Volume 17 Number 17 1989 Nucleic Acids Research

# Rapid detection and sequencing of alleles in the 3′ flanking region of the interleukin-6 gene

Anne M.Bowcock*, Anuradha Ray[1], Henry Erlich[2] and Pravinkumar B.Sehgal[1]

Department of Genetics, Stanford University, Stanford, CA 94305, [1]Rockefeller University, New York, NY 10021 and [2]Cetus Corporation, Department of Human Genetics, 1400 53rd Street, Emeryville, CA 94608, USA

Received June 28, 1989; Revised and Accepted July 27, 1989

## ABSTRACT

The 3′ flanking region of the interleukin 6 gene is polymorphic due to insertions of different size. Within this region lies a sequence of approximately 500 base pairs that is AT rich. Based on flanking sequence information we have constructed oligonucleotides which prime the polymerase chain reaction (PCR) and amplify this AT rich region. The amplification products visualized by agarose gel electrophoresis gave fragment sizes for both homozygous and heterozygous individuals that were concordant with those observed by conventional genomic blotting techniques. Alleles that could not be typed by Southern analysis were resolved with this approach. These results illustrate the value of PCR for the rapid detection of length polymorphisms such as those due to variable numbers of tandem repeats. In contrast to RFLP analysis this procedure takes less than a day to perform, is cheaper, avoids the use of radioactivity and requires far less substrate DNA. Three different human alleles were sequenced, and differences were detected that were due to both large duplications and loss of one or two bases, suggesting that AT rich regions identify highly polymorphic loci. The same primers also amplified non-human primate DNA, allowing a comparison of the human sequence with that of the common chimpanzee and baboon.

## INTRODUCTION

The human interleukin-6 (IL6) gene is located at 7p15-p21 (1,2,3). At least four alleles are observed in the 3′ flanking region of the gene (3) and at least ten enzymes can be used to detect this polymorphism, implying that it is due to insertions of variable size. Within or close to the rearranged region a highly AT rich sequence has been found that extends for approximately 500 base pairs (3). The analysis of restriction fragment length polymorphisms (RFLPs) such as the one described above involves restriction endonuclease digestion of substrate DNA, agarose gel electrophoresis, transfer to a solid support (4), hybridization and autoradiography. We wished to develop a system using the polymerase chain reaction (PCR)(5,6,7) that would rapidly detect the polymorphism described above. Until recently, PCR has only been used to detect base pair substitutions with allele specific oligonucleotides following amplification of substrate DNA (8). Many highly informative DNA polymorphisms are due to insertion/deletion events, such as those due to variable numbers of tandem repeats (VNTRs)(9). The detection of length polymorphisms by PCR followed by agarose gel electrophoresis to resolve the length of the products, would considerably simplify the analysis. However it was not known if PCR would faithfully amplify regions of DNA that contained repeated units. There are three potential problems: 1) when primers anneal to the DNA template the repeated units on complementary strands could anneal to each other in an 'out of register' fashion 2) when extending the primed DNA strand the polymerase could switch strands; in a heterozygous individual where two

© IRL Press 

fragment lengths are expected, extra fragments of the wrong size could arise 3) since the PCR products derived from different VNTR alleles are of different lengths, the shorter fragments may be preferentially amplified.

We describe a PCR system for detecting the IL6 3′ flanking region alleles. This is applicable to disease association studies, population studies, linkage studies and forensic studies. Rapid amplification facilitates rapid sequencing and the sequences of three human alleles and two non-human primate alleles in this region are described.

## MATERIALS AND METHODS

### *Sequencing*

The sequence of the AT rich region in the 3′ flanking region of the IL6 gene has been described previously (3). This terminated in an RsaI site. The sequence of an oligonucleotide primer 5′ to this region was obtained from the previously published sequence (10). To obtain sufficient sequence in order to construct an oligonucleotide primer specific for the 3′ flanking region of the AT repeat, a 0.9 kb HindIII fragment containing the AT rich sequence was cloned between the HindIII sites of the M13 mp19 vector. Two clones each were chosen that were in opposite orientation, and sequenced according to the dideoxy chain termination procedure (11).

