Polymorphisms in the Human High Sulfur Hair Keratin-associated Protein 1, KAP1, Gene Family*

Yutaka Shimomura‡, Noriaki Aoki, Jürgen Schweizer§, Lutz Langbein¶, Michael A. Rogers‡, Hermelita Winter§, and Masaaki Ito

From the Department of Dermatology, Niigata University School of Medicine, Niigata 951-8510, Japan and the Divisions of *Cell Biology and ¶Tumor Cell Regulation, German Cancer Research Center, Heidelberg D-69120, Germany

Hair fiber differentiation and maturation involves the close interaction between hair keratins and their associated proteins, KAPs. Recently, a cluster of seven human KAP multigene families has been identified on chromosome 17q12-21 among which were four hKAP1 genes (hKAP1.1B, hKAP1.3, hKAP1.4, and hKAP1.5). In addition, there were previous as well as recent reports on four additional hKAP1 genes (hKAP1.1A, hKAP1.2, hKAP1.6, and hKAP1.7) with unknown chromosomal location. In this study, we have analyzed these eight hKAP1 genes in unrelated Japanese and Caucasian individuals and discovered that hKAP1.1A, hKAP1.6, and hKAP1.7 represent size polymorphisms of the hKAP1.1B gene. In addition, we show that hKAP1.2 as well as three hitherto unknown genes (hKAP1.8A, hKAP1.8B, and hKAP1.9) are size polymorphisms of the hKAP1.3 gene. In contrast, no polymorphic alleles were found for the hKAP1.4 and hKAP1.5 genes. We provide evidence that the polymorphic hKAP1.1B and hKAP1.3 alleles arose mainly by intragenic deletion and/or duplication events of distinct pentapeptide repeats typical for hKAP1 genes. We also demonstrate the occurrence of both frequent and rare population-specific hKAP1.1B and hKAP1.3 alleles, which were obviously generated after the divergence of the Caucasian and Japanese lineages.

In addition, by means of a pan-hKAP1 antibody, we confirm the previous hKAP1 family mRNA localization data in the middle to upper cortex of the human anagen hair follicle.

The major structural proteins of mammalian hair are the hair keratins and their associated proteins. The hair keratins belong to the large keratin multigene family, which mainly comprises genes that are expressed in various types of epithelia. Hair keratins form the intermediate filament (IF) network in trichocytes, i.e. cells that populate the central hair-forming compartment (hair matrix, cortex, and cuticle) of the anagen hair follicle. In the hair cortex, hair keratin IFs are embedded in an interfilamentous matrix, consisting of hair keratin-associated proteins (KAPs), which are essential for the formation of a rigid and resistant hair shaft through their extensive disulfide bond cross-linking with abundant cysteine residues of hair keratins (1, 2).

Originally, KAPs, mostly of ovine origin, have been classified on the basis of their amino acid composition as high sulfur (16–30 mol % cysteine), ultra-high sulfur (>30 mol % cysteine), and high glycine/tyrosine proteins (2). In 1993, Rogers and Powell (3) introduced a species-independent nomenclature using the abbreviations KAP1n through KAP8n for the eight members known at that time with “n” referring to a number identifying individual members fitting into a given family on the basis of sequence homology and the nature of repeat structures often present in these proteins (3). Since then, the number of new KAPs from sheep, rabbit, mouse, and humans have steadily increased (2), and to date, a total of 23 KAP families are known. Of these, families 1–3, 10–16, and 23 represent high sulfur KAPs (2, 4–17), families 4, 5, 9, and 17 are ultra high sulfur KAPs (2, 16, 18–23), and families 6–8 and 18–22 constitute high glycine/tyrosine KAPs (2, 24–26).

In humans, two KAP1 genes were initially described on a single genomic clone and named hB2A and hB2B, respectively, based on their sequence homology with two earlier described sheep proteins (9). Recently, our laboratory was able to characterize a large cluster of human KAP genes located within the type I keratin gene domain on chromosome 17q12-21. This cluster contained four novel hKAP1 genes, which were designated hKAP1.1B, hKAP1.3, hKAP1.4, and hKAP1.5 (16). On the basis of sequence differences among these genes and the previously described hB2A and hB2B genes, we renamed the latter hKAP1.1A (hB2A) and hKAP1.2 (hB2B), respectively (16). In addition, we identified two novel KAP1 members, hKAP1.6 and hKAP1.7, by PCR amplification of genomic DNA from a Japanese individual and by screening a Caucasian scalp cDNA library (17). However, the chromosomal assignment of the hKAP1.6 and hKAP1.7 genes as well as that of the hKAP1.1A and hKAP1.2 genes (9) remained to be determined.

