Biochemical and Biophysical Research Communications **274,** 149–152 (2000)
doi:10.1006/bbrc.2000.3110, available online at http://www.idealibrary.com on IDEAL®

Genomic Organization and Amplification of the Human Epidermal Type II Keratin Genes K1 and K5

Neil V. Whittock,[1] Robin A. J. Eady, and John A. McGrath
Department of Cell and Molecular Pathology, St. John's Institute of Dermatology, The Guy's, King's College, and St. Thomas' Hospitals' Medical School, London, United Kingdom

Received June 20, 2000

Keratins are a family of structurally related proteins that form the intermediate filament cytoskeleton in epithelial cells. Mutations in K1 and K5 result in the autosomal dominant disorders epidermolytic hyperkeratosis/bullous congenital ichthyosiform erythroderma and epidermolysis bullosa simplex, respectively. Most disease-associated mutations are within exons encoding protein domains involved in keratin filament assembly. However, some mutations occur outside the mutation hot-spots and may perturb intermolecular interactions between keratins and other proteins, usually with milder clinical consequences. To screen the entire keratin 1 and keratin 5 genes we have characterized their intron–exon organization. The keratin 1 gene comprises 9 exons spanning approximately 5.6 kb on 12q, and the keratin 5 gene comprises 9 exons spanning approximately 6.1 kb on 12q. We have also developed a comprehensive PCR-based mutation detection strategy using primers placed on flanking introns followed by direct sequencing of the PCR products.

Key Words: **intermediate filaments; skin; nails; hair.**

The intermediate filaments of epithelial cells are formed by keratins (1), a family of approximately 30 structurally related proteins, which are expressed in pairs of acidic type I and basic type II polypeptides in a tissue and differentiation specific manner (2). The rod domain of epidermal keratins consists of four α-helical regions (i.e., 1A, 1B, 2A, and 2B) that possess a repeating heptad amino acid motif (a-b-c-d-e-f-g)n with the potential to form a two-chain coiled–coil with a similar sequence. Two very highly conserved regions at the start of the 1A region and the end of the 2B region are termed the helix initiation peptide (HIP) and helix termination peptide (HTP) respectively, and it is postulated these regions overlap during filament assembly (reviewed in 3, 4). The α-helical regions are interrupted by non α-helical regions called linkers (i.e., L1, L12, and L2) conferring flexibility to the rod. In addition, type II keratins possess H1 and H2 domains that flank the ends of the HIP and HTP of the rod domain. Mutations in several keratin genes have been implicated in a range of genetic disorders (5, 6).

Keratin 1 is expressed with its type I partner, keratin 10, in suprabasal epidermal keratinocytes and mutations in their respective genes, K1 and K10, result in autosomal dominant epidermolytic hyperkeratosis (EH, otherwise known as bullous congenital ichthyosiform erythroderma, BCIE) (MIM 113800). In contrast, keratin 5 is predominantly expressed with its type I partner, keratin 14, in keratinocytes of the basal layer and mutations in their respective genes underlie the autosomal dominant epidermolysis bullosa simplex (EBS) (MIM 131760/131800/131900).

Mutations within the keratin genes are concentrated around regions that encode the highly conserved domains involved in keratin filament assembly (i.e., H1, HIP in 1A, L12, and HTP in 2B). Limited sequence data are available flanking the mutation hot-spots for both K1 and K5. Therefore, in order to permit more thorough screening of the K1 and K5 genes we have characterized the entire genomic organization of these genes and present conditions for the amplification of their individual exons.

MATERIALS AND METHODS

Characterization of human type II keratin 1 and 5 genomic structure. Primers based on the cDNA sequence of keratin 1 (GenBank No. NM 006121) and keratin 5 (GenBank No. NM 000424) were used to amplify keratin 1 and 5 introns from control human genomic DNA (Roche Molecular Diagnostics, Lewes, England). Approximately 200 ng of genomic DNA was added to a premix containing PCR buffer (67 mM Tris–HCl, pH 8.8, 16.6 mM $(NH_4)_2SO_4$, 1.5 mM $MgCl_2$, 0.17 mg/ml bovine serum albumin (Sigma, Poole, England), and 10 mM 2-mercaptoethanol), 10 nmol of each dNTP, 20 pmol of each primer in a total volume of 50 μl. After an initial denaturation at 95°C for 2

[1] To whom correspondence should be addressed at Department of Cell and Molecular Pathology, St. John's Institute of Dermatology, St. Thomas' Hospital, Lambeth Palace Road, London, SE1 7EH England. Fax: 44-20-7922-8175. E-mail: neil.2.whittock@kcl.ac.uk.

