphic infections (Lee and Van der Ploeg, 1987; Michels *et al.* 1983; Timmers *et al.* 1987).

The dramatic difference in the mode of *VSG* switching between monomorphic and pleomorphic lines is indicative of a dedicated switch mechanism in the pleomorphic lines that is reduced in, or absent from, the monomorphic lines. This hypothesis for gene conversion, proposed by Barry (1997a), which was based on preliminary observations on fly-transmitted trypanosomes, is consistent with the data presented in this thesis. This model suggests that the duplicative transposition is initiated by a DSB in the *BES* 70 bp repeats, and is similar in principle to the mating type switch observed in *Saccharomyces cerevisiae* (Haber, 1998), which also begins with a specific DSB in the recipient locus. In the mating type switch a specific homing endonuclease (HO) mediates the DSB, and a corresponding enzyme activity specific to the trypanosome 70 bp repeats has been hypothesised several times (Pays, 1985); (Pays *et al.* 1994; Barry, 1997a). HO endonuclease cannot cleave its recognition sequence at either of the available mating type donor loci, *HML* or *HMR*, as these sites appear to be occluded by nucleosomes in silent DNA (Haber, 1998). Similarly, it is possible that the non-*BES* 70 bp repeats could be subject to chromatin repression, preventing the specific enzyme from cutting at random in the genome. Possible support for this hypothesis is provided by the observation that transcriptional activity was repressed in *BES*, or ribosomal promoters that were introduced into a *VSG* gene chromosome internal loci (Horn and Cross, 1997) (see section 6.6 (page 149) for more details of this experiment).

It is proposed that formation of the DSB would then be followed by a repair-mediated conversion from a silent *VSG* gene, using a strand invasion process similar to that seen in the mating type switch, and would terminate at the short homology blocks at the downstream end of the gene. A low level of point mutations detected in newly generated ELCs (Kamper and Barbet, 1992; Graham and Barry, 1996) could be indicative of a DNA repair enzyme operating during ELC formation. Presumably, the putative invading strand from the *BES* barren region would represent a diverse set of the repeats (which display degeneracy across their sequence (Aline *et al.* 1985)) enabling a high degree of 70 bp homology to be associated with the relatively few repeats upstream of a particular BC gene. This aspect of the model could explain why such a vast array of 70 bp repeats are contained within the *BES*s.

An alternative model for duplicative transposition has been suggested by Borst *et al.* (1996), and invokes non-specific DNA strand breaks as initiating the process in the 5' flank of the *VSG* gene. Again, this model proposes a repair-mediated conversion from a silent *VSG* gene, using a strand invasion process similar to that seen in the mating type switch. This theory is consistent with the observed imprecision in the upstream limit of *VSG* duplication events in monomorphic trypanosomes, which only locate precisely to the 70 bp repeats in about half of the reported cases (Campbell *et al.* 1984; de Lange *et al.* 1985; Florent *et al.* 1987; Lee and Van der Ploeg, 1987). The other events appear to result from fortuitous similarity between the donor *VSG* and recipient *BES* (Donelson *et al.* 1983; Lee and Van der Ploeg, 1987; Michiels *et al.* 1983; Pays *et al.* 1983c). It has also been demonstrated that the removal or reversal of the 70 bp region of a *BES* in monomorphic trypanosomes had no effect on the incidence of *VSG* duplications into that *BES* (McCulloch *et al.* 1997). Although the Borst model provides some explanation for the lack of specificity at the upstream limit of conversion apparent in monomorphic trypanosomes, it does not explain the reduced overall switch rate, and the low relative frequency of duplicative transposition seen in these lines.

In the relatively few studies performed on pleomorphic lines, the upstream limit of duplicative events consistently maps to the 70 bp region (Matthews *et al.* 1990; Shah *et al.* 1987), and the preliminary results presented in this thesis also appear to demonstrate the involvement of these repeats (section 5.6, page 128). Telomeric *VSG*s lacking 70 bp repeats are activated in the bloodstream with great infrequency

during pleomorphic infections, as such VATs require extensive positive and negative selection (Lu *et al.* 1993). In these cases, either fortuitous sequence similarities appear to be used for duplication into a *BES*, or activation occurs *in situ*, using unusual promoters (Donelson *et al.* 1998). However, it has been demonstrated that genes with just one intact 70 bp repeat can be activated efficiently, and the upstream limit of duplication maps to the repeats (Matthews *et al.* 1990; Shah *et al.* 1987). The contrast in the mode of gene activation and switching rate observed between monomorphic and pleomorphic trypanosomes is indicative of a dedicated switch mechanism that is reduced or absent in monomorphic lines. It appears that pleomorphic trypanosomes are programmed to duplicate *VSG*s, using the 70 bp repeats to initiate the process, and this mechanism is repressed or lost in the monomorphic lines.

The control of antigenic variation remains a completely unresolved area of trypanosome biology, since no switch intermediates or regulatory genes have been discovered thus far. It is conceivable that *VSG* duplication could occur by a repair-mediated conversion from the silent *VSG* gene using a strand invasion process similar to that seen in the mating type switch, as hypothesized in the Barry and Borst models. However, several other modes of general recombination have been conserved throughout eukaryotic evolution (Paques and Haber, 1999; Haber, 1999a), and it is possible that the process could occur by any of these mechanisms. The trypanosome might even utilize a unique system that has not yet been described, and this could possibly involve more than one pathway, as in DSB repair in vertebrates. However, it does seem likely that the switching process involves homologous recombination, since the disruption of *RAD51*, the trypanosome homologue of the bacterial *RecA* gene, leads to a substantial decrease apparently in all types of switching (McCulloch and Barry, 1999). This work was performed using a monomorphic line which, as previously discussed, displays a marked reduction in the overall rate of *VSG* gene switching and a decreased proportion of duplicative activations relative to non-duplicative events. It would therefore be extremely interesting and informative to examine the outcome of a similar knockout experiment in a pleomorphic line.

It is perhaps more than coincidence that monomorphic trypanosomes exhibit both a marked reduction in switch rate and a significant alteration in the life cycle, being unable to perform the pleomorphic cell's differentiation from dividing long slender

cells to non-dividing short stumpy cells. It seems plausible that the individual mechanisms underlying these two distinct phenotypes could be regulated by a single process (or gene) that is lost (or repressed) during the selection to monomorphism. This hypothesis implies a close genetic association between the trypanosome's cell-cycling and recombinational processes.

6.6 Telomeric *VSG* genes are activated most frequently, and predominate in the early stages of infection

Prior to this investigation, other authors have demonstrated the preferential use of telomeric donors in early duplicative events (Liu *et al.* 1985; Myler *et al.* 1984; Pays *et al.* 1983a; Young *et al.* 1983), although most of these studies were performed using monomorphic lines and followed a series of switches, rather than single relapses. This favoured use of telomeres could result from the fact that telomeric *VSG* loci, which encompass long stretches of both the 70 bp and (sub)-telomeric hexanucleotide repeats, share more sequence homology with each other than with the internal *VSG* loci. Additionally, the probability of telomeric *VSG* gene expression could perhaps be enhanced by a general interactivity of chromosome ends. It has been reported previously that telomeric interactions appear to aid their recombination with active *BES*s. For example, some *VSG* genes that are usually expressed infrequently (or emerge late in an infection) can become activated early if they are translocated into a telomeric environment showing more homology with the expression site (Laurent *et al.* 1984). A similar effect has been shown to occur during the expression of the AnTat 1.1 (Antwerp *Trypanozoon* antigen type 1.1) *VSG* gene, which is telomeric but present in the reverse orientation with respect to the chromosome end. This gene is normally expressed late in infection, but becomes predominantly activated when it occurs on a telomere in the same orientation as the chromosome end (Van der Werf *et al.* 1990). In addition, it has been shown in yeast that placing a Y' element (a highly conserved sub-telomeric element found repeated in 0-4 tandem copies adjacent to the telomere repeats) at an internal locus results in at least a 10-fold reduction in ectopic interactions with the telomeric Y's (Louis, 1995). This led to the suggestion that the telomere regions could perhaps be sequestered from the rest of the genome during recombination.

Minichromosomal *VSG* genes appeared to play an important role in antigenic variation during the early stage of the ILTat 1.2 infection presented in this thesis, and there did appear to be hierarchical use within this subset. In the first relapse peak, four of the six VATs possessed minichromosomal copies (ILTats 1.25, 1.68, 1.69 and 1.21) and appeared at similar times in all four rabbit infections, and in the 1.61c high-switching investigation, the 1.25 *VSG* gene was duplicated from a single minichromosomal BC. Two of the VATs from the second relapse peak, ILTats 1.23 (the dominant VAT in this peak) and 1.72, also displayed minichromosomal *VSG* gene copies, but they appeared to be activated less frequently than the VATs observed in the first relapse peak. In a previous study, the ILTat 1.3 minichromosomal *VSG* gene has been shown to possess a long array of 70 bp repeats (Shah *et al.* 1987). The preliminary investigation of the 1.21 minichromosomal upstream flank presented in this thesis is also suggestive of a long 70 bp array of approximately 3.5 kb (section 5.6, page 128). These extensive 70 bp arrays could increase the probability of strong homology arising between the donor and *BES*, and this could explain the preferential use of minichromosomal donors, if the ILTat 1.3 and 1.21 *VSG*s are representative of the minichromosomal repertoire. However, the prominence of minichromosomal telomeres as donors could simply be due to their relative abundance in the genome (Van der Ploeg *et al.* 1985; Weiden *et al.* 1991), rather than any inherent feature.

It has been observed that, during cell division in procyclic trypanosomes, the minichromosomes appear to utilize a partitioning mechanism different from that of larger chromosomes (Ersfeld and Gull, 1997). This results in the two chromosome classes localizing to distinctly different regions of the nucleus during the various phases of the cell cycle (Chung *et al.* 1990; Ersfeld and Gull, 1997). The partitioning of the chromosomes has not been as comprehensively analysed in bloodstream form trypanosomes, but it appears that the distribution of both the minichromosomes and the larger chromosomes is more random and dispersed than in procyclic cells (Chung *et al.* 1990). Perhaps this more heterogeneous chromosome organization increases the amount of interaction between the two size classes of chromosome, which possibly have less opportunity for recombination in procyclic trypanosomes due to their differential compartmentalisation throughout the cell cycle.

The absence of minichromosomes in trypanosomatid flagellates unable to undergo antigenic variation, like *Crithidia* (Van der Ploeg *et al.* 1984b) and *Leishmania*, has led to the suggestion that minichromosomes have evolved to expand the telomeric *VSG* gene repertoire (Borst *et al.* 1993). In *T. vivax*, which has few, if any, minichromosomes (Dickin and Gibson, 1989; Van der Ploeg *et al.* 1985), antigenic variation does occur, but the hierarchical order of *VSG* expression appears to follow a more continuous spectrum than the complicated pattern seen in *T. brucei* (Barry, 1986). It is possible that this comparatively simple hierarchy is due to the smaller pool of telomeric *VSG* genes available in the *T. vivax* repertoire.

