Keratin 19: Predicted Amino Acid Sequence and Broad Tissue Distribution Suggest it Evolved from Keratinocyte Keratins

Pauline C. Stasiak, BSc., Ph.D., Patricia E. Purkis, Irene M. Leigh, BSc., M.B. B.S., MRCP, and E. Birgitte Lane, BSc., Ph.D.

Cell Structure Laboratory, Imperial Cancer Research Fund, Clare Hall Laboratories, Hertfordshire (PCS, EBL), and Experimental Dermatology Labs, London Hospital, London, U.K. (PEP, IML)

The type I keratin 19 is unusual in its tissue distribution in that under normal circumstances it does not seem to be restricted, as the other keratins are, to expression in either stratified or simple epithelia. In addition to the previously reported distribution of keratin 19 in human tissues, we have observed keratin 19 in epidermal basal cells, in a defined region of the hair follicle, and in nipple epidermis. We noticed that expression of keratin 19 appears to be especially characteristic of regions of labile or variable cellular differentiation as indicated by the presence of multiple keratin phenotypes in close proximity to each other. Using a monoclonal antibody recognizing keratin 19 (LP2K) to screen a human placenta cDNA expression library, we have isolated, cloned, and sequenced cDNA coding for full-length human keratin 19, as confirmed by its reactivity with several other known anti-keratin 19 monoclonal antibodies and by the near identity of its sequence with that of the bovine keratin 19 homologue. This similarity extends to both proteins being truncated at the C-terminal end to only 13 amino acids beyond the rod domain. Although the amino acid homology over the N-terminal and helical rod domains is particularly high, the human and bovine proteins diverge substantially over the short C-terminal domain, which suggests that this region has no conserved function. Comparison with other type I keratins indicates that the closest evolutionary neighbors of keratin 19 are keratinocyte keratins, probably 13 and 14, and not the simple epithelial keratin 18. Assessing the histochemistry and sequence data together, we propose that the cell may use this apparently deficient keratin as a "neutral" keratin. While unimpaired in its ability to polymerize (keeping the cell integrated into the epithelial sheet via filament-desmosome networks), keratin 19 expression does not irrevocably commit a cell to any one of the local differentiation options. Such predicted differentiational flexibility may also imply vulnerability to transformation. *J Invest Dermatol 92:707–716, 1989*

The keratin family of intermediate filaments consists of at least 19 individual epithelial keratin polypeptides [1], or 27 including the "hard" hair keratins [2], and each appears to be the product of a different keratin gene. Different combinations of these intermediate filament proteins are expressed in a tissue-specific manner, but it has still not been possible to demonstrate their primary function in an experimental context. The ubiquitous expression of keratins by epithelial cells together with the structural complexity of the transcellular filament-desmosome network in all epithelia suggest that keratin filaments are essential for the function of an epithelium. Yet our current understanding of keratin chemistry and biology is insufficient to explain the need for such extensive heterogeneity.

The smallest keratin, keratin 19, was first detected in squamous cell carcinoma lines [3], and it is identifiable by its characteristically low molecular weight on SDS gels at 40 kD. It is generally regarded as being characteristic of simple epithelia (see Ref 4) such as intestine, kidney collecting ducts, gallbladder, mesothelium, and glandular secretory cells (see Ref 5 for review). Keratin 19 has, in fact, an unusually wide tissue distribution, and has now been reported to be present, in significant quantities and under apparently normal conditions, in both stratified and squamous epithelia [6–12]. The gene for the bovine protein homologous to keratin 19 was recently sequenced [13], and its sequence suggested that this keratin was truncated, accounting for its low molecular weight. Our preliminary observations on the cDNA sequence of human keratin 19 from simple epithelium supported this finding [14], and recently a second report of a sequence for human keratin 19 has been published [15] which is almost identical to ours, from cDNA isolated from stratified epithelial cells (keratinocytes).

This paper describes the use of a monoclonal antibody specific for keratin 19, LP2K, to examine the cellular distribution of this keratin within various tissues, including a population of basal keratinocyte cells within normal hair follicles, structures previously reported as not containing keratin 19 [16, 17, but see 7, 12]. This antibody was used to identify cDNA clones coding for human keratin 19 which were expressed within a λgt11 library of human placenta, and we describe our subsequent analysis of this protein in terms of its relationship to other keratins. Combining the histochemistry data with the sequence analysis, it seems likely that keratin 19, a keratinocyte keratin which is only more distantly related to simple epithelial keratin 18, has been conserved through evolution, especially because of its lack of a tail-domain. The lack of a functional C-terminal domain in keratin 19 may reflect deferred committment to one of multiple differentiation pathways open to the cells expressing

Manuscript received July 13, 1987; accepted for publication November 11, 1988.

This research was supported by the Imperial Cancer Research Fund.

Reprint requests to: Dr. E. B. Lane, Imperial Cancer Research Fund, Clare Hall Laboratories, Blanche Lane, South Mimms, Potter's Bar, Hertfordshire EN6 3LD, U.K.

0022-202X/89/$03.50 Copyright © 1989 by The Society for Investigative Dermatology, Inc.

keratin 19 in these tissues without sacrificing the essential keratin filament function of forming a transcellular desmosome-filament complex for structural integrity of the epithelium.

METHODS

Monoclonal Antibodies The monoclonal antibody LP2K was derived from BALB/c mice as described previously [18] using sonicated cytoskeleton fractions [material insoluble in 1% Nonidet-P40 (N-P40)] from SVK14 cells (SV40-transformed human neonatal keratinocytes: 19) as the immunogen. Positive hybrids were selected using a microimmunofluorescence assay [20] on SVK14 cells and selected cells were cloned in soft agarose [18]. Other monoclonal antibodies used that were reported to react with keratin 19 were BA16 and BA17 [6], A53-B/A2 [21], KM4.62 [22] (BioMakor, Rehovot, Israel) and (not monospecific) KA4 [23], AE1 [24], and anti-IFA [25]. Antibodies LE61 (to keratin 18) and LE41 (to keratin 8) [18] were also used as controls.

Immunoblotting Cytoskeleton extracts were made by washing cultured cells in phosphate buffered saline (PBS), lysing in 0.1% N-P40 (in 10 mM Tris-HCl pH 7.4, 150 mM NaCl, 3 mM EDTA) for 5–10 min and further extracting the pellet in a high salt solution (1.5M KCl, 10 mM Tris-HCl pH7.4, 150 mM NaCl, 3 mM EDTA and 0.05% N-P40) for 5–10 min, and then washing thoroughly in Tris (10 mM, pH7.4) EDTA (3 mM). The resulting pelleted material was used as the starting material for one- and two-dimensional (isoelectric focussing and non-equilibrium pH gradient [26,27] gel electrophoresis. Gels were blotted onto nitrocellulose paper (0.45 μm, Schleicher and Schuell) for 1 h [28] and blocked with 0.05% TWEEN-20 in PBS overnight. Primary antibody incubations were carried out at room temperature for 1 h using antibody-containing culture supernatants diluted 1:5 (except for AE1, supernatant diluted 1:20; KA4, ascites fluid diluted 1:300; KM4.62, commercial stock diluted 1:1000 and anti-IFA, supernatant diluted 1:10) in expired unused tissue culture medium containing 10% fetal calf serum (FCS) and 0.05% sodium azide. Blots were then washed extensively and rapidly in tap water (10 min at room temperature), incubated in a second antibody solution (rabbit anti-mouse immunoglobulins conjugated to horseradish peroxidase, DAKOPatts a/s, diluted 1:100 in expired unused medium containing 10% FCS) for 1 h at room temperature, washed copiously again in tap water (5–10 min), washed in PBS + 0.05% TWEEN-20 + 0.5% bovine serum albumin (BSA) for ten min, rinsed in tap water, then incubated in diamino benzidine (DAB; Sigma) substrate [0.5 mg/ml DAB in PBS, with fresh solutions to .003% aqueous nickel sulphate and .001% hydrogen peroxide (30 vol) added just before use] for 1–5 min. Blots were stained for total protein content with India ink [29].

