MOLECULAR AND CELLULAR BIOLOGY, Feb. 1988, p. 722–736
0270-7306/88/020722-15$02.00/0

Vol. 8, No. 2

A Group of Type I Keratin Genes on Human Chromosome 17: Characterization and Expression

MARJORIE ROSENBERG,[1,2] AMLAN RAYCHAUDHURY,[1,2] THOMAS B. SHOWS,[3] MICHELLE M. LE BEAU,[4] AND ELAINE FUCHS[1,2]*

Departments of Molecular Genetics and Cell Biology,[1] Biochemistry and Molecular Biology,[2] and Medicine,[4] The University of Chicago, Chicago, Illinois 60637, and Roswell Park Memorial Institute, New York State Department of Health, Buffalo, New York 14263[3]

Received 17 August 1987/Accepted 6 November 1987

The human type I keratins K16 and K14 are coexpressed in a number of epithelial tissues, including esophagus, tongue, and hair follicles. We determined that two genes encoding K16 and three genes encoding K14 were clustered in two distinct segments of chromosome 17. The genes within each cluster were tightly linked, and large parts of the genome containing these genes have been recently duplicated. The sequences of the two K16 genes showed striking homology not only within the coding sequences, but also within the intron positions and sequences and extending at least 400 base pairs 5′ upstream and 850 base pairs 3′ downstream from these genes. Despite the strong homologies between these two genes, only one of the genes encoded a protein which assembled into keratin filaments when introduced into simple epithelial cells. While there were no obvious abnormalities in the sequence of the other gene, its promoter seemed to be significantly weaker, and even a hybrid gene with the other gene's promoter gave rise to a much reduced mRNA level after gene transfection. To demonstrate that the functional K16 gene that we identified was in fact responsible for the K16 expressed in human tissues, we made a polyclonal antiserum which recognized our functional K16 gene product in both denatured and filamentous form and which was specific for bona fide human K16.

The keratins form a group of more than 20 related proteins (40 to 67 kilodaltons [kDa]) which comprise 8-nm cytoskeletal filaments in most if not all epithelial cells. These keratins have been assigned numbers according to their molecular size and their isoelectric mobilities (for a review, see Moll et al. [44]). Although keratins can be subdivided into two distinct sequence classes, type I and type II, all keratins have very similar secondary structures (8, 24, 65). The central 310-amino-acid portion is relatively constant in length and consists of several large α-helical domains (for reviews, see Fuchs and Hanukoglu [15], Steinert et al. [67], and Parry et al. [50]). The classification of keratins according to type is based on their sequence relatedness within these domains: while a comparison of any two members of the same type shows 50 to 99% homology in this region, comparisons of members of opposite type reveals only 25 to 35% homology (2, 20, 23, 24, 26–28, 30, 31, 36, 48, 52, 53, 63, 65, 66, 73). Throughout the helical segments are heptad repeats of hydrophobic residues, indicative of their ability to form coiled-coil subunits (40, 41). The nonhelical termini are variable in length and in sequence and may play a role in lateral and end-to-end interactions necessary to pack or stabilize the coiled-coil subunits into the filamentous structure (18, 19, 66, 67). At least one member of each of the two keratin classes is expressed in all tissues, suggesting the importance of the two types of sequences in filament assembly. In vitro assembly studies support this notion (11, 25).

Frequently, keratins are coexpressed as specific type I and type II keratin pairs, which are differentially expressed in epithelial tissues and at various stages of differentiation and development (10, 68). Although the role of the variability in keratin expression is not yet fully understood, these differences may enable the properties of the resulting filaments to be tailored to suit the particular protective requirements of each epithelial cell. As such, the mechanisms that operate and control the tissue-specific expression of keratin genes may be important in epithelial structure or function.

Elucidating the genetic complexity, organization, and regulation of the keratin gene family has been complicated by the fact that not only are there high degrees of homologies among different keratins of the same type, but also multiple genes for what has been assigned a single keratin number (73). Little is known about whether these multiple keratin genes are functional, whether they are differentially expressed, or how they evolved. To begin to examine these issues, we have focused on a group of human genomic clones that were originally isolated on the basis of their ability to cross-hybridize strongly with a cloned cDNA to the human K14, a type I keratin expressed in the basal layer of all stratified squamous epithelia, including epidermis (45).

In this paper, we report the mapping of all of these clones to human chromosome 17, and we provide evidence that large segments of the genome containing these type I genes have been recently duplicated. We demonstrate the existence of three K14 genes and two additional genes which encode K16, a keratin that is constitutively expressed in most stratified squamous epithelia, but only transiently expressed in epidermis during hyperproliferation (76; E. B. Lane, personal communication). Sequence and expression analyses of the two K16 genes reveal that only one of these genes seems to be expressed, while the other appears to be a part of a duplicated gene region that has been targeted for inactivation in the genome.

MATERIALS AND METHODS

Construction of subclones and preparation of radiolabeled probes. (i) From GK-3 DNA. Previously, a 4.3-kilobase (kb) *Bam*HI fragment, containing the 5′ and 3′ portions of a K14-related keratin gene, was subcloned into plasmid pUC9

* Corresponding author.

(N. E. Biolabs) and sequenced (52). A 1,225-base-pair (bp) *Pst*I-*Bam*HI fragment of this 4.3-kb fragment was subcloned into plasmid pGEM1 (Promega Biotec) for use in in situ hybridization analysis of chromosomes and in Southern blot analyses of human-mouse hybrid cells. This fragment, referred to as GK3-3′, contains the sequence beginning within intron 7, continuing through 31 codons of translated region, and terminating 876 bases 3′ to the site of polyadenylation. A 353-bp *Sac*I-*Hin*cII subfragment of GK3-3′ containing the sequence beginning 6 bp 5′ upstream from the TAG stop codon of the keratin gene and extending 210 bp 3′ downstream from the polyadenylation signal was also subcloned into plasmid pGEM1 (Promega Biotec). This plasmid, referred to as pGK3-3′NC, was used in Northern (RNA) and Southern blot analyses.

Radiolabeled cRNA probe was made by linearizing pGK3-3′NC with *Sac*I or pGK3-3′ with *Pst*I and transcribing the DNA with SP6 polymerase in the presence of [^{32}P]UTP and unlabeled nucleotides as described by the manufacturer (Promega Biotec).

The chimeric plasmid pJ2GK3 · Bam was constructed by inserting the 4.3-kb *Bam*HI fragment obtained from GK-3 into the *Bam*HI site of the expression vector pJ2 (J. Morgenstern and R. Weinberg, Massachusetts Institute of Technology). The orientation of the fragment was such that the 72-bp repeats of the simian virus 40 (SV40) enhancer region were approximately 400 bp 5′ upstream from the TATA box of the GK-3 gene. For expression in COS cells (42), the same GK-3 fragment was inserted into plasmid pJ1 (pJ1GK3 · Bam [33]). The plasmid is similar to pJ2 except that it contains the SV40 promoter and origin of replication in addition to the 72-bp repeat enhancer sequence. The SV40 sequences were located approximately 400 bp upstream from the TATA box of the GK-3 gene. These two plasmids were used for all of the DNA transfection studies involving the GK3-encoded keratin gene.

(ii) From GK-6 DNA. A 4.3-kb *Bam*HI fragment, encompassing the 5′ and 3′ portions of a GK-3-related keratin gene contained in the GK-6 genomic clone, was subcloned into plasmid pUC9 for DNA sequence analysis. For expression studies, the chimeric plasmid pJ1GK6 · Bam was constructed by inserting the 4.3-kb *Bam*HI fragment obtained from GK-6 into the *Bam*HI site of the expression vector pJ1 (see above). This plasmid was used for all of the DNA transfection studies involving the GK6-encoded keratin gene.

(iii) From GK-1 DNA. To prepare a pure intron probe for the GK-1 K14 gene, an 831-bp *Hin*dIII-*Stu*I fragment extending 263 bp 3′ from the *gt* junction of intron I to 160 bp 5′ from the *ag* junction of intron I was subcloned into the *Hin*dIII-*Hin*cII sites of pGEM1. This plasmid will be referred to as GK1-int. Radiolabeled cRNA probe was prepared by using T7 polymerase (Promega Biotec) as described above.

Northern and Southern hybridization blots. Polyadenylated [poly(A)$^+$] RNAs were isolated essentially by the guanidinium isothiocyanate procedure of Berk and Sharp (3). RNAs were resolved by formaldehyde-agarose gel electrophoresis and transferred to nitrocellulose paper by blotting (70).

Human genomic DNAs were isolated from bacterial clones (37) or from lymphocytes (32). After digesting with the appropriate restriction endonucleases, the DNA fragments were resolved by agarose gel electrophoresis. Following denaturation in 0.5 M NaOH, gels were transferred to nitrocellulose paper by blotting (64).

Prehybridizations and hybridizations were conducted at 41°C in 50% deionized formamide–750 mM NaCl–75 mM sodium citrate–20 mM HEPES (*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid, pH 7.4)–1 mM EDTA–100 μg of yeast tRNA (Miles) per ml–0.01% bovine serum albumin (BSA)–0.02% Ficoll 400–0.02% polyvinylpyrrolidone 360–10 μg of sonicated, denatured salmon sperm DNA per ml. From 1×10^5 to 5×10^5 dpm of ^{32}P-labeled probe per ml was used for Northern blots, and 10^6 dpm of probe per ml was used for Southern blots. Following hybridizations of Northern and Southern blots with ^{32}P-labeled cRNA probes, blots were washed three times in 0.1% sodium dodecyl sulfate (SDS)–0.1× SSC (0.015 M NaCl, 0.0015 M sodium citrate) at 67°C and then rinsed well in 2× SSC at room temperature. Blots were then incubated in 2× SSC containing RNAse A (0.1 μg/ml) for 15 to 20 min at room temperature (9). The filter was finally washed at 50°C in 0.1% SDS–0.1× SSC for 30 min.

Human-mouse hybrids. Thirty-two cell hybrids involving 15 unrelated human cell lines and four mouse cell lines (60–62) were analyzed. Hybrids were characterized by chromosome analysis and mapped enzyme markers and partly by mapped DNA probes (58, 61, 62).

In situ chromosomal hybridizations. Radiolabeled keratin probes pGK3-3′ and pGK1-int were prepared by nick translation of the entire plasmid with all four ^{3}H-labeled deoxynucleoside triphosphates to a specific activity of 1.2×10^8 dpm/μg. In situ hybridization was performed as described previously (35). Metaphase cells from phytohemagglutinin-stimulated peripheral blood lymphocytes were hybridized at 4.0 and 8.0 ng of probe per ml of hybridization mixture. Autoradiographs were exposed for 11 days.

DNA sequencing. DNA sequence analysis was performed on the 4.3-kb *Bam*HI fragment of GK-6. The M13 dideoxy sequencing strategy (54), and the Bal 31 exonuclease shotgun cloning method of Anderson (1) was used. To obtain overlapping clones for sequencing, plasmid DNA was digested with *Kpn*I, followed by limited Bal 31 exonuclease treatment. At 2-min intervals, portions from the Bal 31 digest were removed and the reactions were stopped with 15 mM EGTA. These fragments were finally digested with *Hin*dIII, inserted into the *Hin*cII and *Hin*dIII sites of M13mp19 DNA, and transformed into *Escherichia coli* JM101. To obtain overlapping clones for sequencing the opposite strand, plasmid DNA was digested with *Hin*dIII, followed by Bal 31 and finally *Kpn*I as above. These fragments were inserted into the *Hin*cII and *Kpn*I sites of M13mp18 and transformed into *E. coli* JM101. Positive clones were screened with ^{32}P-radiolabeled cDNA prepared from the purified 4.3-kb *Bam*HI fragment of GK-6 and subsequently sequenced. The entire *Bam*HI fragment was sequenced in both directions.

