Vol. 189, No. 1, 1992 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
November 30, 1992 Pages 154-164

Animal Histo-blood Group ABO Genes

Yoshihiko Kominato, Patricia D. McNeill, Miyako Yamamoto, Mai Russell, Sen-itiroh Hakomori, and Fumi-ichiro Yamamoto*

The Biomembrane Institute, 201 Elliott Ave W, Seattle, WA 98119; and Dept. of Pathobiology, University of Washington, Seattle, WA

Received September 29, 1992

SUMMARY: Sequences homologous to the human histo-blood group ABO genes are present in the genomic DNA of various mammals. We have PCR-amplified, subcloned, and sequenced a portion of these genes from several species of primates and found high conservation of the nucleotide as well as the deduced amino acid sequences during evolution.

The human histo-blood group ABO system has been a major focus in transfusion medicine. We cloned cDNA encoding A transferase (1) based on the partial a.a. sequence of the purified enzyme (2), and elucidated the molecular genetic basis of the ABO system by subsequent cDNA cloning of B and O allelic cDNAs followed by nucleotide sequencing of the isolated cDNA clones (3). We extended our study to the molecular bases of subtypes (A^1, A^2, A^3, and B^3) (4, 5) and *cis*-AB (6).

ABH substances are known to be present in various living species (7). Bernstein's model of inheritance (8) applies not only to man; it also fits almost every simian primate. New World monkeys and Old World monkeys have A, B, and H substances in their secretions, depending on their genotypes (9). In anthropoid apes such as orangutans and chimpanzees, these substances are present on the red cells, as well as in secretions. (In the gorilla, the quantity of antigens present on the red cells is considerably lower than in the other anthropoid species). In almost all of these primates, antibodies are present in the serum of the animal for which the corresponding antigens are absent: Landsteiner's Law also applies here.

*To whom correspondence should be addressed.

0006-291X/92 $4.00

In humans there are at least two other sequences homologous to the ABO genes. Those are the α1→3 galactosyltransferase pseudogene (α1→3GalT) (10,11) and the hgt4 which we named for the human glycosyltransferase group 4 sequence (12). Based on differences in the degree of homology, a possible evolutionary pathway of these sequences has been postulated.

Here we report the identification of ABO genes in various species of mammals. Furthermore, we sequenced a portion of the ABO genes from several species of primates. The portions of primate ABO genes that we sequenced corresponded with part of the coding sequence in the last coding exon in human ABO genes. Comparison of the nucleotide and deduced amino acid sequences with those of human A and B genes reconfirms the importance of the third and the fourth positions of the four known amino acid substitutions found between human A and B transferases which determine their different nucleotide-sugar specificities (13).

MATERIALS AND METHODS

Materials

Saliva and blood specimens from several individual baboons were obtained from the Regional Primate Research Center, University of Washington. EVO BLOT and PCRable DNA from the other species were purchased from BIOS, Inc. (New Haven, CT). Taq DNA polymerase for PCR was purchased from Perkin-Elmer Cetus (Norwalk, CT). Oligodeoxynucleotides were custom-synthesized at Bio-synthesis, Inc (Lewisville, TX). Geneclean Kit was from BIO 101 (La Jolla, CA). Sequencing vector pT7T3U18 was purchased from Pharmacia-LKB (Piscataway, NJ). *E. coli XL-1-blue* frozen competent bacteria were from Stratagene (La Jolla, CA) and α-^{32}P dATP was from Amersham Corp. (Arlington Heights, IL). DNA sequencing was performed with Sequenase sequencing kits (United States Biochemicals, Cleveland, OH).

Probe preparation and Southern hybridization

The Eco RI fragment from FY-59-5 (1), which contains the entire coding sequence of human A transferase, was labelled using the PCR amplification method with internal primers, fy-127 and fy-106, under the same conditions described below except ^{32}P-dATP was used in place of non-radioactive dATP. EVO BLOTs (BIOS) were prehybridized in 50% formamide, 5X SSPE, 5X Denhardt's, and 0.1% SDS solution at 42°C for 8 hours, then hybridized overnight at 42°C with a ^{32}P PCR-labeled probe. The filters were washed three times in 2X SSC, 0.1% SDS at room temperature and then once in 1X SSC, 0.1% SDS at 68°C for 1 hour.

