PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Sequence Variation and Evolution of Nuclear DNA in Man and the Primates

A. J. Jeffreys and P. A. Barrie

Phil. Trans. R. Soc. Lond. B 1981 **292**, 133-142
doi: 10.1098/rstb.1981.0021

Phil. Trans. R. Soc. Lond. B **292**, 133–142 (1981) [133]
Printed in Great Britain

Sequence variation and evolution of nuclear DNA in man and the primates

By A. J. Jeffreys and P. A. Barrie
Genetics Department, University of Leicester, University Road, Leicester LE1 7RH, U.K.

Recent advances in nucleic acid technology have facilitated the detection and detailed structural analysis of a wide variety of genes in higher organisms, including those in man. This in turn has opened the way to an examination of the evolution of structural genes and their surrounding and intervening sequences. In a study of the evolution of haemoglobin genes and neighbouring sequences in man and the primates, we have investigated gene arrangement and DNA sequence divergence both within and between species ranging from Old World monkeys to man. This analysis is beginning to reveal the evolutionary constraints that have acted on this region of the genome during primate evolution. Furthermore, DNA sequence variation, both within and between species, provides, in principle, a novel and powerful method for determining interspecific phylogenetic distances and also for analysing the structure of present-day human populations. Application of this new branch of molecular biology to other areas of the human genome should prove important in unravelling the history of genetic changes that have occurred during the evolution of man.

Introduction

The biochemical, physiological, morphological and behavioural characteristics of any organism are largely determined by the precise nature of the genetic information inherited from its parents. Ultimately, this information can be specified as a set of nucleotide base pair sequences that together define the DNA sequence of the organism's entire genome. Once these sequences have been experimentally determined, it should be possible, at least in principle, to decipher them in terms of stretches of DNA coding for various RNAs and proteins, that is, structural genes, interspersed with additional DNA elements, whose role, in part or in whole, might be to modulate the expression of these genes. The precise DNA sequence of a given animal's genome depends on its particular parentage and ultimately on the detailed evolutionary history of the species. A comparison of DNA sequences between related species should therefore reveal the rates and detailed characteristics of genomic evolution at the molecular level.

In this paper we shall describe the various approaches to studying molecular evolution in primates, and show how new methods in nucleic acid technology, in particular the use of recombinant DNA, have provided a way to a detailed comparative study of nuclear DNA, both within human populations and between man and the primates.

Approaches to analysing genetic variation at the molecular level

All genetic variation, both within and between species, is ultimately due to variation in nucleotide sequences, primarily in the nuclear DNA. Intraspecific variation, between, for example, different human races, can give important clues to the relationship between racial groups and hence to the recent evolutionary history of man. Similarly, interspecific differences between man and primates can reveal rates of evolution, recent major events in the evolution

of the human genome and possible genomic sequences of human ancestors. While neither approach can ever enable us to specify the detailed nucleotide sequence of a hominid ancestor, we should at least be able to define a constrained series of possible sequence types. Whether or not such possible ancestral sequences can ever be interpreted in terms of overall phenotypic effect will require an ability, as yet non-existent, to recognize control sequences in the genomic DNA of higher organisms and to understand their effect at the cellular and organismal level.

Since it has only very recently become possible to examine nuclear DNA sequences in detail, most work to date has centred on amino acid sequences of proteins, which in turn are specified by only a very limited fraction of total nuclear DNA. Comparisons of proteins in different humans by various electrophoretic and immunological techniques have revealed considerable genetic variation in amino acid sequence in many different proteins (Harris & Hopkinson 1972) and has contributed to our understanding of genetic variation between human races (Bodmer 1975). At the interspecific level, amino acid sequence comparisons of homologous proteins in man and the primates have shown a considerable similarity within the primates, as expected from their evolutionary relatedness (Dayhoff 1972). These similarities make it difficult to use such data to construct detailed molecular phylogenies of recent primate evolution. In an attempt to increase the resolving power of protein divergence, Sarich & Wilson (1968) have used microcomplement fixation assays to estimate immunological distances between complex sets of presumably homologous proteins isolated from various primates, and have used these distances to produce molecular phylogenies and to estimate divergence times of various primate groups in recent evolution.

Evidence is accumulating that non-coding nuclear DNA sequences have evolved more rapidly than structural gene sequences coding for proteins, perhaps as a result of the preferential elimination by selection of structural gene nucleotide substitutions that unfavourably alter the amino acid sequence of a protein (Gummerson & Williamson 1974; Van Ooyen *et al.* 1979; Konkel *et al.* 1979). It therefore follows that any attempt to examine evolution by comparing amino acid sequences will ignore what is probably the major reservoir of genetic variability between species, namely non-coding DNA. Furthermore, these regions are likely to contain control elements that modulate gene activity, and genetic variation within these elements might be crucial in generating altered gene regulation patterns that could result in profound phenotypic variation (King & Wilson 1975; Bruce & Ayala 1978).