DNA transfections. DNA transfections leading to transient expression were carried out by the calcium phosphate precipitation method (22) as described by Gorman (21). A fine precipitate of DNA and calcium phosphate was generated by dropwise addition of a solution containing approximately 0.2 μg of DNA per ml and 0.25 M $CaCl_2$ (pH 5.8) to an equal volume of 40 mM HEPES-buffered saline (pH 7.12) that included 2% (wt/vol) NaH_2PO_4. This mixture was then added at 0.75 μl per 1.5 ml of medium to cultured NIH 3T3 mouse fibroblasts, marsupial PtK2 cells, or monkey kidney epithelial COS cells. After 12 to 15 h of incubation, the DNA-calcium phosphate suspension was removed, and cells were exposed to 15% glycerol in medium for 1 to 2 min depending upon cell type (49). Cells were then refed with normal medium and returned to the incubator for 50 h. When NIH 3T3 cells were transfected, a sodium butyrate step (overnight incubation in medium containing 5 mM sodium

butyrate, pH 7.0) was added after the glycerol shock to enhance transfection efficiency (21).

Intermediate filament protein extractions. Intermediate filament proteins were isolated essentially by the procedure described by Wu et al. (80). Cells (2×10^6 to 10×10^6) were rinsed with phosphate-buffered saline (PBS) and scraped from their dish into a Corex centrifuge tube containing 10 ml of 20 mM Tris hydrochloride (pH 7.4)–0.6 M KCl–1% Triton X-100–0.3 μg of phenylmethylsulfonyl fluoride per ml. After sonication at 4°C, the broken cells were pelleted by centrifugation at 10,000 × *g* for 20 min. The pellet was suspended in the same buffer, and the procedure was repeated five times. The final pellet was suspended in 100 μl of 8 M urea–10% beta-mercaptoethanol, warmed to 37°C, sonicated briefly, and used for gel electrophoresis.

Antibodies. A rabbit polyclonal antibody (5547) was prepared against the tetradecapeptide sequence SSAVRPGPS SEQSS conjugated to keyhole limpet hemocyanin (Cambridge Biochemicals, England). This antiserum shows marked and specific cross-reactivity with human K16. Of this sequence, only the carboxy-terminal QSS corresponds to the GK-3 sequence. At a concentration of 300 μg/ml of antiserum, a synthetic heptapeptide to the carboxy-terminal sequence FSQGQSS of the putative GK-3 keratin (Peninsular Laboratories) quantitatively preadsorbs nearly all of the K16-binding activity of the antiserum. The human keratin K14, which has an internal QSS sequence, shows no cross-reactivity with the antiserum. The only keratin which by immunoblot analysis shows slight cross-reactivity with the 5547 antisera is K6, and since this cross-reactivity was variable, this may be an artifact due to complex formation between type I and type II keratin pairs.

The monoclonal antibody LE41 (34) was used to identify the endogenous keratin network of transfected PtK2 cells. Two antibodies were used in immunoblot analyses: the monoclonal antibody AE1, which is broadly cross-reactive with a number of type I keratins, including K14 and K16, but which does not recognize K17 (72, 78), and a broadly cross-reactive rabbit polyclonal antiserum which was made against a total keratin extract from cultured human epidermal cells (16).

Immunoblot analysis. Immunoblot analysis was carried out essentially as described by Towbin et al. (71). Briefly, proteins resolved by SDS-polyacrylamide gel electrophoresis (PAGE) were transferred to nitrocellulose paper with a Biorad electroblotter as described by the manufacturer. Blots were presoaked in a solution of 5% BSA in PBS. Blots were incubated at room temperature in a solution of 5% BSA–1× PBS containing one of three different antibodies: (i) anti-K16 (5547) diluted 1:30, (ii) a 1:2 dilution of a cell lysate containing the mouse monoclonal antibody AE1 (72), or (iii) a 1:100 dilution of a general rabbit polyclonal antiserum prepared against cultured human epidermal keratins (16). After hybridization for 2 h, the blots were washed three times in 0.1% Triton X-10–10× PBS and then once in PBS. The blot incubated with the monoclonal antibody AE1 was further incubated with a rabbit anti-mouse immunoglobulin G (IgG) (Miles Laboratories) diluted to 10 μg/ml for 1 h and washed thoroughly as above. All blots were then incubated in 10^5 cpm of ^{125}I-*Staphylococcus aureus* Protein A (Amersham) for 1 h, followed by thorough washing as above.

Immunofluorescence. Double immunofluorescence was carried out by incubating methanol-fixed pJ2GK3 · Bam-transfected cells with the 5547 rabbit polyclonal antiserum (1:30 dilution in 10% newborn goat serum in PBS) and the mouse monoclonal simple epithelial keratin antibody LE41 (undiluted) (34) at room temperature for 30 min. After being washed extensively with PBS, cells were incubated for 30 min with a 1:30 to 1:60 dilution of the secondary antibodies fluorescein isothiocyanate-conjugated goat anti-rabbit IgG (Cooper Biomedical) and Texas Red-conjugated goat anti-mouse IgG (TAGO Corp.).

RESULTS

Chromosomal mapping of the K14-related genes. Previously, we reported the isolation and restriction endonuclease mapping of seven different human genomic clones encoding

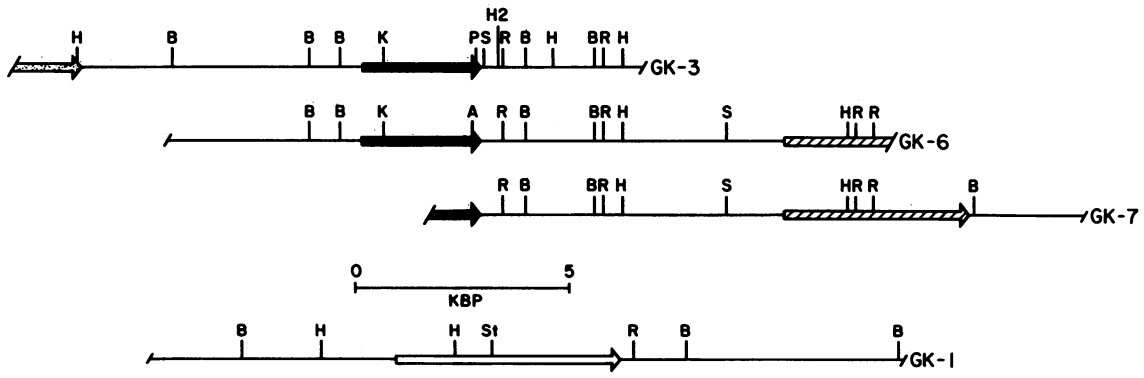

FIG. 1. Restriction maps and chromosomal organization of type I keratin genes located on four human genomic clones. The genomic clones shown were isolated and partially mapped as described previously (52). Additional sites marked here were identified by restriction map analyses and DNA sequencing. The localization of keratin genes was determined by Southern blot analysis with ^{32}P-radiolabeled probe to the type I human K14 keratin cDNA KB-2 (52). The thick arrows mark the approximate boundaries of each keratin gene, and the direction of the arrow indicates the 5′-to-3′ end of the coding strand. The solid arrows represent genes sequenced in this report. The open arrow represents a K14 gene that had been sequenced previously (38, 39) and was determined to encode a K14 keratin protein. The hatched arrows represent regions that hybridized to the KB-2 cDNA probe. Preliminary sequence information indicates that this region is highly homologous to the K14 gene in GK-1. The stippled arrow represents another region which hybridized to the KB-2 cDNA probe. Another genomic clone (not shown) has been isolated which appears to contain the entire gene represented by the stippled arrow, but it has not yet been characterized, nor has the identity of the gene been established. The diagonal lines mark the synthetic *Eco*RI sites of the vector. Restriction endonuclease sites: B, *Bam*HI; H, *Hin*dIII; R, *Eco*RI. The following restriction endonuclease sites appear in the clones more often than indicated: A, *Ava*I; H2, *Hin*cII; K, *Kpn*I; P, *Pst*I; S, *Sac*I; St, *Stu*I. Only the specific sites shown are relevant to the text. The size scale is in kilobase pairs.

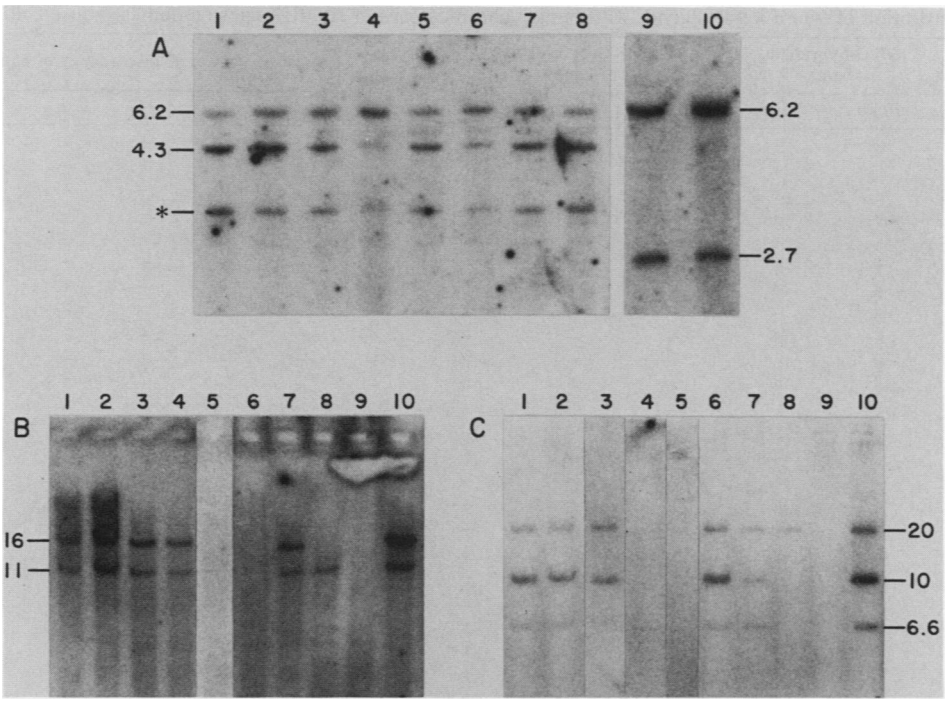

FIG. 2. Human genomic and chromosomal analyses of sequence homologies to GK-3 and GK-6. (A) Southern blot analysis of human genomic DNAs from eight individuals. Human genomic DNAs isolated from the lymphocytes of eight different individuals were digested with a mixture of *Hin*dIII and *Kpn*I restriction endonuclease and resolved by electrophoresis through a 0.8% agarose gel (lanes 1 to 8). In addition, two of these samples were digested with a mixture of *Hin*dIII, *Kpn*I, and *Hin*cII (lanes 9 and 10). DNA fragments were transferred to a nylon filter by blotting (64) and hybridized with ^{32}P-labeled cRNA made from pGK3-3'NC. After being washed, the blot was treated with RNase as described in the Materials and Methods section. Note that for lanes 1 to 8, the two hybridizing DNA bands of 4.3 and 6.2 kb were the expected *Hin*dIII-*Kpn*I bands for the complete genes contained in GK-3 and GK-6, respectively. Note that for lanes 9 and 10, the 4.3-kb *Hin*dIII-*Kpn*I fragment was further digested by *Hin*cII to yield a 2.7-kb fragment, while the 6.2-kb fragment showed no change. These results are also consistent with the restriction maps for GK-3 and GK-6, respectively. Asterisk denotes a 3.3-kb plasmid contamination present in lanes 1 to 8. Lanes 9 and 10 came from samples run on a different gel where no contamination was present. (B and C) Southern blot analysis of genomic DNAs from human-mouse cell hybrids. DNAs isolated from human-mouse cell hybrids (61) were digested with *Hin*dIII, and the DNA fragments were resolved by agarose gel electrophoresis. After transfer to nitrocellulose paper, the resolved DNA fragments were hybridized with radiolabeled probe to pGK3-3' (B) or the *Hin*dIII-*Stu*I intron probe, pGK14-int, from GK-1 (C). DNAs were from the following hybrids. (B) Lanes: 1, DUA-3BSAGA; 2, TSL-2; 3, ICL-15; 4, VTL-6; 5, XER-7; 6, XTR-3BSAgB; 7, ITW; 8, NSL-5; 9, LM/TK (mouse DNA); 10, WI38 (human DNA). (C) Lanes: 1, DUA-3BSAGA; 2, SIR-8; 3, ICL-15; 4, XTR-3BSAgB; 5, XER-7; 6, TSL-2; 7, ITW; 8, NSL-5; 9, RAG (mouse DNA); 10, WI38 (human DNA). A complete list of the chromosomes contained in each hybrid can be found in Table 1. Sizes (in kilobase pairs) are indicated and were determined by using *Hin*dIII-digested lambda DNA markers.

type I keratins (52). One clone, GK-1, contained a single gene encoding an expressed basal epidermal keratin K14 (38, 39, 52). Although the identity of the genes contained in two additional clones, GK-3 and GK-6, was not unequivocally established, each contained one homologous and seemingly overlapping keratin gene and one separate and clearly nonoverlapping keratin gene.

