DNA Rearrangements Involving the Genes for Variant Antigens in *Trypanosoma brucei*

P. Borst, A. C. C. Frasch,* A. Bernards, L. H. T. Van der Ploeg, J. H. J. Hoeijmakers,† A. C. Arnberg,‡ and G. A. M. Cross§

Section for Medical Enzymology and Molecular Biology, Laboratory of Biochemistry, University of Amsterdam, Jan Swammerdam Institute, 1005 GA Amsterdam, The Netherlands; ‡ Biochemical Laboratory, State University Groningen, 9747 AG Groningen, The Netherlands; and § Department of Immunochemistry, The Wellcome Research Laboratories, Beckenham, Kent BR3 3BS, England

Trypanosomes are unicellular, flagellate protozoa responsible for some of the most widespread parasitic diseases of humans and domestic animals (Hoare 1972). In itself, this would explain the increasing popularity of trypanosomes as research objects. However, trypanosomes also have a number of intriguing attributes that are of scientific interest: (1) The mitochondrial DNA of trypanosomes is the most bizarre DNA in nature; it consists of a large network of catenated circles, usually called kinetoplast DNA (see Borst and Hoeijmakers 1979). (2) Trypanosome glycolysis is unique since it is organized in a separate organelle, the glycosome (see Opperdoes and Borst 1977). (3) Trypanosomes can completely suppress mitochondrial biogenesis (Hoare 1972), a property they share only with yeast. (4) Although trypanosomes are primitive eukaryotes with much less genetic complexity than that of *Escherichia coli*, they share some of the typical aspects of eukaryotic genome organization, i.e., simple sequence satellites and a large fraction of intermediate repetitive DNA (Borst et al. 1980b). (5) Finally, the African trypanosomes are able to evade host antibodies by antigenic variation; this unusual property is the subject of this paper.

Several aspects of antigenic variation are now well understood (see Gray and Luckins 1976; Cross 1978; Vickerman 1978). *Trypanosoma brucei* is completely covered with a thick surface coat (Vickerman 1969), and this coat is the only part of the trypanosome seen by host antibodies. By producing a repertoire of coats that are antigenically distinct, the parasite escapes immunodestruction. Cross and his co-workers have provided strong evidence that the surface coat consists of a single glycoprotein, called variant surface glycoprotein or VSG (Cross 1978). VSGs from different cloned variants of *T. brucei* have molecular masses in the 50–60 kilodalton (kD) range but differ in amino acid composition and sequence. Probably, the N-terminal half of the protein makes up the outer layer of the surface coat, whereas the C terminal is associated with the cell membrane and carries a carbohydrate side chain (Cross and Johnson 1976; Johnson and Cross 1979; Holder and Cross 1981). One trypanosome clone can make at least 100 different VSGs (Capbern et al. 1977; Cross 1978).

The ability of trypanosomes to produce such an impressive repertoire of VSGs raises two related questions: (1) What is the origin of VSG diversity? Does the trypanosome have more than 100 genes for VSG, or is some kind of localized hypermutagenesis or scrambling mechanism involved (cf. immunoglobulin genes)? (2) How does the trypanosome switch from one VSG to the next? Is this controlled at the level of transcription initiation, transcript processing, or translation?

Recombinant DNA technology provides an obvious approach to answering these questions. The results obtained thus far are summarized in this paper.

Experimental Approach and Characterization of Complementary DNA Clones

We have cloned DNA complementary to the messenger RNAs (mRNAs) for four VSGs in *E. coli* and used these complementary DNA (cDNA) clones to study VSG genes and their expression. The four variants used—called 117, 118, 121, and 221—all arose from a single cloned stock of *T. brucei* strain 427. The variants were cloned and subcloned in irradiated mice, and large batches of each variant were grown in rats. Such rodent-adapted, syringe-passaged strains are so virulent that they kill the host in a few days, before antibodies can arise. This makes it possible to get a large batch of trypanosomes in which more than 99% express a single VSG. Total poly(A)$^+$ RNA was isolated from each variant and DNA complementary to this RNA was joined to plasmid pBR322 and cloned by standard techniques. We selected the clones containing RNA complementary to VSG mRNA by differential hybridization. Since antibodies against the protein part of any of the four VSGs do not crossreact with the other three, we reasoned that the mRNAs for these VSGs would have little homology either. Moreover, since all four variants come from the same trypanosome and are grown under identical conditions, the major, if not only, difference in the mRNA populations isolated from these four variants should be in the mRNA for VSG. Hence, cDNA clones only hybridizing with homologous RNA (from the same variant) and not with heterologous RNA (from the other three variants) should contain DNA complementary to VSG mRNA. This turned out to be the case. We have

* Permanant address: Instituto de Química Biológica, Facultad de Medicina, Universidad de Buenos Aires, Paraguay 2155 (5° piso), Buenos Aires, Argentina; † Present address: Laboratory of Medical Microbiology, University of Amsterdam, Mauritskade 57A, 1092 AD Amsterdam, The Netherlands.

Table 1. Apparent Sizes of VSG mRNA and Cloned cDNA

Variant	mRNA (nucleotides)	cDNA	
		longest insert (bp)	total sequence in cDNA (bp)
117	2250	1575	1750
118	2050	1450	1600
121	2150	1150	1150
221	1950	1200	1800

The apparent sizes of the mRNAs were determined from experiments like that shown in Fig. 1. The apparent sizes of cDNA inserts were estimated from their mobility in agarose gels relative to markers. The sizes given include dG-dC tails of unknown length (but probably smaller than 30 bp). The total cDNA sequence present in our clone collection was calculated from maps of overlapping cDNA clones. Data are from Hoeijmakers et al. (1980a) and unpubl.

checked 31 of such homo-specific clones, and in each case the clone contained DNA complementary to VSG mRNA.

None of our clones contain complete cDNA copies of the VSG mRNAs, and several lack the 3′ end of the mRNA. We attribute this to technical problems in the isolation of intact mRNA and in the construction and cloning of full-length cDNA. By using more than one clone as a probe, however, we can detect sequences complementary to most of the mRNA with three of the four variants (Table 1). Details of this work have been published by Hoeijmakers et al. (1980a).

There Is No Major Sequence Homology between the Four Different VSG mRNAs Studied, and the Expression of These VSG Genes Is Not Controlled at the Translation Level

Our ability to isolate cDNA by differential hybridization already indicated that the mRNAs for different VSGs do not have much sequence homology. This is now supported by two types of experiments: (1) There is no detectable cross-hybridization between the VSG cDNAs of different variants at about T_m−30 °C (Hoeijmakers et al. 1980a). (2) The cDNA of each variant only hybridizes to homologous RNA, as shown for the 117 cDNA in Figure 1.

The result in Figure 1 also shows that the control of VSG gene expression is not at the translational level, because no latent VSG mRNA can be detected by hybridization in a variant that does not express the corresponding VSG. Since we do not see precursors of VSG mRNAs in these blots either, even after long exposures, the expression of VSG genes appears to be primarily controlled at transcription initiation.