### *Amplification of DNA*

Template DNA (1 microgram) was incubated with 25 pmoles of each primer, 50mM KCl, 1.5mM $MgCl_2$, 10mM Tris-HCl (pH 7.4), deoxyribonucleotide triphosphates (dATP, dCTP, TTP, and dGTP; 200 micromolar each) and 2.5 units of Taq DNA polymerase (Perkin Elmer-Cetus Instruments). The samples were subjected to 28 cycles of amplification. The cycling reaction was performed in a programmable heat block (DNA Thermal Cycler, PECI). The substrate DNA was denatured at 94°C for 15 sec annealed to the primers at 55°C for 30 sec and the annealed primers were allowed to extend at 72°C for 60 sec.

### *Analysis of amplified products*

One twentieth of the reaction was subjected to electrophoresis in a 2.5% Nusieve agarose gel at 3 V/cm for four to five hours. The gel was stained with ethidium bromide and photographed under UV light. Transfer to nitrocellulose and hybridization procedures were performed as described previously (12). The DNA probe pB2.3 was the 900bp HindIII fragment from the 3′ flanking region of IL6 (Figure 1). This was $^{32}P$ labelled using the hexamer primer reaction (13). $10^5$ cpm/ml of probe were used for hybridization. The filter was exposed to Kodak XAR-5 film for 16 hours.

### *Comparison with conventional genomic blotting, hybridization and autoradiography*

DNA from eight members of one family (2.5 micrograms each) was digested with 7.5 units of BstNI (New England Biolabs) and fractionated in a 1.2% agarose gel at 1 V/cm for 18 hours. The gel was transferred to Zetabind, and hybridized with probe pBeta2.15 (3). The blot was exposed to Kodak XAR-5 film for three days.

### *Sequencing of amplified products*

This was performed in two ways: 1) The PCR products from three individuals homozygous for different IL6 alleles were purified from agarose gels, digested with RsaI/HindIII, and cloned between the SmaI/HindIII sites of M13 mp18 and mp19 vectors. Sequencing was performed as described above. 2) Single copy sequences from DNA of the homozygous individuals, the common chimpanzee and baboon were obtained by fractionating the PCR products in agarose, isolating the fragment from agarose, and subjecting one twentieth of the products to a second round of amplification using only one oligonucleotide primer.

```
    425bp                            540bp
    ┌────┐                           ┌──────────────────┐
                                     HINDIII

EXON V......GCAACTTTGAGTGTGTCACGTGAAGCTT.......AT RICH

RSAI

GTACACACAC ACATATATAT ATACATATAA GCACCTACTA CATGCCAGGC

ATCATTAAAT GTGTTGCATC CATCACGTCA TTTAACCCCA GACTTGCACA

CTCCTTTCTG GTTGTGGAAG ACTAAGTAAT TTATCTAAGT CACCCAGCTG

GAAGGTCAGG CAGGGACCCA GATTTGAAAT CCAAGATCTA CCTACCTACA

GGTCCCCTAC TCTTAACCTG TAGGTCCCAC TGCCTACCCA GGAACTGAGG

GATGATGTAG AAAATCCCAA AACGATGTTA ATATAGGGAA TACCTATAAA

GCATGGCAAT CAAAGTCTTT GGGACTATAC AACCACTGTA TAAAGCATAA

CAATGGTACA AGCTT

        HINDIII
```

**Figure 1.** *Sequence of the 3′ flanking region of the IL6 gene.* The location of the fifth exon relative to the first HindIII site (10) and the next 540 nucleotides which includes the AT rich region (3) are shown. The sequence of the region 3′ to the AT rich repeat was obtained as described in 'Methods'. The sequences homologous to the oligonucleotide primers are underlined.

Alternatively single copy amplification products were obtained by using asymmetric amounts of primers in the amplification process as recently described (14). In both cases single copy DNA sequences were purified and sequenced as described previously (14).

## RESULTS

*Sequence of 3′ flanking region*

The sequence 3′ to the AT rich repeat is shown in Figure 1. The sequences which were used to obtain the left and right oligonucleotide primers (5′-GCAACTTTGAGTGTGTCACG-3′and 5′-GACGTGATGGATGCAACAC-3′ respectively) are shown.