A *Hin*dIII site and a *Hin*cII site distinguished the seemingly overlapping portion of the GK-3 clone from the GK-6 clone (Fig. 1). Careful restriction analyses revealed additional differences. Despite these differences in restriction endonuclease sites, a radiolabeled probe to the 3' noncoding segment (GK3-3'NC) of the homologous gene in GK-3 cross-hybridized under stringent conditions with the gene contained in GK-6. Therefore, these analyses alone could not distinguish whether the restriction endonuclease site differences in the GK-3 and GK-6 clones represent polymorphic variations or whether the two highly similar clones contain recently duplicated but distinct genomic fragments.

To resolve this issue, we conducted Southern blot analyses on genomic DNAs from eight different individuals (Fig. 2A). Hybridization of *Hin*dIII-*Kpn*I-digested DNAs with ^{32}P-labeled cRNA prepared to GK3-3'NC revealed two distinct bands of 4.3 and 6.2 kb in every sample (lanes 1 to 8). When the *Hin*dIII-*Kpn*I fragments were digested with *Hin*cII prior to Southern blot analysis, the 6.2-kb fragment was unchanged, the 4.3-kb fragment disappeared, and a new 2.7-kb fragment was detected (lanes 9 and 10). These observations are consistent with the notion that the clones GK-3 and GK-6 represent two nonoverlapping and nonallelic genomic fragments even though their restriction maps are largely similar.

If the genomic clones GK-3 and GK-6 represent distinct genes, then they contain a large region of human DNA which either has been duplicated very recently or has been subjected to extreme evolutionary conservation. To determine whether these closely related sequences are genetically linked in the human genome, we used a series of human-mouse somatic cell hybrids in which different human chromosomes were lost. By correlating the presence of the retained human chromosomes with the Southern blot hybridization patterns obtained with *Hin*dIII-digested DNAs probed with radiolabeled GK3-3', the putative GK-3 and GK-6 genes could be mapped (Fig. 2B). Two distinct GK-3-

TABLE 1. Segregation of type I keratin genes with human chromosomes in *Hin*dIII-digested human-mouse cell hybrid DNA[a]

Hybrid[b] and % discordancy	GK-3-hybridizing bands		GK-1-hybridizing bands			Segregation with or % discordancy of human chromosome:							
	16 kb	11 kb	20 kb	10 kb	6.6 kb	1	2	3	4	5	6	7	8
ATR-13	+	+	+	−	+	+	+	+	+	+	+	+	+
DUA-3BSAGA*	+	+	+	+	+	−	+	−	−	−	−	+	+
DUM-13	+	+	+	+	+	+	+	+	−	+	+	+	−
ICL-15*	+	+	+	+	+	−	−	−	−	−	−	−	+
ITW	+	+	+	+	+	−	−	−	−	−	−	−	−
JSR-2	−	−	−	−	−	−	−	+	+	−	−	+	−
JSR-17S	+	+	+	+	+	+	+	+	−	+	−	t	+
LNR-5	−	−	−	−	−	−	−	−	−	−	−	−	−
NSL-5*	−	+	+	−	−	+	−	−	−	−	−	−	+
REW-7	+	+	+	+	+	+	+	+	+	+	+	+	+
REW-8D	+	+	+	+	+	−	−	−	+	−	−	−	+
REX-11BSAgB	−	−	−	−	−	−	−	+	−	−	−	−	−
REX-11BBSHF	−	−	−	−	−	−	−	+	−	−	−	−	−
SIR-8	+	+	+	+	+	+	+	+	+	+	−	+	+
SIR-11	−	−	−	−	−	−	−	−	−	−	−	+	−
TSL-2	+	+	+	+	+	−	+	t	−	+	+	−	−
VTL-6	+	+	+	+	+	−	+	−	−	−	+	+	+
VTL-17*	+	+	+	+	+	−	−	−	−	+	−	+	−
WIL-17	+	+	+	+	+	−	−	−	−	−	−	−	+
WIL-2*	+	+	+	+	+	−	−	−	−	−	−	−	+
WIL-5*	+	+	+	+	+	−	−	−	+	−	−	−	+
WIL-6	+	+	+	+	+	−	+	−	+	+	+	+	+
WIL-7	+	+	+	+	+	−	+	+	−	+	+	−	+
WIL-8	+	+	+	+	+	+	+	+	+	+	+	+	+
WIL-14*	+	+	+	+	+	−	−	+	−	−	−	+	+
WIL-15*	+	+	+	+	+	−	+	+	+	−	+	+	−
XER-7	−	−	−	−	−	+	+	+	+	+	+	+	+
XER-11	+	+	+	+	+	+	−	+	+	+	+	+	+
XOL-6	+	+	+	+	+	t	−	−	−	+	+	+	−
XOL-9	+	+	+	+	+	t	+	+	+	−	+	−	−
XOL-13*	+	+	+	+	+	t	−	−	+	+	−	+	−
XTR-3BSAgB	−	−	−	−	−	−	−	t	−	−	−	−	−
% Discordancy													
GK-3 16-kb band						55	37	53	47	37	41	39	31
GK-3 11-kb band						52	41	57	50	41	44	42	28
GK-1 20-kb band						52	41	57	50	41	44	42	28
GK-1 10-kb band						59	41	57	50	41	44	42	34
GK-1 6.6-kb band						55	37	53	47	37	41	39	31

hybridizing human DNA fragments of approximately 11 and 16 kb were detected in digests from every somatic cell hybrid containing human chromosome 17 (for representative examples, see lanes 1 to 4 and 7). In all, 24 somatic cell hybrids prepared from the DNAs of 15 different humans all contained at least one complete copy of human chromosome 17 and all contained the 11- and 16-kb *Hin*dIII fragments that hybridized with the GK3-3′ probe (Table 1). In contrast, of seven somatic cell hybrids that did not contain a part or all of one human chromosome 17, not a single genomic fragment was detected with the GK3-3′ probe.

The human chromosomes retained by the somatic cell hybrids listed in Table 1 were identified by karyotypic analysis and chromosome-specific enzyme markers (59, 62). The percent discordant segregation indicates that no chromosome other than human chromosome 17 cosegregated with all of the samples showing GK-3 hybridization. Thus, the two genomic clones GK-3 and GK-6 appear to contain segments that are both located on chromosome 17. Since each of these clones contains a portion of additional tightly linked type I keratin genes (Fig. 1), there appear to be at least four human type I keratin genes on chromosome 17.

To determine whether additional K14-related type I keratin genes might be located on chromosome 17, we repeated the experiment described above, except that in this case the probe was an intron-specific ^{32}P-labeled cDNA (GK1-int) to the K14 gene contained in GK-1 (Fig. 1). Under stringent conditions, hybridization was seen with three human genomic *Hin*dIII fragments of approximately 6.6, 10, and ≥20 kb (Fig. 2C, lane 10). With one exception, all three bands were detected in DNAs from all somatic cell hybrids containing at least one copy of human chromosome 17 (Fig. 2C, lanes 1 to 3 and 6 to 7; Table 1). In the ATR-13 hybrid (Table 1), the 10-kb band was missing. Although karyotyping showed this hybrid to have 18 chromosome 17s out of 30 metaphase spreads examined, it is possible that the hybrid ATR-13 might have a previously undetected deletion on chromosome 17.

Since all three bands appeared in hybrids involving 15 unrelated human cell lines, there seem to be three separate K14 genes which share a high degree of homology, extending even to their intron sequences. One of the three genes clearly corresponds to the K14 gene contained in GK-1 (38, 39). Hybridization and partial sequence analyses (not shown) indicate that another of these fragments corresponds to a pseudogene which is contained in its entirety within GK-7 (Fig. 1). The origin of the third gene is presently unknown.

One of the somatic cell hybrids, NSL-5, had a translocation in which a portion of the short arm of human chromo-

TABLE 1—*Continued*

Segregation with or % discordancy of human chromosome:															Trans-location(s)
9	10	11	12	13	14	15	16	17	18	19	20	21	22	X	
+	+	−	+	+	+	+	+	+	+	+	−	−	+	t	5/X
−	−	−	−	+	+	−	−	+	−	−	−	−	−	−	
−	+	+	+	−	+	t	+	+	+	+	+	+	+	t	X/15, 15/X
−	−	−	+	−	−	−	−	+	−	−	+	+	−	−	
−	−	−	−	−	−	−	−	+	−	−	−	−	−	−	
−	−	−	−	+	+	−	−	−	−	−	−	−	−	+	
+	+	+	+	+	+	+	+	+	+	−	+	+	+	−	7/9
−	−	−	−	−	−	−	−	−	−	−	−	+	+	−	
t	+	−	+	−	+	−	+	t	+	−	+	−	−	−	17/9, 12q+
−	+	+	+	+	+	+	−	+	+	+	+	+	+	+	
−	−	−	−	−	+	−	−	+	−	−	+	+	+	+	
−	+	−	−	−	+	+	−	−	+	−	−	−	−	−	
−	+	−	−	−	+	−	−	−	+	−	−	−	t	t	22/X
+	+	+	+	+	+	+	+	+	+	−	−	+	+	+	
−	−	−	−	+	−	−	−	−	−	−	−	+	+	+	
−	+	−	+	−	−	−	−	t	+	−	+	+	−	+	17/3, 3/17
−	+	+	−	−	−	+	−	+	−	+	+	+	+	−	
−	+	+	−	+	+	−	−	+	−	−	+	+	−	−	
−	−	−	+	−	+	−	−	+	+	−	−	+	−	+	
−	−	−	+	−	−	+	−	+	−	−	−	+	−	+	
−	+	−	−	−	−	−	−	+	+	−	−	+	−	+	
−	+	+	−	−	+	−	−	+	−	+	+	+	−	+	
−	+	+	−	+	+	−	−	+	+	−	−	+	−	+	
−	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
−	+	−	−	−	+	+	−	+	−	−	−	−	−	+	
−	+	+	+	+	+	+	−	+	+	−	+	+	−	+	
+	+	t	+	+	+	+	−	−	+	+	−	+	−	+	11/X
+	+	t	+	−	−	+	+	+	+	+	+	+	+	t	11/X, X/11
−	+	+	+	−	+	−	−	+	−	+	+	−	+	t	1/X
−	−	−	+	−	−	+	−	+	+	+	−	+	+	+	X/1
−	+	+	+	−	−	−	−	+	−	−	+	−	−	+	X/1
+	t	−	+	−	−	−	−	−	−	−	+	+	−	t	3/X, 10q−
71	35	37	38	56	44	45	59	0	47	50	38	31	48	35	
71	32	40	34	59	41	48	56	0	44	53	34	34	52	38	
71	32	40	34	59	41	48	56	0	44	53	34	34	52	38	
74	39	33	41	59	47	48	63	3	50	53	34	28	52	35	
71	35	37	38	56	44	45	59	0	47	50	38	31	48	35	

[a] DNA probes for the GK-3 (GK3-3′) and GK-1 (GK1-int) genes were hybridized to Southern blots containing *Hin*dII-digested DNA from the human-mouse hybrids listed in the table. Scorings were determined by the presence (+) or absence (−) of hybridizing bands in the hybrid DNAs; t, a translocated piece of the chromosome was present, but no intact chromosome. t values were not tabulated for percent discordancy.

[b] Hybrids showing only one chromosome 17 in any of the cell spreads examined by karyotyping are marked (*). Concordant hybrids have either retained or lost the keratin bands together with a specific chromosome. Discordant hybrids either retained the genes but not a specific chromosome or vice versa. Percent discordancy indicates the degree of discordant segregation for a marker and a chromosome. A 0% discordancy is the basis for chromosome assignment. Hybrid NSL-5 has a translocation (17/9) containing 17p11.2 → 17qter::9q12 → 9qter.

some 17 was missing (Table 1). Surprisingly, the GK-3 band of 11 kb (Fig. 2B, lane 8) and the K14 band of 20 kb (Fig. 2C, lane 8) were detected in the DNA from this hybrid, but the GK-6 band of 16 kb and the K14 bands of 10 and 6.6 kb were not. These data suggest that the GK-3-hybridizing band of 11 kb and the GK-1-hybridizing band of 20 kb are located on the p11.2→qter region of chromosome 17 and that the GK-3-hybridizing band of 16 kb and the GK-1-hybridizing bands of 10 and 6.6 kb are located on the p11.2→pter region of chromosome 17.