Activation of a VSG Gene Is Accompanied by the Appearance of a Novel Expression-linked Copy of That Gene

When the cDNAs of these four VSG mRNAs are hybridized with restriction digests of the nuclear DNA of one variant, each cDNA hybridizes with a separate set of bands. Figure 2 shows that the 117 cDNA probe sees a large series of bands, derived from a family of related genes (see below). The 118 cDNA probe mainly sees one gene, and the 121 cDNA probe sees three to four genes (but in this blot, they give the same bands in most digests). As expected from the preceding section, the four different VSG gene families have no obvious bands in common. We can now ask whether the expression of a VSG gene is accompanied by rearrangements in DNA, by hybridizing one cDNA with restriction digests of the nuclear DNAs from the four variants. The results for two cDNAs are illustrated in Figure 3. The 117 VSG cDNA hybridizes with the same fragments in all nuclear DNAs; the only exception to this is one extra band found in the *Msp*I digest of homologous (117) nuclear DNA. Such extra bands in some digests of homologous nuclear DNA have also been found with the 118 (Fig. 3)

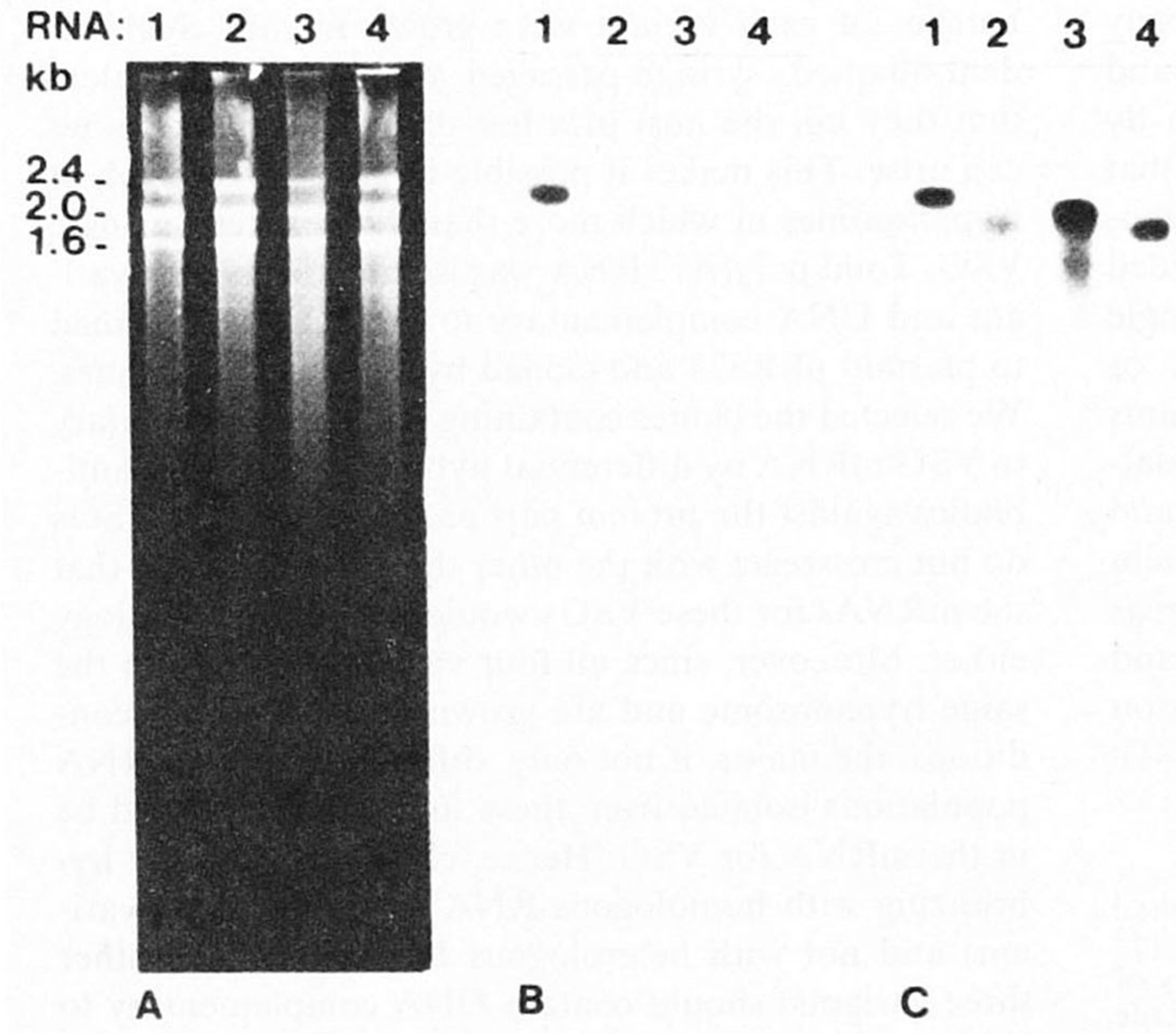

Figure 1. Hybridization of cloned cDNA with poly(A)$^+$ RNA from the homologous variant and other variants. Glyoxal-treated poly(A)$^+$ RNA (2 μg) from each of the four trypanosome variants was electrophoresed through a 1.75% agarose gel in Tris-borate buffer (Van Ommen et al. 1979). After transfer of the RNA to diazobenzyloxymethyl paper (Alwine et al. 1977), the filters were hybridized with nick-translated recombinant plasmid DNA containing sequences complementary to VSG mRNA (Hoeijmakers et al. 1980a). (*A*) Photograph of the ethidium-stained gel after electrophoresis: (lane *1*) variant 117 RNA; (lane *2*) 118 RNA; (lane *3*) 121 RNA; (lane *4*) 221 RNA. The large ribosomal RNA contains an internal break and yields two bands. (*B*) Autoradiograph of the RNA in *A* after transfer to diazobenzyloxymethyl paper and hybridization with a recombinant plasmid containing DNA complementary to VSG 117 mRNA (TcV117-1). (*C*) As *B*, but each RNA hybridized with the homologous cDNA probe: (lane *1*) TcV117-1; (lane *2*) TcV118-2; (lane *3*) TcV121-3; (lane *4*) TcV221-12. Reprinted, with permission, from Hoeijmakers et al. (1980b).