*Results of amplification*

Figure 3A shows a 2% agarose gel stained with ethidium bromide and containing PCR amplification products from nine unrelated human individuals (lanes 1–9). Figure 3B shows a Southern transfer of this gel demonstrating that the fragments observed were homologous to a probe from the 3′ region of IL6 (Figure 2, pB2.3). This confirmed that the 3′ flanking region of the IL6 gene had been amplified. The additional lower fragments observed in Figure 3B, but not seen on the gel in Figure 3A are probably single stranded PCR products or other PCR artefacts, since similar results have been seen at other loci (21).

The BstNI alleles described previously of 1.6, 1.5, 1.44 and 1.37kb, corresponded to alleles B1, B2, B3 and B4 (3). For each sample a complete correlation between the PCR and RFLP polymorphisms was observed. PCR fragments of 0.76, 0.68, 0.64 and 0.61kb corresponded to the B1, B2, B3 and B4 alleles. Lane 1 (Fig 3A) shows the fragment length obtained after amplifying DNA from an individual previously shown to be homozygous for the B1 allele. Lane 2 contains DNA amplified from a B3 homozygote, and lane 3 contains DNA amplified from a B4 homozygote. Lanes 4,5,6,7 and 8 contain DNA amplified from the following heterozygotes: B1/B3, B1/B4, B3/B4, B2/B4 and B2/B3. The individual in lane 9 is heterozygous for allele B1 and a novel allele with a length between those of alleles B2 and B3. This, and a number of novel alleles that had previously

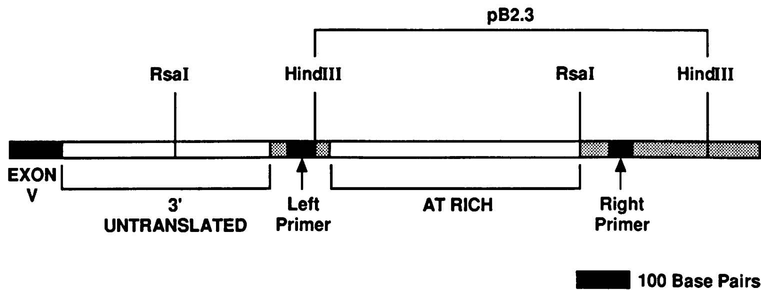


**Figure 2.** *The 3' flanking region of the IL6 gene.* The locations of primers used in the polymerase chain reaction and the region to which probe pB2.3 hybridizes are shown.

been suspected to exist with Southern analysis were confirmed to exist by PCR analysis. In order to resolve alleles B2, B3 and B4 by Southern analysis, it was necessary to run gels under special conditions. Analysis of these alleles with PCR allowed us to rapidly resolve these alleles that differ by less than 0.2kb in size.

The primers and the conditions described to amplify human DNA also amplified DNA from four primate species: common chimpanzee (*Pan troglodytes*— lane 10, 0.46kb); gorilla (*Gorilla-gorilla*— lane 11, 0.48kb and 0.50kb); orang-utan (*Pongo pygmaeus* 0.16kb) and baboon (*Papio cynocephalus*—0.16kb)(orang-utan and baboon are not shown). This region is polymorphic in chimpanzee and gorilla; lane 11 has two PCR products because the gorilla used was heterozygous at this locus.

Figure 4 demonstrates that this polymorphism is inherited in a Mendelian fashion. Both PCR products (4A) and fragments detected by Southern analysis (4B) are shown. Alleles