Polymerase chain reaction and nucleotide sequencing

PCR (14) was performed per manufacturer's protocol, using a DNA-Thermal Cycler from Perkin-Elmer Cetus. One μg of genomic DNA was used as a template for the amplification reaction. The pairs of synthetic oligodeoxynucleotide primers used for PCR amplification were (fy-81 and fy-113/fy-156) and (fy-127 and fy-106). The nucleotide sequences of these oligos are:

fy-81: CGGAATTCA(A/T)G(T/C)ACTTCATGGT(G/T)GGCCA,
fy-106: CGGAATTCGAACCTCAGCTTCCTCAGGA,
fy-113: CGGAATTCACCTCTTGCACCGACCC,
fy-127: CGGAATTCCTGGTGTGCGTGGAC, and
fy-156: CGGAATTCACCTCCTGCACCGACCC.

All the Eco RI sites are artificial to facilitate the following subcloning. fy-81 is an oligodeoxynucleotide primer degenerated at the three positions in parentheses. Five μl of 20pmol/μl oligos each were added to the heat-denatured DNA, followed by addition of the reaction mixture (45 μl H_2O, 10 μl 10x reaction buffer, 8 μl 2.5mM dNTP mixture, and 0.5 μl 2.5 units/μl Taq DNA polymerase). Two drops of paraffin oil were overlaid to prevent evaporation. Amplification was performed in a step-cycle mode of 40 rounds of 94°C for 2 min, 50°C for 2 min, and 72°C for 3 min, followed by one round of 94°C for 2 min, 55°C for 3 min, and 72°C for 10 min. The samples were then left at 10°C until processing occurred. Amplified DNA was extracted with 100 μl of a phenol:chloroform:iso-amyl alcohol mixture (25:24:1), and the aqueous fraction was transferred into Eppendorf tubes with 12 μl 3M sodium acetate (pH 7.5) and 250 μl ethanol. After centrifugation, the pellet was dried, resuspended, and then subjected to restriction enzyme digestion with Eco RI. DNA was then electrophoresed through a 2% agarose gel for size fractionation, recovered from gel fragments with Geneclean Kit, ligated with Eco RI-digested, BAP-treated pT7T3U18 sequencing vector, and used for DNA transformation of *E. coli XL-1-blue* strain competent bacteria. DNA from transformant clones was analyzed for the inserts. DNA from a multiple number of correct constructs was individually alkaline-denatured, and used for nucleotide sequencing by the Sanger dideoxy termination method with the Sequenase kit (15).

RESULTS

We examined whether sequences corresponding to the human ABO genes existed in other species of animals. Southern hybridization of the Zoo blot was performed with a ^{32}P-labelled human ABO gene probe under conditions where only the ABO genes, and not the other two homologous sequences (α1→3 GalT pseudogene and hgt4), could be hybridized in human genomic DNA. This result is shown in figure 1. Hybridization was observed with genomic DNA from marmoset, hamster, rat, mouse, sheep, cow, rabbit, cat, and dog. However, no hybridization was detected in any of the animals examined which are considered lower than mammals on the evolutionary tree, except for the obscure result with chicken genomic DNA.

We determined the partial nucleotide and deduced amino acid sequences of the primate ABO genes because of our interest in evolutionary genetics. We employed a PCR approach based on the fact that amplification can be made if the sequences chosen for the primers are well conserved. We tried several combinations of oligos and found that two pairs of oligos (fy-81 and fy-113/156, and fy-127 and fy-106) worked well. These two amplified fragments cover nucleotides, numbered from 435 to 813 and from 634 to 1003 respectively, of the human A transferase coding sequence and overlap each other (1,4). We first analyzed genomic DNA from several baboons because we had easy access to their blood and saliva specimens. In baboon, a species of Old World monkey, the ABO blood group substances are reported to exist in secretions but not on the red cells. Therefore, ABO typing was performed by testing monkey saliva for its ability to inhibit anti-A and anti-B reagents, and by testing monkey sera after absorption with human O red cells, for its ability