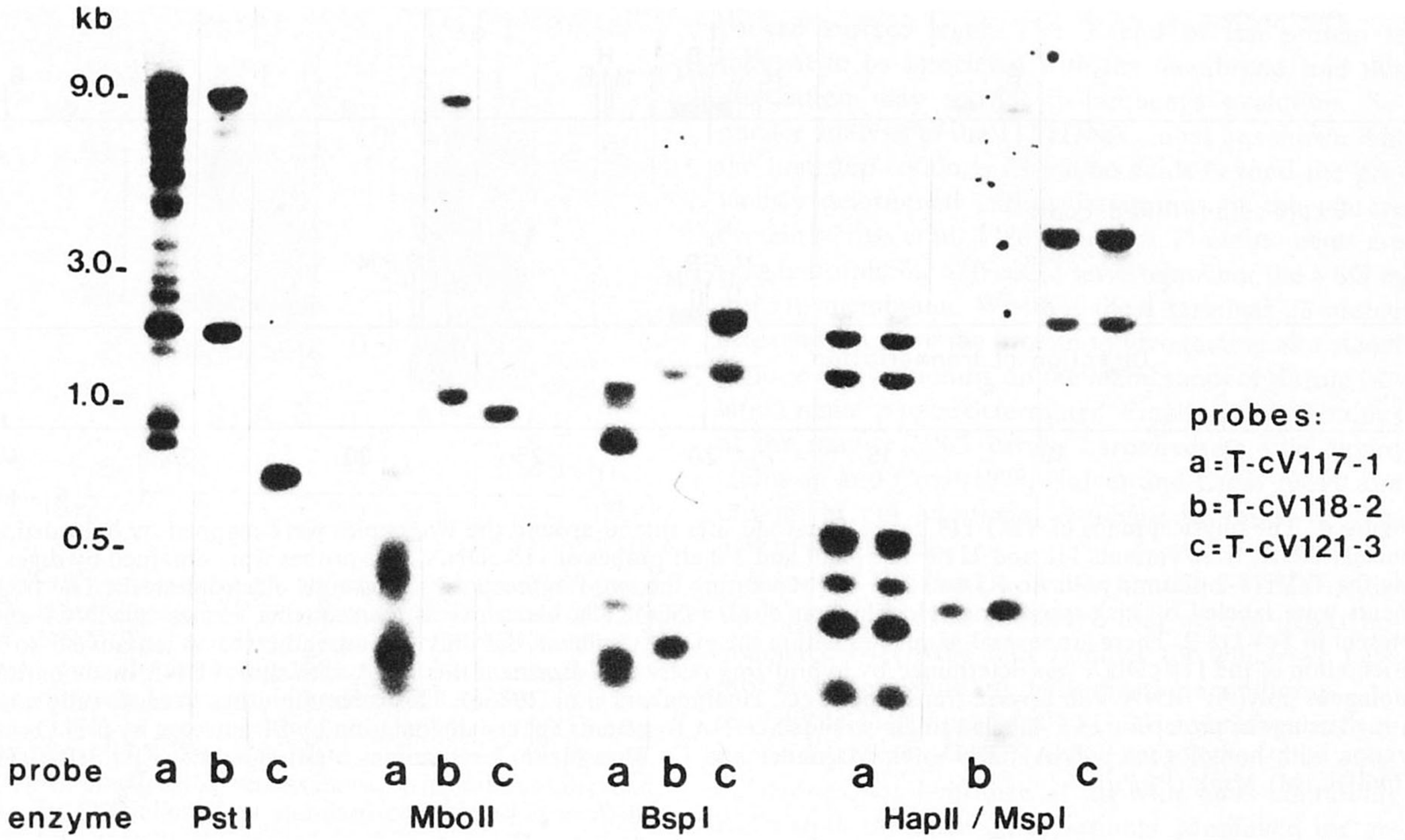

Figure 2. Autoradiograph of variant-specific cDNA probes hybridized to restriction enzyme digests of nuclear DNA from *T. brucei* variant 118. Nuclear DNA was isolated by a modification of the kinetoplast DNA isolation procedure (Borst and Fase-Fowler 1979), digested under standard conditions (Borst and Fase-Fowler 1979) with the enzyme indicated, and electrophoresed through a 0.6% agarose gel (Borst and Fase-Fowler 1979). The gel was blotted onto nitrocellulose filter strips as described elsewhere (Hoeijmakers et al. 1980a) and hybridized with nick-translated DNA from the following recombinant plasmids: (*a*) TcV117-1; (*b*) TcV118-2; (*c*) TcV121-3. See legend to Fig. 1 for further details. Reprinted, with permission, from Hoeijmakers et al. (1980a).

and 121 cDNAs. The fact that each of these three probes sees the same set of bands in the nuclear DNA from all four variants strongly indicates that each variant contains the whole repertoire of VSG genes and that antigenic variation is based on differential expression of preexistent genes rather than on the creation of new genes by DNA rearrangements or local hypermutagenesis. We attribute the presence of extra bands in homologous nuclear DNA to a novel expression-linked copy (ELC) of the VSG gene.

The VSG 118 gene family is rather simple, and, under stringent hybridization conditions, all fragments seen by the 118 cDNA probes in heterologous nuclear DNA can be attributed to a single gene, the basic copy of the 118 VSG gene.

It was therefore relatively easy in this case to con-

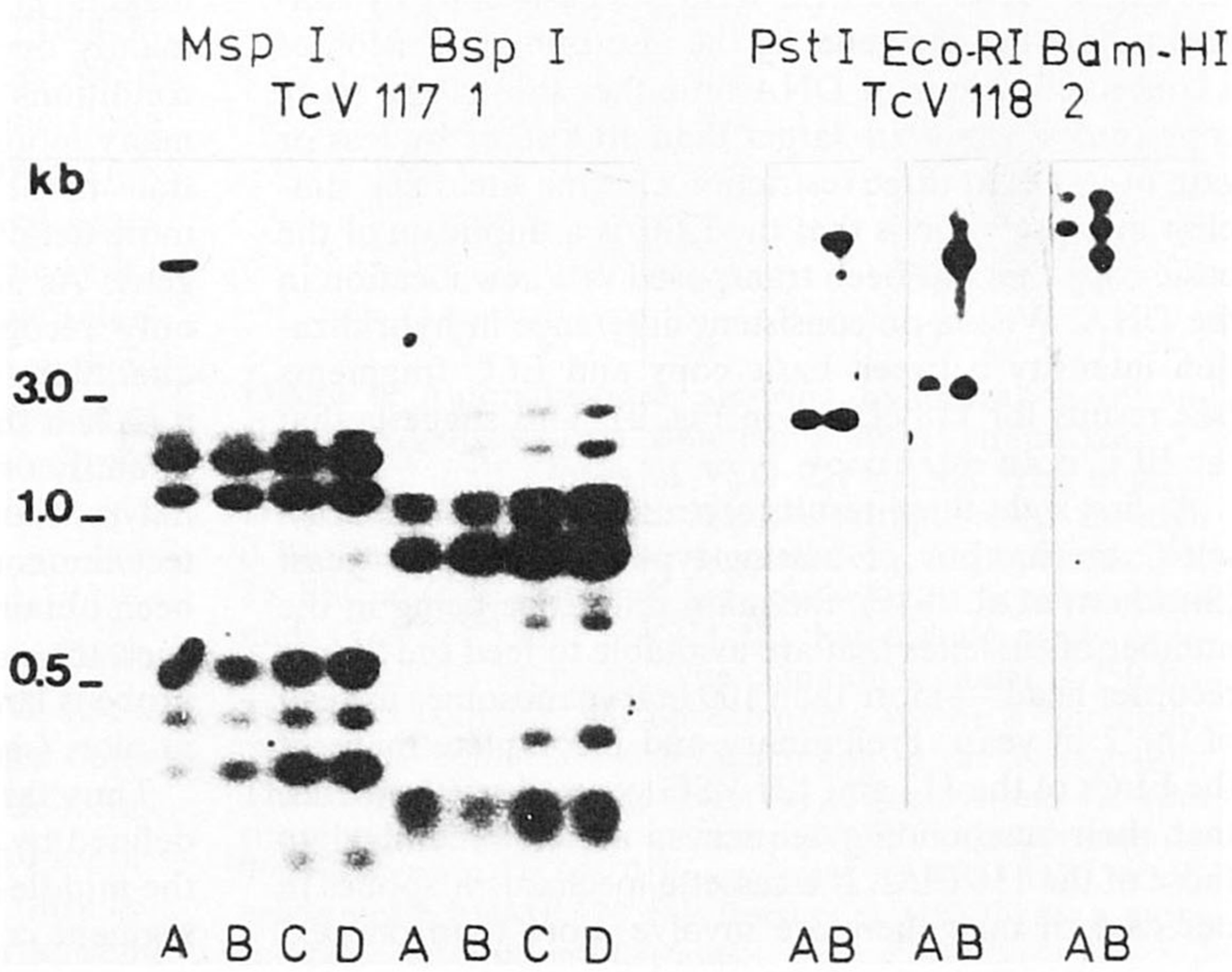

Figure 3. Autoradiograph showing hybridization of VSG-specific cDNA probes with homologous and heterologous nuclear DNAs. Nuclear DNA digests from the two variants indicated were electrophoresed through a 1.2% agarose gel, blotted, and hybridized with the cDNA probes indicated. See legend to Fig. 2 for further details.