Immunohistochemistry Skin biopsies were obtained from normal human volunteers or from surgical trimmings; other epithelial samples were obtained from surgical trimmings or from autopsies. Fetal material was obtained from the Royal Marsden Tissue Bank, London. Tissues were snap-frozen in liquid nitrogen; unfixed cryostat sections were stored at −70°C until used for staining. Antibody incubations and enzyme visualization procedures were as for immunoblotting (above). Counterstaining was with Harris's haematoxylin and specimens were air-dried and mounted in DPX mountant (BDH).

Selection of cDNA Clones by Expression Approximately 6×10^5 plaque-forming units (pfu) of a cDNA λgt11 expression library constructed from human placental mRNA (Clontech) were screened with the monoclonal antibody LP2K. The screening procedure of Young and Davis [30] was followed for the induction of fusion protein in λgt11 and transfer of colonies to nitrocellulose filters (BA85, 0.45 μm, Schleicher and Schuell). After washing the filters in PBS for 5–10 min, immunolocalization of clones expressing keratin 19 fusion proteins was carried out following the antibody incubation and washing procedures described above for immunoblotting. Positive reactions were detected in 3×10^{-3} % of the colonies screened. Positive clones were checked with BA17 (anti-keratin 19 monoclonal antibody), and purified clones were further checked for reactivity with other anti-19 antibodies as above. After several rounds of screening, pure bacteriophage DNA samples were prepared. For each clone, a 15 cm plate of the *Escherischia coli* bacterial host strain Y1090 was infected with approximately 10^5 pfu of bacteriophage and the phage lysis allowed to proceed to completion; the plate was overlayed with buffer (100 mM NaCl; 10 mM $MgSO_4$; 50 mM Tris pH 7.0) at 4°C for 2 d. The overlay was then removed and spun in a caesium chloride equilibrium gradient at 55,000 rpm for 14 h. The bacteriophage band was recovered and spun in a two-step caesium chloride density gradient at 35,000 rpm for 1 h. The bacteriophage band was extracted, denatured with formamide, and the DNA was precipitated with ethanol. The inserts from positive clones were also subcloned into the plasmid expression vector pUR292 [31], and subsequent reactivity of multiple keratin 19-specific antibodies with the β-galactosidase fusion proteins expressed in these constructs confirmed that the clones selected were indeed coding for keratin 19.

Sequencing and Analysis of Human Keratin 19 Clones Two of the largest LP2K-positive λgt11 clones, λPSE12 (1335 bp) and λPSE21 (1394 bp), were chosen for further analysis. Purified insert DNA was self-ligated, sonicated into random fragments, blunt-ended and subcloned into the *Sma*I site of M13mp9 and mp18 [32,33], and sequenced by the dideoxy chain-termination method [34]. The majority of the sequence was covered by overlapping fragments and included sequence of the complementary strand; any uncertain areas were verified using oligonucleotide primers constructed to regions within the sequence. The sequence of the PSE21 clone provided the complete coding region including the initiating methionine and 32 bp of the 5′ untranslated region. Restriction enzymes were obtained from New England Biolabs. Calf intestinal alkaline phosphatase was obtained from Boehringer Mannheim. [α^{35}S]ATP was obtained from Amersham International and deoxy- and dideoxynucleotriphosphates from Pharmacia. Computer analysis programs were used to assemble the cDNA sequence [35], predicting amino acid sequence and subsequent protein structural conformation [36,37]. Multiple sequence alignment and comparison of the rod regions of type I keratins were done manually.

RESULTS

Biochemical Characterization of Monoclonal Antibody LP2K Figure 1 shows that the monoclonal antibody LP2K reacts by immunoblotting consistently and specifically with a component of molecular weight 40,000 Daltons in the keratin preparations, identified as keratin 19 according to the numerical classification of Moll et al [1]. Positive reactivity of LP2K with cell lines either by immunoblotting or by immunofluorescence is completely correlated with the presence of this low molecular weight keratin as detected biochemically.

Immunohistochemical Reactivity of LP2K: Hair Follicles Immunohistochemical staining of the antibody LP2K on frozen, unfixed sections of a variety of body sites of normal skin indicated the presence of keratin 19 in hair follicles (Fig 2) in a defined zone of basal keratinocytes in the deep outer root sheath. These cells could be detected with all the antibodies to keratin 19, although BA16 antibody staining was always weaker than the others. Positive cells were seen where the outer root sheath widens and the *arrector pili* muscle attaches to the hair follicle, below the sebaceous gland duct and below the follicle isthmus. These cells are thus located adjacent to the transition from epidermis-like differentiation in the follicle isthmus to a less well differentiated domain in the deep follicle (keratin 10 is only expressed in the upper infundibulum region and not in the lower portion [7]). They have the appearance of normal keratinocytes, cuboidal in shape and situated directly on the basal lamina, forming an integral part of the basal cell layer of the outer hair root sheath. Staining of basal and occasionally scattered suprabasal cells was seen sporadically from this point down to the hair bulb, but it was never seen in cells further out towards the skin