Telomeric genes appear to play an important role in antigenic variation of several protozoon parasites and other pathogens. The *VSG* genes of *T. brucei*, the *var* and *rif* genes of *Plasmodium falciparum*, the *VSP* genes of *Giardia lamblia* and the *vmp* genes of *Borrelia hermsii* can all be found at telomeric loci. This has led to the suggestion that telomeric interactions, and the associated genetic flexibility, are an essential feature of antigenic variation in many protozoans (reviewed in Lanzer *et al.*, 1995).

In other organisms, telomeric interactions are now being seen as having important functions, and the growing amount of interest and research in this field is beginning to elucidate the various proteins associated with telomeric function. An important example is the involvement of the Ku and Sir (silent information repressor) proteins in telomere organization, and regulation of this has been studied in some detail (Shore, 1998; Haber, 1999b). It had previously been observed that, in the absence of Ku, mammalian cells were unable to repair double-strand breaks or perform recombination of the immunoglobulin V(D)J region (Jeggo *et al.* 1995), and it has since been discovered that Ku protein plays an important role in non-homologous end-joining (NHEJ) in both vertebrates and yeast (Kanaar *et al.* 1998; Critchlow and Jackson, 1998), and in telomere length regulation and silencing in yeast (Boulton and Jackson, 1998). Sir proteins are also known to be involved in silencing at telomeres and at *HM* mating type loci in yeast (Laurenson and Rine, 1992), and recent studies have revealed that there appears to be a close association between Ku and Sir proteins (Tsukamoto *et al.* 1997). The Ku and Sir proteins normally reside at the telomeric and subtelomeric regions (Martin *et al.* 1999), but delocalize when a DSB is introduced, and subsequently associate with the damaged region. Following the redistribution of Ku and Sir proteins, the genes near telomeres that were

previously epigenetically silenced become more strongly transcribed (Martin *et al.* 1999). It has also been demonstrated that telomere silencing and the subnuclear organization of telomeres appear to be closely associated. Laroche *et al.* (1998) have shown, to some extent, that the mutation of yeast Ku genes result in the loss of telomeric silencing, and an alteration in the localization of the telomeric DNA, which normally clusters around the nuclear periphery at several loci in wild-type cells. In addition, it has been observed that telomere-silencing mutations in fission yeast affect the pattern of chromosome alignment and migration during meiosis, and meiotic recombination appears to be severely reduced (Cooper *et al.* 1998; Nimmo *et al.* 1998).

Transcriptional repression in yeast appears to be associated with specialized chromatin structure at the affected region. In *Drosophila*, telomeric silencing is also linked to local chromatin structure, which seems to be dependent on nuclear organization (Cryderman *et al.* 1999). It is known that transcriptional silencing in mammals and *Drosophila* can be delimited by specialized boundary elements, and negated by insulating elements (Gerasimova and Corces, 1996), and recent studies are beginning to suggest that analogous elements also exist on yeast telomeres (Fourel *et al.* 1999; Pryde and Louis, 1999).

Horn and Cross (1997) have demonstrated that *T. brucei* also displays position-dependent transcriptional silencing. This investigation revealed that inserting reporter genes (under the control of either an *rRNA* or *VSG BES* promoter) into a silent *BES* in bloodstream form trypanosomes resulted in the silencing of the reporter, and this effect was more pronounced closer to the telomere. The promoter repression was also seen when the constructs were introduced at a non-telomeric *VSG* locus in bloodstream form cells. However, in procyclic trypanosomes all the inserted *rRNA* promoters displayed transcriptional activity, while the *BES* promoters remained repressed. These results are indicative of a positional effect, related to *VSG* proximity or some feature of the *BES*, which prevents transcription from promoters introduced into bloodstream form cells. It is thought that the *rRNA* promoter-specific derepression in procyclic cells could be attributed to sequences that occur within rRNA promoters, but are absent from *VSG BES* promoters.

Telomeric interaction and transcriptional silencing are known to be important in trypanosome antigenic variation, although the mechanisms controlling these processes have yet to be determined. In other organisms, there is growing evidence

that chromosomal interaction and transcriptional control are closely associated (Henikoff, 1997), and it is possible that similar systems exist in *T. brucei*. It is tempting to speculate that Ku and Sir proteins could be involved in trypanosome telomeric regulation, but this theory can only be confirmed by forthcoming research.

6.7 The occurrence of a dominant *BES*

In every duplicative activation detected in the first relapse peak of the ILTat 1.2 infection presented in this thesis, the ELC was accompanied by the loss of the 1.8 Mb ILTat 1.2 *VSG* fragment, indicating that the same *BES* was utilised in each case. The 600s PFGE separation established that this *BES* was present on the largest of the five comigrating 1.8 Mb chromosomes. However, this study only examined the switches resulted in a phenotypic change, and there could also have been duplicative transpositions into silent *BES*s, as previously demonstrated (Myler *et al.* 1988a). The preferential utilization of a dominant *BES* has also been observed in monomorphic studies (Liu *et al.* 1985).

It is possible that the repeated use of the active *BES* may be because the transcriptionally active state makes it a more accessible target for recombinational events. Actively transcribed telomeres have been shown to grow slightly faster than inactive telomeres (Pays *et al.* 1983b; Myler *et al.* 1988b), and this is perhaps indicative of differences in chromatin structure between the active and inactive sites. In support of this theory, it has been observed that only the active *BES* is highly sensitive to DNAaseI (Pays *et al.* 1983b) and single-strand specific endonucleases (Greaves and Borst, 1987). Additionally, transcriptional silencing around subtelomeric *VSG* genes is associated with the presence of β-D-glucosyl-hydroxymethyluracil (base J), which is reversibly lost during transcriptional reactivation of a *BES* (Bernards *et al.* 1984; van Leeuwen *et al.* 1997) (see section 1.9, page 25 for further details). It is unclear, however, whether changes associated with this modification are the cause or consequence of transcriptional silencing. The extent of the modification appears to vary with the length of the telomeric repeat

sequence and the proximity of the telomere end (Bernards *et al.* 1984; van Leeuwen *et al.* 1996; van Leeuwen *et al.* 1997).

Base J modification renders various restriction enzyme sites indigestible in some DNA molecules, and this can result in partial products being observed in genomic digests. Indirect evidence for base J modification of the minichromosomal telomere containing the silent ILTat 1.21 *VSG* BC was provided by the *Pst*I and *Sal*I digests presented in this thesis (section 4.5, page 91; section 5.2, page 112). Both these restriction sites contain a thymidine that could be used as a substrate for base J modification. Partial products were generated from the ILTat 1.21 minichromosome (demonstrated by the ILTat 1.21 cDNA probing) when 1.2 DNA was digested with these two enzymes. It was apparent that the partial digestion resulted from modified DNA, rather than an incomplete enzyme reaction, since no partial products were detected when the same filters were hybridized with the *RAD51* control probe. Base J modification has also been observed in the 177 bp repeats that make up the central part of the minichromosomes (van Leeuwen *et al.* 1997).

6.8 The ILTat 1.21 minichromosomal *VSG* upstream flank

The ILTat 1.21 minichromosomal clone characterized in this thesis encompassed the entire cotransposed region and included three and a third copies of the 70 bp repeats at its upstream end. BLAST and CLUSTAL W analysis revealed that the 1700 bp cotransposed region contained a 1200 bp *ESAG3* full length pseudogene. The presence of this *ESAG3* homologous sequence downstream of the 70 bp repeats suggests that minichromosomal telomeric sequences result from recombinational events with other telomeric regions within the genome. It also seems likely that the cotransposed region is not utilized as a recognition sequence with the corresponding region of a *BES* for the ILTat 1.21 duplication event, since such a large proportion of this BC domain is comprised of the *ESAG3* sequence. The structure of the ILTat 1.21 cotransposed region implies that this domain is merely "stuffer" sequence, and does not perform any dedicated function in the *BES*. This theory contrasts with the view of Davies *et al.* (1997), who observed an increased rate in *BES* switching following the deletion of the cotransposed region in the active *BES*, and hypothesised that this domain could play an important role in *BES* regulation.

The preliminary investigation of the ILTat 1.21 upstream duplication boundary suggests that the 70 bp repeats extend for approximately 3.5 kb, and the duplication boundary lies somewhere beyond the first 1 kb, upstream of the *VSG* (section 5.6, pages 128-135). It also seems that the 70 bp and 177 bp minichromosomal repeat elements are arranged at close proximity, with no (or very little) non-repetitive sequence between them, since the entire minichromosome appeared to be unaffected by the frequently cleaving restriction enzyme, *Hae*III. The 3.5 kb array of the 70 bp repeats observed in the ILTat 1.21 minichromosome is similar to the stretch of repeats detected in ILTat 1.3 (Shah *et al.* 1987). It is possible that this long array of 70 bp repeats could be an inherent feature of the minichromosomes, but this can only be confirmed by the characterization of more minichromosomal *VSG* 5' flanks.

6.9 Future work

In order to confirm the hypothetical ILTat 1.2 upstream duplication boundary, it will be necessary to repeat the last analysis presented in this thesis (section 5.6, page 128) using newly prepared, minimally sheared, genomic DNA to overcome the degradation problems. An alternative method for manufacturing the probe (such as gamma-radiolabelling) could also be adopted to attempt to enhance the hybridization specificity. If the repeat hybridization is consistent with the previous results, then the following approach could be utilized to construct a full length clone of the ILTat 1.21 upstream flank (summarized in Figure 52).

It would first be necessary to clone the 3.5 kb *Alu*I upstream fragment (Chapter 5.6, page 133), and this could be achieved by constructing an *Alu*I/ *Hin*dIII size-selected minichromosomal library in pBluescript. Digesting the DNA initially with *Hin*dIII, and then with *Alu*I would produce a *Hin*dIII specific overhang at the 3' end of the fragment, which would facilitate later cloning procedures. The minichromosomal DNA would be isolated by PFGE, although a substantial amount of material would be required for the size-selection step, which would remove the smaller restriction products generated by the frequently cleaving *Alu*I enzyme. Following cloning, this fragment would then be sequenced to characterize the full extent of the 5' flank.

The next phase of this procedure would be to create a full length clone of the ILTat 1.21 minichromosomal flank encompassing the whole *VSG* gene and the entire 70 bp

array. The full *VSG* sequence would be introduced into the original 1.21 minichromosomal *VSG* clone, using the 1.21 cDNA clone sequence. By using unfavourable digestion conditions, it should be possible to produce a *Hinc*II partial digest of the cDNA clone. When used in conjunction with a second enzyme that cuts within the plasmid polylinker, this digest should liberate the 1473 bp fragment extending from the first *Hinc*II site to the end of the cDNA clone. This fragment would then be introduced into the *Hinc*II site in the 5' terminal *VSG* sequence of the 1.21 minichromosomal clone. Once this stage is completed it would be possible to ligate this whole fragment to the other, contiguous, upstream fragment (isolated from the *Hin*dIII/ *Alu*I size-selected library) using the *Hin*dIII site present in both clones. The individual steps required to produce this ILTat 1.21 full-length minichromosomal clone are indicated in Figure 52.