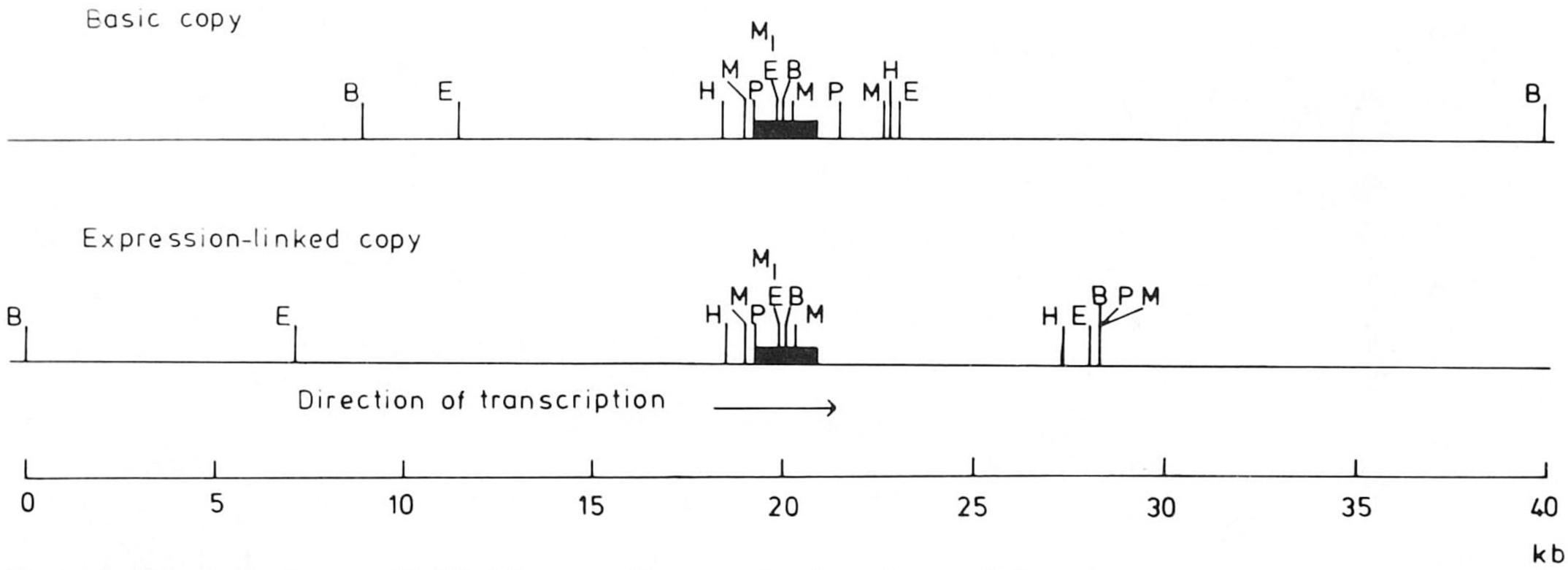

Figure 4. The physical maps of VSG 118 genes. Cleavage sites in and around the two copies were mapped by hybridizing nuclear DNAs from variants 118 and 221 with 5′-half and 3′-half probes of 118 cDNA. The probes were obtained by digesting the TcV118-2 plasmid with *Eco*RI and *Pst*I and separating the two fragments by agarose gel electrophoresis. The fragments were labeled by nick-translation (Hoeijmakers et al. 1980a). The black box is the sequence cloned as cDNA and present in TcV118-2. There are several *Msp*I sites within the cDNA segment, but only two are indicated. A tentative 5′ to 3′ orientation of the 118 cDNA was determined by hybridizing restriction digests of this DNA with short cDNA made on homologous poly(A)$^+$ RNA with reverse transcriptase (cf. Hoeijmakers et al. 1980a). This orientation has been directly confirmed using the protection of 5′-labeled single-stranded cDNA fragments against degradation by S1 nuclease by prehybridization with homologous poly(A)$^+$ RNA (H. Majunder and C. Weissmann, pers. comm.). (B) *Bam*HI; (E) *Eco*RI; (H) *Hin*dIII; (M) *Msp*I; (P) *Pst*I.

struct physical maps of the basic copy and the ELC. The comparison of the two maps presented in Figure 4 leads to the following conclusions: (1) The relative positions of the restriction sites in the region covered by the cDNA and 1 kb in the 5′ direction are identical. Since the cDNA is only about 500 nucleotides shorter than the 118 VSG mRNA (Table 1), this indicates that the basic copy and ELC can code for the complete leader sequence of this mRNA. The nearest site at the 3′ end of the gene that has been found to differ in basic copy and ELC is the *Pst*I site 300 nucleotides from the end of the region corresponding to the cDNA. This site could be outside the gene, and therefore both genes may contain the complete mRNA sequence. (2) The sequence on either side of the ELC differs from the corresponding sequence in the basic copy. This difference is such that one cannot derive the ELC from the basic copy by simple inversion of the gene, by the insertion or excision of a continuous block of DNA on either side of the basic copy (unless these are larger than 10 kb), or by loss or gain of less than three restriction enzyme sites. The simplest interpretation is that the ELC is a duplicate of the basic copy that has been transposed to a new location in the DNA. We see no consistent difference in hybridization intensity between basic copy and ELC fragments (see results for 118 cDNA in Fig. 3). This suggests that the ELC is an extra copy.

At first sight these results are reminiscent of the "cassette" mechanism of mating-type switching in yeast (Strathern et al. 1979), the main difference being in the number of cassettes that are available to feed the "tape-recorder head"—more than 100 in trypanosomes instead of the 2 in yeast. Preliminary and incomplete maps of the ELCs of the 117 and 121 VSG genes show, however, that their surrounding sequences are not identical to those of the 118 ELC. If a cassette mechanism applies in our case, it may therefore involve more than one expression site. Having a separate expression site for each gene family could even provide a mechanism to prevent two members of the same family from being expressed in sequence. It is not excluded, however, that the insertion of VSG gene copies into the expression site and/or their inactivation/removal could lead to sequence alterations around a single expression site.

Organization of VSG Genes in Families

Under standard hybridization conditions, the 117 cDNA hybridizes to a large number of bands of heterologous nuclear DNA in Southern blots (see Fig. 2), suggesting that it recognizes a family of related genes. By raising the stringency of the hybridization, the family is reduced to a single gene, as mentioned before (Hoeijmakers et al. 1980b). With the 118 cDNA probes, mainly one gene is seen under standard hybridization conditions (Fig. 2), but when the stringency is relaxed, many more fragments become visible. The family relations in the 118 VSG gene family have been analyzed in more detail with probes for the 5′ half and 3′ half of the gene. As shown in Figure 5, the family members are only recognized by the 3′-half gene probe. From a quantitative comparison of hybridization experiments, it is clear that this result is not due to the fact that we see a family of genes with the same 5′ half and diverged 3′ halves. We rather see 3′ halves without being able to detect the corresponding 5′ half. Analogous results have been obtained with the 221 family; in this family, however, the multiplicity of genes recognized by the 3′-half probe is large and involves more than 50 separate bands in blots (see below).