There have been various approaches to studying primate evolution at the DNA level. Comparative karyotypic analyses of man and the primates have shown an evolutionary conservatism in gross chromosome organization, although interspecific differences, revealed particularly by chromosome banding techniques, have given useful phylogenetic information (De Grouchy *et al.* 1978). To study the molecular evolution of nuclear DNA sequences, Kohne (1970) has compared the relative sequence homologies of total nuclear DNA isolated from various primates. Nuclear DNAs from two different species were denatured, and the resulting single-stranded DNA preparations were mixed and annealed to produce interspecific hybrid duplex DNA. The reduction in melting point (ΔT_m) of these heteroduplexes, compared with perfectly matched double-stranded DNA isolated from one species, gave a measure of the reduced homology and therefore sequence divergence between nuclear DNAs from two species. It was found that related species had closely homologous DNAs, and that the overall degree of sequence divergence between two species varied in a phylogenetically consistent fashion. From such measurements, a molecular phylogeny of primate evolution could be constructed.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Both DNA ΔT_m analysis and estimates of immunological distance of proteins by micro-complement fixation revealed, surprisingly, what appeared to be a reduced rate of molecular divergence in recent primate evolution. This has been interpreted as evidence that macro-molecular sequences do not evolve in a clock-like mode, and that for some reason the rate of occurrence and/or fixation of genetic variation has indeed declined in recent primate evolution. Alternatively, it is possible that divergence times for various primate subgroups have been estimated incorrectly from palaeontological evidence, and that the evolutionary clock rate has been constant within primate species, which have diverged in evolution much more recently than has hitherto been supposed (Simons 1969; Wilson *et al.* 1978).

Estimation of homologies between primate DNAs by interspecific nucleic acid reannealing (Kohne 1970) gives no information about the detailed character of sequence divergence between related genomes. To overcome this problem, simple subfractions of cellular DNA have been compared in detail among various primates. Brown *et al.* (1979) have mapped the cleavage sites for 11 different restriction endonucleases in mitochondrial DNA isolated from man and from various Old World monkeys. They find clear interspecific homologies between the maps of this simple (16400 base pairs) circular organelle DNA, and deduce that no gross rearrange-ments of the mitochondrial genome have occurred within the group of species examined. They use cleavage site map differences to calculate the overall degree of DNA sequence divergence and find that the interspecific differences in mitochondrial DNA sequences are considerably greater (4–17-fold) than in nuclear DNA. Why mitochondrial DNA should evolve so rapidly is unknown, and therefore the use of mitochondrial DNA as a model for nuclear DNA evolution should be treated with caution. Furthermore, the observed interspecific divergence of this DNA appears to be relatively insensitive to phylogenetic distance.

To date, the only comparisons of specific nuclear DNA sequences between various primates have been made on simple, highly repetitive DNAs (Manuelidis & Wu 1978; Donehower & Gillespie 1979; Singer & Donehower 1979). These enigmatic tandem repetitive elements show considerable variation between primate species and can also vary within a single block of tandem repeats in one species. Since mechanisms probably exist that enable such sequence families to expand and contract rapidly in evolution, these repetitive elements are unlikely to be suitable models for overall nuclear DNA evolution.

Analysis of single copy structural genes in human DNA

The enormous complexity of human DNA, with a haploid genome size of 3×10^9 base pairs, has until recently precluded any analysis of structural genes present in one copy per haploid genome. However, recent advances in recombinant DNA technology have enabled us to examine in detail small selected parts of the human genome and to begin to characterize both structural genes and the DNA sequences that separate them.

To date, the best characterized set of human genes are those that code for the globin poly-peptides of haemoglobin. In the human foetus, the major species of haemoglobin is HbF (a mixture of $\alpha_2^G\gamma_2$ and $\alpha_2^A\gamma_2$ tetramers). At birth, HbF is replaced by adult haemoglobin (HbA, $\alpha_2\beta_2$, plus a low level of HbA_2, $\alpha_2\delta_2$). These various related haemoglobins are encoded by corresponding α-, β-, γ- and δ-globin genes. The α-globin genes remain expressed throughout foetal and adult life; in contrast, there is a switch from $^G\gamma$- and $^A\gamma$- to β- and δ-globin gene expression towards the end of gestation (see Weatherall & Clegg 1979).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Messenger RNAs for α-, β- and γ-globins have been isolated from foetal and adult reticulocytes and have been copied into double-stranded complementary DNA (cDNA) by means of reverse transcriptase. These globin cDNAs have been inserted into plasmid vectors and cloned to provide recombinant globin cDNAs, each plasmid containing a totally pure DNA copy of α-, β- or γ-globin messenger RNA (Little *et al.* 1978). Since a globin cDNA sequence is homologous with its corresponding structural gene, these cloned cDNAs have been of enormous use in the detection of human globin genes.

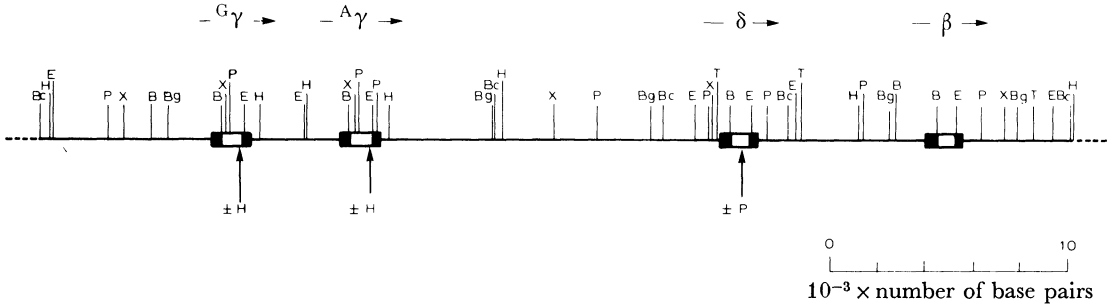

FIGURE 1. A physical map of restriction endonuclease cleavage sites around the human γ-, δ- and β-globin genes. Data for this map were taken from Flavell *et al.* (1978), Little *et al.* (1979), Bernards *et al.* (1979) and Jeffreys (1979). The positions of coding sequences are indicated by closed boxes, and those of major intervening sequences by open boxes. The direction of transcription is shown by horizontal arrows. Cleavage sites that show no polymorphic variation in northern Europeans are shown above the map for restriction endonucleases *Bam* HI (B), *Bcl* I (Bc), *Bgl* II (Bg), *Eco* RI (E), *Hind* III (H), *Pst* I (P), *Taq* YI (T) and *Xba* I (X). Endonuclease *Taq* YI cleavage sites near the ${}^{G}\gamma$- and ${}^{A}\gamma$-globin genes have not yet been mapped. It should be stressed that this map was deduced from Southern blot analyses of globin DNA fragments present in restriction endonuclease digests of total human DNA, and therefore only cleavage sites that generate fragments containing globin genes are shown here. Additional cleavage sites that show a presence/absence of polymorphic variation in northern Europeans are indicated by vertical arrows (Jeffreys 1979).