Finally, it would be possible to manipulate this construct, which would subsequently be integrated into ILTat 1.2 trypanosome genomic DNA via homologous recombination. A series of 177 bp minichromosomal repeats (approximately 500 bp in length) would be introduced at the upstream end of the construct to provide homology at this end. It would also be necessary to introduce a promoter and selectable marker gene, such as hygromycin resistance gene, within this flank to allow the transformed cells to be selected. The promoter would have to be inserted in the reverse orientation to the *VSG* gene to prevent any artifactual transcription of the *VSG* occurring. Five potential constructs are illustrated in Figure 53: (1) unaltered 70 bp repeats; (2) deleted 70 bp repeats; (3) reversed full array of 70 bp repeats; (4) full array replaced with 1-2 70 bp repeats, in forward orientation; (5) full array replaced with 1-2 70 bp repeats, in reverse orientation. The transformed trypanosomes would then be used in single relapse investigations to discover whether the 70 bp repeat manipulations affect the activation frequency of the ILTat 1.21 *VSG* gene. Unmodified ILTat 1.2 trypanosomes would be used as the control. Construct (1) (in which the 70 bp repeats remain unaltered) should demonstrate that any changes in the ILTat 1.21 activation frequency were due to the manipulation of the flanks, rather than unpredicted changes resulting from the introduction of the construct itself.

Figure 52. Construction of the full length ILTat 1.21 minichromosomal clone. (1) The 3.5 kb "barren" region upstream of the *Nsi*I site would be isolated from a *Hin*dIII/ *Alu*I size selected minichromosomal library cloned into pBluescript. The ILTat 1.21 cDNA clone (2) would then be digested with *Hinc*II under unfavourable conditions to produce the 1473 bp partial fragment extending from the first *Hinc*II site to the end of the clone. This fragment would then be cloned into the *Hinc*II site of the original ILTat 1.21 minichromosomal clone (3). Fragment (1) would then be ligated to this product (2+3) to produce the full length clone (4). This diagram also displays the introduction of the 177 bp minichromosomal repeat sequence, and the promoter and selectable marker at the upstream end of the fragment. Abbreviations: A, *Alu*I; Hc, *Hinc*II; Hd, *Hin*dIII; N, *Nsi*I.

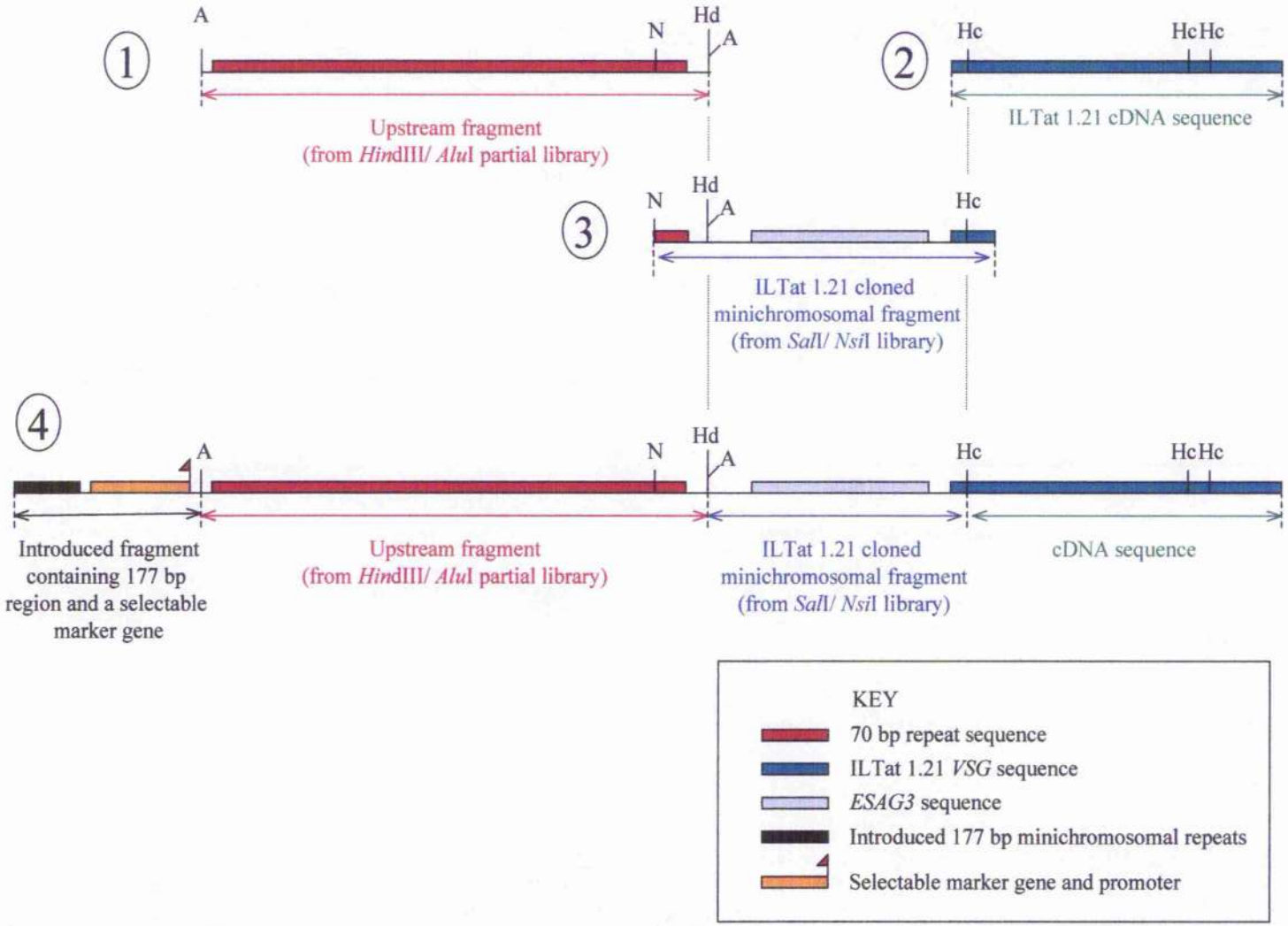
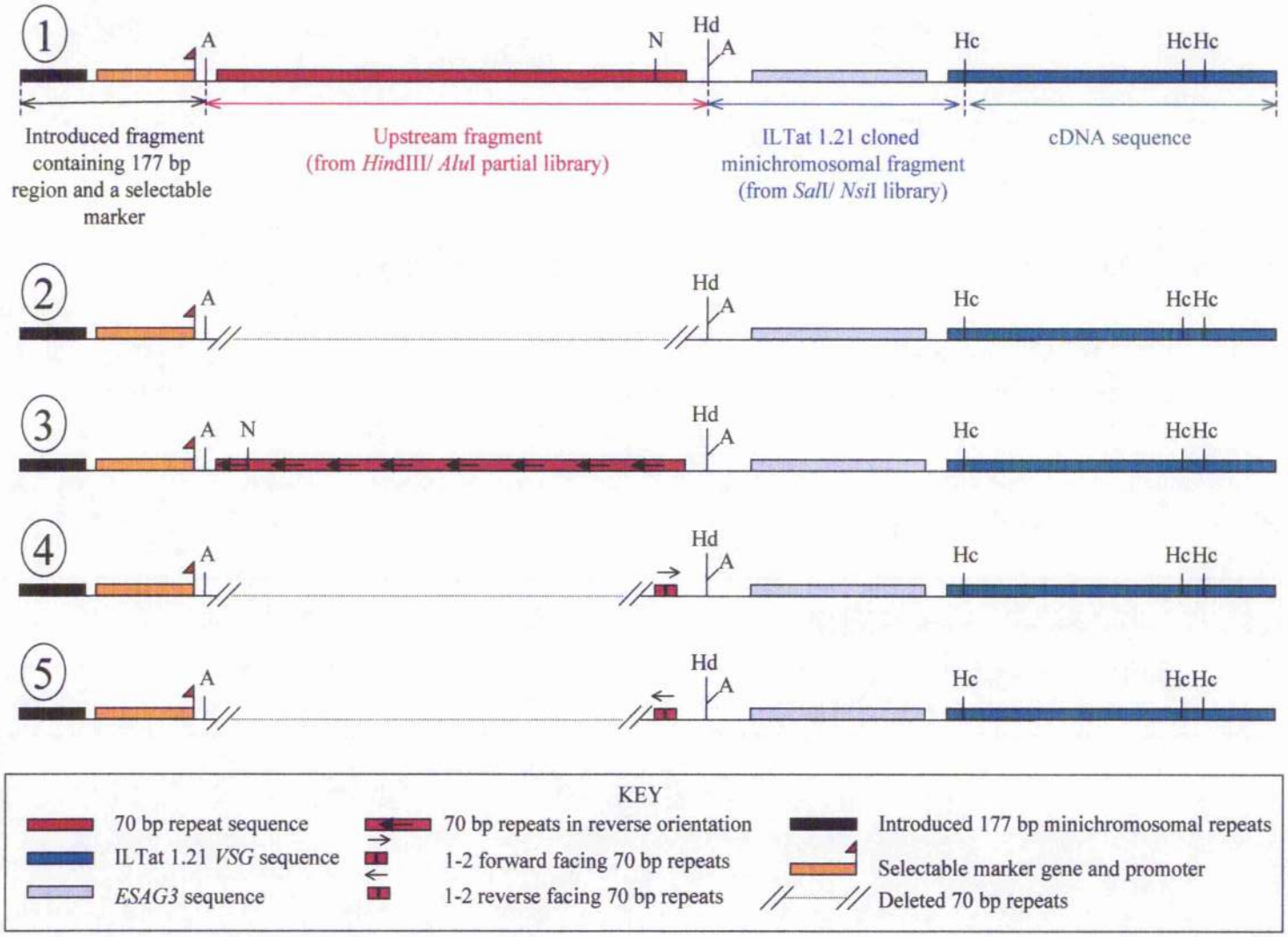

1
A
N
Hd
A
Upstream fragment
(from HindIII/ AluI partial library)
2
Hc
Hc Hc
ILTat 1.21 cDNA sequence
3
N
Hd
A
Hc
ILTat 1.21 cloned
minichromosomal fragment
(from SalI/ NsiI library)
4
A
N
Hd
A
Hc
Hc Hc
Introduced fragment
containing 177 bp
region and a selectable
marker gene
Upstream fragment
(from HindIII/ AluI partial library)
ILTat 1.21 cloned
minichromosomal fragment
(from SalI/ NsiI library)
cDNA sequence
KEY
70 bp repeat sequence
ILTat 1.21 VSG sequence
ESAG3 sequence
Introduced 177 bp minichromosomal repeats
Selectable marker gene and promoter

Figure 53. The five suggested constructs for manipulation of the 70 bp repeat region: (1) unaltered 70 bp repeats; (2) deleted 70 bp repeats; (3) reversed full array of 70 bp repeats; (4) full array replaced with 1-2 70 bp repeats, in forward orientation; (5) full array replaced with 1-2 70 bp repeats, in reverse orientation. These constructs would be integrated into ILTat 1.2 trypanosome genomic DNA via homologous recombination. The transformed trypanosomes would then be used in single relapse investigations to discover whether the 70 bp repeat manipulations affect the activation frequency of the ILTat 1.21 *VSG* gene. Unmodified ILTat 1.2 trypanosomes would be used as the control. Construct (1) (in which the 70 bp repeats remain unaltered) should demonstrate that any changes in the ILTat 1.21 activation frequency were due to the manipulation of the flanks, rather than unpredicted changes resulting from the introduction of the construct itself. Abbreviations: A, *Alu*I; Hc, *Hinc*II; Hd, *Hin*dIII; N, *Nsi*I.