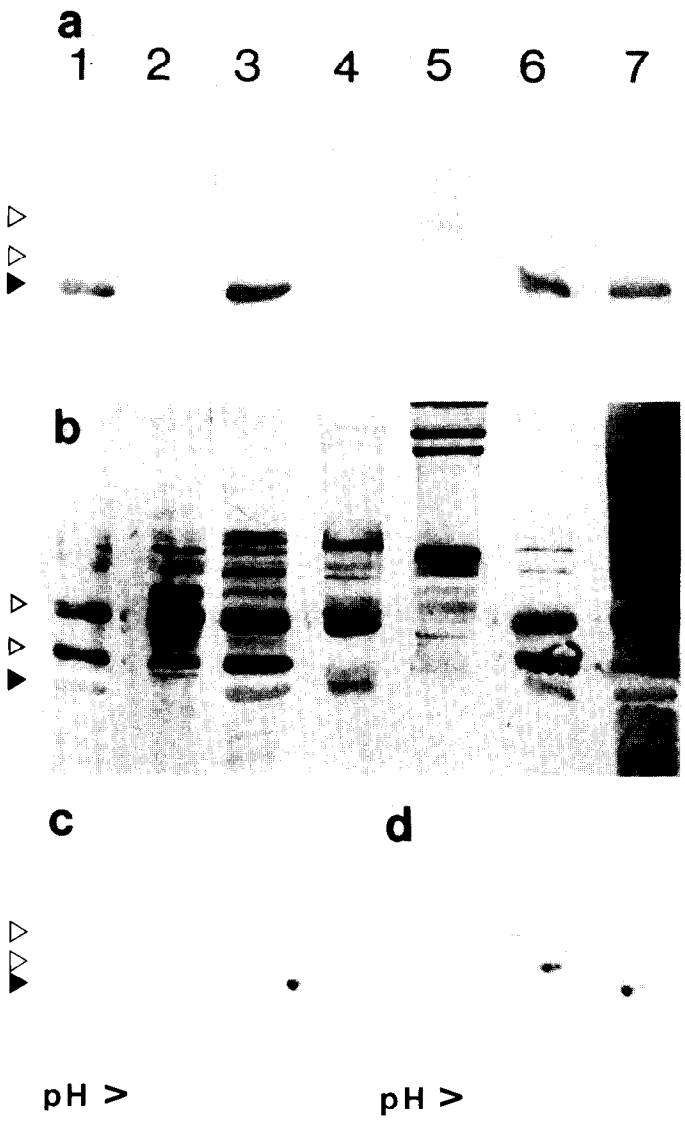

Figure 1. Specificity of monoclonal antibody LP2K for keratin 19 as demonstrated by immunoblotting. *a* and *b:* Specificity of LP2K immunoreactivity only on the cell lines known to express keratin 19: *(a)* immunoblot and *(b)* protein stain of same blot using India ink. 1 = T47D (mammary carcinoma line), 2 = A431 (vulval carcinoma line), 3 = BT20 (mammary carcinoma line), 4 = PtK1 (kidney cell line, marsupial origin), 5 = normal epidermis, 6 = MCF-7, and 7 = CAMA-1 (also mammary carcinoma lines). *c* and *d:* Immunoperoxidase localization of LP2K reacting with keratin 19 only *(c)* on nitrocellulose blot of two-dimensional gel electrophoretic separation of MCF-7 breast carcinoma cell line keratins, and *(d)* protein staining of the same blot using India ink; *pH>* indicates direction of reducing pH or isoelectric point of proteins. *Arrowheads* show Mr (relative migration) of internal molecular weight standards keratins 8 (52.5 kD: *upper, open arrowhead*) and 18 (45 kD: *lower, open arrowhead*) and keratin 19 (40 kD: *lowest, closed arrowhead*).

surface. Though few in number, the keratin 19-containing cells constituted a predictable feature of the hair follicle.

In forearm hair follicles a small subpopulation of cells showed reactivity with LP2K [7]. In the hair follicles from scalp, scrotum, and eyelid skin the same type of staining was seen (Fig 2), although the keratin 19-containing zone was much more extensive in the large hair follicles of the scalp than in forearm. In non-scalp skin, positive groups of cells sometimes appeared as little knobs protruding into the connective tissue from the otherwise smooth cross-sectional profile of the follicle.

Keratin 19 is present in human hair follicles at all stages of development. Hair buds and hair pegs from fetal skin of 14 weeks gestational age express keratin 19 in all their cells (Fig 3*a*); as the follicles grow, the staining becomes progressively restricted until the adult pattern is attained. Figure 2*c,d* shows staining by BA16 of a small cluster of cells in the appropriate position 18-week fetal skin hair follicles; this antibody was previously interpreted as not staining hair follicles [7,12]. The lower swelling on the side of the developing follicle (containing the positive cells) will become the attachment site of the smooth muscle, while the upper swelling (keratin 19 negative) will form the sebaceous gland.

Basal cells showing LP2K-positive reactivity for keratin 19 were unstained by monoclonal antibodies specific for keratin 8 (LE41) or 18 (LE61), distinguishing them unequivocally from Merkel cells, which do react with these and other antibodies to keratins 8 and 18 [7,38].

Immunohistochemical Reactivity of LP2K: Other Tissues In common with the staining in the hair follicles, all the places where keratin 19 was detected were areas where epithelial cells expressing different keratin phenotypes coexisted in close proximity. Our results regarding tissues expressing keratin 19 were completely consistent with previous reports [12]. Positive reactivity was observed on sweat gland and mammary gland ductal and secretory cells, bile ducts, gastrointestinal tract, and homogeneously through all cell layers in bladder urothelium, which expresses a particularly complex profile of keratins. Basal layer keratinocytes were stained with varying degrees of heterogeneity in oral epithelia, ectocervical epithelium [39], including endo-ectocervical junction and oesophagus. In the oral epithelia, basal cell staining with LP2K was strikingly correlated with areas of variable keratin phenotype of the suprabasal cells [40,41] (Fig 3*b*). Hepatocytes and interfollicular body epidermis were negative, but patches of positive basal cells were often observed in nipple skin, close to the ducts (Fig 3*c*). Extent of nipple skin staining was highly variable on an individual level. Keratin 19-positive basal cells in nipple skin were not stained with antibodies to keratins 8 and 18 (LE41,LE61), again distinguishing them from Merkel cells and also from the cell type seen in Paget's disease [23,42]. In "mixed" epithelia such as gland ducts (sweat gland, mammary gland), where luminal and basal cells appear to have distinct keratin phenotypes [43], staining was observed in both cell layers but most strongly in the luminal cells, which are true simple epithelial cells. Patches of keratin 19-negative luminal cells have been reported to occur beyond the branch points of mammary gland ducts in the region where pluripotent progenitor cells might be predicted [8].

In all samples examined, LP2K staining was compared with that of at least two other antibodies reported as being specific for keratin 19, and no discrepancies were observed other than quantitative ones: culture supernatants containing LP2K gave more intense staining than BA16, but usually less than A53-B/A2. No useful staining with LP2K was obtained if tissues were fixed in formaldehyde.

Identification and Sequencing of cDNA Clones for Keratin 19 A number of λgt11 clones expressing fusion proteins reactive with LP2K and other antibodies to keratin 19 were identified and purified from the placental cDNA library, using an immunoperoxidase staining protocol that was originally devised for immunoblotting from SDS polyacrylamide gels. All pure clones were checked for their reactivity with several other monoclonal antibodies to keratin 19, either on the modest amounts of β-galactosidase fusion proteins expressed by the λgt11 clones or on the high levels of fusion protein expressed by the pUR plasmid constructs, which supported their identification as true keratin 19-encoding clones. Sequencing of the clone PSE21 gave the full-length coding sequence for human keratin 19 [14], as became apparent when the predicted amino acid sequence for which it coded was deduced, as shown in Fig 4.

The sequenced keratin 19 cDNA is 1394 bp long (Fig 4), containing an open reading frame of 1200 bp starting with the initiation codon ATG at nucleotide 33 and ending with the termination codon at nucleotide 1233. The ATG is in an appropriate context for