Each of the eight hKAP1 members shared a generally high homology in both nucleotide and amino acid sequence (16, 17, 27). However, comparisons among the hKAP1.1A/B, hKAP1.6, and hKAP1.7 genes not only yielded a particularly high amino acid homology but also an extraordinary conservation of their 3'-noncoding region and, with the exception of the hKAP1.1A gene (9), their 5'-noncoding region (17). These findings promp-
Polymorphisms in hKAP1 Genes

Identification of Polymorphisms in hKAP1.1B and hKAP1.3 Genes—Peripheral leukocyte DNA was prepared from consenting Japanese and Caucasian individuals using standard protocols. The hKAP1.1B gene and its polymorphic variants were analyzed by PCR using three different upstream primers and one common downstream primer indicated in Table I. The hKAP1.1A- and hKAP1.1B-specific upstream primers KAP1.1A\^5-1 and KAP1.1B\^5-1, respectively, were derived from the differing 5′-noncoding regions of the genes (9, 16). In contrast, primers KAP1.5-11 and KAP1.3-1 were derived from regions common to the hKAP1.1A/B, hKAP1.6, and hKAP1.7 genes (9, 16, 17). PCR was performed using Advantage™ 2 DNA polymerase (Clontech, Tokyo, Japan) and the amplification conditions indicated in Table I. The amplification conditions were the same as those used for the PCR of genomic DNA (Table I). The PCR products were sequenced as described above. The amplified fragments were electrophoresed on 6% polyacrylamide gels. Because the amplified fragments were separated by only ~30 bp, they were subcloned into the pCR\^II-TOPO vector (Invitrogen) by ΑΤ cloning and then sequenced separately. In case single fragments were observed, they were sequenced directly after extraction of the fragments. The resulting sequences were used to search for DNA homologies with genes registered in the GenBank™ database using the BLASTN programs. The hKAP1.2, hKAP1.4, and hKAP1.5 genes were also analyzed by PCR using specific primer combinations for each gene (Table I).

RT-PCR—Total RNA was isolated from 15 freshly plucked anagen hair follicles of consenting Japanese individuals using the Isogen kit (Nippongene, Tokyo, Japan) according to the manufacturer’s recommendations. The RNA was digested for 10 min with RNase-free DNase (Roche Molecular Biochemicals) to remove contaminating genomic DNA and reverse-transcribed using an oligo(dT) primer. The cDNAs were amplified by PCR using hKAP1.3-specific primers (Table I). Amplification conditions were the same as those used for the PCR of genomic DNA (Table I). The PCR products were sequenced as described above.

3′-RACE—To identify the sequence of the 3′-noncoding regions of both hKAP1.8A and hKAP1.8B mRNAs, 3′-RACE was performed using a standard 3′-RACE kit according to the manufacturer’s instructions (Takara, Tokyo, Japan). After the synthesis of the first strand cDNAs using an oligo(dT)-adaptor primer derived from total RNA of anagen hair follicles of a Japanese individual who was heterozygous for hKAP1.8A and hKAP1.8B alleles, PCR was performed using the adaptor primer (5′-GGTTCCCGTCACGAC-3′) and KAP1.3-5′-1 primer designed at the 5′-noncoding region of hKAP1.3 (Table I). The PCR product was directly cloned into the pCR\^II-TOPO vector. The cDNA clones of hKAP1.8A and hKAP1.8B were sequenced.

ISH and IIF—ISH was carried out on cryostat sections of human hair follicles of a Japanese individual who was heterozygous for the expression of these genes in the human hair follicle. The resulting sequences were used to search for DNA homologies with genes registered in the GenBank™ database. The expression of these genes in the human hair follicle. The amplified fragments were electrophoresed on 6% polyacrylamide gels. Because the amplified fragments were separated by only ~30 bp, they were subcloned into the pCR\^II-TOPO vector (Invitrogen) by ΑΤ cloning and then sequenced separately. In case single fragments were observed, they were sequenced directly after extraction of the fragments. The resulting sequences were used to search for DNA homologies with genes registered in the GenBank™ database using the BLASTN programs. The hKAP1.2, hKAP1.4, and hKAP1.5 genes were also analyzed by PCR using specific primer combinations for each gene (Table I).

RT-PCR—Total RNA was isolated from 15 freshly plucked anagen hair follicles of consenting Japanese individuals using the Isogen kit (Nippongene, Tokyo, Japan) according to the manufacturer’s recommendations. The RNA was digested for 10 min with RNase-free DNase (Roche Molecular Biochemicals) to remove contaminating genomic DNA and reverse-transcribed using an oligo(dT) primer. The cDNAs were amplified by PCR using hKAP1.3-specific primers (Table I). Amplification conditions were the same as those used for the PCR of genomic DNA (Table I). The PCR products were sequenced as described above.