1
A
N
Hd
A
Hc
HcHc
Introduced fragment containing 177 bp region and a selectable marker
Upstream fragment (from HindIII/ AluI partial library)
ILTat 1.21 cloned minichromosomal fragment (from SalI/ NsiI library)
cDNA sequence
2
3
4
5
KEY
70 bp repeat sequence
ILTat 1.21 VSG sequence
ESAG3 sequence
70 bp repeats in reverse orientation
1-2 forward facing 70 bp repeats
1-2 reverse facing 70 bp repeats
Introduced 177 bp minichromosomal repeats
Selectable marker gene and promoter
Deleted 70 bp repeats

REFERENCE LIST

Alarcon, C.M., Pedram, M. and Donelson, J.E. (1999) Leaky transcription of variant surface glycoprotein gene expression sites in bloodstream African trypanosomes. *J.Biol.Chem.* **274,** 16884-16893.

Alexandre, S., Guyaux, M., Murphy, N.B., Coquelet, H., Pays, A., Steinert, M. and Pays, E. (1988) Putative genes of a variant-specific antigen gene-transcription unit in Trypanosoma brucei. *Mol.Cell.Biol.* **8**, 2367-2378.

Alexandre, S., Paindavoine, P., Hanocq-Quertier, J., Paturiaux-Hanocq, F., Tebabi, P. and Pays, E. (1996) Families of adenylate cyclase genes in *Trypanosoma brucei*. *Mol.Biochem.Parasitol.* **77**, 173-182.

Aline, R., Macdonald, G., Brown, E., Allison, J., Myler, P., Rothwell, V. and Stuart, K. (1985) $(TAA)_n$ within sequences flanking several intrachromosomal variant surface glycoprotein genes in *Typanosoma brucei*. *Nucleic Acids Res.* **13**, 3161-3177.

Allen, G. and Gurnett, L.P. (1983) Locations of the 6 disulfide bonds in a variant surface glycoprotein (VSG 117) From *Trypanosoma brucei*. *Biochem.J.* **209**, 481-487.

Ansorge, I., Steverding, D., Melville, S., Hartmann, C. and Clayton, C. (1999) Transcription of 'inactive' expression sites in African trypanosomes leads to expression of multiple transferrin receptor RNAs in bloodstream forms. *Mol.Biochem.Parasitol.* **101**, 81-94.

Baltz, T., Giroud, C., Bringaud, F., Eisen, H., Jacquemot, C. and Roth, C.W. (1991) Exposed epitopes on a trypanosoma-equiperdum variant surface glycoprotein altered by point mutations. *EMBO J.* **10**, 1653-1659.

Barbet, A.F. (1985) Conservation of structure detected in 2 trypanosome surface glycoproteins by amino-acid sequence alignment. *Mol.Biochem.Parasitol.* **14**, 175-185.

Barbet, A.F. and Kamper, S.M. (1993) The importance of mosaic genes to trypanosome survival. *Parasitol.Today* **9**, 63-66.

Barbour, A.G. (1990) Antigenic variation of a relapsing fever borrelia species. *Annu.Rev.Microbiol.* **44**, 155-171.

Barbour, A.G. (1993) Linear DNA of *Borrelia* species and antigenic variation. *Trends Microbiol.* **1**, 236-239.

Barry, J.D. (1986) Antigenic variation during *Trypanosoma vivax* infections of different host species. *Parasitology* **92**, 51-65.

Barry, J.D. (1997b) The biology of antigenic variation in African trypanosomes. In: Hide, G., Mottram, J.C., Coombs, G.H. and Holmes, P.H., (Eds.) *Trypanosomiasis and Leishmaniasis: biology and control*, pp. 89-107. Oxford: CAB International]

Barry, J.D. (1997a) The relative significance of mechanisms of antigenic variation in African trypanosomes. *Parasitol.Today* **13**, 212-218.

Barry, J.D., Crowe, J.S. and Vickerman, K. (1983) Instability of the *Trypanosoma brucei rhodesiense* metacyclic variable antigen repertoire. *Nature* **306**, 699-701.

Barry, J.D., Crowe, J.S. and Vickerman, K. (1985) Neutralization of individual variable antigen types in metacyclic populations of *Trypanosoma brucei* does not prevent their subsequent expression in mice. *Parasitology* **90**, 79-88.

Barry, J.D. and Emery, D.L. (1984) Parasite development and host responses during the establishment of *Trypanosoma brucei* infection transmitted by tsetse fly. *Parasitology* **88**, 67-84.

Barry, J.D., Graham, S.V., Fotheringham, M., Graham, V.S., Kobryn, K. and Wymer, B. (1998) VSG gene control and infectivity strategy of metacyclic stage *Trypanosoma brucei*. *Mol.Biochem.Parasitol.* **91**, 93-105.

Barry, J.D., Graham, S.V., Matthews, K.R., Shiels, P.G. and Shonekan, O.A. (1990) Stage-specific mechanisms for activation and expression of variant surface glycoprotein genes in *Trypanosoma brucei*. *Biochem.Soc.Trans.* **18**, 708-710.

Barry, J.D. and Turner, C.M.R. (1991) The dynamics of antigenic variation and growth of African trypanosomes. *Parasitol.Today* **7**, 207-211.

Bernards, A., Van der Ploeg, L.H.T., Frasch, A.C.C., Borst, P., Boothroyd, J.C., Coleman, S. and Cross, G.A.M. (1981) Activation of trypanosome surface glycoprotein genes involves a duplication-transposition leading to an altered 3' end. *Cell* **27**, 497-505.

Bernards, A., Van Hartenloosbroek, N. and Borst, P. (1984) Modification of telomeric DNA in *Trypanosoma brucei* - a role in antigenic variation. *Nucleic Acids Res.* **12**, 4153-4170.

Blum, M.L., Down, J.A., Gurnett, A.M., Carrington, M., Turner, M.J. and Wiley, D.C. (1993) A structural motif in the variant surface glycoproteins of *Trypanosoma brucei*. *Nature* **362**, 603-609.

Borst, P. (1991b) Molecular genetics of antigenic variation. *Immunoparasitology Today-A Combined Issue Of Immunology Today And Parasitology Today* A 29-A 33

Borst, P. (1991a) Transferrin receptor, antigenic variation and the prospect of a trypanosome vaccine. *Trends Genet.* **7**, 307-309.

Borst, P., Bitter, W., Blundell, P.A., Chaves, I., Cross, M., Gerrits, H., van Leeuwen, F., McCulloch, R., Taylor, M. and Rudenko, G. (1998) Control of VSG gene expression sites in *Trypanosoma brucei*. *Mol.Biochem.Parasitol.* **91**, 67-76.

Borst, P., Bitter, W., McCulloch, R., van Leeuwen, F. and Rudenko, G. (1995) Antigenic variation in malaria. *Cell* **82**, 1-4.

Borst, P. and Fairlamb, A.H. (1998) Surface receptors and transporters of *Trypanosoma brucei*. *Annu.Rev.Microbiol.* **52**, 745-778.

Borst, P., Fase-Fowler, F. and Gibson, W.C. (1981) Quantitation of genetic differences between *Trypanosoma brucei gambiense* , *rhodesiense* and *brucei* by restriction enzyme analysis of kinetoplast DNA. *Mol.Biochem.Parasitol.* **3**, 117-131.

Borst, P., Gommers-Ampt, J.H., Ligtenberg, M.J.L., Rudenko, G., Kieft, R., Taylor, M.C., Blundell, P.A. and van Leeuwen, F. (1993) Control of antigenic variation in african trypanosomes. *Cold Spring Harbor Symp.Quant.Biol.* **58**, 105-114.

Borst, P., Gommers-Ampt, J.H., Lutgerink, J., Ouellette, M., ten Asbroek, A.L.M.A. and Zomerdijk, J.C.B. (1990) Control of surface-antigen genes in african trypanosomes. *Biochem.Soc.Trans.* **18**, 741-742.

Borst, P. and Rudenko, G. (1994) Antigenic variation in African trypanosomes. *Science* **264**, 1872-1873.

Borst, P., Rudenko, G., Taylor, M.C., Blundell, P.A., van Leeuwen, F., Bitter, W., Cross, M. and McCulloch, R. (1996) Antigenic variation in trypanosomes. *Arch.Med.Res.* **27**, 379-388.

Borst, P., Van der Ploeg, M., Van Hoek, J.F.M., Tas, J. and James, J. (1982) On the DNA content and ploidy of trypanosomes. *Mol.Biochem.Parasitol.* **6**, 13-23.

Boulton, S.J. and Jackson, S.P. (1998) Components of the Ku-dependent non-homologous end-joining pathway are involved in telomeric length maintenance and telomeric silencing. *EMBO J.* **17**, 1819-1828.

Bringaud, F. and Baltz, T. (1993) Differential regulation of two distinct families of glucose transporter genes in *Trypanosoma brucei*. *Mol.Cell.Biol.* **13**, 1146-1154.

Campbell, D.A., Vanbree, M.P. and Boothroyd, J.C. (1984) The 5'-limit of transposition and upstream barren region of a trypanosome VSG gene - tandem 76 base-pair repeats flanking $(TAA)_{90}$. *Nucleic Acids Res.* **12**, 2759-2774.

Capbern, A., Giroud, C., Baltz, T. and Mattern, P. (1977) *Trypanosoma equiperdum*: étude des variations antigéniques au cours de la trypanosomose expérimentale du lapin. *Exp.Parasitol.* **42**, 6-13.

Carrington, M., Miller, N., Blum, M., Roditi, I., Wiley, D. and Turner, M. (1991) Variant specific glycoprotein of *Trypanosoma brucei* consists of 2 domains each having an independently conserved pattern of cysteine residues. *J.Mol.Biol.* **221**, 823-835.

Chaves, I., Rudenko, G., Dirks-Mulder, A., Cross, M. and Borst, P. (1999) Control of variant surface glycoprotein gene expression sites in *Trypanosoma brucei*. *EMBO J.* **18**, 4846-4855.

Chen, Q.J., Fernandez, V., Sundstrom, A., Schlichtherle, M., Datta, S., Hagblom, P. and Wahlgren, M. (1998) Developmental selection of *var* gene expression in *Plasmodium falciparum*. *Nature* **394**, 392-395.

Chung, H.M.M., Shea, C., Fields, S., Taub, R.N. and Van der Ploeg, L.H.T. (1990) Architectural organization in the interphase nucleus of the protozoan Trypanosoma brucei - location of telomeres and minichromosomes. *EMBO J.* **9**, 2611-2619.

Clayton, C. (1992) Developmental regulation of nuclear gene expression in *Trypanosoma brucei*. *Prog.Nucleic Acid Res.Mol.Biol.* **43**, 37-66.