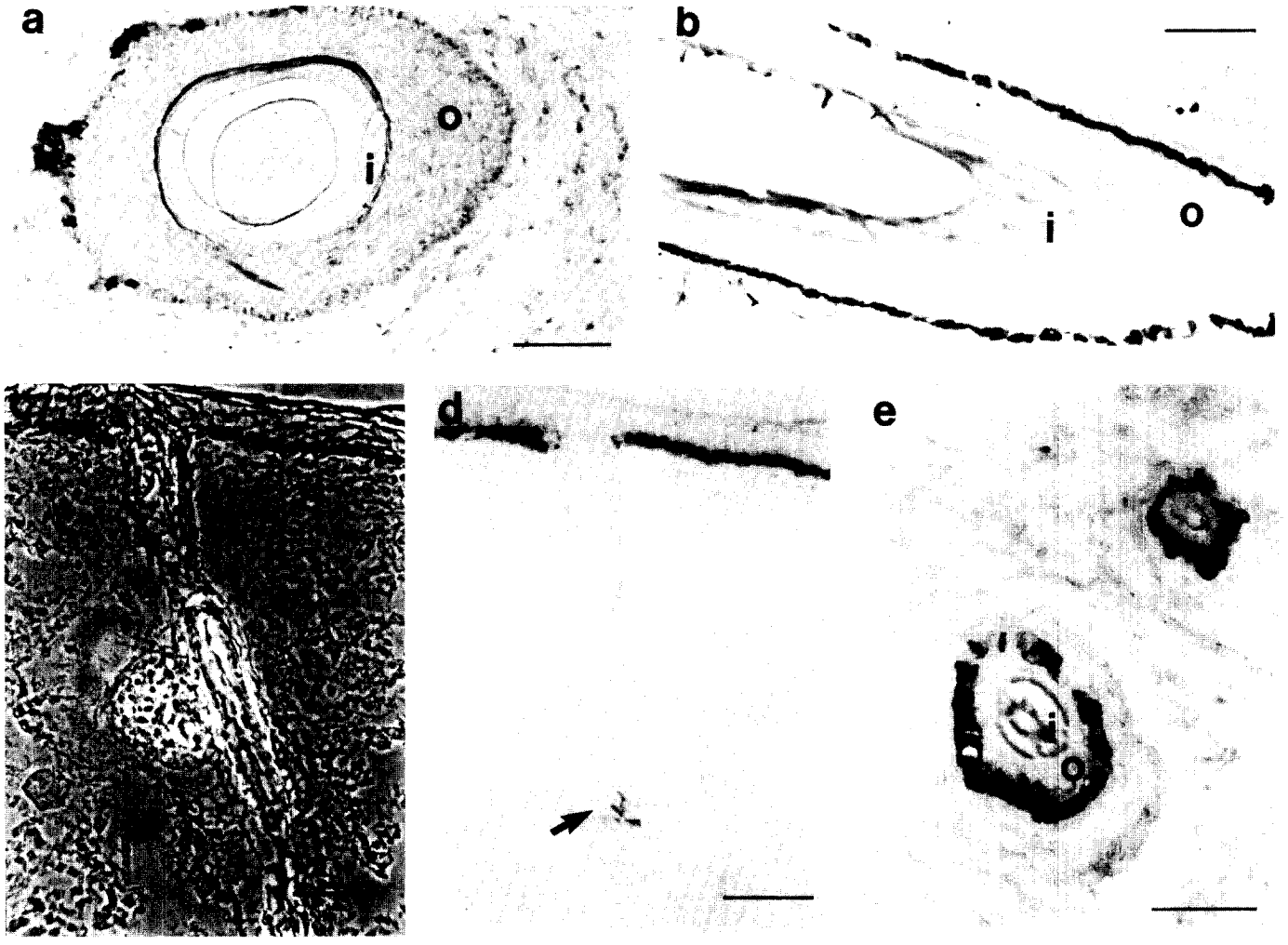

Figure 2. Keratin 19 is expressed in hair follicles. Unfixed frozen sections of hair follicles through attachment point of *arrector pili* muscle, below isthmus, where inner root sheath is still substantial. Stained by immunoperoxidase with LP2K *(a,b,e)* or BA16 *(c,d)*: reaction product is black. *a:* Scrotum skin (transverse sections: patchy staining); *b:* Scalp (oblique section: extensive positive zone); *c,d:* foetal body skin at 18.5 weeks gestation, longitudinal section seen with phase contrast *(c)* and bright field *(d)* microscopy: interfollicular basal cells are positive (negative in adult) and most of the follicle is negative, but the few cells in the lower protrusion *(arrow)* which will form the site of arrector pili attachment are in the same relative position as the positive cells seen in adult follicles. *e:* Eyelid skin with cross-sections through two follicles. With *(a)* or without *(b–e)* haematoxylin counter stain. (*o*: outer root sheath; *i*: inner root sheath; *Scale bars:* 100 μm.)

the start of translation [44]. The 159 bp 3′ untranslated region includes a 34 nucleotide poly-A tail. A putative polyadenylation signal AATAAA lying in the consensus sequence CCAATAAAAATT [45] is present 17 bp upstream of the start of the poly-A tail. This 1394 bp cDNA clone is smaller than the expected mRNA size of 1585 bp reported by Eckert and Green [46] and the approximately 1.5 kb size mRNA species with which the PSE21 clone hybridizes in Northern blot analysis of MCF-7 RNA (data not shown); obviously this clone does not include the complete 5′ untranslated region and probably does not have a complete poly-A tail.

The open reading frame codes for a 400 amino acid protein, including the initiating methionine. The sequence predicts a protein of 44.1 kD molecular weight, which is larger than the estimated molecular weight of 40 kD as determined by electrophoretic mobility on SDS polyacrylamide gel electrohoresis. It is a typical member of the type I intermediate filament family, with a non-helical amino-terminal head domain and an alpha-helical rod domain, but as was found for the homologous bovine protein the tail region is abnormally short. Only 13 amino acids occur beyond the TYR[X]LLEG[$^{Q}/_{E}$] sequence, which marks the end of the helical domain in other intermediate filaments.

Secondary structure predictions from the amino acid sequence show the consensus α-helical rod domain from amino acid positions 83–387. It is possible that the helical configuration starts even further towards the amino terminus at position 74 and could extend right to the carboxy-terminal end of the sequence. The whole of this region shows the predominance of hydrophobic residues at the first and fourth positions of each seven residues, which is characteristic of proteins that form coiled-coils [47] due to the hydrophobic "backbone" of the coiled monomer. There are two indications, apart from the linkers L1 and L2, of irregularities in this heptad repeat pattern around two phenylalanine residues at amino acid positions 271 and 331. The first one following position 271 may be accomodated as a couple of wider turns in the helix, or it may indicate the presence of a helix break or linker region (linker "L2" [48]) from positions 259–274. The second irregularity, around position 331, is probably the polarity reversal feature that appears to be well conserved among intermediate filaments.

Finally, analysis of the predicted amino acid sequence with the PCGENE program indicated the occurrence of a possible glycosylation site on human keratin 19 at amino acid position 393.

The sequence for human keratin 19 recently published by Eckert [15] is virtually identical to ours. The apparent discrepancies between his and ours as first published [14] can probably be explained as sequencing errors. Figure 4 presented here has changes at positions 268–273 bp and 1056–1057 bp from that published pre-

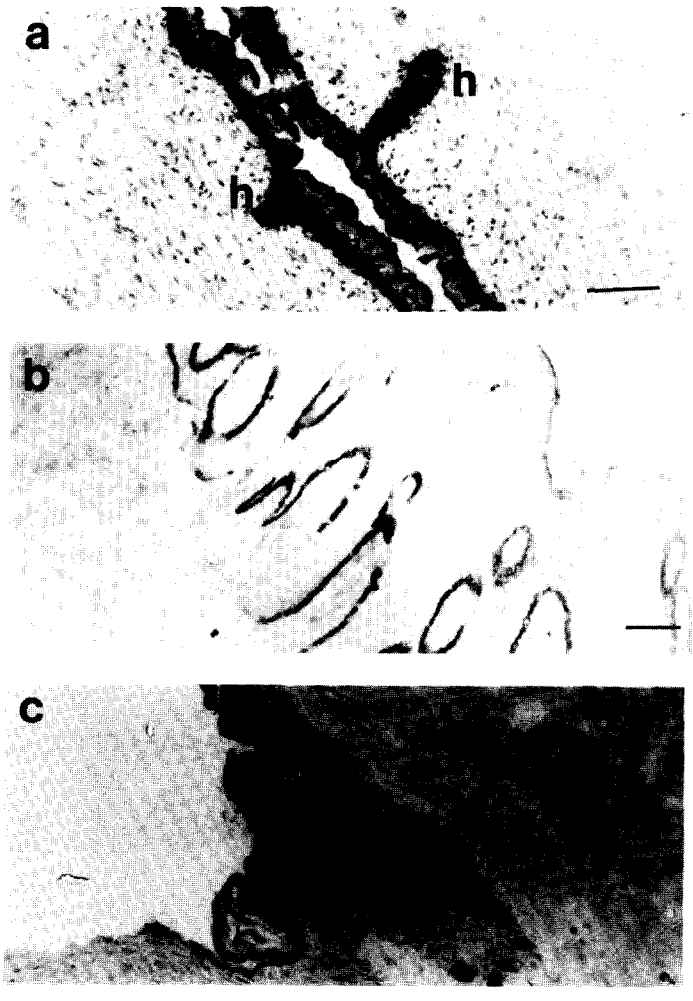

Figure 3. Keratin 19 expression in transitional regions. Immunoperoxidase staining with LP2K; reaction product appears black [unfixed frozen sections with *(a)* or without *(b,c)* haematoxylin counterstaining]. *a:* Foetal body skin of 14 weeks gestation (specimen folded over so two epidermal surfaces face each other); extensive staining seen strongly in hair follicle pegs, basal cells, and periderm layer, weak to moderate in intermediate layers. *b:* Oral sulcular epithelium (see Refs 42 and 43); patchy staining of basal cells. *c:* Variable patchy staining of basal cells often seen in epidermis of nipple; in some individuals this was a continuous basal layer. (*h:* hair follicle pegs; *Scale bars:* 100 μm.)