Globin genes can be detected in total human DNA by means of cloned cDNAs as nucleic acid hybridization probes. Human DNA is cleaved with a restriction endonuclease and denatured, and the complex series of digest products are fractionated according to size by electrophoresis through an agarose gel. DNA fragments are transferred from the gel to a nitrocellulose filter by blotting (Southern 1975) and the filter is hybridized with ^{32}P-labelled cloned globin cDNA. The probe binds only to the corresponding structural gene sequence, and the restriction endonuclease fragment of human DNA that contains this globin gene can be detected by subsequent autoradiography of the filter. By means of this approach, a cloned β-globin cDNA is found to detect human DNA fragments containing the closely homologous β- and δ-globin genes. Similarly, γ-globin cDNA detects ${}^{G}\gamma$- and ${}^{A}\gamma$-globin gene fragments. Comparison of these fragments in human DNA digested with various combinations of different restriction endonucleases enables us to arrange these fragments into a physical map of cleavage sites around these non-α-globin genes (Flavell *et al.* 1978; Little *et al.* 1979; Bernards *et al.* 1979).

Figure 1 shows the physical map of the human ${}^{G}\gamma$-, ${}^{A}\gamma$-, δ- and β-globin genes. All four genes are closely linked to each other and each is interrupted by at least one intervening sequence. There are also substantial tracts of DNA separating these genes; the function of this intergenic DNA is unknown, although it seems likely that these sequences, in part or in whole, will be involved in globin gene expression and regulation of, for example, the γ to β switch at birth (Fritsch *et al.* 1979).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

This arrangement of the human γ- and β-globin genes has been confirmed by examining cloned genomic DNA. Human DNA cleaved with a restriction endonuclease has been inserted into a λ bacteriophage vector. Subsequent growth in *Escherichia coli* of a sufficient number of different λ–human DNA recombinants gives a 'library' of cloned human DNA fragments covering most of the human genome. The library can then be screened for bacteriophage containing the human DNA sequence of interest. By this approach, fragments of human DNA containing β-related globin genes have been purified (Lawn *et al.* 1978). This purification opens the way to a detailed fine-structure analysis of this small region of the human genome.

Variations in DNA sequence between individuals

As already noted, virtually all genetic variation within a population can ultimately be traced back to individual variation of nuclear DNA sequences. It is therefore important to establish just how much individual variation occurs at this fundamental level. This could be determined by cloning β-globin genes from various individuals and directly comparing their sequences, at the moment an exceedingly onerous task. Alternatively we could screen for DNA sequence variants that by chance create or destroy restriction endonuclease cleavage sites near globin genes; such variants would cause an alteration of globin restriction fragment size, which could be detected in a restriction endonuclease digest of total human DNA, with cloned globin cDNAs as probes for globin genes. An extensive screening of 60 unrelated individuals, mainly northern Europeans, showed that almost all the restriction endonuclease cleavage sites round the β-related globin genes shown in figure 1 are invariant (Jeffreys 1979). Of 52–54 different cleavage sites examined, only three showed a presence/absence polymorphic variation in northern Europeans. Curiously, all three variant sites were within intervening sequences (figure 1), suggesting that genetic variation might be accumulating preferentially in these sections of the genome.

This survey indicated that, on average, approximately one base pair in every 100 in this region of the genome varies polymorphically in the population screened, although few of these variants will create or destroy a detectable restriction enzyme site. Also, the actual level of sequence divergence between two randomly selected chromosomes will be less than one base pair per 100, since both chromosomes will possess identical alleles at the majority of variable loci.

These variant restriction endonuclease cleavage sites provide a major new type of human genetic marker. This type of screening analysis could be repeated for any section of single copy human DNA, with random cloned fragments of human DNA, even those of completely unknown function, as probes. This approach would provide enormous numbers of variant sequence markers in all parts of the human genome. We are currently investigating racial variation in gene frequencies of the globin markers; it is likely that, as with proteins, gene frequencies will be markedly race-dependent and will provide important clues to the relationships between various human races, and their origins.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Globin genes and surrounding sequences in primates

The β-globin genes of man have now been well characterized, both in terms of gene structure and organization and in terms of genetic variability within man (albeit to date only within northern Europeans). We were therefore interested in comparing this area of the human genome with the homologous region of primate DNAs.