Cooper, J.P., Watanabe, Y. and Nurse, P. (1998) Fission yeast Taz1 protein is required for meiotic telomere clustering and recombination. *Nature* **392**, 828-831.

Coppens, I., Baudhuin, P., Opperdoes, F.R. and Courtoy, P.J. (1988) Receptors for the host low density lipoproteins on the hemoflagellate *Trypanosoma brucei* - purification and involvement in the growth of the parasite. *Proc.Natl.Acad.Sci.USA* **85**, 6753-6757.

Cornelissen, A.W.C.A., Bakkeren, G.A.M., Barry, J.D., Michels, P.A.M. and Borst, P. (1985b) Characteristics of trypanosome variant antigen genes active in the tsetse fly. *Nucleic Acids Res.* **13**, 4661-4676.

Cornelissen, A.W.C.A., Johnson, P.J., Kooter, J.M., Van der Ploeg, L.H.T. and Borst, P. (1985a) 2 simultaneously active vsg gene-transcription units in a single *Trypanosoma brucei* variant. *Cell* **41**, 825-832.

Critchlow, S.E. and Jackson, S.P. (1998) DNA end-joining: from yeast to man. *Trends In Biochemical Sciences* **23**, 394-398.

Cross, G.A.M. (1984) Release and purification of Trypanosoma brucei variant surface glycoprotein. *J.Cell.Biochem.* **24**, 79-90.

Cross, G.A.M. (1990a) Cellular and genetic aspects of antigenic variation in trypanosomes. *Annu.Rev.Immunol.* **8**, 83-110.

Cross, G.A.M. (1990b) Glycolipid anchoring of plasma membrane proteins. *Annu.Rev.Cell Biol.* **6**, 1-39.

Cross, G.A.M. (1996) Antigenic variation in trypanosomes - secrets surface slowly. *Bioessays* **18** , 283-291.

Cross, G.A.M., Wirtz, L.E. and Navarro, M. (1998) Regulation of *vsg* expression site transcription and switching in *Trypanosoma brucei*. *Mol.Biochem.Parasitol.* **91**, 77-91.

Crowe, J.S., Barry, J.D., Luckins, A.G., Ross, C.A. and Vickerman, K. (1983) All metacyclic variable antigen types of *Trypanosoma congolense* identified using monoclonal antibodies. *Nature* **306**, 389-391.

Cryderman, D.E., Morris, E.J., Biessmann, H., Elgin, S.C.R. and Wallrath, L.L. (1999) Silencing at *Drosophila* telomeres: nuclear organization and chromatin structure play critical roles. *EMBO J.* **18**, 3724-3735.

Cully, D.F., Ip, H.S. and Cross, G.A.M. (1985) Coordinate transcription of variant surface glycoprotein genes and an expression site associated gene family in *Trypanosoma brucei*. *Cell* **42**, 173-182.

Davies, K.P., Carruthers, V.B. and Cross, G.M. (1997) Manipulation of the vsg co-transposed region increases expression- site switching in *Trypanosoma brucei*. *Mol.Biochem.Parasitol.* **86**, 163-177.

de Lange, T., Kooter, J.M., Luirink, J. and Borst, P. (1985) Transcription of a transposed trypanosome surface antigen gene starts upstream of the transposed segment. *EMBO J.* **4**, 3299-3306.

de Lange, T., Kooter, J.M., Michels, P.A.M. and Borst, P. (1983) Telomere conversion in trypanosomes. *Nucleic Acids Res.* **11**, 8149-8165.

Deitsch, K.W., Moxon, E.R. and Wellems, T.E. (1997) Shared themes of antigenic variation and virulence in bacterial, protozoal, and fungal infections. *Microbiol.Mol.Biol.Rev.* **61**, 281-294.

Delauw, M.F., Laurent, M., Paindavoine, P., Aerts, D., Pays, E., Le Ray, D. and Steinert, M. (1987) Characterization of genes coding for 2 major metacyclic surface antigens in *Trypanosoma brucei*. *Mol.Biochem.Parasitol.* **23**, 9-17.

Delauw, M.F., Pays, E., Steinert, M., Aerts, D., Van Meirvenne, N. and Le Ray, D. (1985) Inactivation and reactivation of a variant-specific antigen gene in cyclically transmitted *Trypanosoma brucei*. *EMBO J.* **4**, 989-993.

Dickin, S.K. and Gibson, W.C. (1989) Hybridization with a repetitive DNA probe reveals the presence of small chromosomes in *Trypanosoma vivax*. *Mol.Biochem.Parasitol.* **33**, 135-142.

Donelson, J.E. (1995) Mechanisms of antigenic variation in *Borrelia hermsii* and African trypanosomes. *J.Biol.Chem.* **270**, 7783-7786.

Donelson, J.E., Hill, K.L. and Elsayed, N.A. (1998) Multiple mechanisms of immune evasion by African trypanosomes. *Mol.Biochem.Parasitol.* **91**, 51-66.

Donelson, J.E., Murphy, W.J., Brentano, S.T., Rice-Ficht, A.C. and Cain, G.D. (1983) Comparison of the expression-linked extra copy (ELC) and basic copy (BC) genes of a trypanosome surface antigen. *J.Cell.Biochem.* **23**, 1-12.

Ersfeld, K. and Gull, K. (1997) Partitioning of large and minichromosomes in *Trypanosoma brucei*. *Science* **276**, 611-614.

Esser, K.M. and Schoenbechler, M.J. (1985) Expression of two variant surface glycoproteins on individual African trypanosomes during antigen switching. *Science* **229**, 190-193.

Esser, K.M., Schoenbechler, M.J. and Gingrich, J.B. (1982) *Trypanosoma rhodesiense* blood forms express all antigen specificities relevant to protection against metacyclic (insect form) challenge. *J.Immunol.* **129**, 1715-1718.

Fairlamb, A.H., Carter, N.S., Cunningham, M. and Smith, K. (1992) Characterization of melarsen-resistant *Trypanosoma brucei brucei* with respect to cross-resistance to other drugs and trypanothione metabolism. *Mol.Biochem.Parasitol.* **53**, 213-222.

Ferrante, A. and Allison, A.C. (1983) Alternative pathway activation of complement by African trypanosomes lacking a glycoprotein coat. *Parasite Immunol.* **5**, 491-498.

Fischer, K., Horrocks, P., Preuss, M., Wiesner, J., Wunsch, S., Camargo, A.A. and Lanzer, M. (1997) Expression of *var* genes located within polymorphic subtelomeric domains of *Plasmodium falciparum* chromosomes. *Mol.Cell.Biol.* **17**, 3679-3686.

Florent, I., Baltz, T., Raibaud, A. and Eisen, H. (1987) On the role of repeated sequences 5' to variant surface glycoprotein genes in African trypanosomes. *Gene* **53**, 55-62.

Fourel, G., Revardel, E., Koering, C.E. and Gilson, E. (1999) Cohabitation of insulators and silencing elements in yeast subtelomeric regions. *EMBO J.* **18**, 2522-2537.

Game, J.C. (1993) DNA double-strand breaks and the *rad50-rad57* genes in *Saccharomyces*. *Seminars In Cancer Biology* **4**, 73-83.

Gerasimova, T.I. and Corces, V.G. (1996) Boundary and insulator elements in chromosomes. *Curr.Op.Gen.Dev.* **6**, 185-192.

Gibson, W.C. and Borst, P. (1986) Size-fractionation of the small chromosomes of trypanozoon and nannomonas trypanosomes by pulsed field gradient gel-electrophoresis. *Mol.Biochem.Parasitol.* **18**, 127-140.

Gibson, W.C., Swinkels, B.W. and Borst, P. (1988) Post-transcriptional control of the differential expression of phosphoglycerate kinase genes in *Trypanosoma brucei*. *J.Mol.Biol.* **201**, 315-325.

Gommers-Ampt, J., Lutgerink, J. and Borst, P. (1991) A novel DNA nucleotide in *Trypanosoma brucei* only present in the mammalian phase of the life cycle. *Nucleic Acids Res.* **19**, 1745-1751.

Gommers-Ampt, J.H., van Leeuwen, F., Debeer, A.L.J., Vliegenthart, J.F.G., Dizdaroglu, M., Kowalak, J.A., Crain, P.F. and Borst, P. (1993) β-d-glucosyl-hydroxymethyluracil - a novel modified base present in the DNA of the parasitic protozoan *Trypanosoma brucei*. *Cell* **75**, 1129-1136.

Gottesdiener, K., Chung, H.M., Brown, S.D., Lee, M.G.S. and Van der Ploeg, L.H.T. (1991) Characterization of VSG gene expression site promoters and promoter-associated DNA rearrangement events. *Mol.Cell.Biol.* **11**, 2467-2480.

Gottesdiener, K.M., Goriparthi, L., Masucci, J.P. and Van der Ploeg, L.H.T. (1992) A proposed mechanism for promoter-associated DNA rearrangement events at a variant surface glycoprotein gene expression site. *Mol.Cell.Biol.* **12**, 4784-4795.

Grab, D.J. and Bwayo, J.J. (1982) Isopycnic isolation of African trypanosomes on percoll gradients formed *in situ*. *Acta Trop.* **39**, 363-366.

Grab, D.J., Russo, D., Naessens, J. and Verjee, Y. (1992) Transferrin binding proteins in *Trypanosoma brucei*. *FASEB Journal* **6**, A1900

Graham, S.V. and Barry, J.D. (1991) Expression Site Associated Genes transcribed independently of Variant Surface Glycoprotein genes in *Trypanosoma brucei*. *Mol.Biochem.Parasitol.* **47**, 31-41.

Graham, S.V. and Barry, J.D. (1995) Transcriptional regulation of metacyclic variant surface glycoprotein gene expression during the life cycle of *Trypanosoma brucei*. *Mol.Cell.Biol.* **15**, 5945-5956.

Graham, S.V., Wymer, B. and Barry, J.D. (1998) Activity of a trypanosome metacyclic variant surface glycoprotein gene promoter is dependent upon life cycle stage and chromosomal context. *Mol.Cell.Biol.* **18**, 1137-1146.

Graham, V.S. and Barry, J.D. (1996) Is point mutagenesis a mechanism for antigenic variation in *Trypanosoma brucei*? *Mol.Biochem.Parasitol.* **79**, 35-45.

Gray, A.R. (1965) Antigenic variation in a strain of *Trypanosoma brucei* transmitted by *Glossina morsitans* and *G.palpalis*. *J.Gen.Microbiol* **41**, 195-214.

Greaves, D.R. and Borst, P. (1987) Trypanosoma brucei variant-specific glycoprotein gene chromatin is sensitive to single-strand-specific endonuclease digestion. *J.Mol.Biol.* **197**, 471-483.

Gull, K., Alsford, S. and Ersfeld, K. (1998) Segregation of minichromosomes in trypanosomes: implications for mitotic mechanisms. *Trends Microbiol.* **6**, 319-323.

Haber, J.E. (1998) Mating-type gene switching in *Saccharomyces cerevisiae*. *Annu.Rev.Genetics* **32**, 561-599.