viously [14], but maintains differences from Eckert's sequence at positions 503 bp and 1080–1082 bp. The virtual identity of the two sequences arrived at independently and from different tissue sources (i.e., our from simple epithelium of placenta and Eckert's from a stratified squamous epithelium-derived cell line) argues for the presence of only one active gene for keratin 19 in *Homo sapiens.*

Sequence Analysis

(i) Comparison with the Bovine Homologue: Figure 5 shows a comparison of the bovine and human keratin 19 amino acid sequences. As has been found previously when other homologous intermediate filament sequences have been compared across species, the similarity of the two sequences is extremely high, with 86.3% identity in the head region (amino acids 1–73 with respect to human keratin 19) and 91% identity in the alpha-helical rod region overall (amino acids 74–387). At the DNA level, the 3′ untranslated region is also highly homologous with that of bovine keratin 19, showing 73% identity (not shown): similarity between the 3′ untranslated regions of homologous genes across a species barrier has been noted before [49]. However, great divergence is seen when the sequences of the tail domain are compared, because 8 out of the 13 amino acids occurring beyond the conserved rod domain termination sequence TYR[X]LLEG[$^{Q}/_{E}$] are non-identical. Within any 13 amino acid stretch over the whole sequence comparison of bovine and human keratin 19, there are never less than nine identical amino acids (i.e., 69% identity); this divergence is thus statistically significant.

(ii) The Non-Helical Head Domain: Although the central α-helical rod domain of intermediate filament proteins is well conserved within each intermediate filament class, the selection pressures acting upon the non-helical head and tail domains appear to be different, because non-helical domain sequence features are often shared between type I and type II keratins coexpressed in the same tissue [50,51]. Therefore, in order to try and assess the relatedness between human keratin 19 and the other type I keratins, the helical and non-helical domains are considered separately.

The 73 amino acid long head domain of human keratin 19 contains a number of glycine-rich stretches, for example "GGYGG-GYGG," together with many serine residues. Glycine and serine alone account for 52% of amino acids in this domain. Glycine- and serine-rich repeats have been noted in the termini of a number of keratinocyte keratins, expressed in stratified epithelia, such as human keratin 14 [52], the mouse keratin 10 homologue [53], and *Xenopus laevis* larval epidermal keratins DG81.A1 and DG81.B2 (54). In contrast, very few such repeats are present in the non-helical domain sequences of type II keratins 7 [55], and 8 [56,57], and in Type I keratin 18 [58–60], which are the three keratins specifically expressed in simple epithelia.

(iii) The Central α-Helical Rod Region—Areas of Conservation: Within the rod domain, keratin 19 shows significant homology only with the other type I keratins. The great number of (mostly partial, mostly C-terminal) sequences of type I keratins that have been identified allows an analysis of the degree of sequence conservation between the different subdomains of the type I keratin helices, shown in Fig 4. The consensus boundaries of the helical subdomains have been taken as amino acid positions 83–116 (helix 1A), 126–225 (helix 1B), and 245–387 (helix 2), inclusive with respect to human keratin 19, based on Geisler and Weber's analysis of several kinds of intermediate filaments [61] and our interpretation of type I keratin sequences. Thus the helical rod domain for this analysis begins with the highly conserved peptide TMQ[X]LNDRLA[X]YL and ends with the shorter conserved peptide TYR[X]LLEG[$^{Q}/_{E}$].

From this figure two points may be made. First, the keratins which are most homologous to human keratin 19, apart from the nearly identical bovine keratin 19, are the human keratinocyte keratins 14, 15 [65], and 16 [66] (formerly thought to be keratin 17), and the mouse homologue [63] of human keratin 13, suggesting a high degree of relatedness in spite of the species differences.

Second, the divergence of keratin 18 from all the keratinocyte keratins, as well as from keratin 19, is strikingly apparent from the consistently low scores of the human [58] and mouse [60] keratin 18 sequences.

DISCUSSION

It is clear that human keratin 19, like the homologous bovine protein, does not have a non-helical C-terminal domain in the sense that other keratins do. When the bovine and human species sequences are compared, there is a dramatic loss of any significant homology over the last 13 amino acids extending beyond the consensus helical domain termination sequence. Because the rest of the sequence is particularly highly conserved, this divergence indicates either a reduction of selection pressure on this stretch, or that this region of the polypeptide has a completely different sequence requirement in the two species. The latter explanation seems unlikely because no such divergence has been seen in any other pair of homologous keratins sequenced from different species, and the non-helical tail domains are usually no less well conserved than the N-terminal domains. We believe the divergence is strong evidence for absence of function.