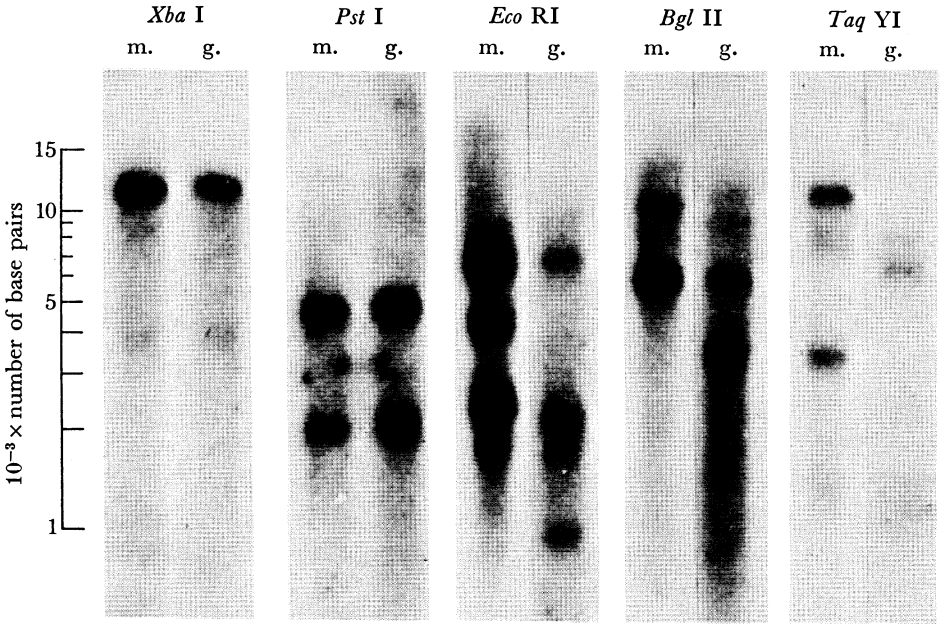

Figure 2. A comparison of β- and δ-globin DNA fragments produced by restriction endonuclease cleavage of human and gorilla DNA. Samples (15 μg) of human (m.) and western lowland gorilla (*Gorilla gorilla gorilla*) (g.) white blood cell DNA were cleaved with the indicated restriction endonucleases, alkali-denatured and electrophoresed through a 1% (by mass) agarose gel; DNA fragments were then transferred, by blotting, to a nitrocellulose filter. Fragments containing the β- and δ-globin genes were labelled by hybridizing the filter with ^{32}P-labelled cloned human β-globin cDNA (plasmid pHβG1; Little *et al.* 1978) and labelled fragments were detected by autoradiography (see Jeffreys (1979) for experimental details).

We first compared human and gorilla globin genes. Total human and gorilla DNAs were cleaved with a restriction endonuclease, DNA fragments were separated by agarose gel electrophoresis and blotted onto a nitrocellulose filter. Fragments of β- and δ-globin DNA were detected by hybridization with ^{32}P-labelled cloned human β-globin cDNA (figure 2). Certain restriction endonucleases (*Xba* I, *Pst* I) gave patterns of β- and δ-globin DNA fragments that were indistinguishable between man and gorilla. This immediately shows that the gorilla also has one β- and one δ-globin gene, and that the arrangement of these genes is, within the resolution of the system, identical in the two species. Other restriction enzymes generate patterns of β- and δ-globin DNA fragments that differ to a greater or lesser extent between man and gorilla. This is evidence of sequence divergence accumulating between the two species' DNA, sufficient to create or destroy a fraction of the restriction endonuclease cleavage sites. Detailed mapping of gorilla β-globin genes gives the physical map shown in figure 4. As can be seen, sufficient divergence has occurred to cause about 16% of cleavage sites to differ between the two species (table 1).

We extended this analysis to the yellow baboon, a representative of the Old World monkeys.

We were concerned that human and baboon DNAs have diverged so far that human β-globin cDNA would fail to detect the homologous baboon genes in total baboon DNA. Figure 3 shows that human and baboon β-globin DNA sequences have in fact been well conserved. By subjecting filters to post-hybridization washes of increased stringency (by progressively lowering the ionic strength of the washing solution), we can examine the stability of DNA hybrids made