0006-291X/00 $35.00

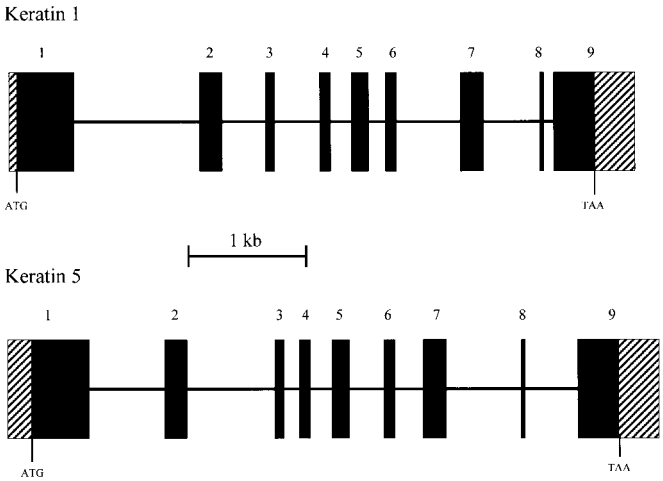

FIG. 1. Intron–exon organization of the human keratin 1 and 5 genes. Exons are represented by vertical boxes, introns by horizontal lines. The keratin 1 gene consists of 9 exons, 35–859 bp in size, spanning 5.6 kb of genomic DNA on 12q, and the keratin 5 gene consists of 9 exons, 35–946 bp in size, spanning 6.1 kb of genomic DNA on 12q (Table 1). The positions of the translation-initiation and termination codons are indicated. The keratin 1 and keratin 5 gene sequences have been deposited in the GenBank database under Accession Nos. AF237621 and AF274874, respectively.

min, 2.5 units of *Taq* polymerase (Promega, Madison, WI) was added followed by 35 cycles of 95°C for 10 s, annealing temperature for 10 s, 72°C for 1 min, with a final incubation of 72°C for 5 min. The PCR products were examined by 3% agarose gel electrophoresis, purified using spin columns (Qiagen, Crawley, England) and directly sequenced using Big Dye terminators on an ABI 310 genetic analyzer (Perkin–Elmer, Foster City, CA) (Table 1).

PCR amplification of genomic DNA. For direct amplification of the nine keratin 1 exons and the nine keratin 5 exons (see results) from genomic DNA, primer pairs were positioned within the introns flanking the exonic sequences (Table 2). For PCR, 200 ng of genomic DNA was added to a premix containing PCR buffer (as above), 10 nmol of each dNTP, 20 pmol of each primer in a total volume of 50 μl. After an initial denaturation at 95°C for 2 min, 2.5 units of Taq polymerase (Promega) was added followed by 35 cycles of 95°C for 10 s, annealing temperature for 10 s, 72°C for 30 s, with a final incubation of 72°C for 5 min. The annealing temperatures for each primer pair are displayed in Table 2. The PCR products were examined by 3% agarose gel electrophoresis, purified and directly sequenced as above.

RESULTS

Genomic Cloning of Human Keratins 1 and 5

To initiate genomic cloning of keratins 1 and 5, primers positioned within the respective cDNAs were used to amplify introns that were then sequenced. Analysis of sequence data revealed that keratin 1 consisted of 9 exons spanning approximately 5.6 kb of genomic DNA (Fig. 1). The exons varied in size from 35 bp (exon 8) to 859 bp (exon 9), and the introns ranged from 92 bp (intron 8) to 1000 bp (intron 1) (Table 1). Analysis of sequence data revealed that keratin 5 also consisted of 9 exons spanning approximately 6.1 kb of genomic DNA (Fig. 1). The exons varied in size from 35 bp (exon 8) to 946 bp (exon 1), and the introns ranged from 110 bp (intron 3) to 804 bp (intron 7) (Table 1). Apart from exons 1 and 9 which code for the variable (V1 and V2)