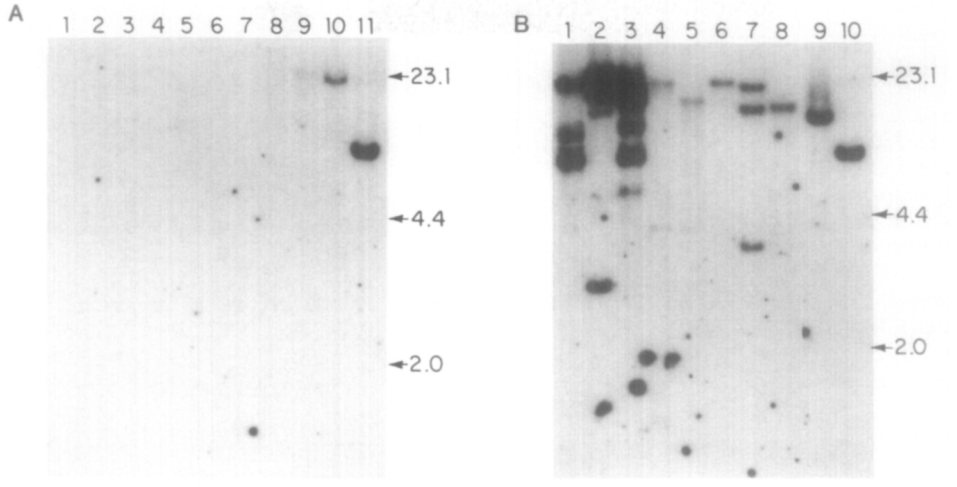

Figure 1. Southern hybridization
BIOS EVO BLOTS were hybridized with the human ABO gene probe. The amount of DNA used was 50ng, 0.5μg, 1.25μg, 2.5μg for bacteria, yeast, nematode, and fly, respectively, and 8μg for all the other species. DNA was Eco RI digested and size fractionated by electrophoresis through 1.0% agarose gel. The lane numbers and names of species are as follows:
(A) Genetic model blot
1: bacteria (*Escherichia coli*), 2: yeast (*Saccharomyces cerevisiae*), 3: nematode (*Caernorhabditis elegans*), 4: fly (*Drosophila melanogaster*), 5: frog (*Xenopus laevis*), 6: sea urchin (*Strongylocentrotus purpuratus*), 7: clam (*Mercenaria mercenaria*), 8: lobster (*Homarus americanus*), 9: chicken (*Gallus domesticus*), 10: mouse (*Mus musculus*), and 11: human (*Homo sapiens*).
(B) Mammalian blot
1: dog (*Canis familiaris*), 2: cat (*Felus catus*), 3: rabbit (*Oryctolagus cuniculatus*), 4: cow (*Bovis domesticus*), 5: sheep (*Ovis aries*), 6: mouse, 7: rat (*Rattus norvegicus*), 8: hamster (*Mespcricetus auratus*), 9: marmoset (*Saiguinus oedipus*), and 10: human.
The positions of the marker bands (lambda DNA-Hind III digest) are indicated.

to agglutinate human groups A and B red cells. DNA from baboons with three phenotypes (A, B, and AB) was used for separate amplifications. The nucleotide sequences of these amplified fragments were determined after being subcloned into the sequencing vector and correlated to the ABO genotypes. The results are shown as baboon A and B in figure 2. As we were unable to obtain the blood and saliva specimens from other species of primates, we used commercially available genomic DNA (PCRable DNA, BIOS). Because information on the ABO genotypes of these animals was not available, we determined the nucleotide sequences in the same region of the ABO genes, corresponding to part of the last coding exon in human ABO genes, from one individual each of chimpanzee, gorilla, orangutan (all anthropoid apes), and macaque (Old Word monkey) for comparison, rather than trying to correlate with ABO genotypes. These results are also shown in figure 2. The two genes on the two homologous chromosomes were distinguishable for the chimpanzee and the orangutan, numbered 1 and 2, showing that these individual apes were heterozygous at this

```
human A     CCGTGTCCACTACTATGTCTTCACCGACCAGCCGGCCGCGGTGCCCCGCGTGACGCT    491
human B
chimp.1                                     A     A
chimp.2                                     A     A
gorilla
orang.1
orang.2
macaque       G           C                       T         G       G
baboon A      C           C                       T         G       G
baboon B      C           C                       T         G       G

human A       R  V  H  Y  Y  V  F  T  D  Q  P  A  A  V  P  R  V  T  L    164
human B
chimp.1
chimp.2
gorilla
orang.1
orang.2
macaque                                                             A
baboon A                                                            A
baboon B                                                            A

human A     GGGGACCGGTCGGCAGCTGTCAGTGCTGGAGGTGCGCGCCTACAAGCGCTGGCAGGA    548
human B                                     G
chimp.1                     G
chimp.2                     G
gorilla                     G
orang.1                     G       G
orang.2                     G       G
macaque                     G     T G             T
baboon A                    G     T G             T
baboon B                    G     T G             T

human A       G  T  G  R  Q  L  S  V  L  E  V  R  A  Y  K  R  W  Q  D    183
human B                                     G
chimp.1
chimp.2
gorilla
orang.1                                 G
orang.2                                 G
macaque                                 G
baboon A                                G
baboon B                                G

human A     CGTGTCCATGCGCCGCATGGAGATGATCAGTGACTTCTGCGAGCGGCGCTTCCTCAG    605
human B
chimp.1                                       C
chimp.2                                       C
gorilla
orang.1
orang.2
macaque                                 C           A
baboon A                                C
baboon B                                C
```