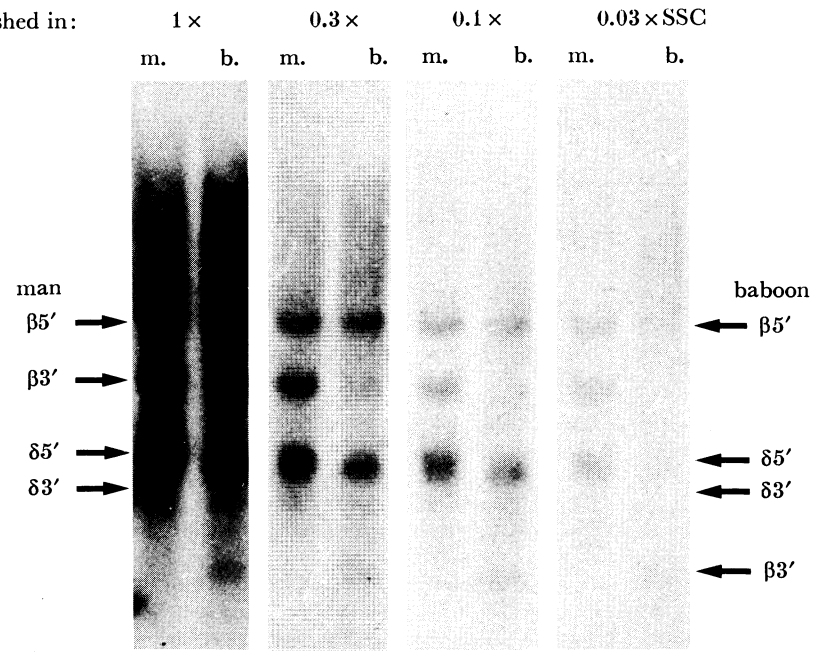

FIGURE 3. Detection of β-globin DNA fragments in restriction endonuclease digests of baboon DNA. DNA samples prepared from man and from a yellow baboon (*Papio cynocephalus*) were digested with restriction endonuclease *Eco* RI. Human (m.) and baboon (b.) digests were alkali-denatured, electrophoresed through a 1% (by mass) agarose gel and transferred to a nitrocellulose filter. The filter was hybridized with ^{32}P-labelled pHβG1 DNA at 65 °C in 3 × SSC (1 × SSC = saline sodium citrate: 0.15 M NaCl, 15 mM trisodium citrate (pH 7.0)). Unbound label was then removed by washing in 1 × SSC at 65 °C. Identical filters were given a subsequent higher stringency wash in 1 ×, 0.3 ×, 0.1 × or 0.03 × SSC at 65 °C and remaining labelled globin DNA fragments were detected by autoradiography. Since endonuclease *Eco* RI cleaves within both the δ- and β-globin gene, each gene gives rise to two hybridizable fragments (see figure 1). The autoradiographic positions of the resultant δ5′, δ3′, β5′ and β3′ globin DNA fragments are indicated; all fragments except the β3′ fragment appear identical in the two species (see figure 4).

between baboon globin genes and the human β-globin cDNA probe. Only at the highest stringency (0.03 × SSC at 65 °C, see figure 3) is there preferential loss of heterologous man–baboon DNA hybrids; at lower stringencies the human probe is capable of detecting specifically the baboon genes. At the lowest stringency (1 × SSC at 65 °C) additional human and baboon DNA fragments become labelled; these fragments contain ^{G}γ- and ^{A}γ-globin genes which show partial homology to β-globin cDNA.

All β-globin DNA fragments detected in baboon and human DNA cleaved with *Eco* RI are identical except for one. This again establishes that the baboon contains both a β- and a δ-globin gene. Detailed mapping of baboon DNA shows that these genes are arranged as in man and gorilla, but that a substantial divergence of cleavage sites has occurred between the baboon and gorilla/man (figure 4; table 1). Extension of this analysis to the ^{G}γ- and ^{A}γ-globin

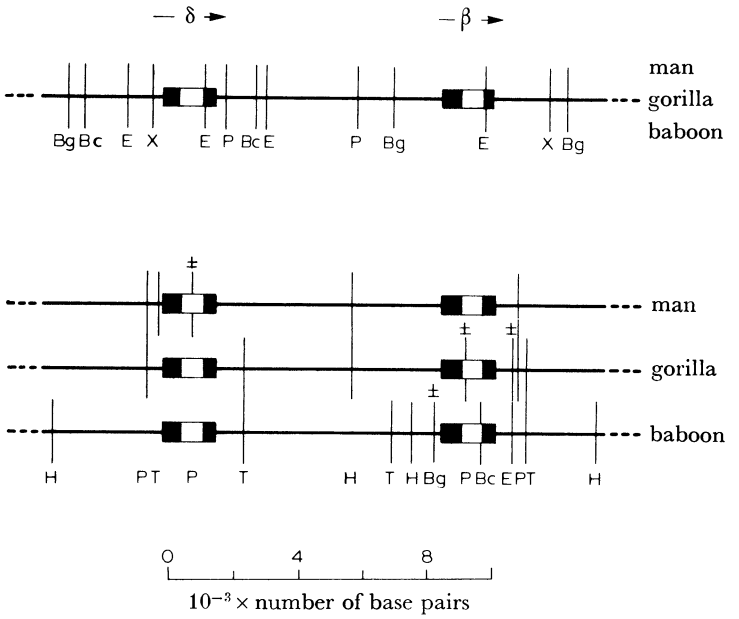

FIGURE 4. A comparison of the physical maps of restriction endonuclease cleavage sites in and around the δ- and β-globin genes of man, gorilla and baboon. DNA was prepared from blood obtained from four western lowland gorillas (*Gorilla gorilla gorilla*) and from liver removed from a single yellow baboon (*Papio cynocephalus*). Cleavage sites for restriction endonucleases *Bcl* I (Bc), *Bgl* II (Bg), *Eco* RI (E), *Hind* III (H), *Pst* I (P), *Taq* YI (T) and *Xba* I (X) were mapped near the β- and δ-globin genes, by analysing β- and δ-globin DNA fragments present in primate DNAs cleaved with various combinations of restriction endonucleases (figures 2, 3).

The cleavage sites shown are divided into conserved sites, which map to indistinguishable positions around the δ- and β-globin genes in all three species, and non-conserved sites, each of which is absent in at least one of the three species examined. Cleavage sites that have been shown to vary within a given species are indicated thus: ±. Additional cleavage sites in the human map shown in figure 1 have not yet been checked across all three species, and are therefore omitted.

TABLE 1. DNA SEQUENCE DIVERGENCE IN THE β-GLOBIN LOCI OF VARIOUS PRIMATES

	percentage difference in: β- and δ-globin genes		total genome:	
	cleavage sites †	DNA sequence‡	DNA sequence §	
man–man‖	3	< 1	—	
man–gorilla	16	3	2.6	man–chimpanzee
man–baboon	44	10	10.5	man–green monkey
gorilla–baboon	33	7	10.8	chimpanzee–green monkey

† Percentage difference in cleavage sites taken from figures 1 and 4. A polymorphic site is scored as a half difference.

‡ Calculated from the cleavage site variation by the method of Upholt (1977).

§ Total genomic DNA sequence divergence of corresponding man–great ape–Old World monkey pairs derived from ΔT_m measurements (data taken from Kohne (1970)).

‖ Cleavage site and DNA sequence variation found within man (northern European) (Jeffreys 1979), scoring polymorphic sites as half differences.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

genes has shown that the entire topology of the β-related globin gene region shown in figure 1 is indistinguishable between man, gorilla and baboon.