Haber, J.E. (1999a) DNA repair - Gatekeepers of recombination. *Nature* **398**, 665-667.

Haber, J.E. (1999b) Sir-Ku-itous routes to make ends meet. *Cell* **97**, 829-832.

Hajduk, S.L., Cameron, C.R., Barry, J.D. and Vickerman, K. (1981) Antigenic variation in cyclically transmitted *Trypanosoma brucei*. Variable antigen type composition of metacyclic trypanosome populations from the salivary glands of *Glossina morsitans*. *Parasitology* **83**, 595-607.

Henikoff, S. (1997) Nuclear organization and gene expression: homologous pairing and long-range interactions. *Curr.Opin.Cell Biol.* **9**, 388-95

Herbert, W.J. and Lumsden, W.H.R. (1976) *Trypanosoma brucei:* a rapid 'matching' method for estimating the host's parasitaemia. *Exp.Parasitol.* **40**, 427-431.

Hoeijmakers, J.H., Frasch, A.C., Bernards, A., Borst, P. and Cross, G.A.M. (1980) Novel expression-linked copies of the genes for variant surface antigens in trypanosomes. *Nature* **284**, 78-80.

Holder, A.A. and Cross, G.A.M. (1981) Glycopeptides from variant surface glycoproteins of *Trypanosoma brucei* - C-terminal location of antigenically cross-reacting carbohydrate moieties. *Mol.Biochem.Parasitol.* **2**, 135-150.

Horn, D. and Cross, G.M. (1997) Position-dependent and promoter-specific regulation of gene expression in *Trypanosoma brucei*. *EMBO J.* **16**, 7422-7431.

Jeggo, P.A., Taccioli, G.E. and Jackson, S.P. (1995) Ménage-à-trois - double-strand break repair, V(D)J recombination and DNA-PK. *Bioessays* **17**, 949-957.

Johnson, J.G. and Cross, G.A.M. (1979) Selective cleavage of variant surface glycoproteins from *Trypanosoma brucei*. *Biochem.J.* **178**, 689-697.

Johnson, P.J., Kooter, J.M. and Borst, P. (1987) Inactivation of transcription by UV-irradiation of *T brucei* provides evidence for a multicistronic transcription unit including a VSG gene. *Cell* **51**, 273-281.

Kamper, S.M. and Barbet, A.F. (1992) Surface epitope variation via mosaic gene formation is potential key to long-term survival of *Trypanosoma brucei*. *Mol.Biochem.Parasitol.* **53**, 33-44.

Kanaar, R., Hoeijmakers, J.H.J. and Van Gent, D.C. (1998) Molecular mechanisms of DNA double-strand break repair. *Trends.Cell.Biol.* **8**, 483-489.

Kitten, T. and Barbour, A.G. (1990) Juxtaposition of expressed variable antigen genes with a conserved telomere in the bacterium borrelia-hermsii. *Proc.Natl.Acad.Sci.USA* **87**, 6077-6081.

Koomey, M. (1997) Bacterial pathogenesis: A variation on variation in Lyme disease. *Curr.Biol.* **7**, R538-R540

Kooter, J.M. and Borst, P. (1984) Alpha amanitin insensitive transcription of variant surface glycoprotein genes provides further evidence for discontinuous transcription in trypanosomes. *Nucleic Acids Res.* **12**, 9457-9472.

Kooter, J.M., Van der Spek, H.J., Wagter, R., d'Oliveira, C.E., Van der Hoeven, F., Johnson, P.J. and Borst, P. (1987) The anatomy and transcription of a telomeric expression site for variant specific surface antigens in *T. brucei*. *Cell* **51**, 261-272.

Kyes, S.A., Rowe, J.A., Kriek, N. and Newbold, C.I. (1999) Rifins: A second family of clonally variant proteins expressed on the surface of red cells infected with *Plasmodium falciparum*. *Proc.Natl.Acad.Sci.USA* **96**, 9333-9338.

Lamont, G.S., Tucker, R.S. and Cross, G.A.M. (1986) Analysis of antigen switching rates in *Trypanosoma brucei*. *Parasitology* **92**, 355-367.

Lanzer, M., Fischer, K. and Leblancq, S.M. (1995) Parasitism and chromosome dynamics in protozoan parasites - is there a connection. *Mol.Biochem.Parasitol.* **70**, 1-8.

Laroche, T., Martin, S.G., Gotta, M., Gorham, H.C., Pryde, F.E., Louis, E.J. and Gasser, S.M. (1998) Mutation of yeast Ku genes disrupts the subnuclear organization of telomeres. *Curr.Biol.* **8**, 653-656.

Laurenson, P. and Rine, J. (1992) Silencers, silencing, and heritable transcriptional states. *Micr.Rev.* **56**, 543-560.

Laurent, M., Pays, E., Van der Werf, A., Aerts, D., Magnus, E., Van Meirvenne, N. and Steinert, M. (1984) Translocation alters the activation rate of a trypanosome surface- antigen gene. *Nucleic Acids Res.* **12**, 8319-8328.

Lee, M.G.-S. and Van der Ploeg, L.H. (1997) Transcription of protein-coding genes in trypanosomes by RNA polymerase I. *Annu.Rev.Microbiol.* **51**, 463-489.

Lee, M.G. and Van der Ploeg, L.H.T. (1987) Frequent independent duplicative transpositions activate a single VSG gene. *Mol.Cell.Biol.* **7**, 357-364.

Lenardo, M.J., Esser, K.M., Moon, A.M., Van der Ploeg, L.H.T. and Donelson, J.E. (1986) Metacyclic variant surface glycoprotein genes of *Trypanosoma brucei* subsp *rhodesiense* are activated *in situ*, and their expression is transcriptionally regulated. *Mol.Cell.Biol.* **6**, 1991-1997.

Liang, F., Han, M.G., Romanienko, P.J. and Jasin, M. (1998) Homology-directed repair is a major double-strand break repair pathway in mammalian cells. *Proc.Natl.Acad.Sci.USA* **95**, 5172-5177.

Lips, S., Geuskens, M., Paturiaux-Hanocq, F., Hanocq-Quertier, J. and Pays, E. (1996) The esag-8 gene of *Trypanosoma brucei* encodes a nuclear-protein. *Mol.Biochem.Parasitol.* **79**, 113-117.

Liu, A.Y.C., Michels, P.A.M., Bernards, A. and Borst, P. (1985) Trypanosome variant surface glycoprotein genes expressed early in infection. *J.Mol.Biol.* **182**, 383-396.

Liu, A.Y.C., Van der Ploeg, L.H.T., Rijsewijk, F.A.M. and Borst, P. (1983) The transposition unit of variant surface glycoprotein gene 118 of *Trypanosoma brucei* - presence of repeated elements at its border and absence of promoter-associated sequences. *J.Mol.Biol.* **167**, 57-75.

Longacre, S. and Eisen, H. (1986) Expression of whole and hybrid genes in *Trypanosoma equiperdum* antigenic variation. *EMBO J.* **5**, 1057-1063.

Louis, E.J. (1995) The chromosome ends of *Saccharomyces cerevisiae*. *Yeast* **11**, 1553-1573.

Lu, Y., Alarcon, C.M., Hall, T., Reddy, L.V. and Donelson, J.E. (1994) A strand bias occurs in point mutations associated with variant surface glycoprotein gene conversion in *Trypanosoma rhodesiense*. *Mol.Cell.Biol.* **14**, 3971-3980.

Lu, Y., Hall, T., Gay, L.S. and Donelson, J.E. (1993) Point mutations are associated with a gene duplication leading to the bloodstream re-expression of a trypanosome metacyclic VSG. *Cell* **72**, 397-406.

Luo, M.Z. and Cella, R. (1994) A reliable amplification technique with single-sided specificity for the isolation of 5' gene-regulating regions. *Gene* **140**, 59-62.

MacLeod, A. (1999) Genetic analysis of *Trypanosoma brucei*. *Ph.D.thesis, University of Glasgow*

Magowan, C., Wollish, W., Anderson, L. and Leech, J. (1988) Cytoadherence by *Plasmodium falciparum*-infected erythrocytes is correlated with the expression of a family of variable proteins on infected erythrocytes. *J.Exp.Med.* **168**, 1307-1320.

Majumder, H.K., Boothroyd, J.C. and Weber, H. (1981) Homologous 3'-terminal regions of mRNAs for surface antigens of different antigenic variants of *Trypanosoma brucei*. *Nucleic Acids Res.* **9**, 4745-4753.

Marchetti, M.A., Tschudi, C., Silva, E. and Ullu, E. (1998) Physical and transcriptional analysis of the *Trypanosoma brucei* genome reveals a typical eukaryotic arrangement with close interspersion of RNA polymerase. *Nucleic Acids Res.* **26**, 3591-3598.

Martin, S.G., Laroche, T., Suka, N., Grunstein, M. and Gasser, S.M. (1999) Relocalization of telomeric Ku and SIR proteins in response to DNA strand breaks in yeast. *Cell* **97**, 621-633.

Matthews, K.R. and Gull, K. (1994) Evidence for an interplay between cell cycle progression and the initiation of differentiation between life cycle forms of African trypanosomes. *J.Cell Biol.* **125**, 1147-1156.

Matthews, K.R., Shiels, P.G., Graham, S.V., Cowan, C. and Barry, J.D. (1990) Duplicative activation mechanisms of two trypanosome telomeric VSG genes with structurally simple 5' flanks. *Nucleic Acids Res.* **18**, 7219-7227.

Matthyssens, G., Michiels, F., Hamers, R., Pays, E. and Steinert, M. (1981) Two variant surface glycoproteins of *Trypanosoma brucei* have a conserved C-terminus. *Nature* **293**, 230-233.

McCulloch, R. and Barry, J.D. (1999) A role for RAD51 and homologous recombination in *Trypanosoma brucei* antigenic variation. *Genes Dev.* **13**, 2875-2888

McCulloch, R., Rudenko, G. and Borst, P. (1997) Gene conversions mediating antigenic variation in *Trypanosoma brucei* can occur in variant surface glycoprotein expression sites lacking 70 base-pair repeat sequences. *Mol.Cell.Biol.* **17**, 833-843.

McNeillage, G.J.C., Herbert, W.J. and Lumsden, W.H.R. (1969) Antigenic type of first relapse variants arising from a strain of *Trypanosoma* (*Trypanozoon*) *brucei*. *Exp.Parasitol.* **25**, 1-7.

Meyer, T.F., Gibbs, C.P. and Haas, R. (1990) Variation and control of protein expression in *Neisseria*. *Annu.Rev.Microbiol.* **44**, 451-477.

Michels, P.A.M., Liu, A.Y.C., Bernards, A., Sloof, P., Vanderbijl, M.M.W., Schinkel, A.H., Menke, H.H., Borst, P., Veeneman, G.H., Tromp, M.C. and Vanboom, J.H. (1983) Activation of the genes for variant surface glycoprotein-117 and glycoprotein-118 in *Trypanosoma brucei*. *J.Mol.Biol.* **166**, 537-556.

Michiels, F., Matthyssens, G., Kronenberger, P., Pays, E., Dero, B., Van Assel, S., Darville, M., Cravador, A., Steinert, M. and Hamers, R. (1983) Gene activation and re-expression of a *Trypanosoma brucei* variant surface glycoprotein. *EMBO J.* **2**, 1185-1192.