Although it is obvious that the helical rod domain of the polypep-

```
                                                                      56
      CGGGGGTTGCTCCGTCCGTGCTCCGCCTCGCC ATG ACT TCC TAC AGC TAT CGC CAG
                                       Met Thr Ser Tyr Ser Tyr Arg Gln
                                         1
                                  83                                  110
TCG TCG GCC ACG TCG TCC TTC GGA GGC CTG GGC GGC GGC TCC GTG CGT TTT GGG
Ser Ser Ala Thr Ser Ser Phe Gly Gly Leu Gly Gly Gly Ser Val Arg Phe Gly
 9

                                 137                                  164
CCG GGG GTC GCT TTT CGC GCG CCC AGC ATT CAC GGG GGC TCC GGC GGC CGC GGC
Pro Gly Val Ala Phe Arg Ala Pro Ser Ile His Gly Gly Ser Gly Gly Arg Gly
 27

                                 191                                  218
GTA TCC GTG TCC TCC GCC CGC TTT GTG TCC TCG TCC TCC TCG GGG GGC TAC GGC
Val Ser Val Ser Ser Ala Arg Phe Val Ser Ser Ser Ser Ser Gly Gly Tyr Gly
 44

                                 245                                  272
GGC GGC TAC GGC GGC GTC CTG ACC GCG TCC GAC GGG CTG CTG GCG GGC AAC GAG
Gly Gly Tyr Gly Gly Val Leu Thr Ala Ser Asp Gly Leu Leu Ala Gly Asn Glu
 63

       |-Helix 1A ->               299                                  326
AAG CTA ACC ATG CAG AAC CTC AAC GAC CGC CTG GCC TCC TAC CTG GAC AAG GTG
Lys Leu Thr Met Gln Asn Leu Asn Asp Arg Leu Ala Ser Tyr Leu Asp Lys Val
 81          *               *                *               *       *

                                 353                                  ->|
CGC GCC CTG GAG GCG GCC AAC GGC GAG CTA GAG GTG AAG ATC CGC GAC TGG TAC
Arg Ala Leu Glu Ala Ala Asn Gly Glu Leu Glu Val Lys Ile Arg Asp Trp Tyr
 99      *               *                *               *       *

                                 407 |-Helix 1B ->                    434
CAG AAG CAG GGG CCT GGG CCC TCC CGC GAC TAC AGC CAC TAC TAC ACG ACC ATC
Gln Lys Gln Gly Pro Gly Pro Ser Arg Asp Tyr Ser His Tyr Tyr Thr Thr Ile
117                              *               *                       *

                                 461                                  488
CAG GAC CTG CGG GAC AAG ATT CTT GGT GCC ACC ATT GAG AAC TCC AGG ATT GTC
Gln Asp Leu Arg Asp Lys Ile Leu Gly Ala Thr Ile Glu Asn Ser Arg Ile Val
135      *               *               *               *       *

                                 515                                  542
CTG CAG ATC GAC AAC GCC CGT CTG GCT GCA GAT GAC TTC CGA ACC AAG TTT GAG
Leu Gln Ile Asp Asn Ala Arg Leu Ala Ala Asp Asp Phe Arg Thr Lys Phe Glu
153      *               *               *               *       *

                                 569                                  596
ACG GAA CAG GCT CTG CGC ATG AGC GTG GAG GCC GAC ATC AAC GGC CTG CGC AGG
Thr Glu Gln Ala Leu Arg Met Ser Val Glu Ala Asp Ile Asn Gly Leu Arg Arg
171  *               *               *               *       *

                                 623                                  650
GTG CTG GAT GAG CTG ACC CTG GCC AGG ACC GAC CTG GAG ATG CAG ATC GAA GGC
Val Leu Asp Glu Leu Thr Leu Ala Arg Thr Asp Leu Glu Met Gln Ile Glu Gly
189  *               *               *               *       *
```

```
                                 677                                  704
CTG AAG GAA GAG CTG GCC TAC CTG AAG AAG AAC CAT GAG GAG GAA ATC AGT ACG
Leu Lys Glu Glu Leu Ala Tyr Leu Lys Lys Asn His Glu Glu Glu Ile Ser Thr
 *  208          *               *                       *               *

 ->|             731                                  758
CTG AGG GGC CAA GTG GGA GGC CAG GTC AGT GTG GAG GTG GAT TCC GCT CCG GGC
Leu Arg Gly Gln Val Gly Gly Gln Val Ser Val Glu Val Asp Ser Ala Pro Gly
 *  226

       |-Helix 2 ->                785                                  812
ACC GAT CTC GCC AAG ATC CTG AGT GAC ATG CGA AGC CAA TAT GAG GTC ATG GCC
Thr Asp Leu Ala Lys Ile Leu Ser Asp Met Arg Ser Gln Tyr Glu Val Met Ala
243      *               *               *               *       *

                                 839                                  866
GAG CAG AAC CGG AAG GAT GCT GAA GCC TGG TTC ACC AGC CGG ACT GAA GAA TTG
Glu Gln Asn Arg Lys Asp Ala Glu Ala Trp Phe Thr Ser Arg Thr Glu Glu Leu
261      *               *                *Δ*             *       *

                                 893                                  920
AAC CGG GAG GTC GCT GGC CAC ACG GAG CAG CTC CAG ATG AGC AGG TCC GAG GTT
Asn Arg Glu Val Ala Gly His Thr Glu Gln Leu Gln Met Ser Arg Ser Glu Val
279          *               *               *               *       *

                                 947                                  974
ACT GAC CTG CGG CGC ACC CTT CAG GGT CTT GAG ATT GAG CTG CAG TCA CAG CTG
Thr Asp Leu Arg Arg Thr Leu Gln Gly Leu Glu Ile Glu Leu Gln Ser Gln Leu
297      *               *               *               *       *

                                1001                                 1028
AGC ATG AAA GCT GCC TTG GAA GAC ACA CTG GCA GAA ACG GAG GCG CGC TTT GGA
Ser Met Lys Ala Ala Leu Glu Asp Thr Leu Ala Glu Thr Glu Ala Arg Phe Gly
315      *               *               *               *               *Δ

                                1055                                 1082
GCC CAG CTG GCG CAT ATC CAG GCG CTG ATC AGC GGT ATT GAA GCC CAG CTG GCG
Ala Gln Leu Ala His Ile Gln Ala Leu Ile Ser Gly Ile Glu Ala Gln Leu Ala
333      *               *               *               *       *

                                1109                                 1136
GAT GTG CGA GCT GAT AGT GAG CGG CAG AAT CAG GAG TAC CAG CGG CTC ATG GAC
Asp Val Arg Ala Asp Ser Glu Arg Gln Asn Gln Glu Tyr Gln Arg Leu Met Asp
351  *               *               *               *       *

                                1163                                 1190
ATC AAG TCG CGG CTG GAG CAG GAG ATT GCC ACC TAC CGC AGC CTG CTC GAG GGA
Ile Lys Ser Arg Leu Glu Gln Glu Ile Ala Thr Tyr Arg Ser Leu Leu Glu Gly
369  *               *               *               *       *

 ->                              1217                                 1247
CAG GAA GAT CAC TAC AAC AAT TTG TCT GCC TCC AAG GTC CTC TGAGGCAGCAGGCTC
Gln Glu Asp His Tyr Asn Asn Leu Ser Ala Ser Lys Val Leu
 *               *       Ω       *                       400

TGGGGCTTCTGCTGTCCTTTGGAGGGTGTCTTCTGGGTAGAGGGATGGGAAGGAAGGGACCCTTACCCCCG
GCTCTTCTCCTGACCTGCCAATAAAAATTTATGGTCCAAGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAA  (1394)
```

Figure 4. cDNA sequence of human keratin 19 with amino acid translation. Anotated to show the following: predicted a-helical regions (coil 1A, coil 1B, coil 2) with *asterisks* below residues at key positions in heptad repeat; putative glycosylation site (Ω) at amino acid position 392; probable loops or discontinuities in helix (Δ) around amino acid positions 270 and 330. Initiation codon, stop codon, and polyadenylation signal are underlined.

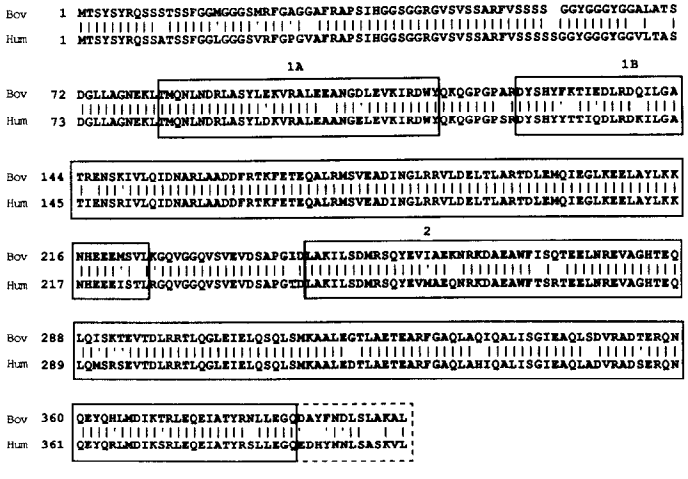

Figure 5. Amino acid sequence homology between bovine and human keratin 19. Identical residues: *vertical bars:* conservative substitutions: *vertical dash.* Helical domains are indicated by *boxes (1A, 1B, 2). Dashed box outlining post-helical domain terminal extension:* homology is 90% in the helix and N-terminal head region, but drops to 33% (only 5/13 shared residues) after the TYRSLLEGE consensus helix termination peptide.

tide must form the fundamental construction unit of intermediate filaments, little else is known about specific domain functions. Biochemical cleavage studies indicate that the amino terminal non-helical domain is required, together with the α-helix, for correct filament polymerization, but that the carboxy-terminal domain is not [72]; keratin 19 appears to polymerize normally, at least with keratin 8 [13]. It has been suggested that the non-helical domains of intermediate filaments may extend laterally from the body of the filament, and there is good evidence that this is the case for the carboxy-terminal domain of neurofilament proteins [73]. Sequence characteristics of head and tail domains can be shared between type I/type II pairs of keratins that are coexpressed in the same tissue, while their helical domain sequences bear little relationship to each other [55,51], so that some aspect of the tissue specificity of keratins is reflected in the C-terminal domain. Extending from the filament, and with minimal involvement in polymerization, the tail domain could be a prime site for tissue-specific interactions with other cytoplasmic components.