Despite the small number of primate individuals examined, intraspecific variant cleavage sites have been detected (figure 4), including a *Pst* I cleavage site within the major intervening sequence of the β-globin gene in gorilla. There is a curious polymorphism in gorilla for the presence/absence of an *Eco* RI cleavage site to the 3′ side of the β-globin gene. This site is present in all species of Old World monkey so far examined, but is absent in man. The gorilla would thus appear to represent a species in some way intermediate for this site between monkeys and man.

The degree of cleavage site divergence can be used to calculate an approximate extent of overall sequence divergence between two species (table 1). As seen, the degree of interspecific divergence is considerably greater than that seen within a human population and seems to vary in a phylogenetically consistent fashion; thus man and gorilla show a greater map homology than does either species compared with baboon. Furthermore, the divergence of sequence homology near the β- and δ-globin genes is similar to overall genomic DNA sequence divergence calculated from ΔT_m measurements (Kohne 1970). The β-globin DNA region of the primates seems therefore to have evolved at a rate similar to that of the entire genome.

The future

This comparison of primate β-globin genes has shown a region of DNA whose topology has remained invariant over tens of millions of years, yet has accumulated genetic divergence that should prove useful in estimating phylogenetic distances. This analysis could be repeated with any single copy region of the genome to produce refined and testable divergence estimates for primate phylogenies. We are also extending this analysis further back to prosimians, to search for events such as gene duplication that have been crucial in moulding the contemporary arrangement of these genes.

The advantage of studying DNA over proteins is already apparent: the amino acid sequences of primate globin polypeptides show little variation, yet by analysing globin DNA, we can detect genetic variation within and between genes that is probably never expressed at the protein level. We should stress, however, that the phenotypic significance of these interspecies variants is not clear; it is possible that these variants have become fixed in evolution by chance and do not contribute to the phenotypic difference between two species. Current DNA technology makes it at least theoretically possible to sequence, for example, the entire human and gorilla genomes. Within the two sequences will be differences that are instrumental in determining the characteristics of a man or a great ape. Whether we could ever be able to recognize such sequences would require an understanding of DNA sequences that we do not yet possess. Ultimately, it might be possible to test for important sequence differences by introducing selected human DNA sequences into a great ape zygote and observing the effects of this genetic alteration on the phenotype of the offspring; however, such an approach is likely to be precluded on both technical and ethical grounds.

P. A. B. is the recipient of a research scholarship from the University of Leicester.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

References

Bernards, R., Little, P. F. R., Annison, G., Williamson, R. & Flavell, R. A. 1979 Structure of the human $^{G}\gamma$- $^{A}\gamma$- δ -β-globin gene locus. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4827–4831.

Bodmer, W. F. 1975 Genetic markers, evolution and selection. In *Transfusion and Immunology Plenary Session Lectures of the Fourteenth Congress of the International Society of Blood Transfusion and the Tenth Congress of the World Federation of Hemophilia* (ed. E. E. Ikkala & A. Nykanen), pp. 35–46. Helsinki: Vammala.

Brown, W. M., George, M. & Wilson, A. C. 1979 Rapid evolution of animal mitochondrial DNA. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1967–1971.

Bruce, E. J. & Ayala, F. J. 1978 Humans and apes are genetically very similar. *Nature, Lond.* **276**, 264–265.

Dayhoff, M. O. (ed.) 1972 *Atlas of protein sequence and structure*, vol. 5. Silver Spring, Maryland: National Biomedical Research Foundation.

De Grouchy, J., Turleau, C. & Finaz, C. 1978 Chromosome phylogeny of the primates. *A. Rev. Genet.* **12**, 289–328.

Donehower, L. & Gillespie, D. 1979 Restriction site periodicities in highly repetitive DNA of primates. *J. molec. Biol.* **134**, 805–834.

Flavell, R. A., Kooter, J. M., De Boer, E., Little, P. F. R. & Williamson, R. 1978 Analysis of the β-δ-globin gene loci in normal and Hb Lepore DNA: direct determination of gene linkage and intergene distance. *Cell* **15**, 25–41.

Fritsch, E. F., Lawn, R. M. & Maniatis, T. 1979 Characterisation of deletions which affect the expression of fetal globin genes in man. *Nature, Lond.* **279**, 598–603.

Gummerson, K. S. & Williamson, R. 1974 Sequence divergence of mammalian globin messenger RNA. *Nature, Lond.* **247**, 265–267.

Harris, H. & Hopkinson, D. A. 1972 Average heterozygosity per locus in man: an estimate based on the incidence of enzyme polymorphisms. *Ann. hum. Genet.* **36**, 9–19.

Jeffreys, A. J. 1979 DNA sequence variants in the $^{G}\gamma$-, $^{A}\gamma$-, δ- and β-globin genes of man. *Cell* **18**, 1–10.

King, M. C. & Wilson, A. C. 1975 Evolution at two levels in humans and chimpanzees. *Science, N.Y.* **188**, 107–116.

Kohne, D. E. 1970 Evolution of higher-organism DNA. *Q. Rev. Biophys.* **3**, 327–375.

Konkel, A. D., Maizel, J. V. & Leder, P. 1979 The evolution and sequence comparison of two recently diverged mouse chromosomal β-globin genes. *Cell* **18**, 865–873.

Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. 1978 The isolation and characterization of linked δ- and β-globin genes from a cloned library of human DNA. *Cell* **15**, 25–41.

Little, P. F. R., Curtis, P., Coutelle, C., Van Den Berg, J., Dalgleish, R., Malcolm, S., Courtney, M., Westaway, D. & Williamson, R. 1978 Isolation and partial sequence of recombinant plasmids containing human α-, β- and γ-globin cDNA fragments. *Nature, Lond.* **273**, 640–643.