3′-RACE—To identify the sequence of the 3′-noncoding regions of both hKAP1.8A and hKAP1.8B mRNAs, 3′-RACE was performed using a standard 3′-RACE kit according to the manufacturer’s instructions (Takara, Tokyo, Japan). After the synthesis of the first strand cDNAs using an oligo(dT)-adaptor primer derived from total RNA of anagen hair follicles of a Japanese individual who was heterozygous for hKAP1.8A and hKAP1.8B alleles, PCR was performed using the adaptor primer (5′-GGTTCCCGTCACGAC-3′) and KAP1.3-5′-1 primer designed at the 5′-noncoding region of hKAP1.3 (Table I). The PCR product was directly cloned into the pCR\^II-TOPO vector. The cDNA clones of hKAP1.8A and hKAP1.8B were sequenced.

ISH and IIF—ISH was carried out on cryostat sections of human scalp biopsies taken for medical reasons (kindly provided by Dr. Bernard Cribier, Strasbourg, France) or plucked beard hairs as previously described in detail (28, 29). hKAP1.4 transcripts were detected using a specific PCR fragment, which encompassed 249 bp of the 3′-untranslated regions of the hKAP1.4 mRNA (16). The fragment was cloned into vector pCR2.1 (Invitrogen). Using this plasmid, [35S]-radiolabeled hKAP1.4 RNA probes were generated by in vitro transcription and hybridization overnight at 42 °C. Sections were washed with 2× SSC, 50% formamide, 20 mM DTT, 1× SSC, 50% formamide, 20 mM DTT, and 1× SSC, 50% formamide, 0.1% SDS at 50 °C for 30 min each digested with 100 U/ml RNase A (30 min at 37 °C), followed by washes in 0.5× SSC, 50% formamide, 20 mM DTT at 50 °C. After dipping in photoemulsion (NTB-2, Eastman Kodak Co.) and drying, sections were mostly exposed for 2–3 days, stained with hematoxylin, and embedded. For the recording of theISH signals by reflection microscopy, the confocal laser-scanning microscope LSM 510 was used, which allows simultaneous visualization of ISH in epi-illumination for the detection of reflection signals and transmitted light in bright field for hematoxylin staining. The two signal channels were combined by an overlay in pseudocolor (transmission image in green and electronically changed into black/white using the ZEISSL-SMiB software; reflection image, i.e. IHS signals, in red).

For immunohistochemistry, a pan-hKAP1 antisera was generated in guinea pigs using the synthetic oligopeptide QEGSSGAVSTRIRWCR coupled to keyhole limpet protein (Peptide Specialty Laboratories, Heidelberg, Germany) as antigen. This oligopeptide was derived from the central non-repetitive domain and is common to all hKAP1 proteins (16, 17). After the third booster injection, the antisera was used at a dilution of 1:1,500. As secondary antibodies, Cy3-coupled goat anti-guinea pig IgGs (Dianova, Hamburg, Germany) were used at a dilution of 1:50. IIF on cryostat sections of human scalp or plucked beard hairs was carried out essentially as described previously (28, 29) with the following modifications aimed at generating reductive conditions for the IIF procedure. Cryostat sections were fixed in methanol (−20 °C, 5 min) or used unfixed. The sections were incubated (30 min, RT) in PBS, which was supplemented with 10 mM DTT and 1 mM EDTA and flowed through with gaseous argon (PBSDEAr) for ~2 h. Sections were subsequently permeabilized with 0.1% Triton X-100 in PBSDEAr for 5 min. After three times of rinsing of specimens with argonated PBS supplemented with 1 mM EDTA (PBSDEAr), the primary antisera was applied overnight at 4 °C followed by three rinses in PBSDEAr (5 min each). The secondary antibodies were applied for 30 min following by washing in PBS, and the sections were rinsed in ethanol, dried, and mounted in fluoromount-G (Southern Biotechnology Associates, Birmingham, AL). In the case of unfixed cryostate sections, fixation was carried out subsequent to the application of the secondary antibodies. Visualization and documentation were performed with a photomicroscope (Axioskop II, Carl Zeiss, Jena/Oberkochen, Germany).