Miller, E.N., Allan, L.M. and Turner, M.J. (1984) Topological analysis of antigenic determinants on a variant surface glycoprotein of *Trypanosoma brucei*. *Mol.Biochem.Parasitol.* **13**, 67-81.

Miller, E.N. and Turner, M.J. (1981) Analysis of antigenic types appearing in first relapse populations of clones of *Trypanosoma brucei*. *Parasitology* **82**, 63-80.

Muñoz-Jordán, J.L., Davies, K.P. and Cross, G.A.M. (1996) Stable expression of mosaic coats of variant surface glycoproteins in *Trypanosoma brucei*. *Science* **272**, 1795-1797.

Murphy, W.J., Watkins, K.P. and Agabian, N. (1986) Identification of a novel Y-branch structure as an intermediate in trypanosome mRNA processing - evidence for *trans* splicing. *Cell* **47**, 517-525.

Myler, P.J., Aline, R.F., Scholler, J.K. and Stuart, K.D. (1988b) Changes in telomere length associated with antigenic variation in *Trypanosoma brucei*. *Mol.Biochem.Parasitol.* **29**, 243-250.

Myler, P.J., Aline, R.F., Scholler, J.K. and Stuart, K.D. (1988a) Multiple events associated with antigenic switching in *Trypanosoma brucei*. *Mol.Biochem.Parasitol.* **29**, 227-241.

Myler, P.J., Allen, A.L., Agabian, N. and Stuart, K. (1985) Antigenic variation in clones of Trypanosoma brucei grown in immune-deficient mice. *Infect.Immun.* **47**, 684-690.

Myler, P.J., Allison, J., Agabian, N. and Stuart, K.D. (1984) Antigenic variation in African trypanosomes by gene replacement or activation of alternate telomeres. *Cell* **39**, 203-211.

Myler, P.J., Audleman, L., Devos, T., Hixson, G., Kiser, P., Lemley, C., Magness, C., Rickel, E., Sisk, E., Sunkin, S., Swartzell, S., Westlake, T., Bastien, P., Fu, G.L., Ivens, A. and Stuart, K. (1999) *Leishmania major* Friedlin chromosome 1 has an unusual distribution of protein-coding genes. *Proc.Natl.Acad.Sci.USA* **96**, 2902-2906.

Nantulya, V.M., Musoke, A.J. and Moloo, S.K. (1986) Apparent exhaustion of the variable antigen repertoires of *Trypanosoma vivax* in infected cattle. *Infect.Immun.* **54**, 444-447.

Nantulya, V.M., Musoke, A.J., Rurangirwa, F.R. and Moloo, S.K. (1984) Resistance of cattle to tsetse-transmitted challenge with *Trypanosoma brucei* and *Trypanosoma congolense* after spontaneous recovery from syringe-passaged infections. *Infect.Immun.* **43**, 735-738.

Nash, T.E. (1997) Antigenic variation in *Giardia lamblia* and the host's immune response. *Phil.Trans.Roy.Soc.B* **352**, 1369-1375.

Navarro, M. and Cross, G.A.M. (1996) DNA rearrangements associated with multiple consecutive directed antigenic switches in *Trypanosoma brucei*. *Mol.Cell.Biol.* **16**, 3615-3625.

Navarro, M., Cross, G.A.M. and Wirtz, E. (1999) Trypanosoma brucei variant surface glycoprotein regulation involves coupled activation/inactivation and chromatin remodeling of expression sites. *EMBO J.* **18**, 2265-2272.

Navarro, M. and Cross, G.M. (1998) In situ analysis of a variant surface glycoprotein expression site promoter region in *Trypanosoma brucei*. *Mol.Biochem.Parasitol.* **94**, 53-66.

Newbold, C.I. (1999) Antigenic variation in *Plasmodium falciparum*: mechanisms and consequences. *Curr.Op.Microbiol.* **2**, 420-425.

Nilsen, T.W. (1995) trans-splicing - an update. *Mol.Biochem.Parasitol.* **73**, 1-6.

Nimmo, E.R., Pidoux, A.L., Perry, P.E. and Allshire, R.C. (1998) Defective meiosis in telomere-silencing mutants of *Schizosaccharomyces pombe*. *Nature* **392**, 825-828.

Olafson, R.W., Clarke, M.W., Kielland, S.L., Pearson, T.W., Barbet, A.F. and McGuire, T.C. (1984) Amino-terminal sequence homology among variant surface glycoproteins of African trypanosomes. *Mol.Biochem.Parasitol.* **12**, 287-298.

Onyango, R.J., Van Hoeve, K. and De Raadt, P. (1966) The epidemiology of Trypanosoma rhodesiense sleeping sickness in Alego location, Central Nyanza, Kenya. I. Evidence that cattle may act as reservoir hosts of trypanosomes infective to man. *Trans.R.Soc.Trop.Med.Hyg.* **60**, 175-182.

Overath, P., Chaudhri, M., Steverding, D. and Ziegelbauer, K. (1994) Invariant surface proteins in bloodstream forms of *Trypanosoma brucei*. *Parasitol.Today* **10**, 53-58.

Paques, F. and Haber, J.E. (1999) Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. *Microbiol.Mol.Biol.Rev.* **63**, 349

Pays, E. (1985) Gene conversion in trypanosome antigenic variation. *Prog.Nucleic Acid Res.Mol.Biol.* **32**, 1-26.

Pays, E., Coquelet, H., Pays, A., Tebabi, P. and Steinert, M. (1989a) *Trypanosoma brucei* - posttranscriptional control of the variable surface glycoprotein gene expression site. *Mol.Cell.Biol.* **9**, 4018-4021.

Pays, E., Delauw, M.F., Laurent, M. and Steinert, M. (1984) Possible DNA modification in GC dinucleotides of *Trypanosoma brucei* telomeric sequences - relationship with antigen gene transcription. *Nucleic Acids Res.* **12**, 5235-5247.

Pays, E., Delauw, M.F., Van Assel, S., Laurent, M., Vervoort, T., Van Meirvenne, N. and Steinert, M. (1983a) Modifications of a *Trypanosoma b. brucei* antigen gene repertoire by different DNA recombinational mechanisms. *Cell* **35**, 721-731.

Pays, E., Guyaux, M., Aerts, D., Van Meirvenne, N. and Steinert (1985b) Telomeric reciprocal recombination as a possible mechanism for antigenic variation in trypanosomes. *Nature* **316**, 562-564.

Pays, E., Houard, S., Pays, A., Van Assel, S., Dupont, F., Aerts, D., Huetduvillier, G., Gomes, V., Richet, C., Degand, P., Van Meirvenne, N. and Steinert, M. (1985a) *Trypanosoma brucei* - the extent of conversion in antigen genes may be related to the DNA coding specificity. *Cell* **42**, 821-829.

Pays, E., Laurent, M., Delinte, K., Van Meirvenne, N. and Steinert, M. (1983b) Differential size variations between transcriptionally active and inactive telomeres of *Trypanosoma brucei*. *Nucleic Acids Res.* **11**, 8137-8147.

Pays, E., Tebabi, P., Pays, A., Coquelet, H., Revelard, P., Salmon, D. and Steinert, M. (1989b) The genes and transcripts of an antigen gene expression site from *T. brucei*. *Cell* **57**, 835-845.

Pays, E., Van Assel, S., Laurent, M., Darville, M., Vervoort, T., Van Meirvenne, N. and Steinert, M. (1983c) Gene conversion as a mechanism for antigenic variation in trypanosomes. *Cell* **34**, 371-381.

Pays, E., Van Assel, S., Laurent, M., Dero, B., Michiels, F., Kronenberger, P., Matthyssens, G., Van Meirvenne, N., Le Ray, D. and Steinert, M. (1983d) At least two transposed sequences are associated in the expression site of a surface antigen gene in different trypanosome clones. *Cell* **34**, 359-369.

Pays, E., Vanhamme, L. and Berberof, M. (1994) Genetic controls for the expression of surface antigens in African trypanosomes. *Annu.Rev.Microbiol.* **48**, 25-52.

Plasterk, R.H.A., Simon, M.I. and Barbour, A.G. (1985) Transposition of structural genes to an expression sequence on a linear plasmid causes antigenic variation in the bacterium *Borrelia hermsii*. *Nature* **318**, 257-263.

Pryde, F.E. and Louis, E.J. (1999) Limitations of silencing at native yeast telomeres. *EMBO J.* **18**, 2538-2550.

Revelard, P., Lips, S. and Pays, E. (1993) Alternative splicing within and between alleles of the ATPase gene 1 locus of *Trypanosoma brucei*. *Mol.Biochem.Parasitol.* **62**, 93-101.

Rice-Ficht, A.C., Chen, K.K. and Donelson, J.E. (1981) Sequence homologies near the C-termini of the variable surface glycoproteins of *Trypanosoma brucei*. *Nature* **294**, 53-57.

Rijkers, T., Van den Ouweland, J., Morolli, B., Rolink, A.G., Baarends, W.M., VanSloun, P.P.H., Lohman, P.H.M. and Pastink, A. (1998) Targeted inactivation of mouse *RAD52* reduces homologous recombination but not resistance to ionizing radiation. *Mol.Cell.Biol.* **18**, 6423-6429.

Roditi, I. and Pearson, T.W. (1990) The procyclin coat of African trypanosomes (or the not-so-naked trypanosome). *Parasitol.Today* **6**, 79-82.

Roth, C., Bringaud, F., Layden, R.E., Baltz, T. and Eisen, H. (1989) Active late-appearing variable surface antigen genes in *Trypanosoma equiperdum* are constructed entirely from pseudogenes. *Proc.Natl.Acad.Sci.USA* **86**, 9375-9379.

Rudenko, G., Blundell, P.A., Dirksmulder, A., Kieft, R. and Borst, P. (1995) A ribosomal DNA promoter replacing the promoter of a telomeric VSG gene expression site can be efficiently switched on and off in *Trypanosoma brucei*. *Cell* **83**, 547-553.

Rudenko, G., Blundell, P.A., Taylor, M.C., Kieft, R. and Borst, P. (1994) VSG gene expression site control in insect form *Trypanosoma brucei*. *EMBO J.* **13**, 5470-5482.

Rudenko, G., McCulloch, R., Dirksmulder, A. and Borst, P. (1996) Telomere exchange can be an important mechanism of variant surface glycoprotein gene switching in *Trypanosoma brucei*. *Mol.Biochem.Parasitol.* **80**, 65-75.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular cloning: a laboratory manual. 2 edn, Cold Spring Harbor Laboratory Press.

Scherf, A., Hernandez-Rivas, R., Buffet, P., Bottius, E., Benatar, C., Pouvelle, B., Gysin, J. and Lanzer, M. (1998) Antigenic variation in malaria: in situ switching, relaxed and mutually exclusive transcription of *var* genes during intra-erythrocytic development in *Plasmodium falciparum*. *EMBO J.* **17**, 5418-5426.