To verify that there are two sites for the type I epidermal keratin genes on chromosome 17, we performed in situ chromosomal hybridization (35). Radiolabeled probes specific for the 3′ sequence of the GK-3 gene (GK3-3′) and the intron I sequence of the GK-1-related genes (GK1-int) were hybridized to normal human metaphase chromosomes. Of 100 metaphase cells examined after hybridization with the GK3-3′ probe, 26 were labeled on one or both chromosomes 17. The distribution of labeled sites on this chromosome is illustrated in Fig. 3A. The labeled sites were clustered in two major regions: one on the long arm (q12-21) and one on the short arm (p11-12). Similar results were obtained in the hybridization with the GK1-int probe (Fig. 3B). These data agree with the somatic cell hybridization data and confirm that there are two distinct sites containing GK-1- and GK-3-related sequences on chromosome 17; the two sites are separated by about 9,000 to 15,000 kb of DNA.

The sequences of the genes contained in GK-6 and GK-3 are highly homologous. Previously, the sequence of the complete keratin gene encoded by GK-3 was determined (52). To examine the extent of homology between the GK-6 and GK-3 genes, we sequenced the equivalent 4.3-kb *Bam*HI fragment of GK-6. The sequences of the two genes are shown in Fig. 4. The putative coding portion of this gene was identified by aligning the gene sequence with the coding sequence of the gene in GK-3 (52). The nucleotide sequence within the coding portion of the GK-6 gene had a high degree of homology (94%) with the GK-3 gene. At a position

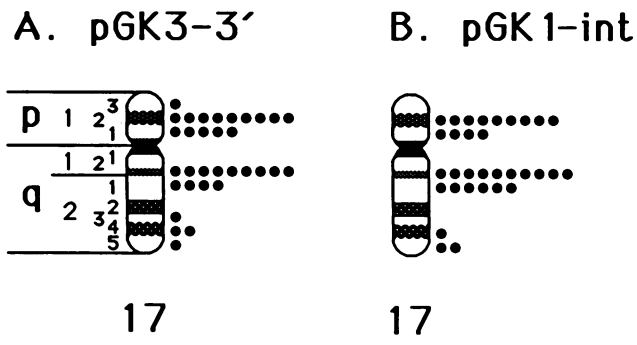

FIG. 3. Distribution of labeled sites on chromosome 17 after in situ hybridization with specific probes for the complete genes contained in GK-3 and GK-1. Normal metaphase cells were hybridized with ^{3}H-labeled probes to pGK3-3' (A) or pGK1-int (B) as described in the Materials and Methods section. Of 155 labeled sites observed in 100 metaphase cells that were examined after hybridization with the pGK3-3' probe (A), 32 were located on chromosome 17. These sites were clustered on bands 17p11-12 and 17q12-21 and were the major sites for pGK3-3' hybridization. The clusters represent 9 and 8.4% of all labeled sites ($P < 0.0005$), respectively. Smaller clusters of grains were also noted on the long arm of chromosome 14 at bands q31-32 (7 of 155, 6.5%; $P < 0.0005$) and on the long arm of chromosome 15 at band q22-24 (7 of 155, 6.5%; $P < 0.0005$). Although Southern blot analyses of the DNA from somatic cell hybrids containing these chromosomes showed no hybridization with the GK3-3' probe under stringent conditions, these may be sites of more distantly related keratin genes. Of 100 metaphase cells examined after hybridization with the pGK1-int probe (B), 28 were labeled on one or both chromosomes 17. In this case as well, we noted specific labeling at bands 17p11-12 and 17q12-21. These clusters represented 8.8 and 11%, respectively, of all labeled sites ($P < 0.0005$). Small clusters of grains were also noted on the long arm of chromosome 14 at bands q31-32 (9 of 147, 6.1%; $P < 0.0005$) and on the long arm of chromosome 21 (6 of 147, 4.1%; $P < 0.0005$). As suggested above, these may be sites of more distantly related keratin genes.

corresponding to the translation initiation site of the GK-3 gene, an ATG was found in the GK-6 sequence. With the inclusion of a single nucleotide, which, because of a compression, could not be distinguished as real or artifactual (marked by an asterisk in Fig. 4), an open reading frame followed for 473 amino acid residues (50,906 Da) until the first in-frame TAG stop codon. The apparent predicted amino acid sequence of GK-6 showed 89% identity with that of the GK-3 sequence. The most significant difference between the two sequences resided in exon VII, where 87.2% identity existed at the nucleotide level and 73% identity at the amino acid level.

In part, the structures of the putative GK-6 and GK-3 keratins were similar to those described for other keratins (for a review, see Fuchs et al. [17]). Residues 114 to 425 of the GK-6 and GK-3 keratins are predicted to be largely α-helical, and they have the heptad repeats of hydrophobic residues characteristic of all intermediate filament sequences (40). Surprisingly, however, they also seem to have a proline residue in the middle of helical domain II (residue 188). Although compression within this sequence allowed unequivocal ordering of the nucleotides for only one of the two strands of each gene, the sequence on that strand showed clear evidence of a proline encoded by both genes (due to a reading error, this finding was not previously realized [52]). Interestingly, the sequences flanking this putative proline residue are so strongly helix-promoting (5, 6) that a proline at this position is predicted to cause only a small perturbation in overall secondary structure. This is in contrast to the two more pronounced helix-breaking regions (residues 147 to 164 and 260 to 281) characteristic of all intermediate filament proteins (18, 23, 24).

At a position 105 bp 5' from the putative translation initiation codon ATG of the GK-6 gene was a putative promoter sequence, TATAAA (Fig. 4). The sequence homology with the GK-3 DNA extended to at least 425 bp 5' upstream from this regulatory element (Table 2). At a position 130 bp downstream from the TAG stop codon was a polyadenylation signal, AATAAA. The sequence homology with the GK-3 DNA extended to at least 875 bp 3' downstream from this regulatory element (Table 2). Similar to GK-3, the GK-6 gene had normal exon-intron junctions, containing the *gt* 5' border and the *ag* 3' border, with a stretch of C- and T-rich sequences near the 3' end of each intron. Introns were identically positioned and of similar sizes and sequences to those of the GK-3 gene (Table 2). Homologies between the introns of GK-6 and GK-3 ranged from 99% (intron 6) to 89% (intron 3). Thus, by sequence criteria, the GK-6 gene appears to be functional, and it seems to encode a sequence which is highly homologous to the putative keratin encoded by the GK-3 gene.

Identification of the genes contained in GK-3 and GK-6 as human K16 genes. Previously, we demonstrated that the subcloned GK3-3' cDNA specifically selected an mRNA encoding a 46-kDa keratin expressed in cultured human epidermal cells (52). On the basis of these data, we predicted that the genes contained in GK-3 and GK-6 might encode keratin K17, one of three keratins found in cultured epidermal cells but not in epidermis (14, 44, 69). Recently, however, we discovered that under certain conditions of SDS-PAGE, the 46-kDa keratin K17 is not always resolved from

FIG. 4. Comparison of the nucleotide and predicted amino acid sequences of the complete genes contained in the human genomic clones GK-3 and GK-6. The sequence of the gene contained in GK-6 and its 5'- and 3'-flanking regions are shown in comparison to the previously reported sequence (52) of the GK-3 gene. The positions of the introns corresponded precisely to those of the GK-3 gene and are indicated by solid triangles. Intron-exon junctions and pyrimidine consensus sequences are shown for each intron in lowercase letters. Complete intron sequences (not shown) were determined, and their exact sizes were compared with those of the GK-3 gene (Table 2). The exons for the two genes were identified by comparison with the coding sequence of the K14 gene (38, 39). The predicted amino acid sequence is shown in one-letter code above the nucleotide sequence. Amino acid residues encoded by the GK-3 gene that are different from those predicted for the GK-6 gene are shown below the nucleotide sequences. The gray boxes mark the α-helical domains predicted from the GK-6 amino acid sequence (5, 6). Throughout these domains are the heptad repeats of hydrophobic residues, which identify the portions of the polypeptide that are involved in coiled-coil interactions with a second keratin. Note that there are two major changes in the sequence of GK-3 over that which was reported previously (52). First, the addition of a G residue led to a reading frame change for amino acid residues 449 to 457. This change was verified by the presence of a *TaqI* restriction endonuclease site (TCGA) which was not predicted by the previous sequence. The additional insertion of previously omitted residues AG restored the original reading frame at residue 458. Second, the correction of reading errors led to the change of residues 187 to 189 from HAL to QPI (the correction leading to a proline codon in helix II is marked by brackets). These changes were verified by recloning and resequencing. Finally, the asterisk over the trinucleotide-encoding amino acid residue 306 indicates a nucleotide which was observed in the sequencing gels for both strands, but due to compression could not be unequivocally assigned as real.

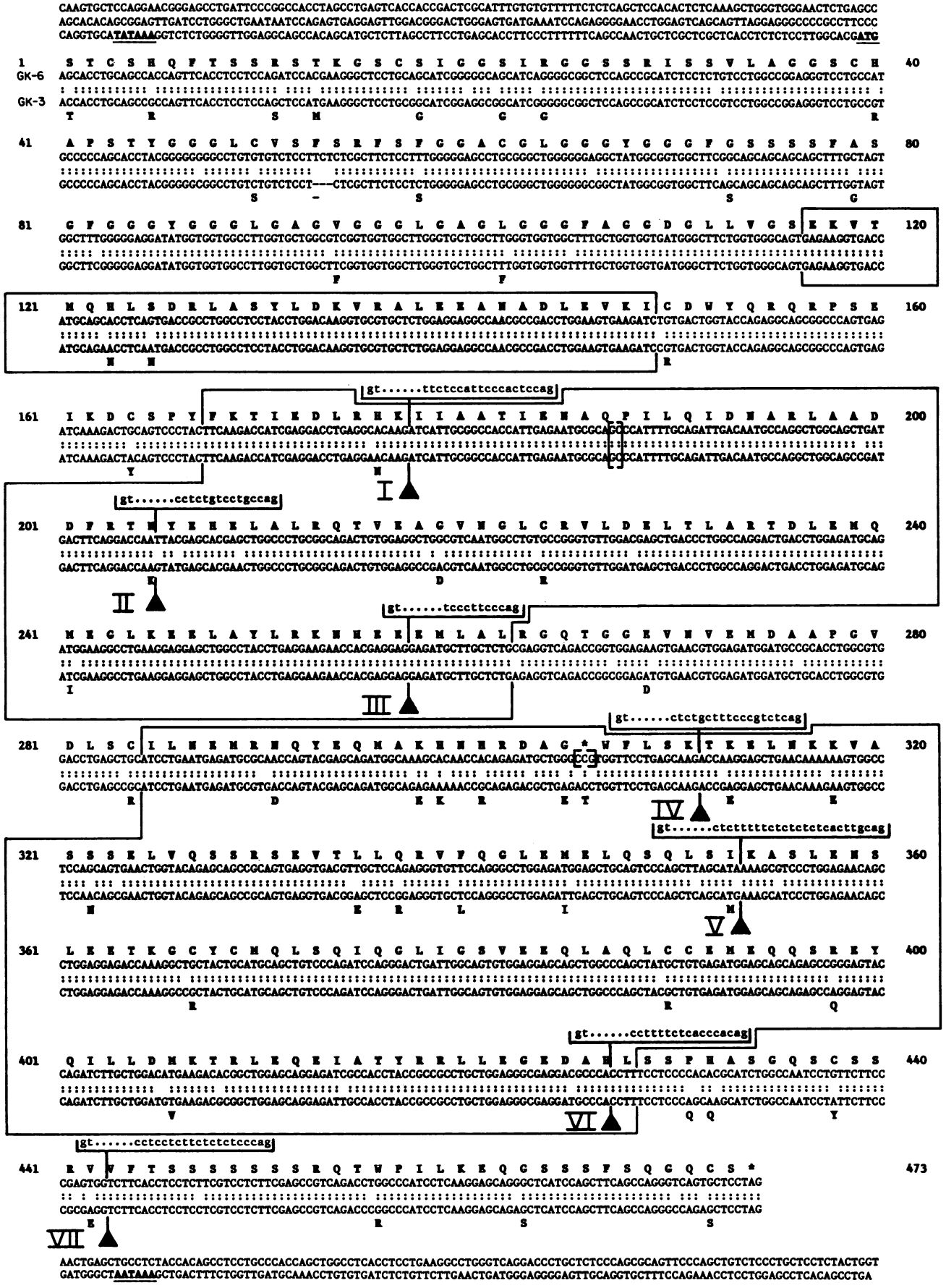
gt.....ttctccattcccactccag
gt.....cctctgtcctgccag
gt.....tcccttcccag
gt.....ctctgctttcccgtctcag
gt.....ctctttttctctctcteacttgcag
gt.....cctttttctcacccacag
gt.....cctcctctttctctctcccag
I
II
III
IV
V
VI
VII