TABLE 1

Exon–Intron Boundaries of the Human Keratin 1 and 5 Genes

Exon	3′ splice acceptor sequence[a]	Exon size (bp)	5′ splice donor sequence[a]	Intron size (bp)
		Keratin 1		
1	Not determined	591	GACAAG**gt**gagtttct	1000
2	gtgtttttt**ag**GTGAGG	215	GAACAA**gt**aagggacc	318
3	tttcctgc**ag**GTATGA	61	AAGAAG**gt**aagcaaat	417
4	gttcctct**ag**GATGTG	96	CAAGCA**gt**aagtcttc	169
5	tgtcttgc**ag**GAGTTG	165	AGCAAG**gt**gagtggct	130
6	ccatcttt**ag**TATGAA	126	AAGCAG**gt**atgtgctt	563
7	tctttttc**ag**ATCTCC	221	AAGCAG**gt**gaggaagg	531
8	tctcctac**ag**GATGTC	35	GTGTGT**gt**aagtacaa	92
9	cctcttgc**ag**CTGTGA	859	Poly (A)	
		Keratin 5		
1	Not determined	946	GACAAG**gt**gagctacg	581
2	tccccctcc**ag**GTGAGG	215	GAACAA**gtg**agttggg	777
3	tcaccacc**ag**GTATGA	61	AAGAAG**gt**gcgtgtgg	110
4	gtttcccc**ag**GATGTA	96	GATGCG**gt**aagaaact	142
5	ctccctgc**ag**GAGCTG	165	ACCAAG**gt**gggtgctc	358
6	gctgcatc**ag**TATGAG	126	AAACAG**gt**agggtgag	249
7	ctatctgt**ag**TGCGCC	221	ATGCAG**gt**gagtagac	804
8	ccccttttc**ag**ACTCAG	35	ACATCT**gt**aagtagct	555
9	tttcctgc**ag**CTGTGA	664	Poly (A)	

[a] The consensus ag/gt splice sequences are in bold.

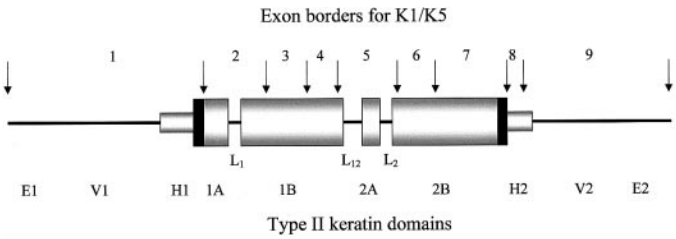

FIG. 2. Schematic diagram displaying the consensus exon borders of K1 and K5 with relation to the type II keratin protein domains. The exons are numbered 1–9, and exon borders are represented by arrows. Boxed domains represent coiled–coil helical segments of the central rod domain, and the lines represent the nonhelical regions. The areas shaded in black in the 1A and 2B domains represent the helix initiation peptide (HIP) and the helix termination peptide (HTP), respectively.

domains respectively, all exons are the same size with highly conserved exon–intron borders (Table 1). The consensus exon–intron borders for the K1 and K5 genes are demonstrated in Fig. 2. The mutation "hot-spots" are covered by the following exons: H1 (exon 1), HIP in 1A (exon 1), L_{12} (exon 5), and HTP in 2B (exon 7).

Development of a Mutation-Detection Strategy

To facilitate identification of pathogenic mutations in patients with EH/BCIE (K1) and EBS (K5) we designed a strategy for detection of sequence variants. The approach consisted of PCR amplification of exons directly using genomic DNA as template (Fig. 3), with primers based on intron-specific sequences followed by direct sequencing. To allow PCR amplification of exons directly from genomic DNA, balanced primer pairs were designed on the basis of sequences close to the keratin intron–exon regions. The primer sequence information, the optimized amplification conditions, and the expected sizes of the PCR products are indicated in Table 2.

GenBank Accession Numbers

Human keratin 1 and keratin 5 genes have been deposited in the GenBank sequence database under Accession Nos. AF237621 and AF274874, respectively.