Thus far, 3′ halves and 5′ halves have mainly been defined by the use of cDNA probes cut more or less in the middle at a convenient restriction site. The repeated segment could therefore be smaller than the 3′ half of

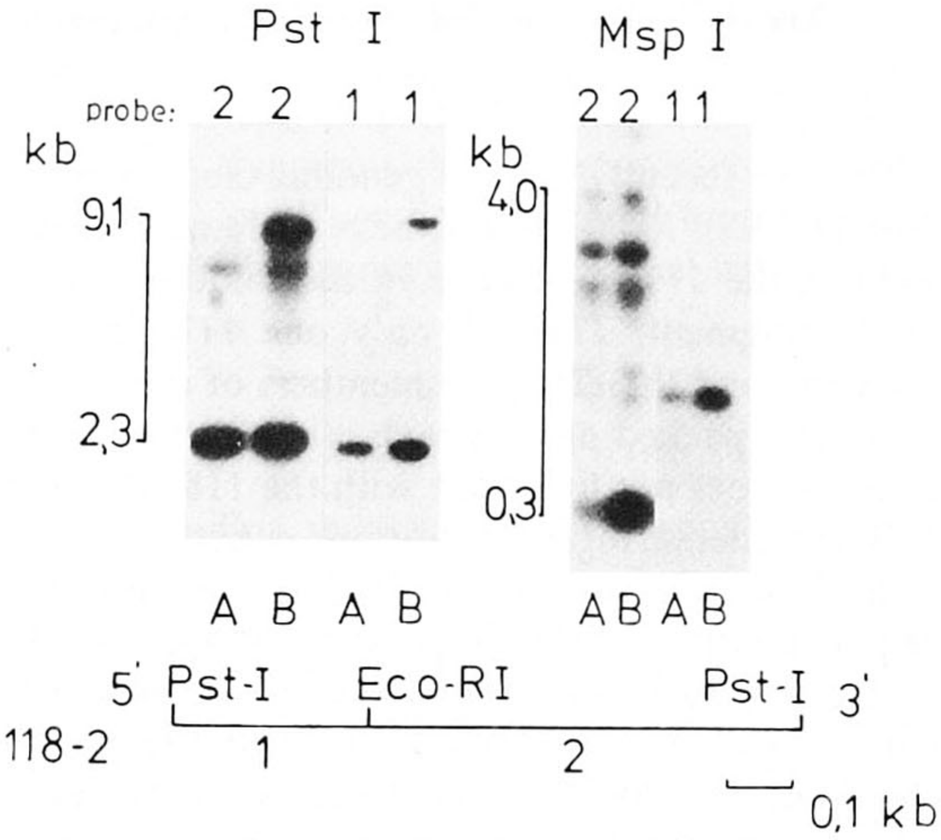

Figure 5. Autoradiograph showing hybridization of nuclear DNA digests with 3′-half and 5′-half probes of VSG 118 cDNA. The nuclear digests ([*A*] 117 DNA; [*B*] 118 DNA) were electrophoresed through 0.7% agarose, blotted onto nitrocellulose, and hybridized in 3× SSC (0.45 M NaCl, 0.045 M sodium citrate, at pH 7.0) at 65 °C in the presence of 10% dextran sulfate as described (Wahl et al. 1979). In practice, this results in less stringent hybridization than 3× SSC alone (our standard conditions) even if the blots are exhaustively washed at 65 °C in 3× SSC without dextran sulfate. The labeled half-gene probes were obtained as described in the legend to Fig. 4.

the gene. We find, however, that a 117 cDNA plasmid that does not contain the DNA segment corresponding to the last 29 amino acids and the 3′ trailing region of this gene still hybridizes with the repeated segment of the 117 family. This segment is therefore not limited to the 3′ end of this gene family.

We do not yet know whether the VSG gene segments that are detected by 3′-half probes at low stringency are part of complete and functional VSG genes. Our evidence on this comes mainly from results with the 117 family. In this family, we see about 20 to 30 fragments with our 3′-half probe under standard conditions. Some of these fragments are also seen with the 5′-half probe, but this hybridization is more easily lost in stringent washes than hybridization with the 3′-half probe. Analogous results have been obtained with a genomic clone of a 117-related gene in the phage λ vector Charon 21A. We infer from this that most of the VSG-related genes detected only with 3′-half probes are complete VSG genes. This leads us to the following speculative interpretation: The present VSG gene families have evolved by gene duplication followed by sequence divergence of genes by mutations. The 5′ halves of the genes evolve faster than the 3′ halves, and hence, 3′-half probes still see family members that are not recognized any more by the 5′-half probes.

This speculative interpretation agrees rather well with the limited information available on the VSG itself. The N-terminal half of the protein forms the outer layer of the surface coat (Cross and Johnson 1976; Johnson and Cross 1979; Holder and Cross 1981), and it presumably is allowed to evolve rapidly since the only restraint on this segment is the ability to form a well-packed surface layer. The 3′ end of the protein is thought to be associated with the membrane, and this association may restrict its sequence evolution. Sequence analysis of the 117 cDNA clones has shown that the first stop codon is 23 amino acids beyond the previously determined carboxyl terminus of the mature protein (Cross et al. 1980). The last 15 amino acids are very hydrophobic and could serve to anchor the VSG in the cell membrane. Whether these terminal 23 amino acids are cut from the protein in vivo (acting as a signal peptide for positioning on the membrane) or during isolation remains to be determined. Finally, the C terminus of the mature VSG carries carbohydrate side chains (Johnson and Cross 1979; Holder and Cross 1981), and this might put additional sequence restraints on that part of the protein.

It should be noted, however, that the 3′ half of the VSG genes is not so conserved that we can detect homology between the four families for which we have probes under standard hybridization conditions. Although the 3′-half probe of VSG 221 hybridizes with an impressive array of bands in blots of nuclear DNA (Fig. 6), it does not hybridize at all with blots containing RNA from the three other variants. Moreover, the genomic clone that contains the VSG gene corresponding to our 117 cDNA does not see significantly more bands in nuclear DNA blots than 117 cDNA itself. This shows