TABLE 2. Sequence comparisons between the 4.3-kb *Bam*HI fragments of GK-3 and GK-6

Sequence	Length (bp)		% Homology, GK-3:GK-6
	GK-3	GK-6	
5′ upstream[a]	429	436	91.0
5′ untranslated[b]	82	82	91.4
Exon 1	528	531	95.1
Intron 1	428	435	93.1
Exon 2	83	83	98.8
Intron 2	135	137	89.8
Exon 3	157	157	94.9
Intron 3	120	114	89.1
Exon 4	162	162	90.7
Intron 4	89	88	94.3
Exon 5	126	126	91.3
Intron 5	298	302	93.3
Exon 6	221	221	95.9
Intron 6	94	94	98.9
Exon 7	47	47	87.2
Intron 7	159	157	94.8
Exon 8	96	96	94.8
3′ untranslated[c]	132	133	88.7
3′ downstream[d]	876	886	93.1

[a] Region between the upstream *Bam*HI site and the putative transcription initiation site.
[b] Region between the putative transcription initiation site and the translation initiation site.
[c] Region between the translation termination codon and the polyadenylation signal.
[d] Region between the polyadenylation signal and the downstream *Bam*HI site.

the 48-kDa keratin K16 (74). In these cases, both K16 and K17 showed an apparent size of 46 kDa. Thus, additional information is required to unequivocally establish the identity of the GK-3 gene.

To satisfactorily resolve this issue, we conducted Northern blot analyses with a 3′ noncoding portion probe to test for the presence of GK-3 and GK-6 mRNAs in the human cervical carcinoma line HeLa (expressing K17 but not K16 [12]) and in cultured human epidermal cells (expressing both K16 and K17 [69]). A strongly hybridizing band of 1.6 kb was detected in epidermal RNA (Fig. 5, lane 1) but not in HeLa RNA (lane 2). The failure of HeLa cell mRNA to hybridize with the GK3-3′NC probe suggested that the transcript produced by GK-3 or GK-6 was likely to encode K16 rather than K17. This assignment was confirmed in subsequent investigations (see below).

The GK-3 but not the GK-6 gene is a functional K16 gene. To determine whether one or both of our sequenced genes were functional and gave rise to K16 proteins, we used gene transfection to express their encoded proteins. To optimize our chances of detecting expression, we prepared the chimeric plasmids pJ1GK3 · Bam and pJ1GK6 · Bam, containing their respective 4.3-kb *Bam*HI fragments inserted 3′ to the SV40 promoter, enhancer, and origin of replication (see Fig. 6A for details). These plasmids were transfected into the kidney epithelial (COS) cells (42), which express the SV40 T antigen, thereby enabling plasmids with an SV40 origin of replication to be amplified several hundredfold. The two chimeric plasmids were transfected into COS cells, and at 50 h posttransfection, the cells were radiolabeled with [^{35}S] methionine. Intermediate filament proteins were extracted and resolved by gel electrophoresis (Fig. 7A). The pJ1 GK3 · Bam-transfected cells produced an intermediate filament protein of 46 to 48 kDa (lane 1). This protein was not detected in the extracts of pJ1GK6 · Bam-transfected or untransfected COS cells (lanes 2 and 3). Total protein extracts of pJ1GK6 · Bam-transfected and untransfected COS cells also showed no differences, suggesting that if the pJ1GK6 · Bam construct encoded a protein, it was expressed at significantly lower levels than the pJ1GK3 · Bam construct. The difference did not seem to be due to inequalities in transfection efficiencies, since mixtures of the two chimeric plasmids had an efficiency of transfection similar to that of pJ1GK3 · Bam alone.

To examine whether the GK-3 expressed protein had an electrophoretic mobility which was identical to that of bona fide K16, we used two-dimensional nonequilibrium-pH PAGE (47) to resolve the radiolabeled extract from pJ1GK3 · Bam-transfected cells (Fig. 7B). The radiolabeled 46- to 48-kDa protein comigrated with the human K16 which was present in a mixture of cold, added keratin markers from cultured human epidermal cells. These data were in good agreement with the predicted pI (4.9) of the GK-3 gene product and suggest that the GK-3 gene is the one which is typically expressed in a variety of different epithelial tissues (44). In contrast, the predicted pI (5.4) of the putative GK-6 gene product was more basic than that of any 46- to 48-kDa keratins that have thus far been reported.

To examine whether the pJ1GK6 · Bam-transfected COS cells were expressing a stable mRNA, we isolated total RNAs at 65 h posttransfection (Fig. 8). Nothern blot analysis with a radiolabeled probe to GK-6 revealed the presence of a 1.6-kb hybridizing band in the pJ1GK6 · Bam-transfected (lane 2) but not the untransfected (lane 1) COS cells. However, additional hybridization was seen over a broad range of 1.6 to 2.5 kb, indicating that transcriptional initiation was occurring from both the SV40 and the keratin promoters and possibly from sites between the two promoters as well. In contrast, the same amount of total RNA from pJ1GK3 · Bam-transfected cells showed approximately a 10 to 50 times higher level of the 1.6-kb band, with little or no evidence of utilization of the SV40 promoter (lane 3; see also lane 7 for lower exposure). Moreover, whereas GK-3 gene expression was readily detected in COS cells transfected

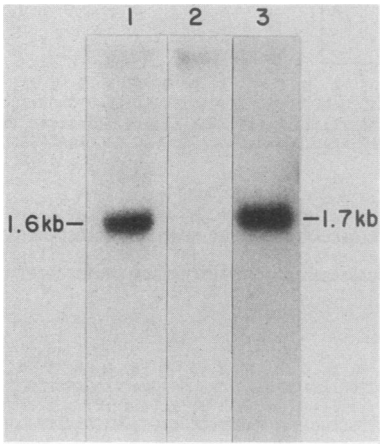

FIG. 5. Expression of the mRNA encoded by GK-3 in cultured human epidermal cells but not HeLa cells. Poly(A)$^+$ mRNAs were resolved by electrophoresis through a formaldehyde-agarose gel and transferred to nitrocellulose paper by blotting (70). The blot was then hybridized with ^{32}P-labeled cRNA probe specific for either the 3′ noncoding portion of GK-3 (pGK3-3′NC; lanes 1 and 2) or, as a positive control, the coding portion of K7 (lane 3) (20). mRNAs were from (lane 1) cultured human epidermal cells or (lanes 2 and 3) HeLa cells.

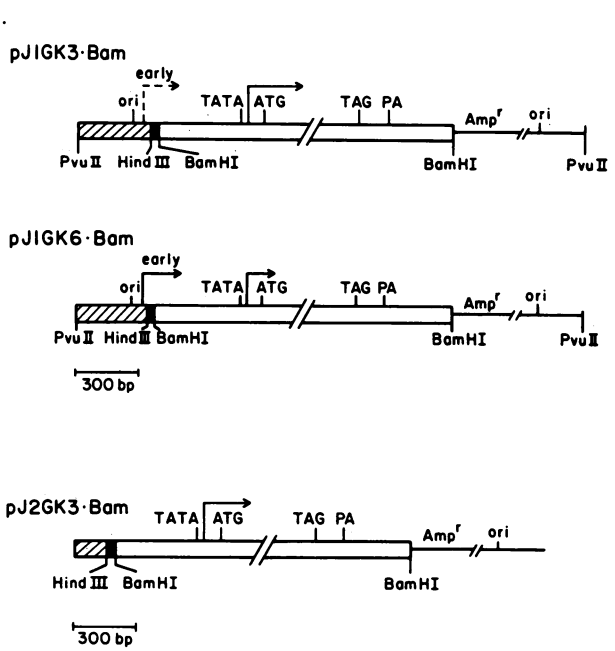

FIG. 6. Genetic maps of pJ1GK3 · Bam, pJ1GK6 · Bam, and pJ2GK3 · Bam. The complete human type 1 genes contained in GK-3 and GK-6 are each located within a 4.3-kb *Bam*HI fragment which also includes 400 bp of 5′ upstream sequences and 1 kb of 3′ downstream sequences. These fragments were inserted into the *Bam*HI site of plasmid pJ1 (A) or plasmid pJ2 (B). Plasmid pJ1 is a bacterial plasmid containing the ampicillin resistance gene and pBR322 origin of replication. In addition, 5′ upstream from a polylinker segment, this vector contains the origin of replication, early promoter, and 72-bp repeat enhancer sequences of the viral SV40 DNA (33). Plasmid pJ2 is similar to pJ1 except that it contains the 72-bp repeat enhancer sequences of SV40 DNA but not its promoter or origin of replication (J. Morgenstern and R. Weinberg, personal communication). The hatched boxes indicate SV40 sequences, the solid boxes indicate the polylinker sequences, and the open boxes indicate GK-3 and GK-6 sequences. The thin line represents bacterial sequences. The TATA box, ATG translation initiation start codon, TAG translation stop codon, and polyadenylation site (PA) of the GK-3 and GK-6 genes are indicated. The solid arrows indicate the putative transcriptional initiation start sites utilized by the cultured cells (as judged by the size of the resulting mRNAs); the dotted line indicates the transcription initiation start site from the SV40 early promoter (pJ1).

with pJ2GK3 · Bam containing only the SV40 enhancer and not the SV40 promoter (see Fig. 6B for details of the construct), GK-6 expression in this construct was below the limits of detection. Collectively, these data suggest that the promoter for the GK-6 gene may be significantly weaker than that of the GK-3 gene.

To test whether the promoter was the only deficient portion of the GK-6 gene, we prepared hybrid GK-3–GK-6 genes linked through their common single *Kpn*I site (see Fig. 1 for details). When transfected into COS cells, pJ1 hybrid plasmids containing the 5′ portion of GK-6 and the 3′ portion of GK-3 (Fig. 8, lane 4) or the 3′ portion of GK-6 and the 5′ portion of GK-3 (lane 5) both produced RNA levels that were significantly lower than that of the GK-3 construct. These data indicate that there must be multiple sites in the GK-6 gene that lead to its markedly reduced functionality. In this context, it is interesting that the keratin gene which is tightly linked 3′ to this GK-6 gene encodes an obvious K14 pseudogene, with multiple abnormal intron-exon splice signals (Rosenberg and Fuchs, unpublished observations). Thus, an entire duplicated keratin gene region on chromosome 17 may have been targeted for inactivation.

A GK-3 keratin antibody can specifically detect human K16 protein in denatured and filamentous forms. While several antibodies are presently available that can distinguish between K16 and K17 (16, 72), none of these are monospecific for K16. To produce a specific K16 antiserum, we injected rabbits with a synthetic peptide which contained the same three carboxy-terminal residues as the GK-3 keratin. As judged in immunoblot analysis, a heptapeptide containing the last seven residues of the GK-3 keratin competed effectively and specifically for antibody binding when present at concentrations as low as 30 μg/ml of antiserum (RayChaudhury and Fuchs, unpublished results). When this antiserum (5547) was used for immunoblot analysis of intermediate filament extracts of pJ2GK3Bam-transfected mouse NIH 3T3 cells, a single 48-kDa band was selected in the transfected (Fig. 9, lane 4) but not the untransfected (lane 5) samples. A cross-reacting band of identical size was also observed in extracts of the squamous cell carcinoma line SCC-13 (lane 1) and of epidermal cells (lane 2). As expected, HeLa cell extracts showed no cross-reaction with the GK-3 antiserum (lane 3). These results indicate that the 5547 antiserum cross-reacts with the K16 protein encoded by GK-3 but it does not cross-react with other closely related keratins, e.g., K14 and K17.