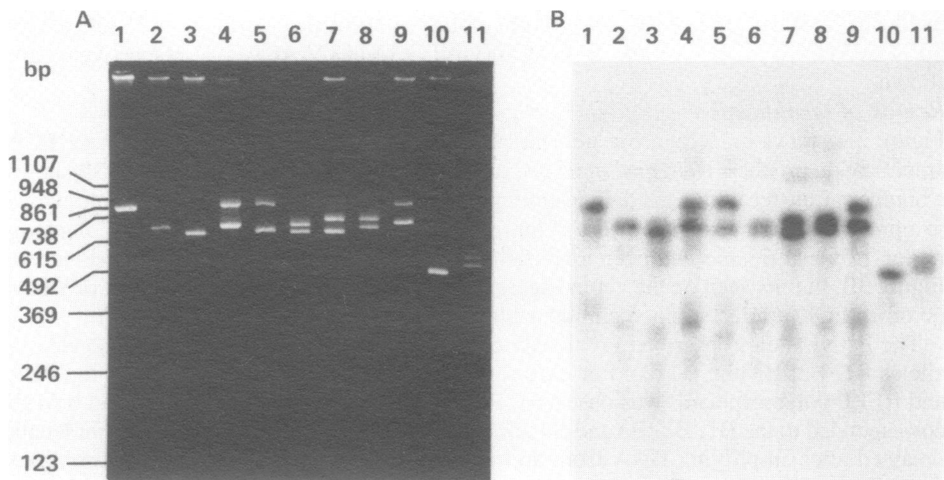


**Figure 3A.** *UV visualization of an ethidium bromide stained 2% agarose gel containing PCR amplification products.* DNA samples from eleven individuals were amplified using the primers shown in Figure 1. as described in 'Methods'. Fragment lengths are shown.

**Figure 3B.** *Autoradiograph showing IL6 specific amplification products.* The gel shown in FIG.3A was blotted as described and hybridized with the probe pB2.3.

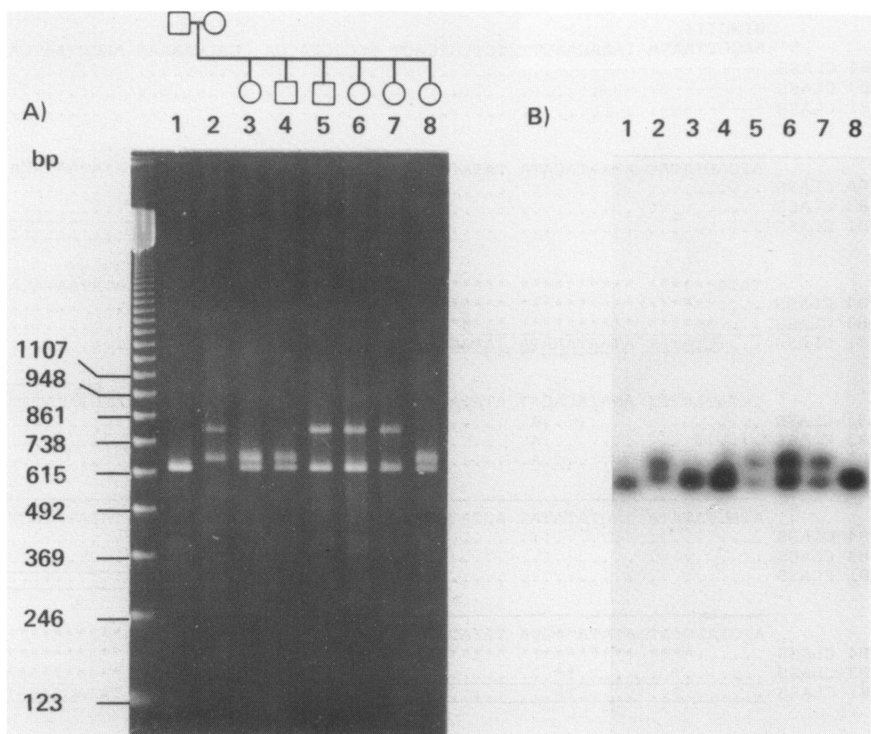


**Figure 4.** *Mendelian inheritance of 3′ IL6 alleles in one family.* The pedigree is drawn above the lanes of the ethidium bromide-stained agarose gel containing PCR products from the different family members (A). B) is traditional Southern blot analysis of the corresponding DNA using the probe pBeta2.15.

of the B1, B3 and B4 class are present in the family whose pedigree is shown in 4A. The father is homozygous for the B4 allele and the mother is a B1/B3 heterozygote. In addition, children in lanes 3,4 and 8 whose genotypes could not be resolved by conventional Southern analysis can be determined to be B3/B4 heterozygotes by PCR.

The PCR products from three individuals homozygous for three different classes of alleles (B1, B3 and B4 , shown in Figure 3, lanes 1,2 and 3 respectively) were sequenced. The sequence of the alleles is given in Figure 5. The sequence originally described (3) is very similar to the sequence of individual P105G who has a B3 allele, however, there are occasional differences due loss or insertion of one or two nucleotides (A and/or T).