Figure 2. Comparison of the nucleotide and deduced amino acid sequences
Genomic DNA from several species of primates was used for separate PCR amplification and the DNA sequence was determined after subcloning the amplified fragments into sequencing plasmid vector pT7T3U18. Comparison of the nucleotide and the deduced amino acid sequences of the genomic DNA surrounded by two primers, fy-81 and fy-106, with those of human A(A^1) and B genes are shown. "Baboon A" and "B" denote A and B alleles from baboons (*Papio cynocephalus*). Chimp. and orang. denote chimpanzee (*Pan troglodytes*) and

```
human A      V  S  M  R  R  M  E  M  I  S  D  F  C  E  R  R  F  L  S    202
human B
chimp.1                                                 Q
chimp.2                                                 Q
gorilla
orang.1
orang.2
macaque                                                    Q
baboon A
baboon B

human A    CGAGGTGGATTACCTGGTGTGCGTGGACGTGGACATGGAGTTCCGCGACCACGTGGG     662
human B                                                          T
chimp.1
chimp.2
gorilla
orang.1                  A                                 T
orang.2                  A                                 T
macaque                                                    T
baboon A                       C
baboon B                       C

human A      E  V  D  Y  L  V  C  V  D  V  D  M  E  F  R  D  H  V  G    221
human B
chimp.1
chimp.2
gorilla
orang.1
orang.2
macaque
baboon A                       A
baboon B                       A

human A    CGTGGAGATCCTGACTCCGCTGTTCGGCACCCTGCACCCCGGCTTCTACGGAAGCAG     719
human B                                          A
chimp.1                                          T
chimp.2                                          T
gorilla
orang.1                                                       C
orang.2                                                       C
macaque                  A                       C     T
baboon A                 A                       C
baboon B                 A                     (C)   (T)

human A      V  E  I  L  T  P  L  F  G  T  L  H  P  G  F  Y  G  S  S    240
human B                                          S
chimp.1
chimp.2
gorilla
orang.1                                                       T
orang.2                                                       T
macaque                                          A
baboon A                                         A
baboon B                                       (A)
```

orangutan (*Pongo pygmaeus*). The numbers 1 and 2 correspond to the two genes on the two homologous chromosomes identified by the differences in the nucleotide sequences. Two separate genes were not distinguishable in gorilla (*Gorilla g. gorilla*) and macaque (*Macaca fascicularis*). Only differences from the human A gene are shown for each case. Polymorphic differences from the human A gene found in baboon B alleles are shown in parentheses.

```
human A     CCGGGAGGCCTTCACCTACGAGCGCCGGCCCCAGTCCCAGGCCTACATCCCCAAGGA     776
human B
chimp.1
chimp.2
gorilla
orang.1                                                   T
orang.2                                                   T
macaque
baboon A
baboon B

human A       R  E  A  F  T  Y  E  R  R  P  Q  S  Q  A  Y  I  P  K  D     259
human B
chimp.1
chimp.2
gorilla
orang.1
orang.2
macaque
baboon A
baboon B

human A     CGAGGGCGATTTCTACTACCTGGGGGGGTTCTTCGGGGGGTCGGTGCAAGAGGTGCA     833
human B                         A     C
chimp.1     T                                 A
chimp.2                                       A
gorilla                         A     C
orang.1                                       A
orang.2                                       A
macaque           T             T                             G
baboon A          T             T                 A           G
baboon B          T             A     C                       G

human A       E  G  D  F  Y  Y  L  G  G  F  F  G  G  S  V  Q  E  V  Q     278
human B                         M     A
chimp.1
chimp.2
gorilla                         M     A
orang.1
orang.2
macaque
baboon A
baboon B                        M     A

human A     GCGGCTCACCAGGGCCTGCCACCAGGCCATGATGGTCGACCAGGCCAACGGCATCGA     890
human B
chimp.1
chimp.2
gorilla
orang.1                   A
orang.2
macaque
baboon A
baboon B
```