RESULTS

Identification of Polymorphisms in the hKAP1.1B Gene—To analyze the relationship between the hKAP1.1A/B, hKAP1.6, and hKAP1.7 genes, we first tried to PCR-amplify the hKAP1.1B gene using genomic DNA of eight unrelated Japanese individuals and the hKAP1.1B-specific primer KAP1.1B-5′-1 as well as primer KAP1.3′-1, which is common to the four genes (Table I). Polycyramide gel electrophoresis of the PCR products revealed the amplification of two fragments from three DNA samples, whereas only a single fragment was amplified from the remaining five DNA samples (Fig. 1A). Direct sequencing showed that the larger 973-bp fragment corresponded to the hKAP1.1B gene and that the smaller 835-bp product corresponded to the hKAP1.6 gene. In addition, we performed the same PCR analysis using genomic DNA from members of a two generation Japanese family (Fig. 1B). In this family, both the hKAP1.1B and hKAP1.6 alleles could be demonstrated in individuals 1, 2, and 3, whereas the hKAP1.6 allele alone was present in individual 4 (Fig. 1B). In toto, these results suggest that hKAP1.6 represents a polymorphic form of hKAP1.1B. The subsequent analysis of a total of 100 Japanese individuals (96 unrelated, 4 representing the members of the family shown in Fig. 1B) resulted in the detection of the single hKAP1.1B fragment in 61 individuals, whereas hKAP1.1B and hKAP1.6 fragments occurred in 36 individuals, and the single hKAP1.6 fragment was present in only three individuals. In contrast, the corresponding analysis of 100 unrelated Caucasian individuals revealed that the single hKAP1.1B fragment was amplified in 99 individuals, whereas hKAP1.1B and hKAP1.6 fragments were each found in only one individual (data not shown).

In contrast to the hKAP1.1B and hKAP1.6 genes, we completely failed to PCR-amplify the hKAP1.1A gene in our collection of Japanese and Caucasian individuals by means of the hKAP1.1A-specific primer KAP1.1A-5′-1 and the KAP1.3′-1 primer (Table I). The same held true for the amplification of the hKAP1.7 gene, which was analyzed in the two populations by means of the KAP1.5′-11 and the KAP1.3′-1 primers common...
to all four genes (Table I). Although the use of this primer pair would not allow a discrimination between the hKAP1.1A and hKAP1.1B genes, it should have, if present, amplified the hKAP1.6 and hKAP1.7 genes. Instead, the overall 973- and 855-bp fragment patterns observed in the two populations were virtually the same as those obtained with the KAP1.1B-5′-1 and KAP1.3-5′-1 primers (see "Results"). Considering, however, that hKAP1.1A has previously been identified on a genomic clone (9) and that hKAP1.7 has been detected by the screening of a human scalp cDNA library (17), we assume that both genes represent rare polymorphic forms of the KAP1.1B gene.

Taken together, our hKAP1 gene analysis in a total of 100 Japanese and 100 Caucasian individuals has shown that in Japanese individuals, 158 of 200 alleles were hKAP1.1B and 42 were hKAP1.6, whereas in Caucasian individuals, only one allele was hKAP1.6 and the remaining genes were hKAP1.1B (Table II).

In the context of these investigations, it is worth mentioning that within the two-generation Japanese family included into our hKAP1.1B polymorphism study (Fig. 1B), individual 3 suffered from a hair disorder. Clinically, the patient exhibited sparse, lusterless, and fragile scalp hairs. Microscopic analysis of the hairs excluded monilethrix, a disorder characterized by regular alterations in the diameter of the hair, and was confirmed by the absence of mutations in the hHb1 and hHb6 hair keratin genes known to be mutated in monilethrix patients (30, 31) (data not shown). The remaining members of the family exhibited clinically normal hairs. As described above, individuals 1, 2, and 3 were heterozygous for hKAP1.1B and hKAP1.6, whereas individual 4 was homozygous for hKAP1.6 (Fig. 1B). The direct DNA sequencing of each allele of all individuals revealed a CAG→TAG point mutation in codon 51, which introduced a premature stop codon in the hKAP1.1B allele (Fig. 1C, arrows) of individual 1 and the affected individual 3 but not in that of individual 2, whereas the hKAP1.6 alleles of all four members were normal.