Keratin 19 may have been specifically conserved for some purpose dependent upon the absence of the carboxy-terminal tail domain. If so, the fact that there only appears to be one "tail-less" keratin (because there is no equivalent type II protein) implies that either a putative tail-determined function can be abrogated by the deletion of the domain from only one of the two keratin types in a

heteropolymer (in which case the two tail domains may function by specifically interactng with each other), or the function of the tail domain in type I keratins is different from their function in type II keratins.

Keratin 19 is Closer to the Keratinocyte Keratins than to Simple Epithelial Keratin 18 Comparing the amino acid sequences of the helical rod domains from a total of 20 type I keratins that have been fully or partially sequenced to date (Fig 6), it is clear that human keratin 19 is very closely related to the keratins expressed by stratified epithelia, whether the comparison is made over the short, highly conserved coil 1A or the long, more divergent coil 2. The homology with keratin 18 is strikingly lower. This is notable because keratin 19 is usually regarded as one of the keratins characteristically expressed by simple epithelia [4,8]; clearly, keratin 19 should rather be thought of as related to the keratinocyte keratins. Divergences between keratin 18 and other type I keratins were noticed by Singer et al [60], as was a lack of expected obvious similarity between keratins 19 and 18 by Bader et al [13]. The divergence of the whole group of keratinocyte keratins from the simple epithelial hallmark keratin 18 suggests that keratin 18 may even belong to a different subgroup, and may possibly be from a more distant, probably older, evolutionary stock.

The degree of homology of human keratin 19 with mouse keratin 13 is considerable in view of the species difference. Once the sequence of human keratin 13 has been determined, the homology between keratins 13 and 19 may well be found to be even higher. This may be relevant because keratins 19 and 13 are often expressed in the same tissues, although at different stages of differentiation because keratin 19 is observed in basal cells and keratin 13 in suprabasal cells. Both these keratins are dramatically inducible by retinoids [46], at least in tissue culture situations. Both the human keratin 19 shown here and the mouse keratin 13 [64] have potential glycosylation sites, the one in keratin 19 at amino acid position 393 (Fig 4) and three in mouse keratin 13 at amino acid positions 290, 291, and 296. Human keratin 13 is known to be glycosylated *in vivo* [74,75], and from its behaviour on gel electrophoresis it seems likely that mouse keratin 13 is also. It is not known whether or not keratin 19 is glycosylated *in vivo;* neither is it known what the function of glycosylation of keratins might be.

Tissue Distribution of Keration 19 Expression Keratin 19 is the only keratin of the 29 separate human keratin species identified to occur in both simple and stratified squamous epithelia in significant amounts and under normal circumstances [1,5,7–12]. Keratin 19 has been detected biochemically in a large number of different epithelial tissues [5], and in many of these tissues the exact cells which express keratin 19 have now been identified using monospecific monoclonal antibodies [6,7,12]. Some of these tissues are illustrated in Figs 2 and 3. Considering the pattern of expression within tissues it appears to be the case that wherever two different epithelial cell phenotypes coexist in close proximity to one another, such that they may have arisen from a common progenitor cell pool, keratin 19 is abundant. Either two morphologically distinct epithelia abut (such as the endo-ectocervical junction), two different cell types coexist within the same epithelium (such as the "mixed" epithelia of glands/gland ducts), or cells with two diverse differentiated keratin phenotypes can be seen in close apposition (as seen in oral epithelia). Furthermore, where the tissue location of the progenitor compartment is known, this location frequently appears to coincide with the distribution of keratin 19: it is in the basal cell layer that keratin 19 is expressed rather than in the committed differentiating suprabasal compartment.

The differentiation versatility, or instability, of cell populations associated with keratin 19 expression is particularly apparent in the following locations. Keratin 19 is widely expressed in oral epithelia, overlapping the many transitional zones. It is found heterogeneously in patches of basal cells in the sulcular (tip of gum) epithelium down to the gingival epithelium [42,43]: non-cornifying sulcular epithelium expresses keratins 4 and 13, but adjacent cornifying gingival epithelium expresses keratins 1/2 and 10. During development of human epidermis, keratin 19 is strongly present in the basal layer and periderm, and weakly in suprabasal layers, at 14 weeks (Figure 3) [7,76]: suprabasal cells simultaneously express keratins 13 and 10 [7,10,76] at this time (thus keratin 19 expression is not limited to keratin 10-negative regions). The mammary glandular epithelium consists of 2 distinct epithelial cell types with different keratin phenotypes [43]: expression of keratin 19 is homogeneous in the luminal cells of most, but not all, parts of the gland [8]. The nipple is the site of the junction of mammary gland ducts with the epidermis, and this epidermis expresses keratin 19 in patches of basal cells in an individually variable manner (Fig 3). Keratin 19 also appears in the basal cells of the ectocervix [39]. Two phenotypic variations are identifiable here: first the transition from endocervix (simple epithelium plus basal reserve cells) to ectocervix (stratified squamous epithelium), and second the variations in keratin pheno-