The PCR products of the common chimpanzee and baboon were also sequenced directly. A comparison of these sequences with the original human sequence (of the B3 class) is shown in Figure 6.

## DISCUSSION

We have been able to successfully amplify the polymorphic 3′ flanking region of the IL6 gene that that is extremely rich in AT dinucleotides. Individuals with a variety of different genotypes have been successfully and reproducibly amplified, facilitating a rapid study of diseases associated with this gene. Successful amplification implies that highly repeated

```
         HINDIII
      5'-AAGCTTAATA TAAACAAGTT TCTTGTCACT GCCACCACCA CGACCAAAAA AAGCTAATCA
B4 CLASS ---------- -------... .......... .......... .......... ..........
B3 CLASS ---------- ---....... .......... .......... .......... ..........
B1 CLASS ---------. .......... .......... .......... .......... ..........

         ATCACTATAT ATAATACATA TATATACTAT ATATAATAAA TATATACTAT ATATAATACA
B4 CLASS .......... .......... .......... .......... .......... ..........
B3 CLASS .......... .......... .......... .......... .......... ..........
B1 CLASS .......... .......... ......[--- ---------- ---------- ----------

         TATA****** ********** ********** ********** **TACACTAT ATATAATACA
B4 CLASS ....****** ********** ********** ********** **........ ..........
B3 CLASS ....****** ********** ********** ********** **........ ..........
B1 CLASS ----[CTATAT ATAATAAATA TATACTATAT ATAATACATA TA]........ ..........

         TATATACTAT ATATACAC*T ATATACTATA TATACACATA TATATTATGA ATGTATATAT
B4 CLASS .......... ........A. .......... .......... .........T ..........
B3 CLASS .......... ........A. .......... .......... .........T ..........
B1 CLASS .......... ........A. .......... .......... .......... ..........

         ATAGTATATA TAGTATATAT ACTATGTATG TATATATA** GTATATATAG TATATATACT
B4 CLASS .......... .......... .......... ........TA .......... ..........
B3 CLASS .......... .......... .......... ........T* .......... ..........
B1 CLASS .......... .......... .........[. ........TA .......... ..........

         ATGTATG*AT ATATA**GTA TATATAGTAT ATATACTATG TATGT***** **********
B4 CLASS ......**** ********** ********** ********** ********** **********
B3 CLASS .......T.. ......**.. .......... .......... .....***** **********
B1 CLASS .......T.. ......TA.. .......... .......... .....]GTATA TATA[GTATAT

         ********** ********** ********** ********** ********** **********
B4 CLASS ********** ********** ********** ********** ********** **********
B3 CLASS ********** ********** ********** ********** ********** **********
B1 CLASS ATAGTATATA TAGTATATAT ACTATGTATG TATATATATA GTATATATAG TATATATACT

         ********GT ATATATAGTA TATATAGTAT ATATAGTATA TATACTATGT ATGTATATAT
B4 CLASS ********.. .......... .......... .......... .......... ..........
B3 CLASS ********.. .......... .......... .......... ......**.. ..........
B1 CLASS ATGTATGT].. .......... .......... .......... .......... ..........

         ATAGTATATA TAGTATATAT ACTGTGTATG TATATATATA GTATATATAT ACTATATATG
B4 CLASS .......... .......... .......... .......... .......... ..........
B3 CLASS .......... .......... .......... .......... .......... ..........
B1 CLASS .......... .......... .......... .......... .......... ..........

         CATACATAGT ATATATGCAT ATATACTATA TATACTATAT ATTTATATAT ACTATATACT
B4 CLASS .......... .......... .......... .......... .......... ..........
B3 CLASS .......... .......... .......... .......... .......... ..........
B1 CLASS .......... .......... .......... .......... .......... ..........

         ATATATACTA TATACTGTAT ATATACTATA TATGTATGTA TACGATATAT ATATATACTA
B4 CLASS .......... .......... .......... .......... .......... .......---
B3 CLASS .......... .......... .......... .......... .......... .......---
B1 CLASS .......... .......... .......... .......... .......... ..........

               RSAI
         TATATGTACA CACACACATA TATATATACA TATAAGCACC TACTACATGC CAGGCATCAT
B4 CLASS ---------- ---------- ---------- ---------- ---------- ----------
B3 CLASS ---------- ---------- ---------- ---------- ---------- ----------
B1 CLASS .......... .......... .......... .......... .......... ...-------
```

units can be faithfully replicated and that slippage of the Taq polymerase does not frequently occur. Regions rich in AT dinucleotides are proposed to be particularly susceptible to slippage of DNA polymerase (15). In addition, detection of only two amplification products from heterozygous individuals implies that the DNA polymerase does not frequently switch from one DNA strand to another during the amplification process; a phenomenon that has occasionally been observed (7). These results suggest that this approach may be generally useful in detecting polymorphisms due to VNTRs.