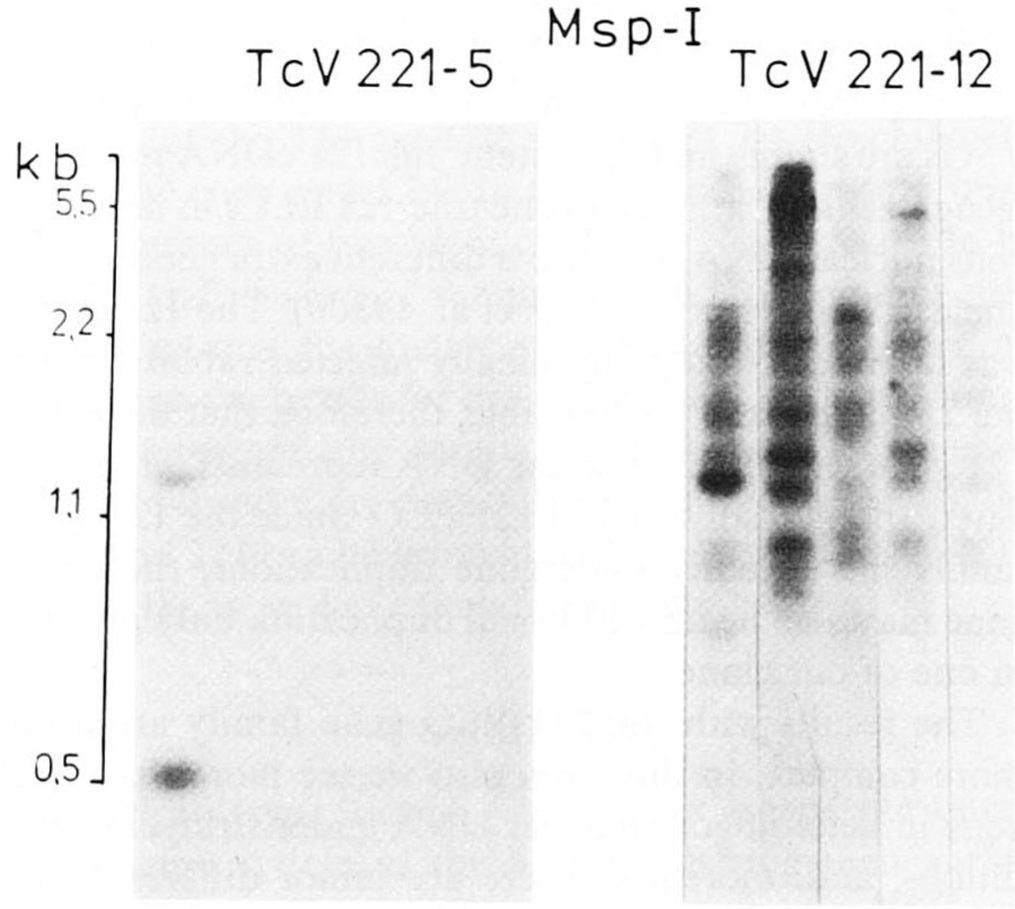

Figure 6. Autoradiograph showing hybridization of nuclear DNA digests from different *T. brucei* strains with 3′-half and 5′-half probes of VSG 221 cDNA. The nuclear DNA was digested with *Msp*I, electrophoresed through 0.7% agarose, blotted onto nitrocellulose, and hybridized in 3× SSC at 65 °C. The half-gene probes used were two plasmids (TcV221-5 and TcV221-12) that contain partly overlapping cDNA segments. The common segment is 750 bp; TcV221-5 contains 390 bp extending in the 5′ direction; TcV221-12 contains 470 bp extending in the 3′ direction. The *T. brucei* strains used were: (*a*) *T. brucei* strain 427 (the strain used in all other experiments); (*b*) *T. equiperdum;* (*c*) *T. brucei* strain 31; (*d*) *T. brucei* strain EATRO 1125; (*e*) *T. brucei* strain LUMP 127. See Borst et al. (1980b) for a more complete description of strains.

that a VSG gene plus surrounding sequences only hybridizes with fragments of the homologous family in nuclear DNA blots. Therefore, there are no elements common to all VSG genes that can be detected under standard hybridization conditions. However, in recent experiments we have observed weak cross-hybridization when the stringency of the hybridization conditions is further relaxed. It is possible that this reflects the faded afterimage of the ancestral gene from which these VSG genes have evolved, but the specificity of this hybridization remains to be proved.

In conclusion, we think that the interpretation given here fits our data and is at present the most plausible one; however, other interpretations are not ruled out. *T. brucei* may be a graveyard full of fossil genes; there may be a collection of unlinked 3′-half and 5′-half genes. The further analysis of the genomic clones now available should settle this.

121 and 221 Genes

The results for the 121 VSG gene family differ in two respects from those obtained for the 117 and 118 families: (1) Whereas in the 117 and 118 families we can distinguish between the homologous basic copy and related genes by stringent hybridization, we see multiple genes (three or four) in the 121 family under stringent conditions. In this case, moreover, the 5′-half probe sees more fragments in blots than the 3′-half probe. We attribute this to duplications of the 121 VSG gene that have occurred recently in the evolution of *T. brucei.* (2) Whereas the results on the ELCs of the 117 and 118 VSGs are simple and consistent, the 121 cDNA not only recognizes an extra VSG gene (the 121 ELC) in homologous nuclear DNA, but also a different extra gene in 221 nuclear DNA (Hoeijmakers et al. 1980b). The 121 clone was obtained from a chronically infected rabbit before the 221 clone, and it is possible, therefore, that the extra 121 gene seen in 221 nuclear DNA represents an inactivated and partially altered 121 ELC. Since the 121 gene family has recently undergone duplications, the extra gene may also be an additional duplication only present in one of our clones.

The results with the 221 VSG gene family are even more complex. In this case, also we see more than one gene in heterologous nuclear DNA under stringent conditions, and, moreover, there are minor differences in the genes found in different nonexpressor variants. In the homologous nuclear DNA, these genes appear to have been replaced by a single different gene. A possible interpretation of these results is that the basic 221 gene is unstable and that the homologous nuclear DNA only contains a 221 ELC, whereas the basic copy has been lost. Evidence for this instability comes from the experiments on the evolution of VSG genes reported below.

The complications encountered with the 121 and 221 VSG genes may explain why Williams and co-workers (Williams et al. 1979 and this volume) have not found an ELC for the single VSG gene studied by them but instead have observed VSG gene rearrangements unrelated to expression.

Arrangement of VSG Genes in Nuclear DNA: Introns

We have recently made a genomic clone bank of nuclear DNA in λgtWES and have isolated a clone that contains the 117 basic gene in the middle of a 8.2-kb *Eco*RI fragment. There is only one 117 gene in this fragment, and therefore the members of this family are not tightly packed next to each other in the DNA. The fragment does not hybridize with the 118, 121, and 221 cDNAs under stringent conditions either.

An R loop made by hybridizing the cloned DNA fragment with homologous poly(A)$^+$ RNA from the 117 variant is shown in Figure 7. The hybrid part of this R loop is continuous, and there is no evidence for introns. The single-stranded RNA tail hanging out on one side is at the 3′ side of the gene, as determined by side-specific half-gene probes on DNA blots of the cloned fragment. The size of the hybrid part in 17 molecules, judged to contain intact RNA, was 1560 ± 83 bp (mean ± S. D.); the size of the tail was 310 ± 80 nucleotides. This tail is unusually long for a poly(A) tail, and part of this tail may be the 3′ trailing region of the mRNA displaced from the hybrid by the homologous DNA strand in this R loop. We cannot exclude, however, that the terminal part of this 3′ trailing region is not present in the basic copy and added onto the gene in the ELC. We note that this genomic clone was isolated from 118 nuclear DNA. This provides additional evidence that the 117 basic