Polymorphisms in Other hKAP1 Genes—The demonstration of polymorphisms for the hKAP1.1B gene prompted us to analyze the remaining hKAP1 family members. We first set out to amplify the hKAP1.3 gene, which is located next to the hKAP1.1B gene on chromosome 17q (16). We used a hKAP1.3-specific primer pair (Table I) and genomic DNA from 10 randomly selected unrelated Japanese individuals. Polyacrylamide gel electrophoresis of the PCR products showed the presence of two fragments that were separated by only ∼30 bp in five individuals, whereas a single fragment of either the larger (719 bp) or the smaller version (689 bp) was amplified in the others (Fig. 2A). DNA sequencing of the smaller 689-bp fragment revealed that it corresponded to the hKAP1.3 gene. In contrast, the larger 719-bp fragment turned out to belong to a novel KAP1 gene. The comparison between hKAP1.3 and this gene revealed a complete nucleotide sequence identity in their coding as well as their 5′- and 3′-noncoding regions with the exception of a 30-nucleotide insertion into the coding region of the novel gene, so that the deduced proteins differed by 10 amino acid residues (see Fig. 6B). These data suggest that the new gene represents a polymorphic form of hKAP1.3, which was termed hKAP1.8. Surprisingly, further sequence analysis of hKAP1.8 in Japanese individuals revealed the existence of two hKAP1.8 variants, hKAP1.8A and hKAP1.8B, because of single nucleotide substitutions in three codons (codon 34, TGC-TCC; codon 55, TGC-TGT; codon 92, GGA-AGA) of their coding regions, which resulted in the substitution of two amino acid residues (Cys-Ser and Gly-Arg) at the protein level (Fig. 3A). The pedigree analysis of two unrelated Japanese families definitively revealed that hKAP1.8A and hKAP1.8B were size polymorphisms of hKAP1.3 and that the respective polymorphic alleles segregated as a normal Mendelian trait (Fig. 2B). The analysis of a total of 100 Japanese individuals demonstrated that 78 of 200 alleles were hKAP1.3, 82 were hKAP1.8A, and 40 were hKAP1.8B (Table II).

The corresponding PCR analysis in unrelated Caucasian individuals also revealed the presence of the hKAP1.3 and hKAP1.8 fragments (Fig. 2C). Surprisingly however, one Caucasian individual elicited an additional fragment, which was distinctly smaller than that of hKAP1.3 (Fig. 2C, lane 9). The sequencing of this 611-bp fragment demonstrated that it was completely identical with hKAP1.3 with the exception of a
78-nucleotide deletion, which lead to a 26-amino acid residue deletion at the protein level. This finding suggests that the underlying gene, which we termed hKAP1.9 (Fig. 3B), represents an additional size polymorphism of hKAP1.3. The final analysis of 200 alleles of unrelated Caucasian individuals demonstrated that 111 were hKAP1.3, 87 were hKAP1.8A, whereas only one allele was either hKAP1.8B or hKAP1.9 (Table II). In this context, it is worth mentioning that the single Caucasian individual harboring the hKAP1.8B allele was also the one exhibiting the only hKAP1.6 allele detected in the Caucasian population investigated (Table II).

Subsequent to the hKAP1.3 gene, both the hKAP1.4 and hKAP1.5 genes were investigated for possible polymorphisms using specific primer pairs (Table I) for the respective genes in 100 Japanese and Caucasian individuals. Consistently, this large scale analysis yielded only single PCR fragments whose sequences corresponded to either hKAP1.4 or hKAP1.5, thus indicating that at least in the investigated population these genes are not polymorphic. Finally, we tried to amplify the hKAP1.2 gene by means of hKAP1.2-specific primers (Table I). This gene was previously identified on the genomic clone, which harbored the hKAP1.1B gene variant hKAP1.1A (9). Just as had been found for the hKAP1.1A gene, the hKAP1.2 gene could not be demonstrated in the investigated Japanese and Caucasian individuals (Table II). In view of its relative orientation to the hKAP1.1B gene variant hKAP1.1A, we suggest that hKAP1.2 represents a rare polymorphism of the hKAP1.3 gene.

Table II
Allele frequencies of the hKAP1.1B and the hKAP1.3 gene and their respective polymorphisms in 100 Japanese and 100 Caucasian individuals

| Alleles | Japanese (200 alleles) | Caucasian (200 alleles) |
|---------|------------------------|-------------------------|
| hKAP1.1B | 158 (0.790)            | 199 (0.995)             |
| hKAP1.6  | 42 (0.210)             | 0 (0.005)               |
| hKAP1.7  | 0 (0.000)              | 0 (0.000)               |
| hKAP1.1A | 0 (0.000)              | 111 (0.555)             |
| hKAP1.3  | 78 (0.390)             | 87 (0.435)              |
| hKAP1.8A | 82 (0.410)             | 87 (0.435)              |
| hKAP1.8B | 40 (0.200)             | 1 (0.005)               |
| hKAP1.9  | 0 (0.000)              | 1 (0.005)               |
| hKAP1.2  | 0 (0.000)              | 0 (0.000)               |
RT-PCR of hKAP1.3 and hKAP1.8A/B—To demonstrate the follicular expression of the hKAP1.8A and hKAP1.8B genes at the mRNA level, we performed RT-PCR as well as 3′-RACE using total RNA from freshly plucked hair follicles of two Japanese individuals. Considering that all known KAP genes consist of only one exon, care was taken to eliminate traces of genomic DNA from the follicular RNA samples by digestions with DNase I. As an additional control for the specificity of the RT-PCR reaction, a housekeeping gene, GAPDH, containing multiple exons was amplified along with the two KAP genes.