Little, P. F. R., Flavell, R. A., Kooter, J. M., Annison, G. & Williamson, R. 1979 Structure of the human foetal globin gene locus. *Nature, Lond.* **278**, 227–231.

Manuelidis, L. & Wu, J. C. 1978 Homology between human and simian repeated DNA. *Nature, Lond.* **276**, 92–94.

Sarich, V. M. & Wilson, A. C. 1968 Immunological time scale for hominid evolution. *Science, N.Y.* **158**, 1200–1202.

Simons, E. L. 1969 The origin and radiation of the primates. *Ann. N.Y. Acad. Sci.* **167**, 319–331.

Singer, D. & Donehower, L. 1979 Highly repeated DNA of the baboon: organization of sequences homologous to highly repeated DNA of the African green monkey. *J. molec. Biol.* **134**, 835–842.

Southern, E. M. 1975 Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. molec. Biol.* **93**, 503–517.

Upholt, W. B. 1977 Estimation of DNA sequence divergence from comparison of restriction endonuclease digests. *Nucl. Acids Res.* **4**, 1257–1265.

Van Ooyen, A., Van Den Berg, J., Mantei, N. & Weissmann, C. 1979 Comparison of total sequence of a cloned rabbit β-globin gene and its flanking regions with a homologous mouse sequence. *Science, N.Y.* **206**, 337–344.

Weatherall, D. J. & Clegg, J. B. 1979 Recent developments in the molecular genetics of human hemoglobin. *Cell* **16**, 467–479.

Wilson, A. C., Carlson, S. S. & White ,T. J. 1978 Biochemical evolution. *A. Rev. Biochem.* **48**, 573–639.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

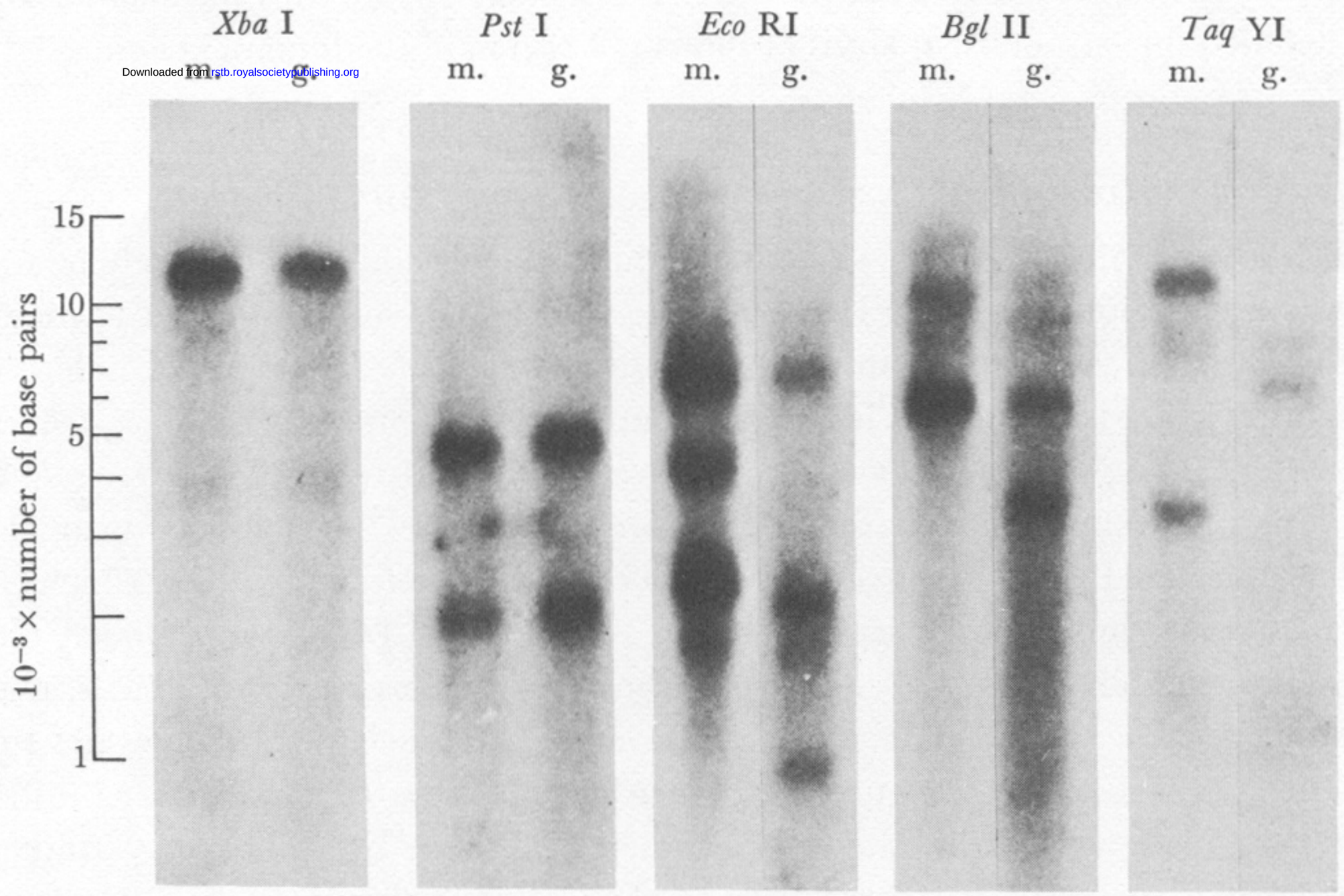

FIGURE 2. A comparison of β- and δ-globin DNA fragments produced by restriction endonuclease cleavage of human and gorilla DNA. Samples (15 μg) of human (m.) and western lowland gorilla (*Gorilla gorilla gorilla*) (g.) white blood cell DNA were cleaved with the indicated restriction endonucleases, alkali-denatured and electrophoresed through a 1% (by mass) agarose gel; DNA fragments were then transferred, by blotting, to a nitrocellulose filter. Fragments containing the β- and δ-globin genes were labelled by hybridizing the filter with ^{32}P-labelled cloned human β-globin cDNA (plasmid pHβG1; Little *et al.* 1978) and labelled fragments were detected by autoradiography (see Jeffreys (1979) for experimental details).