In contrast to the specificity of the 5547 antiserum, the monoclonal antibody AE1 (72) cross-reacted with a number of keratins, including K14, K16, and K19, even though it did

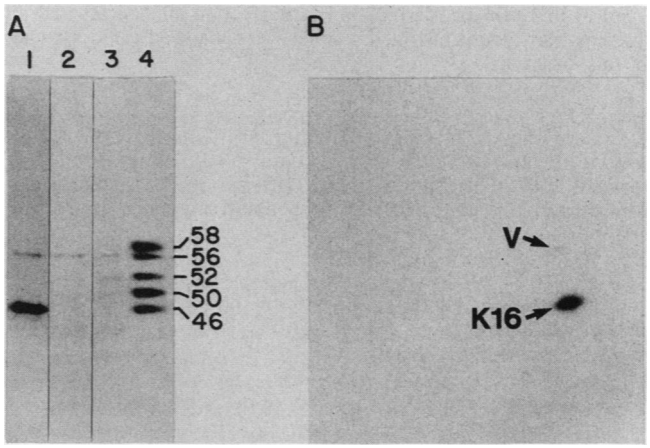

FIG. 7. Transfection of pJ1GK3 · Bam (but not pJ1GK6 · Bam) into COS cells yields an expressed K16 human keratin. (A) Plasmids pJ1GK3 · Bam and pJ1GK6 · Bam were transfected into COS cells, and at 50 h posttransfection, [^{35}S]methionine was added to the culture medium. After a 15-h labeling period, intermediate filament proteins were extracted and resolved by SDS-PAGE. The gel was subjected to fluorography and autoradiography. Extracts: lane 1, pJ1GK3 · Bam-transfected COS cells; lane 2, pJ1GK6 · Bam-transfected COS cells; lane 3, untransfected COS cells; lane 4, cultured human epidermal cells. Sizes (in kilodaltons) are shown at right. (B) Radiolabeled intermediate filament extract from pJ1GK3 · Bam-transfected COS cells was resolved by two-dimensional nonequilibrium-pH gel electrophoresis (47) with nonequilibrium-pH gel electrophoresis in the first dimension and SDS-PAGE in the second dimension. In a control gel (not shown), the radiolabeled extract was combined with 10 to 20 μg of unlabeled protein extract from cultured human epidermal cells prior to analysis. Control proteins were visualized by staining with Coomassie blue. Radiolabeled proteins were visualized after fluorography, autoradiography, and exposure to X-ray film.

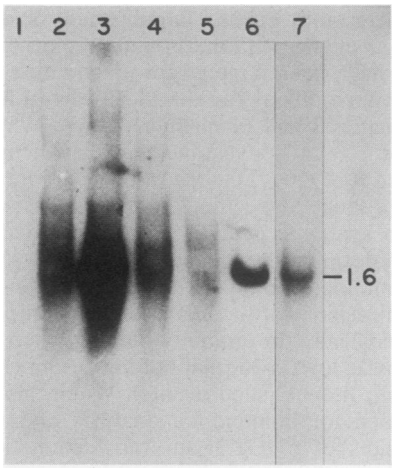

FIG. 8. Northern blot analysis of RNAs extracted from transfected COS cells. Equal amounts of total RNAs from transfected and untransfected COS cells were fractionated by agarose-formaldehyde gel electrophoresis and transferred to nitrocellulose by blotting (70). RNAs: lane 1, untransfected cells; lane 2, pJ1GK6 · Bam-transfected cells; lane 3, pJ1GK3 · Bam-transfected cells; lane 4, pJ1GK6-5′/GK3-3′Bam-transfected cells; lane 5, pJ1GK3-5′/GK6-3′Bam-transfected cells; lane 6, cultured human epidermal cells; lane 7, the same sample as in lane 3, but with a shorter exposure to allow visualization of the discrete mRNA. The blot was hybridized with a ^{32}P-labeled cDNA probe from the purified GK-6 insert. The blot was washed at 65°C in 0.1× SSC–0.1% SDS and then subjected to autoradiography. Lanes 1 to 6 were developed after 3 days of exposure to film. The size of the major hybridizing band is indicated in kilobases. Note: the difference between GK-3 and GK-6 mRNA levels was variable, ranging from about 50 times (shown here, lanes 2 and 3) to 10 times. Variations in transfection efficiencies and in time posttransfection at which samples were analyzed could account for this difference.

not recognize K17 (Fig. 9B, compare lanes 3 and 4). A polyclonal antiserum prepared against total epidermal cell keratins (16) also cross-reacted with a number of keratins, although in this case, its affinity was stronger for K17 than for K16 (Fig. 9C, compare lanes 3 and 4). Thus, while the data in lanes 3 and 4 of Fig. 9B and C demonstrate that two previously available antibodies can be used in an immunoblot analysis to unequivocally identify the transfected GK-3 gene product as a K16 keratin and to distinguish the closely related K16 and K17 keratins, these antibodies would not be useful to identify the presence of K16 in vivo, where other closely related type I keratins (e.g., K14) are usually coexpressed.

To verify that the 5547 antiserum can detect K16 in a filamentous form and that the GK-3 gene product is fully competent for filament assembly, we transfected the marsupial kidney epithelial cell line PtK2 with pJ2GK3 · Bam. At 65 h after transfection, the cells were fixed and subjected to double immunofluorescence (Fig. 10). The 5547 antiserum was used to detect the transfected gene product (Fig. 10A), and a monoclonal antibody LE41 (34) was used to examine the endogenous PtK2 keratin network (Fig. 10B). Our results clearly demonstrate that the GK-3 gene encodes a functional keratin, which incorporates into the endogenous keratin filament network of the transfected PtK2 cells. In addition, it is clear that the carboxy terminus of the K16 protein was not masked as a consequence of filament formation.

DISCUSSION

Sorting out the K16 genes: which one(s) contributes to K16 expression? The mRNAs encoding the four type I keratins of cultured human epidermal cells (K13, K14, K16, and K17) were known to be highly homologous by their ability to cross-hybridize with a cloned K14 cDNA (13, 29). In this paper, we have described the isolation and characterization of two human keratin genes, both of which hybridized strongly with an mRNA encoding a 46- to 48-kDa protein. One of these genes was contained in a 4.3-kb *Bam*HI fragment of a genomic clone GK-3, and we have unequivocally assigned the protein translated from GK-3 hybrid-selected mRNA as K16. We have made an antiserum that is specific for the carboxy-terminal end of K16, and by immunoblot analysis, we have confirmed the expression of the GK-3-encoded K16 protein in cultured human epidermal cells and in a squamous cell carcinoma line (SCC-13). When

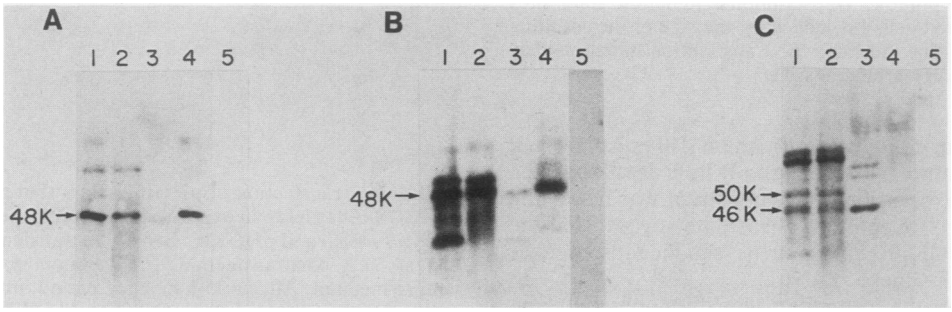

FIG. 9. Immunoblot analysis of human keratins probed with the GK-3-specific antiserum. NIH 3T3 fibroblasts were transfected with plasmid pJ2GK3 · Bam by the calcium phosphate precipitation method of Graham and van der Eb (22). At 65 h posttransfection, the cells were harvested, and intermediate filament proteins were extracted as described in the Materials and Methods section. Intermediate filament proteins were also extracted from untransfected 3T3 cells (keratin negative), from the squamous cell carcinoma line SCC-13 (K5$^+$, K6$^+$, K13$^+$, K14$^+$, K16$^+$, K17$^+$, K19$^+$), from cultured human epidermal cells (K5$^+$, K6$^+$, K14$^+$, K16$^+$, K17$^+$), and from HeLa cells (K7$^+$, K8$^+$, K17$^+$, K18$^+$, K19$^+$). Proteins were resolved by SDS-PAGE (8.5% polyacrylamide) and transferred to nitrocellulose paper by blotting (71). Blots were incubated for 1.5 h with the following antibodies: a polyclonal antiserum (5547) against the carboxy terminus of the GK-3-encoded protein (A); AE1, a monoclonal antibody that cross-reacts with K16 but not K17 (B); and a polyclonal antiserum against cultured human epidermal keratins (C). The antiserum used for C cross-reacts more strongly with K17 than with K16. After washing, all blots were treated with ^{125}I-labeled *S. aureus* protein A (Amersham) to detect bound antibodies. Proteins: lane 1, SCC-13 cells; lane 2, epidermal cells; lane 3, HeLa cells; lane 4, pJ2GK3 · Bam-transfected NIH 3T3 cells; lane 5, untransfected 3T3 cells. Note that the cross-reacting (56K) band in (A) (lanes 1 and 2) is either K6 or an artifact as discussed in Materials and Methods. Note that at much longer exposure times than those shown here, a 5547-cross-reacting protein of 46 to 48 kDa could be detected in the untransfected NIH 3T3 cell extract. Masses are shown in kilodaltons.

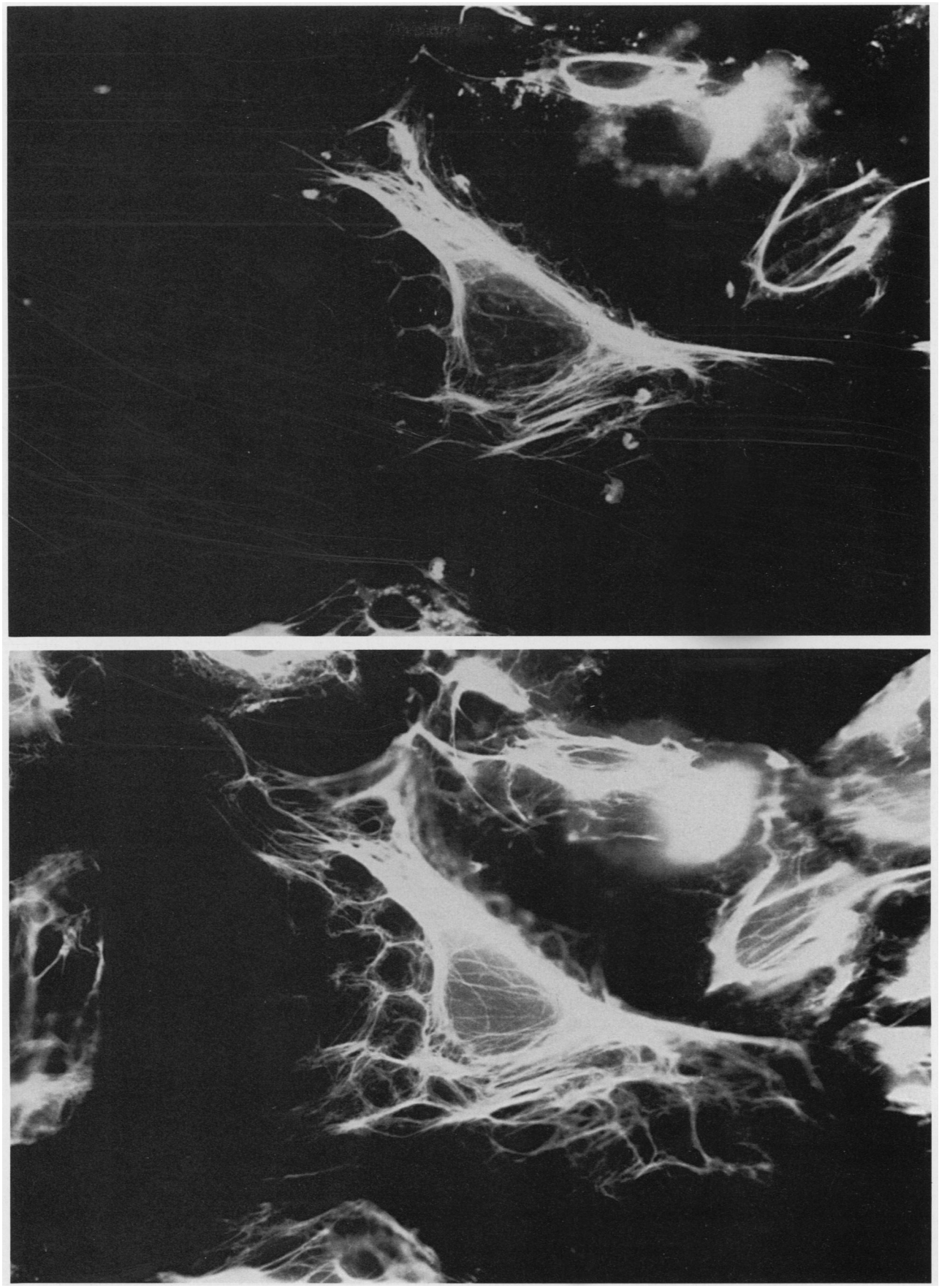

FIG. 10. Integration of the GK-3-encoded human keratin into the endogenous keratin network of cultured PtK2 kidney epithelial cells. PtK2 cells were transfected with pJ2GK3 · Bam DNA and fixed 65 h later in methanol (−20°C). Cells were subjected to double immunofluorescence staining with (A) the 5547 antiserum specific for the GK-3-encoded protein (anti-K16) to detect the transfected gene product and (B) the monoclonal antibody LE41 to detect the endogenous keratin filament network.