Two other human genes have been described which are both polymorphic and extremely AT rich in their 3′ flanking sequence; Apolipoprotein B (APOB, 16) and type II collagen (17). A PCR system has been developed for the APOB gene which has many of the advantages described for this system (18). A PCR system for the detection of a VNTR segment detected by the probe pYNZ22 (HGM locus D17S30) has also recently been employed (23). However, for both the APOB and D17S30 systems other faint bands are reported to sometimes appear after amplification, making discrimination of the alleles difficult. In the case of IL6, only the required fragments were amplified. Length polymorphisms generated in (dC-dA)n rich regions have also been detected with PCR (19,20). However, in these instances, the alleles are small and need to be resolved on polyacrylamide DNA sequencing gels. It has also been demonstrated that hypervariable minisatellite regions can be detected with PCR (21). However, the alleles cannot be visualized directly on ethidium bromide stained agarose gels, but have to be detected with specific minisatellite probes. The IL6 VNTR polymorphism described represents one of the easiest systems to amplify since it is so specific and efficient. The reason for this is not clear, but may relate to the base composition of the region being amplified versus that of the primers.

These results confirm that the different alleles within the highly polymorphic region 3′ to the IL6 gene are due to the variable length of the AT rich sequence. Within the middle of the AT rich region each of the alleles differs due to variable numbers of a tandem repeat sequence. An allele of the B4 class ressembles that of the B3 class, but has lost 37bp of a sequence ($G(TA)^4G(TA)^4G(TA)^4$CTATGTATGT) that could be considered part of a tandemly repeated unit. This sequence, and the preceding 37bp, which are very similar to this unit have been duplicated in tandem in the allele of the B1 class.

In addition, the B1 allele has a duplication at the 5′ end of the AT rich sequence of a 38bp sequence found in B3 and B4. This latter sequence contains a putative poly A addition site. Whether either of these poly A addition sites are functional is not known. The mechanism that gave rise to these alleles is not understood. However, a newly generated allele at the D17S30 locus has been described (22). This resulted from the loss of a repeat unit without exchange of flanking DNA sequences between homologous chromosomes, suggesting that illegitimate recombination between tandemly repeated regions had not occurred, but that the new allele had been generated by another mechanism such as unequal sister chromatid exchange, slippage at DNA replication (15) or loopout deletion.