Fig. 4, lane 1, shows the amplification products obtained from follicular cDNA of a Japanese individual who was heterozygous for hKAP1.3 and hKAP1.8. The sequences of the two well-separated cDNAs were completely consistent with those of the hKAP1.8A and hKAP1.3 genes, respectively. Similar results were obtained with follicular cDNA of a Japanese individual who was heterozygous for hKAP1.8A and hKAP1.8B (data not shown), thus indicating that the hKAP1.8A/B genes are expressed in the hair follicle.

To determine the 3′ ends of hKAP1.8A and hKAP1.8B mRNAs, we performed 3′-RACE using total RNA from a Japanese individual who was heterozygous for the hKAP1.8A and hKAP1.8B alleles and obtained two clones containing a full-length cDNA of each allele. Sequencing of each cDNA revealed that the 3′-noncoding regions of hKAP1.8A and hKAP1.8B were completely identical to that of hKAP1.3 (Fig. 3A). Because plucked hairs from the Caucasian individual harboring the hKAP1.9 allele could not be obtained, the presence of its transcripts in the hair follicle could not be assessed.

Expression of hKAP1 mRNA and Protein in the Hair Follicle—We have previously shown by ISH using a 3′-specific probe for hKAP1.5 that its mRNA is expressed in the middle to upper cortex region of the hair follicle and absent from the cuticle and, if present, from the medulla (16). Subsequently, an identical mRNA expression profile was obtained for hKAP1.1A/B, hKAP1.6, and hKAP1.7 (17) as well as for hKAP1.3, hKAP1.2, hKAP1.8A/B, and hKAP1.9 (27) by means of cDNA probes derived from the respective common 3′-noncoding regions. Here we show, that transcripts of the remaining hKAP1 member, hKAP1.4, are also located in the middle to upper cortex region (Fig. 5A).

Up to now, KAP1 protein expression in hair follicles has not yet been demonstrated by immunohistochemistry. In this study, we generated a pan-hKAP1 antiserum using an oligopeptide derived from the central non-repetitive domain common to all hKAP1 proteins (16, 17) as an antigen. Because of
Fig. 3. Nucleotide sequences of the hKAP1.8A, hKAP1.8B, and hKAP1.9 genes and deduced amino acid sequences. A, hKAP1.8A and hKAP1.8B. B, hKAP1.9. Nucleotides are numbered consecutively on the left-hand side. The derived amino acid sequences appear below the nucleotide sequence in the one-letter code. Asterisks below the nucleotide and amino acid sequence of hKAP1.8A indicate sequence identity to hKAP1.8B. In A, nucleotide and amino acid differences between hKAP1.8A and hKAP1.8B are indicated in bold. The nucleotide and amino acid sequences not present in hKAP1.3 are boxed, and the polyadenylation signals are underlined.

**DISCUSSION**

Recently, we described an ~400-kb gene cluster on human chromosome 17q12-21, which contained seven KAP multigene families, embedded into the type I keratin gene domain. Among these families were four genes of the hKAP1 family that we designated hKAP1.1B, hKAP1.3, hKAP1.4, and hKAP1.5 (16). In a subsequent paper, we identified two additional human KAP1 genes, which were termed hKAP1.6 and hKAP1.7 (17). Together with two previously described human KAP1 genes, originally designated hB2A and hB2B (9) but later renamed hKAP1.1A and hKAP1.2 by us (16), the number of known hKAP1 genes amounted to eight. Because the hKAP1.1A, hKAP1.1B, hKAP1.6, and hKAP1.7 genes did not appear to be part of the KAP gene cluster on chromosome 17q12-21, we suggested their location elsewhere in the human genome (17).