DISCUSSION

The spectrum of pathogenic mutations in the type I and II epidermal keratins is well documented with hot-spots occurring within the H1 domain, the helix initiation peptide in the 1A rod domain, the L12 linker region, and the helix termination peptide in the 2B rod domain (5, 6). Keratin mutations are usually dominantly inherited missense mutations but several cases of autosomal recessive EBS have been reported resulting from homozygous splice-site or nonsense mutations in the K14 gene outside of these hot-spots (7–10). Most dominant or recessive mutations compromise intermediate filament assembly giving the cell a reduced tolerance to mechanical trauma which results in skin fragility (11). However, other mutations outside the hot-spots have been identified that affect intermolecular connections rather than filament assembly. For example, a missense mutation within the variable amino terminal domain (V1) of keratin 1 which replaces a lysine 74 residue for isoleucine has been demonstrated in a case of diffuse nonepidermolytic palmoplantar keratoderma (12). Ultrastructural studies of affected palm epidermis revealed a normal keratin filament network indicating that the mutation did not disrupt filament assembly. However, close morphological inspection showed perturbations in the organization of keratin filaments close to the cornified cell envelope. Indeed, further studies have demonstrated that a highly conserved stretch of approximately 20 amino acid residues within the V1 region of type II epidermal keratins, and in particular the lysine residue, is critical for the formation of intermolecular connections between keratin filaments and proteins of desmosomes and the cornified cell envelope (13, 14). Through transglutaminases, this conserved type II lysine residue has been shown to cross-link to glutamine residues within proteins such as envoplakin, loricrin, involucrin and some of the small proline rich proteins (14). It has been postulated that type II keratins expressed within the basal cells (i.e., keratin 5) form attachments with proteins of the desmosome such as desmoplakin, and as the cell terminally differentiates expressing different

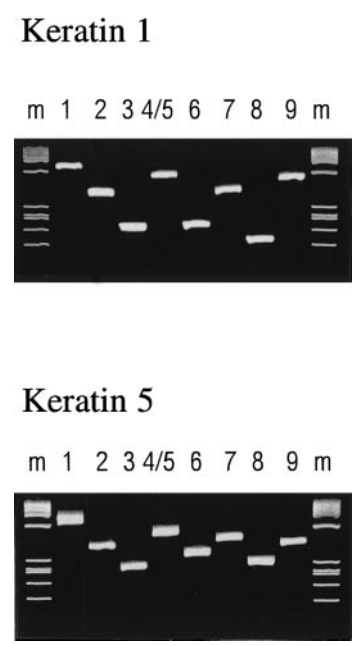

FIG. 3. PCR amplification of exons 1–9 for keratin 1 and exons 1–9 for keratin 5, using genomic DNA as template with the primer pairs shown in Table 2. The PCR products were examined by 3% agarose gel electrophoresis. The lanes m contain φX174 *Hae*III molecular weight markers.

TABLE 2

Primers for the Genomic Amplification of Human Keratins 1 and 5

Exon	Forward (5′ → 3′)	Reverse (5′ → 3′)	Product (bp)	Annealing temp. (°C)
		Keratin 1		
1	ttccaggtctgggtaccgaag (18)[a]	ctcctaggagaccattccac (+140)	714	52
2	ctgtcatggtcagaagatagg (−111)	catgctgcttcatgatcttagc (+84)	410	56
3	gcagtgataggcaatgccac (−91)	ggctctccatatcatggctg (+96)	248	56
4	ctcacaggtaatgagagggac (−67)	gaatcgttgtgggttcaggc (+68)	231	56
5	gcctgaacccacaacgattc (−121)	cacatacatgagataacacagg (−63)	349	56
6	cctgtgttatctcatgtatgtg (−79)	ccattccatacagctgagtg (+41)	246	56
7	caagtggttctgttggactc (−112)	caatcagatggctgcatcttc (+89)	422	56
8	gtgcattggagaagacccag (−114)	ctccgtttccagcccaac (+58)	207	56
9	gttgggctggaaacggag (−52)	gacctcggtcttgccaag (2010)[a]	552	56
		Keratin 5		
1	agctctgttctctccagcac (−71)[b]	cagtctaattcagaacgtgtcc (+169)	795	56
2	gaggcaagccttagtgagttg (−73)	tccatggaaggtatatcctcc (+120)	408	56
3	cagaggttcatgctaccagt (−153)	ggcactgcacacaccgtcac (+90)	304	56
4 + 5	aaccagcagcctgcagctatg (−91)	ccattcttagtgtcgtcatgg (+81)	575	56
6	aggcactagactagctcagc (−145)	actggatctagctgcgtgtg (+108)	379	56
7	cacacgcagctagatccagt (−161)	atgatgtgtcattatcacgcac (+120)	520	56
8	ggtgtacagactcctggctt (−119)	cataagccacattgcttcctgtc (+190)	344	56
9	tagagtgctccaacaccagc (−93)	acatggcttgagcaactgcc (1857)[b]	476	56

[a] In the keratin 1 cDNA (GenBank No. NM 006121) relative to the translation start site.
[b] In the keratin 5 cDNA (GenBank No. NM 000424) relative to the translation start site.

type II keratins (i.e., keratins 1 and 2e) these form connections with proteins of the cell envelope such as involucrin and loricrin. It is, therefore, plausible that mutations affecting other intermolecular connections within the epidermis may occur within other keratins including K1 or K5, thus demonstrating the importance of characterizing the entire genomic structure of these genes for subsequent molecular analyses.