**Figure 5.** *Sequence data for three alleles from the polymorphic flanking region of IL6.* Sequences are represented by dots when identical and by the alternative base where they are different to the published sequence (3) which is an allele of the B3 class. The asterisks represent bases not present. Dashed lines represent sequence not obtained. Arrowheads above the sequence delineate the tandem repeats. The regions in the B1 class of allele that are involved in gene duplication are boxed (broken lines for the first duplication; unbroken lines for the second duplication). The putative poly-A addition sites are underlined.

```
          HINDIII
HUMAN     AAGCTTAATA TAAACAAGTT TCTTGTCACT GCCACCACCA CGACCAAAAA AAGCTAATCA
CHIMP     -------... .......... .......... .......... .......... ..........
BABOON    ---------- ---------- ---....... .......... .C........ ......TAAT

HUMAN     ATCACTATAT ATAATACATA TATATACTAT ATATAATAAA TATATATACT ATATATAATA
CHIMP     .......... .......... ....**.... ........C. **........ ......T..*
BABOON    CAATCACATT TTTTTTATAT ATATATACAT ATATA***** ********** **********

HUMAN     CATATA**** ********** ******TACA CTATATATAA TACATATATA CTATATATAC
CHIMP     *.....CTAT ATATAATACA TATATA**.. .......... .......... ..........
BABOON    ********** ********** ********** ********** ********** **********

HUMAN     ACTATATACT ATATATACAC ATATATATTA TGAATGTATA TATATAGTAT ATATAGTATA
CHIMP     ..******** ********** ********** ********** ********** **********
BABOON    ********** ********** ********** ********** ********** **********

HUMAN     TATACTATGT ATGTATATAT AGTATATATA GTATATATAC TATGTATGAT ATATAGTATA
CHIMP     ********** ********** ********** ********** ********** **********
BABOON    ********** ********** ********** ********** ********** **********

HUMAN     TATAGTATAT ATACTATGTA TGTGTATATA TAGTATATAT AGTATATATA GTATATATAC
CHIMP     ********** ********** ********** ********** ********** **********
BABOON    ********** ********** ********** ********** ********** **********

HUMAN     TATGTATGTA TATATATAGT ATATATAGTA TATATACTGT GTATGTATAT ATATAGTATA
CHIMP     ********** ********** **.C...TAC ......TGTA TATATAT... ...C.TAC..
BABOON    ********** ********** ********** ********** ********** **********

HUMAN     TATATACTAT ATATGCATAC ATAGTATATA TGCATATATA CTATATATAC TATATATTTA
CHIMP     .......... .....T...T ...C...... ..T....... .......GTA ......C.AT
BABOON    ********** ********** ********** ********** ********** **********

HUMAN     TATATACTAT ATACTATATA TACTATATAC TGTATATATA CTATATATGT ATGTATACGA
CHIMP     ATATACTATA TATG...... .......... .A........ .......... ........T.
BABOON    ********** ********** ********** ********** ********** **********

HUMAN     TATATATATA TACTATATAT GTACACACAC ACATATATAT ATACATATAA GCACCTACTA
CHIMP     .......... CTA....... A......... .......... .C........ ..........
BABOON    ********** ********** ********** ********** ********** ..G.......

HUMAN     CATGCCAGGC ATCATTAAAT GTGTTGCATC CATCACGTC
CHIMP     .......--- ---------- ---------- ---------
BABOON    ..*..*.... .......... .GTG...... ..-------
```

**Figure 6.** *Aligned DNA sequences of the flanking region of IL6 for human (Homo sapiens), chimp (Pan troglodoytes) and baboon (Papio cynocephalus).* The functions of dots, dashes and asterisks are as described for Figure 5. The sequence homologous to the right primer is underlined.

Smaller differences between the three IL6 alleles, and between these alleles and the sequence originally described were also detected. These were due to the occasional loss of one or two bases throughout the AT rich repeat which could be due to slippage of the DNA polymerase in-vivo. When two alleles of the B4 class were compared they were also shown to differ due to loss of one base (not shown). It has been demonstrated that regions rich in (CA) repeats can be highly polymorphic, and thus highly informative for linkage studies (19,20) due to variations in repeat number. This study suggests that (AT) rich regions are also likely to be highly polymorphic. As in the case of the polymorphic (CA) repeat containing loci, the alleles would need to be resolved on DNA sequencing gels.

The frequencies of the different classes of alleles differ in different human populations (3). The B4 class is most frequent in Caucasoids with a frequency of 0.84. This class is present at frequencies of between 0.00−0.14 in other populations, where the B3 class of allele is the most common (with frequencies of between 0.68−0.95). The other alleles are rarer, and alleles other than those of the B1, B2, B3 and B4 classes have been seen most commonly in Africa.

Since the human oligonucleotide primers amplified DNA from non-human primates, we determined the sequence of this region in the common chimpanzee and in an olive baboon. Both species were remarkably similar to the human sequence in the regions flanking the AT rich repeat. The chimpanzee sequence showed far more identity with the AT rich human sequence than did that of the baboon. The baboon DNA could be amplified and sequenced directly using the left amplification primer as a sequencing primer, even though the right amplification primer was homologous to the human sequence at only 16 out of 19 bases. The baboon sequence was only 164 bp long, and lacks the repeated units seen in both the human and chimpanzee sequence. The baboon, however, also has an AT rich region. The sequence of the common chimpanzee, like the sequence of the human B1 allele had a duplication in the 5′ region, but the three chimpanzees sequenced had an 'AATACA' sequence in place of the putative poly A addition 'AATAAA' sequence observed in humans.