K16 was expressed as a foreign gene product in GK-3-transfected simple epithelial cells, the protein integrated into the endogenous keratin filament network, presumably drawing upon K8 as a promiscuous partner in filament assembly. Hence, the GK-3 keratin K16 had all of the properties of a functional keratin protein. In contrast, while the GK-6 gene was highly homologous to the GK-3 gene in almost every respect, there seemed to be concealed abnormalities in this gene that could only be detected by the failure of this gene to produce significant levels of stable K16 mRNA and K16 protein.

Most pseudogenes in mammalian genomes seem to arise from either an mRNA-mediated event, leading to an intronless gene (46, 55, 75, 77, 79), or truncation of the gene, presumably as a consequence of an aberrant genetic recombination (77). However, there is certainly precedent for more subtle genetic alterations leading to a loss of gene function. Such examples are prevalent in the beta-globin gene family: both in their introns and in their exons, these pseudogenes frequently show extensive homology with the functional beta-globin gene, with the exception of one or more of the following: point mutations in intron splicing consensus sequences; small deletions or insertions, often giving rise to frameshift mutations in the coding regions; or mutations in the promoter or polyadenylation sequences (4, 7, 43, 51, 56, 57). With the possible exception of the ambiguous nucleotide in exon 4, the GK-6 gene showed no obvious pseudogene mutations, even of the more subtle type seen in the globin gene family. Nonetheless, our expression data demonstrated the presence of defects in both the 5′ and 3′ portions of the GK-6 gene, suggesting that for all practical purposes, it is nonfunctional. Thus, there seems to be only a single active K16 gene in the human genome.

Are the keratin filaments composed of K6/K16 as stable as those made of K5/K14? We were intrigued to discover the apparent presence of a proline residue within the second α-helical domain of the K16 genes. Our in vivo transfection studies demonstrated clearly that K16 was able to incorporate into an existing keratin filament network, a finding which has been confirmed by in vitro investigations (25). Interestingly, however, when cultured human epidermal cell keratins are resolved by nonequilibrium-pH PAGE, K5, K6, K14, and K17 frequently show evidence of dimer/tetramer complex formation in 9 M urea, while K16 consistently shows no tendency to do so (see, for example, Tyner and Fuchs [74]). These results suggest that the presence of a proline residue in the helical region of K16 may lower the stability of the coiled-coil interactions.

A difference in the properties of keratin filaments made with K16 rather than other non-proline-containing type I keratins could provide a possible basis for the curious pattern of expression of K16 in epidermis. In contrast to the closely related keratin K14 (38, 39), K16 is only expressed in epidermis transiently, e.g., during wound healing (E. B. Lane, personal communication), in hyperproliferative diseases of the skin (76), and in cultured epidermal cells (69). One working hypothesis might be that the presence of K16 provides a more dynamic cytoskeleton which could enable the epidermal cell to either divide at a faster rate or to undergo an increased number of cell divisions prior to a commitment to terminally differentiate. Further studies will be necessary to test this notion.

Is the chromosomal organization of the type I keratin genes important for their regulation? Our investigations have revealed that the keratin gene family may have undergone the sort of evolutionary expansion that has been seen previously for other large multigene families. We have found that at least six type I epidermal keratin genes are located on human chromosome 17. In at least two cases, the genes are tightly linked, and at least one region (containing the K16 genes) has been duplicated. Moreover, since the introns and flanking portions of not only the K16 genes but also the K14 genes are highly homologous, it is likely that additional exchange of genetic information via either gene conversion or gene duplication may have taken place, enabling these genes to evolve in concert as a multigene family.

The finding that these large duplicated segments of DNA containing two of the K16 genes and three K14 genes are localized on human chromosome 17 is interesting in light of the coordinate expression of these two keratins in a number of different epithelial tissues, including epiglottis, tongue, hair follicle, and esophagus (44). Additional studies will be necessary to evaluate whether the chromosomal locations of the K14 and K16 genes play a role in their coordinate expression and, in particular, whether the functional genes contained in the GK-1 and GK-3 segments are tightly linked. As more investigations are conducted on the organization and complexity of the human keratin gene family and on defining the promoter and enhancer elements of these genes, we hope to be able to unravel the molecular mechanisms responsible for the coexpression of different sets of keratins in different epithelia and at various stages of development and differentiation.

ACKNOWLEDGMENTS

We thank Ed Fritsch (Genetics Institute, Boston) for providing us with a partial *Hae*III-*Alu*I human genomic library in phage Charon 4A. We are grateful to E. Birgitte Lane (ICRF, London) for her generous gift of the monoclonal antibody LE41 and Henry Sun (New York University School of Medicine) for his generous gift of the monoclonal antibody AE1. Paul Sheppard (Cambridge Research Biochemicals, Cambridge, England) was very helpful in synthesizing the tetradecapeptide linked to keyhole limpet hemocyanin which was used as an antigen in anti-K16 antibody preparation. We acknowledge the helpful assistance of Roger Eddy, Jr. (Roswell Park Memorial Institute), for preparing the Southern blots of human-mouse hybrid DNAs, Rafael Espinosa III (University of Chicago) for his technical assistance in in situ hybridizations, and Grazina Traska (University of Chicago) for her assistance in tissue culture. Finally, special thanks to Douglas Marchuk (Hope College, Michigan) for his initial contributions to this project and to the other members of the Fuchs laboratory for their helpful suggestions.

This work was supported by Public Health Service grant AR 27883 from the National Institutes of Health, the Whirlpool Foundation, and the Cancer Research Foundation. E.F. is a Presidential Young Investigator. M.R. is a National Institutes of Health Predoctoral Trainee (HD07136). M.M.L. is a scholar of the Leukemia Society of America.

LITERATURE CITED

1. **Anderson, S.** 1981. Shotgun DNA sequencing using cloned DNAse I-generated fragments. Nucleic Acids Res. **9**:3015–3026.
2. **Bader, B. L., T. M. Magin, M. Hatzfeld, and W. W. Franke.** 1986. Amino acid sequence and gene organization of cytokeratin no. 19, an exceptional tail-less intermediate filament protein. EMBO J. **5**:1865–1875.
3. **Berk, A., and P. Sharp.** 1978. Spliced early mRNAs of SV40. Proc. Natl. Acad. Sci. USA **75**:1274–1278.
4. **Brunner, A. M., J. C. Schimenti, and C. H. Duncan.** 1986. Dual evolutionary modes in the bovine globin locus. Biochemistry **25**:5028–5035.
5. **Chou, P. Y., and G. D. Fasman.** 1978. Prediction of the secondary structure of proteins from their amino acid sequence. Adv. Enzymol. **47**:45–148.
6. **Chou, P. Y., and G. D. Fasman.** 1979. Prediction of beta-turns.