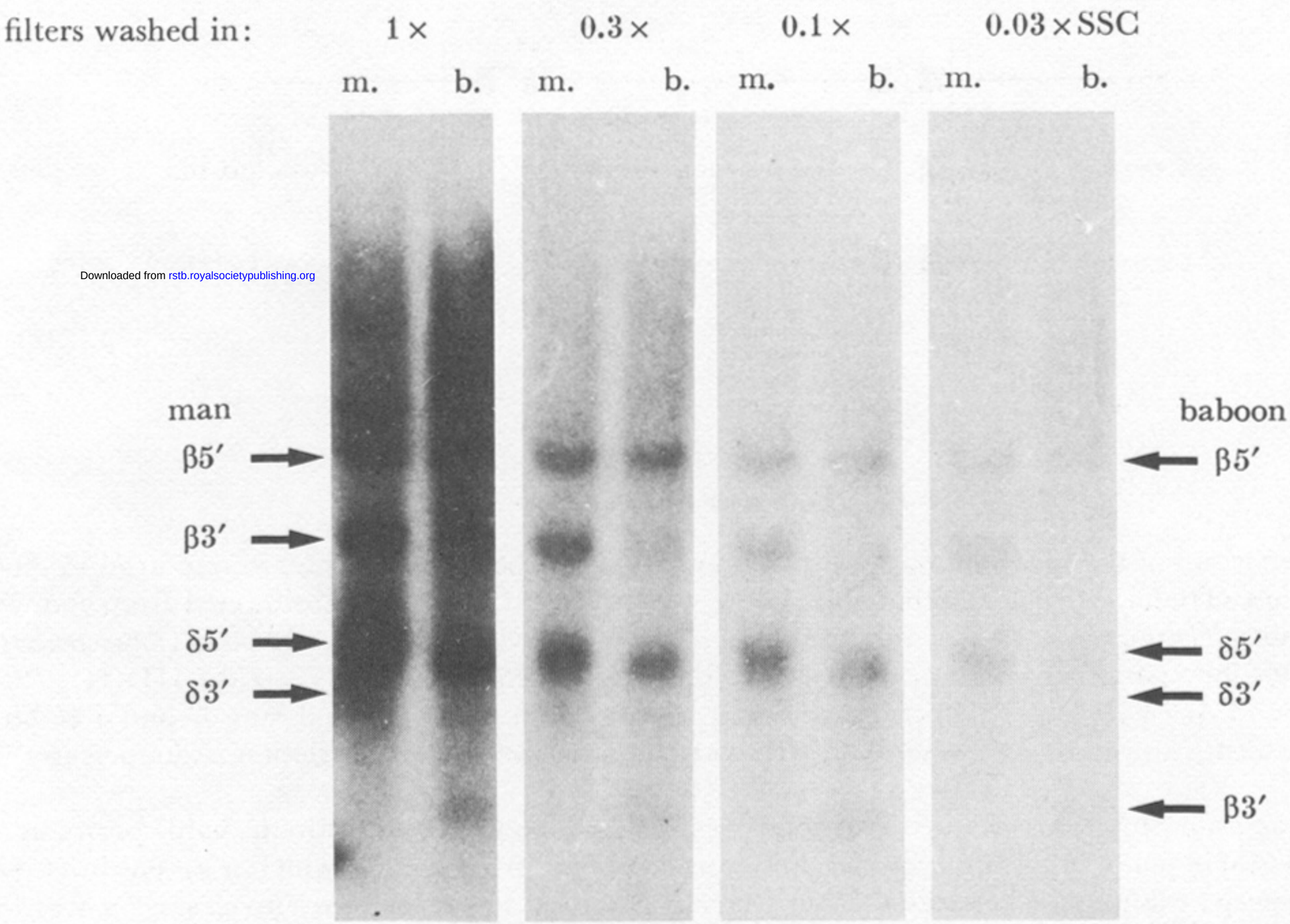

FIGURE 3. Detection of β-globin DNA fragments in restriction endonuclease digests of baboon DNA. DNA samples prepared from man and from a yellow baboon (*Papio cynocephalus*) were digested with restriction endonuclease *Eco* RI. Human (m.) and baboon (b.) digests were alkali-denatured, electrophoresed through a 1% (by mass) agarose gel and transferred to a nitrocellulose filter. The filter was hybridized with ^{32}P-labelled pHβG1 DNA at 65 °C in 3 × SSC (1 × SSC = saline sodium citrate: 0.15 M NaCl, 15 mM trisodium citrate (pH 7.0)). Unbound label was then removed by washing in 1 × SSC at 65 °C. Identical filters were given a subsequent higher stringency wash in 1 × , 0.3 × , 0.1 × or 0.03 × SSC at 65 °C and remaining labelled globin DNA fragments were detected by autoradiography. Since endonuclease *Eco* RI cleaves within both the δ- and β-globin gene, each gene gives rise to two hybridizable fragments (see figure 1). The autoradiographic positions of the resultant δ5′, δ3′, β5′ and β3′ globin DNA fragments are indicated; all fragments except the β3′ fragment appear identical in the two species (see figure 4).