The role of the AT rich repeat sequence and of the different alleles is not known. It is possible that different alleles have different effects on RNA splicing, gene regulation or RNA stability. Although small differences in length are detected within one class of allele, it is suprising that more classes of alleles are not observed, given the proclivity of tandem repeats to undergo rearrangement.

## ACKNOWLEDGMENTS

We acknowledge support for the following grants: NIH grant GM 28428; to P.B.S., research grant AI-16262 from the NIAID and a contract from the National Foundation for Cancer. Primate samples were provided by the Yerkes Regional Primate Research Center and the Southwest Foundation for Biomedical Research. Thanks are due to Mrs. Berta Young for technical assistance and to Luca Cavalli-Sforza for helpful discussions and invaluable support.

*To whom correspondence should be addressed

## REFERENCES

1. Sehgal, P.B., Zilberstein, A., Ruggieri, R., May, L.T., Ferguson-Smith, A., Slate, D.L., Revel, M., and Ruddle, F.H. (1986) Proc. Natl. Acad. Sci., U.S.A., 83, 2529−2537.
2. Ferguson-Smith, A., C., Chen, Y.-F., Newman, M.S., May, L.T., Sehgal, P.B., and Ruddle, F.H. (1988) Genomics, 3, 8−16.
3. Bowcock, A.M., Kidd, J.R., Lathrop, G.M., Daneshvar, L., May, L.T., Ray, A., Sehgal, P.B., Kidd, K.K. and Cavalli-Sforza, L.L. (1988) Genomics, 3, 8−16.
4. Southern, E.M. (1975) J. Mol. Biol., 98, 503−517.
5. Saiki, R., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A. and Arnheim, N. (1985) Science, 230, 1350−1354.
6. Mullis, K.B. and Faloona, F.A. (1987) Methods Enzymol., 155, 335.
7. Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988) Science, 230, 487−494.
8. Saiki, R.K., Bugawan, T.L., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1986) Nature, 324, 163−166.

9. Nakamura, Y., Leppert, M., O'Connell, P., Wolff, R., Holm, T., Culver, M., Martin, C., Fujimoto, E., Hoff, M., Kumlin, E. and White, R. (1986) Science, 235, 1616–1622.
10. Yasukawa, K., Hirano, T., Watanabe, Y., Muratani, K., Matsuda, T., Nakai, S. and Kishimoto, T. (1987) EMBO J., 6, 2939–2945.
11. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. U.S.A., 74, 5463–5467.
12. Feder, J., Yen, L., Wijsman, E., Wang, L., Wilkins, L., Schroder, J., Spurr, N., Cann, H., Blumenberg, M. and Cavalli-Sforza, L.L. (1985) Am. J. Hum. Genet., 37, 635–649.
13. Feinberg, A.P. and Vogelstein, B. (1983) Anal. Biochem., 132, 6–13.
14. Gyllensten, U. B. and Erlich, H.A. (1988) Proc. Natl. Acad. Sci. U.S.A., 85,7652–7656.
15. Kornberg, A. (1974) DNA Replication, Freeman, San Francisco, pp 146–150.
16. Huang,L.-S. and Breslow, J.L. (1987) J. Biol. Chem., 262, 8952–8955.
17. Stoker, N.G., Cheah, K.S.E., Griffin, J.R., Pope, F.M. and Solomon, E. (1985) Nucleic Acids Res., 13, 4613–4622.
18. Boerwinkle, E., Xiong, W., Fourest, E. and Chan, L. (1989) Proc. Natl. Acad. Sci. U.S.A., 86, 212–216.
19. Weber, J.L. and May, P.A. (1989) Am. J. Hum. Genet., 44, 388–396.
20. Litt, M. and Luty, J.A. (1989) Am. J. Hum. Genet., 44, 397–401.
21. Jeffreys, A.J., Wilson, V., Neumann, R. and Keyte, J. (1988) Nucleic Acids Res., 16,10953–10971.
22. Wolff, R.K., Nakamura, Y. and White, R. (1988) Genomics, 3, 347–351.
23. Horn, G.T., Richards, B. and Klinger, K.W. (1989) Nucleic Acids Res., 17, 2140.

**This article, submitted on disc, has been automatically converted into this typeset format by the publisher.**