Figure 2 - Continued

gene locus. We were, however, unable to identify two separate alleles in the gorilla and the macaque in spite of sequencing multiple clones from each amplification, suggesting that the animals whose DNA was used for templates happened to be homozygous at the ABO locus.

```
human A      R  L  T  R  A  C  H  Q  A  M  M  V  D  Q  A  N  G  I  E    297
human B
chimp.1
chimp.2
gorilla
orang.1                  T
orang.2
macaque
baboon A
baboon B

human A    GGCCGTGTGGCACGACGAGAGCCACCTGAACAAGTACCTGCTGCGCCACAAACCCAC     947
human B                                          A
chimp.1
chimp.2
gorilla
orang.1
orang.2
macaque
baboon A
baboon B

human A      A  V  W  H  D  E  S  H  L  N  K  Y  L  L  R  H  K  P  T    316
human B
chimp.1
chimp.2
gorilla
orang.1
orang.2
macaque
baboon A
baboon B

human A    CAAGGTGCTCTCCCCCGAGTACTTGTGGGACCAGCAGCTGCTGGGCTGGCCCGCCG     1003
human B
chimp.1                                                    T
chimp.2                                                    T
gorilla                                                    A
orang.1
orang.2
macaque                        C                           T  G
baboon A                       C                           T  G
baboon B                       C                           T  G

human A      K  V  L  S  P  E  Y  L  W  D  Q  Q  L  L  G  W  P  A      334
human B
chimp.1                                                       S
chimp.2                                                       S
gorilla                                                       T
orang.1
orang.2
macaque
baboon A
baboon B
```

Figure 2 - Continued

Another possibility could be that one allele was more selective for the primers we used and the other was not specifically amplified. A minimum 95% homology was conserved among all these primates including humans in the nucleotide and the deduced amino acid sequences.

DISCUSSION

A and B transferase activities in various mammals have been reported (16). Here we have shown that ABO genes are present in the genomic DNA of various mammals, which may account for those enzymatic activities. Although we have not detected any bands in lower animals, they may have genes that exist with lesser homology considering the universal presence of ABH substances in nature.

We have determined the partial nucleotide sequences of several primate ABO genes. The locations of the four amino acid substitutions which discriminate human A and B transferases (a.a. 176, 235, 266, and 268) are present in the region where the sequences were determined. We were able to correlate baboon A and B alleles to their corresponding nucleotide and deduced amino acid sequences. Polymorphic changes were observed among baboon B alleles and are shown in parentheses in figure 2. Again, only the differences from human A alleles are shown. Except for these polymorphisms, only three common nucleotide substitutions (nt. 796, 803, and 813) were observed between baboon A and B alleles in this region. Among these three, two result in amino acid substitutions. These positions (a.a. 266 and 268) correspond to the third and fourth positions of amino acid substitutions found between A and B transferases in humans. The amino acid residues of baboon A and B transferases at the locations of the first and second amino acid substitutions of human A and B transferases are identical to one another (arginine and alanine, respectively), which suggests that these amino acid residues are not important in determining different nucleotide-sugar specificities between A and B transferases. In addition, amino acid residues at the third and fourth amino acid substitutions were found to be conserved in both A and B alleles between humans and baboons (leucine and glycine in A transferase for both species and methionine and alanine in B transferase for both species). These results led us to conclude that these amino acid substitutions are crucial for the different donor nucleotide-sugar specificities between A (GalNAc) and B (galactose) transferases. We previously obtained the same conclusion through the construction of 16 kinds of A-B transferase combinatorial chimeras at these four amino acid substitution locations and their expression in DNA transfected HeLa cells (13). Except for baboons, we do not know the ABO phenotype, much less genotype, of each monkey whose genomic DNA was used for PCR amplification. However, based on the amino acid residues at the third and fourth amino acid substitutions, we can correctly designate these alleles to encode A or B transferase, unless they are non-functional O alleles. (There may be mutations somewhere outside of these determined sequences, which impair the activity of transferases.) For example, the gorilla used for this experiment may have a B allele but no A allele. It is noteworthy that the B allele of the gorilla is more homologous to the human B allele than the A allele of the chimpanzee is to