Seifert, H.S. (1996) Questions about gonococcal pilus phase- and antigenic variation. *Mol.Microbiol.* **21**, 433-440.

Shah, J.S., Young, J.R., Kimmel, B.E., Iams, K.P. and Williams, R.O. (1987) The 5' flanking sequence of a *Trypanosoma brucei* variable surface glycoprotein gene. *Mol.Biochem.Parasitol.* **24**, 163-174.

Shore, D. (1998) Perspectives: Cell biology - Telomeres - Unsticky ends. *Science* **281**, 1818-1819.

Sloof, P., Menke, H.H., Caspers, M.P.M. and Borst, P. (1983) Size fractionation of *Trypanosoma brucei* DNA - localization of the 177-bp repeat satellite DNA and a variant surface glycoprotein gene in a mini-chromosomal DNA fraction. *Nucleic Acids Res.* **11**, 3889-3901.

Steverding, D., Stierhof, Y.D., Fuchs, H., Tauber, R. and Overath, P. (1995) Transferrin-binding protein complex is the receptor for transferrin uptake in *Trypanosoma brucei*. *J.Cell Biol.* **131**, 1173-1182.

Su, X.Z., Heatwole, V.M., Wertheimer, S.P., Guinet, F., Herrfeldt, J.A., Peterson, D.S., Ravetch, J.A. and Wellems, T.E. (1995) The large diverse gene family *var* encodes proteins involved in cytoadherence and antigenic variation of *Plasmodium falciparum*-infected erythrocytes. *Cell* **82**, 89-100.

Tait, A., Barry, J.D., Wink, R., Sanderson, A. and Crowe, J.S. (1985) Enzyme variation in *T. brucei* ssp. II. Evidence for *T. b. rhodesiense* being a set of variants of *T. b. brucei*. *Parasitology* **90**, 89-100.

Takata, M., Sasaki, M.S., Sonoda, E., Morrison, C., Hashimoto, M., Utsumi, H., Yamaguchi-Iwai, Y., Shinohara, A. and Takeda, S. (1998) Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. *EMBO J.* **17**, 5497-5508.

Tetley, L., Turner, C.M.R., Barry, J.D., Crowe, J.S. and Vickerman, K. (1987) Onset of expression of the variant surface glycoproteins of *Trypanosoma brucei* in the tsetse fly studied using immunoelectron microscopy. *J.Cell Sci.* **87**, 363-372.

Thon, G., Baltz, T. and Eisen, H. (1989) Antigenic diversity by the recombination of pseudogenes. *Genes Dev.* **3**, 1247-1254.

Thon, G., Baltz, T., Giroud, C. and Eisen, H. (1990) Trypanosome variable surface glycoproteins: composite genes and order of expression. *Genes Dev.* **9**, 1374-1383.

Timmers, H.T.M., de Lange, T., Kooter, J.M. and Borst, P. (1987) Coincident multiple activations of the same surface antigen gene in *Trypanosoma brucei*. *J.Mol.Biol.* **194**, 81-90.

Tsukamoto, Y., Kato, J. and Ikeda, H. (1997) Silencing factors participate in DNA repair and recombination in *Saccharomyces cerevisiae*. *Nature* **388**, 900-903.

Turner, C.M., Barry, J.D. and Vickerman, K. (1986) Independent expression of the metacyclic and bloodstream variable antigen repertoires of *Trypanosoma brucei rhodesiense*. *Parasitology* **92**, 67-73.

Turner, C.M.R. (1997) The rate of antigenic variation in fly-transmitted and syringe-passaged infections of *Trypanosoma brucei*. *FEMS Microbiol.Lett.* **153**, 227-231.

Turner, C.M.R. and Barry, J.D. (1989) High frequency of antigenic variation in *Trypanosoma brucei rhodesiense* infections. *Parasitology* **99**, 67-75.

Ullu, E., Matthews, K.R. and Tschudi, C. (1993) Temporal order of RNA processing reactions in trypanosomes - rapid *trans* splicing precedes polyadenylation of newly synthesized tubulin transcripts. *Mol.Cell.Biol.* **13**, 720-725.

Van Den Abbeele, J., Claes, Y., Van Bockstaele, D., Le Ray, D. and Coosemans, M. (1999) *Trypanosoma brucei* spp. development in the tsetse fly: characterization of the post-mesocyclic stages in the foregut and proboscis. *Parasitology* **118**, 469-478.

Van der Ploeg, L.H.T., Cornelissen, A.W.C.A., Barry, J.D. and Borst, P. (1985) Chromosomes of Kinetoplastida. *EMBO J.* **3**, 3109-3115.

Van der Ploeg, L.H.T., Liu, A.Y.C. and Borst, P. (1984a) Structure of the growing telomeres of trypanosomes. *Cell* **36**, 459-468.

Van der Ploeg, L.H.T., Schwartz, D.C., Cantor, C.R. and Borst, P. (1984b) Antigenic variation in *Trypanosoma brucei* analyzed by electrophoretic separation of chromosome sized DNA molecules. *Cell* **37**, 77-84.

Van der Ploeg, L.H.T., Valerio, D., de Lange, T., Bernards, A., Borst, P. and Grosveld, F.G. (1982) An analysis of cosmid clones of nuclear DNA from *Trypanosoma brucei* shows that the genes for variant surface glycoproteins are clustered in the genome. *Nucleic Acids Res.* **10**, 5905-5923.

Van der Werf, A., Van Assel, S., Aerts, D., Steinert, M. and Pays, E. (1990) Telomere interactions may condition the programming of antigen expression in *Trypanosoma brucei*. *EMBO J.* **9**, 1035-1040.

Van Dyck, E., Stasiak, A.Z., Stasiak, A. and West, S.C. (1999) Binding of double-strand breaks in DNA by human Rad52 protein. *Nature* **398**, 728-731.

van Leeuwen, F., Wijsman, E.R., Kieft, R., vanderMarel, G.A., Vanboom, J.H. and Borst, P. (1997) Localization of the modified base J in telomeric VSG gene expression sites of *Trypanosoma brucei*. *Genes Dev.* **11**, 3232-3241.

van Leeuwen, F., Wijsman, E.R., Kuyl-Yeheskiely, E., van der Marel, G.A., Van Boom, J.H. and Borst, P. (1996) The telomeric GGGTAA repeats of *Trypanosoma brucei* contain the hypermodified base J in both strands. *Nucleic Acids Res.* **24**, 2476-2482.

Van Meirvenne, N., Janssens, P.G. and Magnus, E. (1975) Antigenic variation in syringe passaged populations of *Trypanosoma (Trypanozoon) brucei*. I. Rationalization of the experimental approach. *Annales De La Societe Belge De Medecine Tropicale* **55**, 1-23.

Van Valen, L.M. (1973) A new evolutionary law. *Evolutionary Theory* **1**, 1-30.

Vanhamme, L. and Pays, E. (1995) Control of gene expression in trypanosomes. *Micr.Rev.* **59**, 223-240.

Vanhamme, L. and Pays, E. (1998) Controls of the expression of the VSG in *Trypanosoma brucei*. *Mol.Biochem.Parasitol.* **91**, 107-116.

Vassella, E. and Boshart, M. (1996) High-molecular mass agarose matrix supports growth of bloodstream forms of pleomorphic *Trypanosoma brucei* strains in axenic culture. *Mol.Biochem.Parasitol.* **82**, 91-105.

Vassella, E., Braun, R. and Roditi, I. (1994) Control of polyadenylation and alternative splicing of transcripts from adjacent genes in a procyclin expression site - a dual role for polypyrimidine tracts in trypanosomes. *Nucleic Acids Res.* **22**, 1359-1364.

Vickerman, K. (1985) Developmental cycles and biology of pathogenic trypanosomes. *British Medical Bulletin* **41**, 105-114.

Wahlgren, M. and Bejarano, M.T. (1999) Malaria - A blueprint of 'bad air'. *Nature* **400**, 506-507.

Webster, P. and Russel, D.G. (1993) The flagellar pocket of trypanosomatids. *Parasitol.Today* **9**, 201-206.

Weiden, M., Osheim, Y.N., Beyer, A.L. and Van der Ploeg, L.H.T. (1991) Chromosome structure - DNA nucleotide sequence elements of a subset of the minichromosomes of the protozoan *Trypanosoma brucei*. *Mol.Cell.Biol.* **11**, 3823-3834.

Williams, R.O., Young, J.R. and Majiwa, P.A.O. (1982) Genomic environment of *T. brucei* VSG genes - presence of a minichromosome. *Nature* **299**, 417-421.

Williams, R.O., Young, J.R., Majiwa, P.A.O., Doyle, J.J. and Shapiro, S.Z. (1980) Contextural genomic rearrangements of variable antigen genes in *Trypanosoma brucei*. *Cold Spring Harbor Symp.Quant.Biol.* **45**, 945-949.

Xong, H.V., Vanhamme, L., Chamekh, M., Chimfwembe, C.E., Van Den Abbeele, J., Pays, A., Van Meirvenne, N., Hamers, R., Debaetselier, P. and Pays, E. (1998) A VSG expression site-associated gene confers resistance to human serum in *Trypanosoma rhodesiense*. *Cell* **95**, 839-846.

Xu, Y.N., Kodner, C., Coleman, L. and Johnson, R.C. (1996) Correlation of plasmids with infectivity of *Borrelia burgdorferi sensu stricto* type strain B31. *Infect.Immun.* **64**, 3870-3876.

Yamaguchi-Iwai, Y., Sonoda, E., Buerstedde, J.M., Bezzubova, O., Morrison, C., Takata, M., Shinohara, A. and Takeda, S. (1998) Homologous recombination, but not DNA repair, is reduced in vertebrate cells deficient in RAD52. *Mol.Cell.Biol.* **18**, 6430-6435.

Young, J.R., Shah, J.S., Matthyssens, G. and Williams, R.O. (1983) Relationship between multiple copies of a *T. brucei* variable surface glycoprotein gene whose expression is not controlled by duplication. *Cell* **32**, 1149-1159.

Zhang, J. and Williams, N. (1997) Purification, cloning, and expression of two closely related *Trypanosoma brucei* nucleic acid binding proteins. *Mol.Biochem.Parasitol.* **87**, 145-158.

Zhang, J.R., Hardham, J.M., Barbour, A.G. and Norris, S.J. (1997) Antigenic variation in Lyme disease borreliae by promiscuous recombination of VMP-like sequence cassettes. *Cell* **89**, 275-285.

Ziegelbauer, K. and Overath, P. (1990) Surface antigen change during differentiation of *Trypanosoma brucei*. *Biochem.Soc.Trans.* **18**, 731-733.

Zomerdijk, J.C.B.M., Kieft, R. and Borst, P. (1992) A ribosomal RNA gene promoter at the telomere of a minichromosome in *Trypanosoma brucei*. *Nucleic Acids Res.* **20**, 2725-2734.

Zomerdijk, J.C.B.M., Ouellette, M., ten Asbroek, A.L.M.A., Kieft, R., Bommer, A.M.M., Clayton, C.E. and Borst, P. (1990) The promoter for a variant surface glycoprotein gene expression site in *Trypanosoma brucei*. *EMBO J.* **9**, 2791-2801.