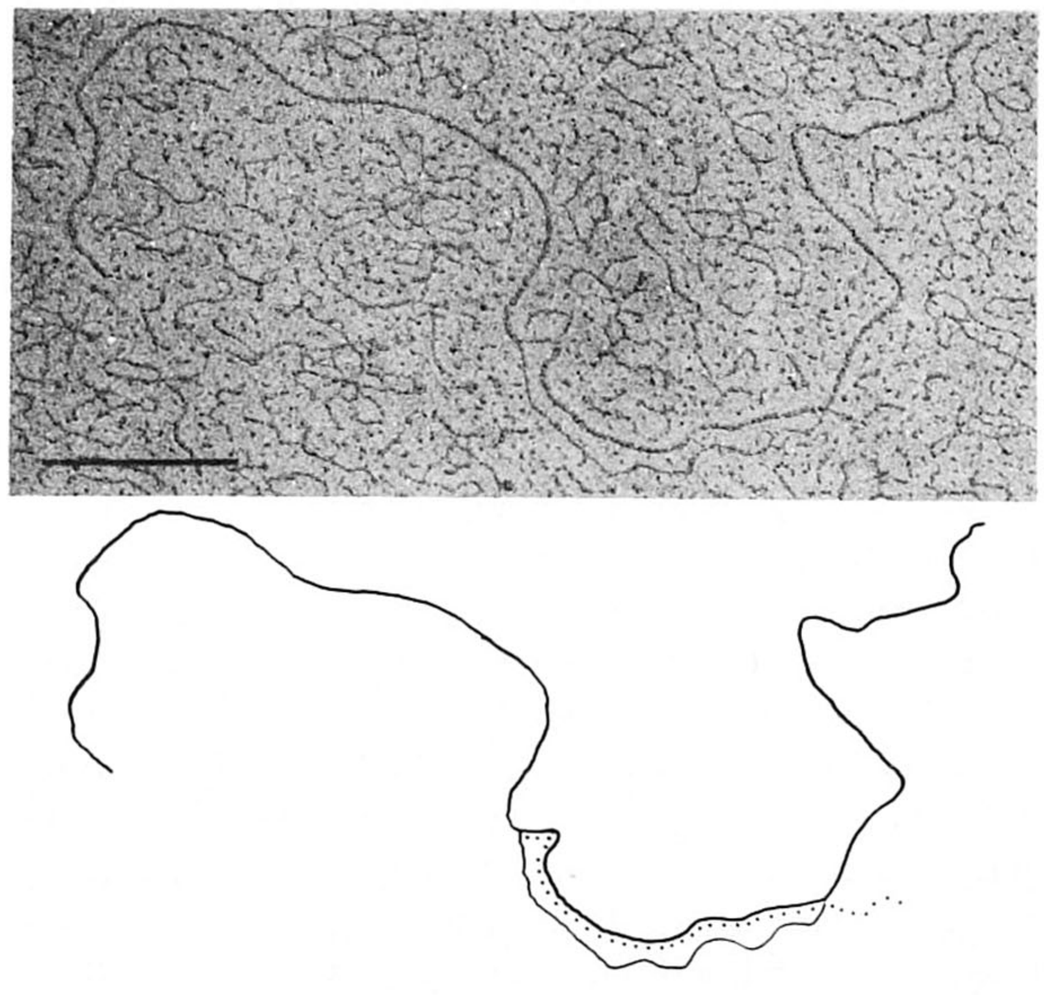

Figure 7. Electron micrograph showing an R loop of VSG 117 mRNA and its homologous basic copy gene. The cloned 8.2-kb *Eco*RI fragment of variant 118 DNA, which contains the 117 gene, was mixed with poly(A)$^+$ RNA from variant 117 and incubated in buffered 70% formamide, 0.1 M PIPES (pH 7.8), 10 mM EDTA, 0.4 M NaCl. Incubation was started at 55 °C and the temperature was lowered at a rate of 4 °C per hr (until 45 °C). The sample was then chilled in ice and spread in a solution containing 68% formamide, 60 mM Tris-HCl, 5.4 mM EDTA, 40 mM PIPES, 0.16 M NaCl, 0.004% cytochrome *c* on H_2O. See Thomas et al. (1976) and Arnberg et al (1980) for further details. The bar is 0.3 μm.

copy is present unaltered in the nuclear DNA of another variant.

With all four cDNA probes, we have found that all internal cDNA fragments checked are present as such in nuclear DNA blots. Together with the R loops, these results show that VSG genes are not riddled with large introns.

Evolution of VSG Genes

In the preceding sections we have argued that each trypanosome has a finite number of complete VSG genes at its disposal, that these genes are grouped in sequence-related families, that families have arisen from an ancestral sequence by duplication followed by mutations, and that the 5′ half of VSG genes evolves faster than the 3′ half. This raises at least two additional questions: (1) Are new VSG genes still being generated by duplication? If so, one would expect to see loss of genes as well, because in a small genome of fixed size, the gene budget should more or less balance. (2) How fast is the evolution of VSG genes? Is there any evidence for a mechanism that would increase the rate of evolution of the 5′ half of the gene beyond that set by the mutagenic events that involve all genes?

To answer these questions, we have looked at our VSG gene families in other stocks of *T. brucei* and *Trypanosoma evansi* (a trypanosome so closely related to *T. brucei* that it is now considered a variant of *T. brucei* rather than a separate species). An example of these experiments is given in Figure 6, which shows that the 5′ half of the 221 gene has been completely lost in all other *T. brucei* strains, whereas the multitude of 3′ halves of this large family is more or less conserved. The other extreme result is obtained with the small 121 family. In this case, our blots show largely the same set of bands for all strains tested. The 118 gene family represents an intermediate case. Here, the complete gene, as seen by the 5′-half probe, is present in some strains and lost in others. However, all strains contain segments of the 3′-half family, and some of these are not detectably altered. The restriction sites in and around the genes of the 117 family are also rather conserved, as judged from conservation of fragments.

We draw two tentative conclusions from these results: (1) The process of gene duplication is still actively continuing in present-day trypanosomes. In a genome of (presumably) constant size, this must be accompanied by loss of genes, and this is probably what we have observed. The multiple 121 genes that we cannot resolve by melting may represent a case of a recent duplication, as mentioned above. The fact that we find such large-scale deletions with two of the four genes studied is significant, because the strains used are very closely related, as indicated by a study (Borst et al. 1980a) of restriction-site polymorphisms in maxicircle DNA, the trypanosome analog of mitochondrial DNA in other organisms. (2) These limited results do not provide evidence for a mutagenic process that actively alters 5′ halves of genes, because we find extensive conservation of restriction sites in and around the genes that are still present in strains other than the *T. brucei* 427 stock. This implies that the time over which the VSG gene families have evolved is much longer than the time that has elapsed since the different strains studied here arose from a common ancestor.

Obviously, these conclusions are tentative, and much more work is required to define the process involved in VSG gene evolution more precisely.

Complications and Caveats

This paper presents work in progress, and we have stressed the clear results rather than the puzzling results (also obtained); we have given the most plausible interpretations available rather than exhaustive lists of alternatives. It should be realized, however, that the system is complex and that we have only scratched the surface.

A brief list of complications and caveats comprises the following points: (1) There is no proof that the ELC is the (only) copy that is being transcribed into mRNA for VSG. (2) There is no proof that the production of an ELC is the only way for the trypanosome to activate a VSG gene. (3) We have titrated the number of basic copies of the 118 gene and find about one per haploid genome. Our blots give no evidence, however, for two alleles of the basic copy of the 117 or 118 gene, although the results for the 121 and 221 genes could be in part explained by heteroalleles. Analysis of the DNA complement/trypanosome in conjunction with our renaturation studies (Borst et al. 1980b) indicate that trypanosomes are diploid (unpubl.). If there are two basic copies of the 118 gene, the apparent 1:1 stoichiometry of basic copy and ELC in the 118 variant would imply the presence of two identical ELCs. This is hard to believe, and in view of the unreliability of some of these quantitations, the matter must be considered open. (4) Whereas the results for the 117 and 118 genes are simple and consistent, the results for 121 and 221 are more complex, and more data will be required to see whether 121 and 221 can be brought under one heading with 117 and 118, as we have tentatively done here.

SUMMARY

African trypanosomes are able to escape the immunodefense system of the host by altering their surface coat. The coat is made up of a single protein, VSG, and by the sequential expression of a large repertoire of radically different VSGs, the trypanosome keeps altering its coat. We have studied whether or not gene rearrangements are involved in this process, and for this purpose we have isolated recombinant plasmids containing DNA complementary to the mRNAs for four different VSGs expressed in the course of a chronic infection of a rabbit with *T. brucei* strain 427. Each plasmid only hybridizes to the mRNA from the homologous antigenic variant, showing that the four VSG mRNAs have no homology detectable by hybridization with the segments cloned as cDNA (more than 50% of each mRNA). Moreover, this experiment shows that the ex-

pression of VSG genes is not controlled at the translational level.