In this study carried out in a large collection of Japanese and Caucasian individuals, we correct this assumption by demonstrating that the hKAP1.1A, hKAP1.6, and hKAP1.7 genes are in fact polymorphic alleles of the hKAP1.1B gene. However, of these four polymorphic alleles, only the hKAP1.1B and hKAP1.6 alleles were found in this study. Although the hKAP1.1B allele was detected with high frequency in both populations investigated, the frequency of the hKAP1.6 allele was distinctly different in Japanese (0.210) and Caucasian individuals (0.005), thus indicating that most probably hKAP1.6 represents a Japanese-specific polymorphism (Table II). Although the hKAP1.1A and hKAP1.7 alleles were not found in this study, each of them might represent a rare polymorphic hKAP1.1B allele.

Similar to the hKAP1.1B gene, we also identified several polymorphic alleles of the hKAP1.3 gene. The hKAP1.3 and hKAP1.8A alleles were commonly found with high frequencies in both Japanese and Caucasian individuals (Table II). In contrast, the hKAP1.8B allele was present in 40 Japanese individuals, whereas it was found in only a single Caucasian individual (Table II). Just like the hKAP1.6 allele, the frequency of the hKAP1.8B allele is so different between the two populations (0.200 in Japanese; 0.005 in Caucasians) that it can also be considered a Japanese-specific polymorphism. The hKAP1.9 allele was identified only in a single Caucasian individual (Table II), suggesting that it represents a rare hKAP1.3 size polymorphism in the Caucasian population. The hKAP1.8A allele and the Japanese-specific hKAP1.8B allele are identical with the exception of three nucleotide differences (G-C, C-T, and G-A, Fig. 3A) in their coding regions. In all of the individuals analyzed, no recombination among these three sites has been observed. Thus, the nucleotides C, T, and A in hKAP1.8B are also considered a Japanese-specific polymorphism (Table II).
Polymorphisms in hKAP1 Genes

with a high frequency. Pending the confirmation of frequent hKAP1.6/hKAP1.8B linkage by genomic sequencing, this could mean that the actual distribution of the hKAP1.6/hKAP1.8B allele goes back to a founder effect in the Japanese population subsequent to the divergence from the Caucasian lineage. Conceptually, the linkage of hKAP1.6 and hKAP1.8B would also help to explain the presence of these two alleles in a single Caucasian individual by unrecognized Japanese ancestry.

We previously divided hKAP1 proteins into five distinct domains: an N-terminal domain, a highly repetitive domain I, a central non-repetitive domain, a less repetitive domain II, and a C-terminal domain (16, 17, 27) (Fig. 6). The multialignment of both hKAP1.1B and hKAP1.3 with their respective polymorphic variants revealed that the differences among them were essentially because of variations in the repetitive domain I (Fig. 6). In hKAP1.1A/B and hKAP1.3, this domain comprises two extended tandem repeats, each consisting of a previously unidentified 16-amino acid segment “FCG(F/Y)PS(C/F)STG-GTC(G/D)SS” followed by 6 and 4 (4 and 4 in hKAP1.3) previously described cysteine-rich pentapeptide repeats, respectively (Fig. 6). Relative to hKAP1.1A/B, the Japanese-specific hKAP1.6 protein has entirely lost the first version of the extended quasi tandem repeats of the repetitive domain I. Based on the numerical dominance of the hKAP1.1B allele in both the Caucasian and Japanese population, we assume that the hKAP1.6 protein arose through the corresponding intragenic deletion in the hKAP1.1B gene. Moreover, a one amino acid exchange (Fig. 6A, Pro to Arg in bold) at the beginning of the second tandem repeat of domain I of the two proteins resulted also from an accompanying single nucleotide polymorphism (CCT to CGT) in the two genes. Compared with hKAP1.3, the hKAP1.8 protein exhibits an insertion of two tandem pentapeptide repeats in the repetitive domain I (Fig. 6B). Because hKAP1.3 and hKAP1.8A exhibit nearly identical allele frequencies (Table II), it is difficult to decide whether this difference is originally attributed to a corresponding intragenic insertion in hKAP1.3 or a deletion in hKAP1.8A. Although the generation of hKAP1.9 can also be traced back to an intragenic deletion of repetitive segments in domain I of either the hKAP1.3 or hKAP1.8A gene (Fig. 6B), notable exceptions from this rule exist for the evolution of the hKAP1.2 and hKAP1.7 proteins.