the A allele of humans, in spite of the fact that the amount of ABH antigens on the red cells in gorilla is rather small, and the gorilla is special among anthropoid apes in this respect (9). We have determined the partial nucleotide sequence of only one individual per species except for baboon. Therefore, a possibility remains that some of the identified differences in the nucleotide sequence are polymorphisms and not allele-specific. The presence of other alleles such as O alleles in chimpanzees and B alleles in orangutans, has also been reported (9). Information on the complete nucleotide sequences of the coding regions of cDNAs for each allele may be necessary in the future.

Because there are more than ten nucleotide positions which are shared among the Old World monkey sequences (such as nt. 438 and 450), it is very probable that the A and B alleles of baboons appeared after the Old World monkey lineage diverged from the hominoid lineage. This result suggests that the A and B alleles of baboons appeared independently from those of humans.

Although the ABO genes are not indispensable, the presence and high conservation of these genes as polymorphisms strongly support the meaningfulness of these genes among the population in mammals. Because ABH antigens on red cells are restricted to anthropoid apes and can more often be found on cells from the gastro-intestinal tract, the functionality of ABO genes, A and B transferases, or ABH antigens, if any, may be found related to these cells rather than red cells.

ACKNOWLEDGMENTS

We are indebted to Dr. Naruya Saitou (National Institute of Genetics, Japan) for his helpful suggestions. We thank Jennifer Stoeck for scientific editing and preparation of the manuscript. This study was supported by funds from The Biomembrane Institute, in part under a research contract from Otsuka Pharmaceutical Co. Y.K. is a postdoctoral fellow supported from Ministry of Education, Science, and Culture, Japan.

REFERENCES

1. Yamamoto, F., Marken, J., Tsuji, T., White, T., Clausen, H., and Hakomori, S. (1990) J. Biol. Chem. **265**, 1146-1151.
2. Clausen, H., White, T., Takio, K., Titani, K., Stroud, M.R., Holmes, E.H., Karkov, J., Thim, L., and Hakomori, S. (1990) J. Biol. Chem. **265**, 1139-1145.
3. Yamamoto, F., Clausen, H., White, T., Marken, J., and Hakomori, S. (1990) Nature **345**, 229-233.
4. Yamamoto, F., McNeill, P.D., and Hakomori, S. Biochem. Biophys. Res. Commun. **187**, 366-374.
5. Yamamoto, F., McNeill, P.D., Yamamoto, M., Hakomori, S., Harris, T., Judd, W.J., and Davenport, R.D. Vox Sang. in press.
6. Yamamoto, F., McNeill, P.D., Kominato, Y., Yamamoto, M., Hakomori, S., Ishimoto, S., Nishida, S., Shima, M., and Fujimura Y. Vox Sang. in press.
7. Kabat E.A. (1956). Blood group substances, Academic Press, New York.
8. Bernstein, F. (1924). Klin. Wochenschr. **3**, 1495-1497.

9. Moor-Jankowski, J., and Wiener, A.S. (1972). In: Pathology of Simian Primates, Part I, pp. 270-317, Karger, Basel.
10. Joziasse, D.H., Shaper, J.H., van den Eijnden, D.H., Van Tunen, A.J., and Shaper, N.L. (1989) J. Biol. Chem. **264**, 14290-14297.
11. Larsen, R.D., Rajan, V.P., Ruff, M.M., Kukowska-Latallo, J., Cummings, R.D., and Lowe, J.B. (1989) Proc. Natl. Acad. Sci. USA **86**, 8227-8231.
12. Yamamoto, F., McNeill, P.D., and Hakomori, S. (1991) Biochem. Biophys. Res. Commu. **175**, 986-994.
13. Yamamoto, F. and Hakomori, S. (1990) J. Biol. Chem. **265**, 19257-19262.
14. Saiki, R.K., Scharf, S.J., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A., and Arnheim, N. (1985) Science **230**, 1350-1354.
15. Sanger, F., Miklen, S., and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. **74**, 5463-5467.
16. Watkins,W.M. (1980) In: Advances in Human Genetics (Harris, H., and Hirschhorn, K., eds.) vol. 10, pp.1-136, Plenum Publishing Corp., New York.