Biophys. J. 26:367–383.
7. **Cleary, M. L., E. A. Schon, and J. B. Lingrel.** 1981. Two related pseudogenes are the result of a gene duplication in the goat beta-globin locus. Cell 26:181–190.
8. **Crewther, W. G., L. M. Dowling, D. A. D. Parry, and P. M. Steinert.** 1983. The structure of intermediate filaments. Int. J. Biol. Macromol. 5:267–282.
9. **DeLeon, D. V., K. H. Cox, L. M. Angerer, and R. C. Angerer.** 1983. Most early-variant histone mRNA is contained in the pronucleus of sea urchin eggs. Dev. Biol. 100:197–206.
10. **Eichner, R., P. Bonitz, and T.-T. Sun.** 1984. Classification of epidermal keratins according to their immunoreactivity, isoelectric point, and mode of expression. J. Cell Biol. 98:1388–1396.
11. **Eichner, R., T.-T. Sun, and U. Aebi.** 1986. The role of keratin subfamilies and keratin pairs in the formation of human epidermal intermediate filaments. J. Cell Biol. 102:1767–1777.
12. **Franke, W. W., D. L. Schiller, R. Moll, S. Winter, E. Schmid, I. Engelbrecht, H. Denk, R. Krepler, and B. Platzer.** 1981. Diversity of cytokeratins: differentiation specific expression of cytokeratin polypeptides in epithelial cells and tissues. J. Mol. Biol. 153:933–959.
13. **Fuchs, E., S. Coppock, H. Green, and D. Cleveland.** 1981. Two distinct classes of keratin genes and their evolutionary significance. Cell 27:75–84.
14. **Fuchs, E., and H. Green.** 1978. The expression of keratin genes in epidermis and cultured epidermal cells. Cell 15:887–897.
15. **Fuchs, E., and I. Hanukoglu.** 1983. Unravelling the structure of the intermediate filament. Cell 34:332–334.
16. **Fuchs, E., and D. Marchuk.** 1983. Type I and type II keratins have evolved from lower eukaryotes to form the epidermal intermediate filaments in mammalian skin. Proc. Natl. Acad. Sci. USA 80:5857–5861.
17. **Fuchs, E., A. L. Tyner, G. J. Giudice, D. Marchuk, A. Ray-Chaudhury, and M. Rosenberg.** 1987. The human keratin genes and their differential expression. Curr. Top. Dev. Biol. 22:5–34.
18. **Geisler, N., E. Kaufmann, and K. Weber.** 1982. Protein chemical characterization of three structurally distinct domains along the protofilament unit of desmin 10 nm filaments. Cell 30:277–286.
19. **Geisler, N., E. Kaufmann, and K. Weber.** 1985. Antiparallel orientation of the two double-stranded coiled-coils in the tetrameric protofilament unit of intermediate filaments. J. Mol. Biol. 182:173–177.
20. **Glass, C., K. H. Kim, and E. Fuchs.** 1985. Sequence and expression of a human type II mesothelial keratin. J. Cell Biol. 101:2366–2373.
21. **Gorman, C.** 1985. High efficiency gene transfer into mammalian cells, p. 143–190. *In* D. M. Glover (ed.), DNA cloning II: a practical approach. IRL Press, Oxford.
22. **Graham, F. L., and A. J. van der Eb.** 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology 52:456–467.
23. **Hanukoglu, I., and E. Fuchs.** 1982. The cDNA sequence of a human epidermal keratin: divergence of sequence but conservation of structure among intermediate filament proteins. Cell 31:242–252.
24. **Hanukoglu, I., and E. Fuchs.** 1983. The cDNA sequence of a type II cytoskeletal keratin reveals constant and variable structural domains among keratins. Cell 33:915–924.
25. **Hatzfeld, M., and W. W. Franke.** 1985. Pair formation and promiscuity of cytokeratins: formation *in vitro* in heterotypic complexes and intermediate-sized filaments by homologous and heterologous recombinations of purified polypeptides. J. Cell Biol. 101:1826–1841.
26. **Johnson, L., W. Idler, X.-M. Zhou, D. Roop, and P. Steinert.** 1985. Structure of a gene for the human epidermal 67-kDa keratin. Proc. Natl. Acad. Sci. USA 82:1896–1900.
27. **Jorcano, J. L., J. K. Franz, and W. W. Franke.** 1984. Amino acid sequence diversity between bovine epidermal cytokeratin polypeptides of the basic (type II) subfamily as determined from cDNA clones. Differentiation 28:155–163.
28. **Jorcano, J. L., M. Reiger, J. K. Franz, D. L. Schiller, R. Moll, and W. W. Franke.** 1984. Identification of two types of keratin polypeptides within the acidic cytokeratin subfamily I. J. Mol. Biol. 179:357–281.
29. **Kim, K. H., J. Rheinwald, and E. Fuchs.** 1983. Tissue specificity of epithelial keratins: differential expression of mRNAs from two multigene families. Mol. Cell. Biol. 3:495–502.
30. **Knapp, B., M. Rentrop, J. Schweizer, and H. Winter.** 1986. Nonepidermal members of the keratin multigene family: cDNA sequences and in situ localization of the mRNAs. Nucleic Acids Res. 14:751–763.
31. **Knapp, B., M. Rentrop, J. Schweizer, and H. Winter.** 1987. Three cDNA sequences of mouse type I keratins: cellular localization of the mRNAs in normal and hyperproliferative tissues. J. Biol. Chem. 262:938–945.
32. **Kunkel, L. M., K. D. Smith, S. H. Boyer, D. S. Borgaonkar, S. S. Wachtel, O. J. Miller, W. R. Breg, H. W. Jones, Jr., and J. M. Rary.** 1977. Analysis of human Y-chromosome-specific reiterated DNA in chromosome variants. Proc. Natl. Acad. Sci. USA 74:1245–1249.
33. **Land, H., A. C. Chen, J. P. Morgenstern, L. F. Parada, and R. A. Weinberg.** 1986. Behavior of *myc* and *ras* oncogenes in transformation of rat embryo fibroblasts. Mol. Cell. Biol. 6: 1917–1925.
34. **Lane, E. B.** 1982. Monoclonal antibodies provide specific intramolecular markers for the study of epithelial tonofilament organization. J. Cell Biol. 92:665–673.
35. **Le Beau, M. M., C. A. Westbrook, M. O. Diaz, and J. D. Rowley.** 1984. Evidence for two distinct c-*src* loci on human chromosomes 1 and 20. Nature (London) 312:70–71.
36. **Magin, T. M., J. L. Jorcano, and W. W. Franke.** 1986. Cytokeratin expression in simple epithelia. II. cDNA cloning and sequence and characteristics of bovine cytokeratin A (no. 8). Differentiation 30:254–264.
37. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
38. **Marchuk, D., S. McCrohon, and E. Fuchs.** 1984. Remarkable conservation of structure among intermediate filament genes. Cell 39:491–498.
39. **Marchuk, D., S. McCrohon, and E. Fuchs.** 1985. Complete sequence of a gene encoding a human type I keratin: sequences homologous to enhancer elements in the regulatory region of the gene. Proc. Natl. Acad. Sci. USA 82:1609–1613.
40. **McLachlan, A.** 1978. Coiled-coil formation and sequence regularities in the helical regions of alpha-keratin. J. Mol. Biol. 124:297–304.
41. **McLachlan, A., and M. Stewart.** 1975. Tropomyosin coiled-coil interactions: evidence for an unstaggered structure. J. Mol. Biol. 98:293–304.
42. **Mellon, P., V. Parker, Y. Gluzman, and T. Maniatis.** 1981. Identification of DNA sequences required for transcription of the human alpha1-globin gene in a new SV40 host-vector system. Cell 27:279–288.
43. **Menon, A. G., and J. B. Lingrel.** 1986. Structural and functional analysis of the goat epsilon-globin genes. Gene 42:141–150.
44. **Moll, R., W. W. Franke, D. L. Schiller, B. Geiger, and R. Krepler.** 1982. The catalog of human cytokeratins: patterns of expression in normal epithelia, tumors and cultured cells. Cell 31:11–24.
45. **Nelson, W., and T.-T. Sun.** 1983. The 50- and 58-kdalton keratin classes as molecular markers for stratified squamous epithelia: cell culture studies. J. Cell Biol. 97:244–251.
46. **Nishioka, Y., A. Leder, and P. Leder.** 1980. Unusual alpha-globin-like gene that has cleanly lost both globin intervening sequences. Proc. Natl. Acad. Sci. USA 77:2806–2809.
47. **O'Farrell, P. Z., H. M. Goodman, and P. H. O'Farrell.** 1977. High-resolution two-dimensional electrophoresis of basic as well as acidic proteins. Cell 12:1133–1142.
48. **Oshima, R. G., J. L. Millan, and G. Cecena.** 1986. Comparison of mouse and human keratin 18: a component of intermediate filaments expressed prior to implantation. Differentiation 33:61–68.
49. **Parker, B. A., and G. R. Stark.** 1979. Regulation of simian virus 40 transcription: sensitive analysis of the RNA species present early in infections by virus or viral DNA. J. Virol. 31:360–369.

50. **Parry, D. A. D., A. C. Steven, and P. M. Steinert.** 1985. The coiled-coil molecules of intermediate filaments consists of two parallel chains in exact axial register. Biochem. Biophys. Res. Commun. **127**:1012–1018.
51. **Proudfoot, N. J., and T. Maniatis.** 1980. The structure of a human alpha-globin pseudogene and its relationship to alpha-globin gene duplication. Cell **21**:537–544.
52. **RayChaudhury, A., D. Marchuk, M. Lindhurst, and E. Fuchs.** 1986. Three tightly linked genes encoding human type I keratins: conservation of sequence in the 5′ untranslated leader and 5′ upstream regions of coexpressed keratin genes. Mol. Cell. Biol. **6**:539–548.
53. **Romano, V., M. Hatzfeld, T. M. Magin, R. Zimbelmann, W. W. Franke, G. Maier, and H. Ponstingl.** 1986. Cytokeratin expression in simple epithelia. I. Identification of mRNA coding for human cytokeratin no. 18 by a cDNA clone. Differentiation **30**: 244–253.
54. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
55. **Scarpulla, R. C.** 1984. Processed pseudogenes of rat cytochrome *c* are preferentially derived from one of three alternate mRNAs. Mol. Cell. Biol. **4**:2279–2288.
56. **Shapiro, S. G., and J. B. Lingrel.** 1984. Identification of a recently evolved goat embryonic beta-globin pseudogene which retains transcriptional activity in vitro. Mol. Cell. Biol. **4**:2120–2124.
57. **Shapiro, S. G., E. A. Schon, and J. B. Lingrel.** 1983. Sequence and linkage of the goat epsilon^{I} and epsilon^{II} beta-globin genes. J. Mol. Biol. **169**:31–52.
58. **Shows, T. B.** 1983. Human genome organization of enzyme loci and metabolic diseases, p. 323–339. *In* M. C. Rattazzi, J. G. Scandalios, and G. S. Whitt (ed.), Isozymes: current topics in biological and medical research, vol. 10. Alan R. Liss, New York.
59. **Shows, T. B., and J. A. Brown.** 1975. Human X-linked genes regionally mapped utilizing X-autosome translocations and somatic cell hybrids. Proc. Natl. Acad. Sci. USA **72**:2125–2129.
60. **Shows, T. B., J. A. Brown, L. L. Haley, M. G. Byers, R. L. Eddy, E. S. Cooper, and A. P. Goggin.** 1978. Assignments of the beta-glucuronidase structural gene to the pter→q22 region of chromosome 7 in man. Cytogenet. Cell Genet. **21**:99–104.
61. **Shows, T. B., R. L. Eddy, L. Haley, M. Byers, M. Henry, T. Fujita, H. Matsui, and T. Taniguchi.** 1984. Interleukin 2 (IL2) is assigned to human chromosome 4. Somat. Cell Mol. Genet. **10**:315–318.
62. **Shows, T. B., A. Y. Sakaguchi, and S. L. Naylor.** 1982. Mapping the human genome, cloned genes, DNA polymorphisms and inherited disease, p. 341–452. *in* H. Harris and K. Hirschhorn (ed.), Advances in human genetics, vol. 12. Plenum Press, New York.
63. **Singer, P. A., K. Trevor, and R. G. Oshima.** 1986. Molecular cloning and characterization of the endo B cytokeratin expressed in preimplantation mouse embryos. J. Biol. Chem. **261**:538–547.
64. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
65. **Steinert, P., D. Parry, E. Racoosin, W. Idler, A. Steven, B. Trus, and D. Roop.** 1984. The complete cDNA and deduced amino acid sequence of a type II mouse epidermal keratin of 60,000 Da: analysis of sequence differences between type I and type II keratins. Proc. Natl. Acad. Sci. USA **81**:5709–5713.
66. **Steinert, P., R. Rice, D. Roop, B. Trus, and A. Steven.** 1983. Complete amino acid sequence of a mouse epidermal keratin subunit and implications for the structure of intermediate filaments. Nature (London) **302**:794–800.
67. **Steinert, P. M., A. C. Steven, and D. R. Roop.** 1985. The molecular biology of intermediate filaments. Cell **42**:411–419.
68. **Sun, T.-T., R. Eichner, A. Schermer, D. Cooper, W. G. Nelson, and R. A. Weiss.** 1984. Classification, expression, and possible mechanisms of evolution of mammalian epithelial keratins: a unifying model, p. 169–180. *In* A. Levine, W. Topp, G. Vande Woude, and J. D. Watson (ed.), The cancer cell, vol. 1: the transformed phenotype. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
69. **Sun, T.-T., and H. Green.** 1978. Keratin filaments of cultured human epidermal cells: formation of intermolecular disulfide bonds during terminal differentiation. J. Biol. Chem. **253**:2053–2060.
70. **Thomas, P.** 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA **77**:5201–5205.
71. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gel onto nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA **76**:4350–4354.
72. **Tseng, S. C. G., M. J. Jarvinen, W. G. Nelson, J.-W. Huang, J. Woodcock-Mitchell, and T.-T. Sun.** 1982. Correlation of specific keratins with different types of epithelial differentiation: monoclonal antibody studies. Cell **30**:361–370.
73. **Tyner, A., M. Eichman, and E. Fuchs.** 1985. The sequence of a type II keratin gene expressed in human skin: conservation of structure among all intermediate filament genes. Proc. Natl. Acad. Sci. USA **82**:4683–4687.
74. **Tyner, A. L., and E. Fuchs.** 1986. Evidence for posttranscriptional regulation of the keratins expressed during hyperproliferation and malignant transformation in human epidermis. J. Cell Biol. **103**:1945–1956.
75. **Vanin, E. F., G. I. Goldberg, P. W. Tucker, and O. Smithies.** 1980. A mouse alpha-globin-related pseudogene lacking intervening sequences. Nature (London) **286**:222–226.
76. **Weiss, R. A., R. Eichner, and T.-T. Sun.** 1984. Monoclonal antibody analysis of keratin expression in epidermal diseases: a 48- and 56-kdalton keratin as molecular markers for hyperproliferative keratinocytes. J. Cell Biol. **98**:1397–1406.
77. **Wilde, C. D., C. E. Crowther, and N. J. Cowan.** 1982. Diverse mechanisms in the generation of human beta-tubulin pseudogenes. Science **217**:549–552.
78. **Woodcock-Mitchell, J., R. Eichner, W. G. Nelson, and T.-T. Sun.** 1982. Immunolocalization of keratin polypeptides in human epidermis using monoclonal antibodies. J. Cell Biol. **95**: 580–588.
79. **Wu, C.-I., W.-H. Li, J. J. Shen, R. C. Scarpulla, K. J. Limbach, and R. Wu.** 1986. Evolution of cytochrome c genes and pseudogenes. J. Mol. Evol. **23**:61–75.
80. **Wu, Y.-J., L. M. Parker, N. E. Binder, M. A. Beckett, J. H. Sinard, C. T. Griffiths, and J. G. Rheinwald.** 1982. The mesothelial keratins: a new family of cytoskeleton proteins identified in cultured mesothelial cells and nonkeratinizing epithelia. Cell **31**:693–703.