Table III

| hKAP1.1B gene | hKAP1.3 gene | Japanese (100 individuals) | Caucasian (100 individuals) |
|---------------|-------------|----------------------------|----------------------------|
| 1.1B/1.1B     | 1.3/1.3     | 13                         | 35                         |
| 1.3/1.8A      | 13/1.8A     | 35                         | 40                         |
| 1.3/1.8A      | 13/1.8A     | 35                         | 40                         |
| 1.3/1.9       | 0/1         | 1                          | 0                          |
| 1.6/1.6       | 1.3/1.8B    | 17                         | 1                          |
| 1.6/1.6       | 1.3/1.8B    | 18                         | 0                          |
| 1.6/1.6       | 1.8/1.8B    | 2                          | 0                          |
| 1.6/1.6       | 1.8/1.8A    | 1*                         | 0                          |
| 1.6/1.6       | 1.8/1.8A    | 1*                         | 0                          |

Fig. 4. RT-PCR expression of hKAP1.3 and hKAP1.8. RT-PCR amplification of hKAP1.3 and hKAP1.8A mRNAs from follicular RNA of a Japanese individual (lane 1). GAPDH mRNA, 597 bp in size, was amplified as a control using primers located in exons 3 and 7 of the GAPDH gene (lane 3). Note that a genomic 1101-bp GAPDH fragment was clearly absent from lane 3. PCR without reverse transcription did not lead to any amplification of the same products (lanes 2 and 4), indicating that there was no genomic DNA contamination in the samples. RT(+) and RT(-) denote reactions with or without reverse transcription. MMW, molecular weight markers.

Fig. 5. Expression of hKAP1.4 mRNA and hKAP1 family proteins in the human hair follicle. A, ISH using a specific hKAP1.4 probe. B, IIF using a pan-hKAP1 antibody. Cu, cuticle; co, cortex; dp, dermal papillae. Bars = 150 μm.
Although relative to hKAP1.3, hKAP1.2 still exhibits the loss of two tandem pentapeptide repeats of the repetitive domain I. It also differs from all known hKAP1 proteins by the unique two tandem pentapeptide repeats of the repetitive domain I. It should be emphasized that the main type of size polymorphisms described here for the hKAP1.1B and hKAP1.3 genes has previously been reported for sheep KAP1 genes. At present, four sheep KAP1 genes, B2A through B2D, are known (32). Because of species-specific amino acid insertions, their orthology with the four human KAP1 genes, hKAP1.1B, hKAP1.3, hKAP1.4 and hKAP1.5, is difficult to assess (16). Remarkably, however, the analysis of the sheep genes in unrelated animals of the same breeding revealed both insertion or deletion polymorphisms of the cysteine-rich pentapeptide tandem-repeats in the repetitive domain I of two proteins, B2A and B2C, but not in the remaining B2B and B2D proteins (32). Thus, the present data in two species clearly indicate a functional tolerance of a varying number of cysteine-rich repeats in distinct KAP1 members.

In conclusion, our study has revealed tandem-repeat pattern polymorphisms in human KAP1 genes. Obviously, despite the concomitant loss of a large number of cysteine residues, which should influence the strength of KAP protein interaction with hair keratin IFs, this phenomenon remains without recognizable phenotypic consequences, although structural microheterogeneities cannot be excluded. We also detected population-specific hKAP1 polymorphisms in Caucasian and Japanese individuals, which however cannot account for the differences in the respective hair phenotypes. On the other hand, it cannot be excluded that in view of the extraordinary high number of similarly repetitive KAPs potentially prone to size polymorphisms, distinct patterns of multiple KAP polymorphisms may influence the structure of the hair in a population-specific manner. Moreover, the extension of this study to other human populations, in particular from Africa and elsewhere in Asia, as well as the inclusion of our closest primate relatives may reveal KAPs to be useful tools regarding the elucidation of the time and geography of evolution of modern human populations.

Acknowledgments—We thank Dr. H. Spring (German Cancer Research Center) for help with confocal laser microscopy and Drs. B. Cribier (Department of Dermatology, University of Strasbourg, Strasbourg, France), and P. Deb (Clinique for Plastic Surgery, Mannheim, Germany) for supply with human scalp samples. We also thank S. Preatzel (German Cancer Research Center) for technical support.

REFERENCES
1. Powell, B. C., Neski, A., and Rogers, G. E. (1991) Ann. N. Y. Acad. Sci. 642, 1–20
2. Powell, B. C., and Rogers, G. E. (1997) in Formation and Structure of Human
Polymorphisms in the Human High Sulfur Hair Keratin-associated Protein 1, KAPI, Gene Family
Yutaka Shimomura, Noriaki Aoki, Jürgen Schweizer, Lutz Langbein, Michael A. Rogers, Hermelita Winter and Masaaki Ito

*J. Biol. Chem. 2002, 277:45493-45501.*
doi: 10.1074/jbc.M206398200 originally published online September 11, 2002

Access the most updated version of this article at doi: 10.1074/jbc.M206398200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 40 references, 9 of which can be accessed free at http://www.jbc.org/content/277/47/45493.full.html#ref-list-1