ACKNOWLEDGMENT

This study was supported by the Special Trustees of St. Thomas' Hospital.

REFERENCES

1. Quinlan, R. A., Hutchison, C. J., and Lane, E. B. (1994) Intermediate filaments. *In* Protein Profiles (Sheterline, P., Ed.), Academic Press, London.
2. Roop, D. R. (1995) Defects in the barrier. *Science* **267,** 474–475.
3. Steinert, P. M. (1993) Structure, function, and dynamics of keratin intermediate filaments. *J. Invest. Dermatol.* **100,** 729–734.
4. Steinert, P. M., North, A. C. T., and Parry, D. A. D. (1994) Structural features of keratin intermediate filaments. *J. Invest. Dermatol.* **103,** 19S–24S.
5. Corden, L. D., and McLean, W. H. I. (1996) Human keratin diseases: hereditary fragility of specific epithelial tissues. *Exp. Dermatol.* **5,** 297–307.
6. Irvine, A. D., and McLean, W. H. I. (1999) Human keratin diseases: The increasing spectrum of disease and subtlety of the phenotype-genotype correlation. *Br. J. Dermatol.* **140,** 815–828.
7. Chan, Y., Anton-Lamprecht, I., Yu, Q. C., Jackel, A., Zabel, B., Ernst, J. P., Fuchs, E. (1994) A human keratin 14 "knockout": The absence of K14 leads to severe epidermolysis bullosa simplex and a function for an intermediate filament protein. *Genes Dev.* **8,** 2574–2587.
8. Rugg, E. L., McLean, W. H., Lane, E. B., Pitera, R., McMillan, J. R., Dopping-Hepenstal, P. J., Navsaria, H. A., Leigh, I. M., and Eady, R. A. (1994) A functional "knockout" of human keratin 14. *Genes Dev.* **8,** 2563–2573.
9. Jonkman, M. F., Heeres, K., Pas, H. H., van Luyn, M. J., Elema, J. D., Corden, L. D., Smith, F. J., McLean, W. H., Ramaekers, F. C., Burton, M., and Scheffer, H. (1996) Effects of keratin 14 ablation on the clinical and cellular phenotype in a kindred with recessive epidermolysis bullosa simplex. *J. Invest. Dermatol.* **107,** 764–769.
10. Corden, L. D., Mellerio, J. E., Gratian, M. J., Eady, R. A., Harper, J. I., Lacour, M., Magee, G., Lane, E. B., McGrath, J. A., and McLean, W. H. (1998) Homozygous nonsense mutation in helix 2 of K14 causes severe recessive epidermolysis bullosa simplex. *Hum. Mutat.* **11,** 279–285.
11. McLean, W. H. I., and Lane, E. B. (1995) Intermediate filaments in disease. *Curr. Opin. Cell. Biol.* **7,** 118–125.
12. Kimonis, V., DiGiovanna, J. J., Yang, J. M., Doyle, S. Z., Bale, S. J., and Compton, J. G. (1994) A mutation in the V1 end domain of keratin 1 in non-epidermolytic palmar–plantar keratoderma. *J. Invest. Dermatol.* **103,** 764–769.
13. Kouklis, P. D., Hutton, E., and Fuchs, E. (1994) Making a connection: direct binding between keratin intermediate filaments and desmosomal proteins. *J. Cell Biol.* **127,** 1049–1060.
14. Candi, E., Tarcsa, E., DiGiovanna, J. J., Compton, J. G., Elias, P. M., Marekov, L. N., and Steinert, P. M. (1998) A highly conserved lysine residue on the head domain of type II keratins is essential for the attachment of keratin intermediate filaments to the cornified cell envelope through isopeptide crosslinking by transglutaminases. *Proc. Natl. Acad. Sci. USA.* **95,** 2067–2072.