Each cDNA plasmid hybridizes to a series of DNA bands in restriction digests of nuclear DNA. The number of bands increases when the hybridization stringency is relaxed; 5′-half gene probes see fewer bands than 3′-half gene probes, and detection of most of these bands requires a lower stringency than with the 3′-half probes. We conclude that VSG genes have evolved by gene duplication/divergence and that the 5′ half evolves faster than the 3′ half. This makes sense because the N terminus of the protein forms the outer layer of the surface coat and is seen by host antibodies, whereas the C terminus is associated with the cell membrane.

Hybridization of these cDNA plasmids with restriction digests of the DNAs from a series of *Trypanosoma* strains that are closely related to *T. brucei* strain 427 shows deletions of (probably) complete VSG genes in two of the gene families. We interpret this as part of a process of gene duplications and deletions that lead to a continuing further evolution of the VSG repertoire at constant genome size. Restriction sites in and around the genes that are still present are largely conserved; this argues against a process inducing localized hypermutagenesis of the 5′ half of VSG genes.

Nuclear DNA digests of two of these variants give identical fragment patterns when hybridized with a single VSG cDNA, with one exception: Each of the VSG cDNAs sees extra DNA fragments in the homologous nuclear DNA (i.e., the variant expressing that VSG) that are absent in heterologous nuclear DNA. We conclude that the whole repertoire of VSG genes is present as such in each variant but that activation of a gene leads to the appearance of an additional copy of that gene, the ELC. We present physical maps of the ELC and the corresponding basic copy of one VSG gene. These maps show that the sequences on both sides of the ELC differ from those of the basic copy. The simplest model to account for our data is that VSG gene activation is accompanied by the duplication of this gene and its transposition to an expression site. There is no evidence as yet that the copy in the expression site (the ELC) is the one transcribed.

ACKNOWLEDGMENTS

We thank Professor C. Weissmann (Institut für Molekularbiologie, Universität Zürich, Switzerland) for his help in the isolation of cDNA plasmids and Miss J. Van den Burg for technical assistance. This work was supported in part by a grant to P. B. from the Foundation for Fundamental Biological Research (BION), which is subsidized by The Netherlands Organization for the Advancement of Pure Research (ZWO). Additional support was provided to A. C. C. F. by a World Health Organization Research Training Grant funded by the UNDP World Bank WHO Special Program for Research and Training in Tropical Diseases (no. M8/181/4/F.78).

REFERENCES

ALWINE, J. C., D. J. KEMP, and G. R. STARK. 1977. Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. *Proc. Natl. Acad. Sci.* **74:** 5350.

ARNBERG, A. C., G. J. B. VAN OMMEN, L. A. GRIVELL, E. F. J. VAN BRUGGEN, and P. BORST. 1980. Some yeast mitochondrial RNAs are circular. *Cell* **19:** 313.

BORST, P. and F. FASE-FOWLER. 1979. The maxi-circle of *Trypanosoma brucei* kinetoplast DNA. *Biochim. Biophys. Acta* **565:** 1.

BORST, P. and J. H. J. HOEIJMAKERS. 1979. Kinetoplast DNA. *Plasmid* **2:** 20.

BORST, P., F. FASE-FOWLER, J. H. J. HOEIJMAKERS, and A. C. C. FRASCH. 1980a. Variations in maxi-circle and mini-circle sequences in kinetoplast DNAs from different *Trypanosoma brucei* strains. *Biochim. Biophys. Acta* **610:** 197.

BORST, P., F. FASE-FOWLER, A. C. C. FRASCH, J. H. J. HOEIJMAKERS, and P. J. WEIJERS. 1980b. Characterization of DNA from *Trypanosoma brucei* and related trypanosomes by restriction endonuclease digestion. *Mol. Biochem. Parasitol.* **1:** 221.

CAPBERN, A., C. GIROUD, T. BALTZ, and P. MATTERN. 1977. *Trypanosoma equiperdum:* Etude des variations antigéniques au cours de la trypanosome expérimentale du lapin. *Exp. Parasitol.* **42:** 6.

CROSS, G. A. M. 1978. Antigenic variation in trypanosomes. *Proc. R. Soc. Lond. B* **202:** 55.

CROSS, G. A. M. and J. G. JOHNSON. 1976. Structure and organisation of the variant-specific surface antigens of *Trypanosoma brucei.* In *Biochemistry of parasites and host-parasite relationships* (ed. H. Van den Bossche), p. 413. North-Holland, Amsterdam.

CROSS, G. A. M., A. A. HOLDER, G. ALLEN, and J. C. BOOTHROYD. 1980. An introduction to antigenic variation in trypanosomes. *Am. J. Trop. Med. Hyg.* (Suppl.) **29:** 1027.

GRAY, A. R. and A. G. LUCKINS. 1976. Antigenic variation in salivarian trypanosomes. In *Biology of the kinetoplastida* (ed. W. H. R. Lumsden and D. A. Evans), vol. 1, p. 493. Academic Press, London.

HOARE, C. A., ed. 1972. *The trypanosomes of mammals: A zoological monograph.* Blackwell, Oxford.

HOEIJMAKERS, J. H. J., P. BORST, J. VAN DEN BURG, C. WEISSMANN, and G. A. M. CROSS. 1980a. The isolation of plasmids containing DNA complementary to messenger RNA for variant surface glycoproteins of *Trypanosoma brucei. Gene* **8:** 391.

HOEIJMAKERS, J. H. J., A. C. C. FRASCH, A. BERNARDS, P. BORST, and G. A. M. CROSS. 1980b. Novel expression-linked copies of the genes for variant surface antigens in trypanosomes. *Nature* **284:** 78.

HOLDER, A. A. and G. A. M. CROSS. 1981. Glycopeptides from variant surface glycoproteins of *Trypanosoma brucei.* C-terminal location of antigenically crossreacting carbohydrate moieties. *Mol. Biochem. Parasitol.* **2:** (in press).

JOHNSON, J. G. and G. A. M. CROSS. 1979. Selective cleavage of variant surface glycoproteins from *Trypanosoma brucei. Biochem. J.* **178:** 689.

OPPERDOES, F. R. and P. BORST. 1977. Localization of nine glycolytic enzymes in a microbody-like organelle in *Trypanosoma brucei:* The glycosome. *FEBS Lett.* **80:** 360.

STRATHERN, J. N., C. S. NEWLON, I. HERSKOWITZ, and J. B. HICKS. 1979. Isolation of a circular derivative of yeast chromosome. III. Implications for the mechanism of mating type interconversion. *Cell* **18:** 309.

THOMAS, M., R. L. WHITE, and R. W. DAVIS. 1976. Hybridization of RNA to double-stranded DNA: Formation of R-loops. *Proc. Natl. Acad. Sci.* **73:** 2294.

VAN OMMEN, G. J. B., G. S. P. GROOT, and L. A. GRIVELL. 1979. Transcription maps of mtDNAs of two strains of *Saccharomyces:* Transcription of strain-specific insertions; complex RNA maturation and splicing. *Cell* **18:** 511.

VICKERMAN, K. 1969. On the surface coat and flagellar adhesion in trypanosomes. *J. Cell Sci.* **5:** 163.

———. 1978. Antigenic variation in trypanosomes. *Nature* **273:** 613.

WAHL, G. M., M. STERN, and G. R. STARK. 1979. Efficient transfer of large DNA fragments from agarose gels to diazobenzyloxymethyl-paper and rapid hybridization using dextran sulfate. *Proc. Natl. Acad. Sci.* **76:** 3683.

WILLIAMS, R. O., J. R. YOUNG, and P. A. O. MAJIWA. 1979. Genomic rearrangements correlated with antigenic variation in *Trypanosoma brucei. Nature* **282:** 847.