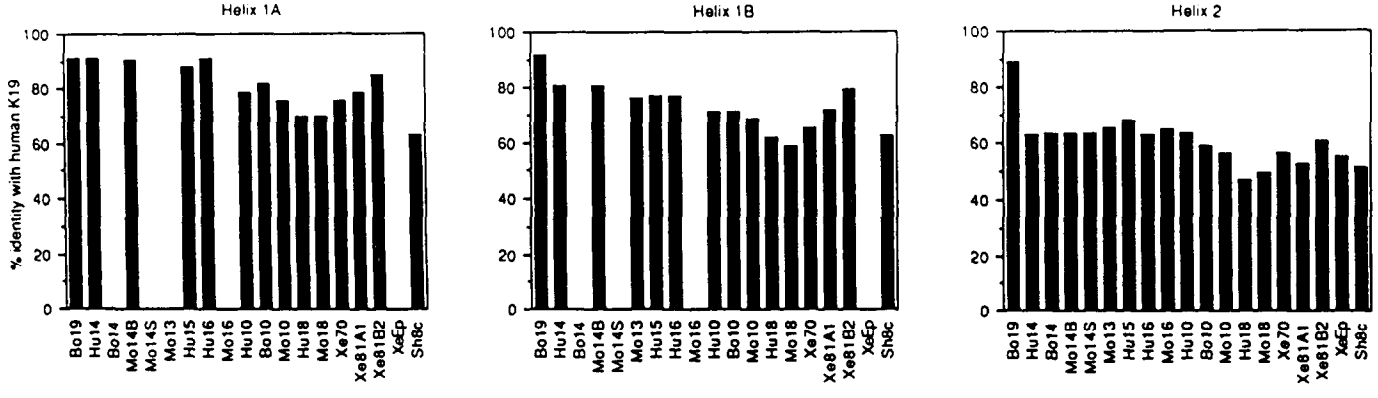

Figure 6. Relatedness of rod domain sequences of other type I keratins to human keratin 19. Expressed as percentage (of available sequence) of amino acids *(aa)* identical to human keratin 19 sequence. Consensus helical domain boundaries were as indicated in Fig 3. Penalties were included for gaps (ranging from 1-11 *aa*) required to obtain an optimal alignment, but not for additional blocks of 4-13 *aa* in the sequence compared to keratin 19. *Bo19:* bovine homologue of keratin 19(13); *Hu14:* human keratin 14(52); *Bo14:* bovine keratin VII, putative homologue of human K14 (only 16 *aa* of helix 2) [62]; *Mo14B:* mouse 50 kD basal-expressed (only 31 *aa* of helix 1A); *Mo14S:* mouse 50 kD basal- and suprabasal-expressed (only 44 *aa* of helix 2), two keratin 14 homologues [63]); *Mo13:* mouse putative keratin 13 homologue [64]; *Hu15:* human keratin 15(65); *Hu16:* human keratin 16 (formerly thought to be keratin 17) [66]; *Mo16:* putative mouse homologue of keratin 16 [63]; *Hu10:* human keratin 10 [67]; *Bo10:* bovine keratin Vlb, putative homologue of human K10/K11 [68]; *Mo10:* mouse, putative homologue of human K10 [53]; *Hu18:* human keratin 18 [58]; *Mo18:* mouse ENDO B, or homologue of human K18 [60]; *Xe70: Xenopus laevis* XK70 pre-metamorphic epidermal keratin [69]; *Xe81A1* and *Xe81B2: X. laevis* pre-metamorphic keratins [54]; *XeEp: X. laevis* epidermal keratin (only 100 *aa* of helix 2), [70]; *Sh8c:* sheep wool keratin 8c [71].

type seen in the suprabasal cells of the ectocervix (patches of suprabasal cells expressing keratins 10/11 have been identified by immunohistochemistry [39]).

The observation (Fig 2) that in skin there exists a keratin 19-positive subpopulation of cells in the hair follicle outer root sheath is of interest in view of the suggestion that a potent stem cell population may exist within the hair follicle [77]. It is known that skin damage to below the level of the hair follicles is much slower to regenerate than shallower damage where hair follicle bases remain [78,79]. Hair follicles had originally been reported as not containing keratin 19 as assayed biochemically [16,17], although later immunohistochemical investigations indicated that there probably was keratin 19 in follicles [7,12]. Earlier negative results were probably due to sampling artifacts. In tissue culture keratin 19 is expressed in most epithelial systems [1,3,5,10,19], and this is certainly a situation where drift is observed in keratin expression. In view of the inducibility of keratin 19 by retinoids [46], though, it is difficult to interpret appearance of keratin 19 in tissue culture in the context of in situ differentiation.

Keratin 19 and the characteristic patches of keratin 19-negative cells that are often observed where keratin 19 is expressed have tentatively been associated with the presence of stem cells [8]. The lack of keratin 19 over extensive regions of body skin such as interfollicular epidermis makes it unlikely that keratin 19 is directly indicative of all stem cells, agreeing with Bartek et al [8]. Our observations, however, do suggest that there is a particular zone of interest, with respect to skin differentiation and regeneration, in the deep outer hair root sheath that warrants closer observation. The trend of the histochemical observations, towards more extensive detection of keratin 19 in the early stages of differentiation and of keratin 19 expression in regions of labile differentiation, does suggest at least that pluripotent progenitor cells may reside within the epithelial locations harboring keratin 19-positive cells. Keratin 19 expression therefore could also indicate the location of particularily transformation-sensitive populations of cells.

Possible Function of Keratin 19 Both Bader et al [13] and Eckert [15] have interpreted the presence of an in-register extension of the helix as indicating that keratin 19 reflects the more primitive state of an ancestral keratin from which other keratins subsequently evolved, because they argue that the good-in register match of the extension would be less likely to arise randomly than to reflect a primary condition that was subsequently lost. However, α-helicity is the most common structural conformation found in proteins, and an α-helix might be selected as a way of compacting any acquired irrelevant C-terminal peptides out of the way if the function of keratin 19 depends positively on having no post-helical domain. Thus it is not improbable that any recently-acquired post-helical extension might evolve as an α-helix.

One good reason for arguing that keratin 19 is not the most primitive keratin is the strong similarity it bears to the keratins of stratified epithelia, which are structurally more elaborate than simple epithelia and arise later in development. If keratin 19 is in any sense an ancestral keratin, it can only be ancestral to some of the type I keratins of stratified epithelia. The simple epithelial keratin 18, which is the earliest embryonic type I keratin, appears to be of a separate and probably older stock.

A hypothesis that would link the tissue distribution pattern with the structural absence of a tail domain would be that keratin 19 has been conserved for its ability to act as a neutral keratin in terms of differentiaion. Keratin 19 is fully competent at polymerization and thus probably can integrate into the tonofilament network fundamental to the structure of an epithelium, but it may not fully restrict the cell expressing it to differentiate along either or any available pathway if it lacks the (as yet unidentified) tail-specified differentiation function. Where keratin 19 is expressed in simple epithelia there is at present little information regarding the temporal sequence of expression of additional keratins, but keratin-19-expressing simple epithelia appear to be potentially heterogeneous. The differentiation-specific keratin pairs of stratified epithelia are expressed once the cells have left the basal layer, but among these suprabasal keratins the type II member of each pair is often, if not always, expressed before its type I keratin partner [80,81] and can be detected in the basal layer. Presence of keratin 19 in the progenitor compartment of complex epithelia may allow a delay in expression of the second (type I) keratin by polymerizing with, and stabilizing, any early-synthesized type II keratins. Assuming that the tissue-specific characteristics of a particular keratin filament type will only be realized when both members of the keratin pair are copolymerized, delayed synthesis of only the type I member of the pair could postpone committment to differentiation as effectively as delaying both keratins. The complex formed between any given type II keratin and keratin 19 would be predicted to be of consistently lower affinity than the complex formed between this type II keratin and its specific type I partner, so that the appropriate type I keratin could always compete away the keratin 19 as the cells differentiated. A likely function for the tail domain, therefore, would be to provide high affinity reinforcement of the specific pair heteropolymer complex formation.

This role for keratin 19 as a "switch" keratin could be a mechanism for suspending post-stem cell amplifying cells in a flexible state of differentiation, pending local demand for one or other cell type. Such an additional level of "decision making" in the regulation of differentiation within complex epithelia may in fact be essential for maintenance of tissue homeostasis within mixed cell populations or across boundaries. The possible existence of a keratin phenotype that is permissive or indicative of a labile state of differentiation may also have clinical implications, in the potential vulnerability to transformation of such cells.

We gratefully acknowledge the assistance of Dr. Josephine Nalbantoglu for advice on sequencing, Dr. John Young for advice on bacteriophage lambda *DNA preparation, Dr. Robert Oshima for helpful discussions. We also thank all who let us use their antibodies (Dr. Uwe Karsten for A53-B/A2, Dr. Jiri Bartek for BA16 and BA17, Dr. Ray Nagle for KA4 and Dr. Tung-Tien Sun for AE1) and Messrs. John Gilbert and George Elia for cutting the frozen sections. Since this manuscript was written, Heid et al [Differentiation 37:137–157, 1988] reported variable staining of keratin 19 in hair follicles of human scalp using monoclonal antibody A53-B/A2 (=Ks19.1); this does not conflict with our own observations.*

REFERENCES

1. Moll R, Franke WW, Schiller DL, Geiger B, Krepler R: The catalog of human cytokeratins: patterns of expression in normal epithelia, tumors and cultured cells. Cell 31:11–24, 1982
2. Heid HW, Moll I, Franke WW: Patterns of expression of trichocytic and epithelial cytokeratins in mammalian tissues. II. Concomitant and mutually exclusive synthesis of trichocytic and epithelial cytokeratins in diverse human and bovine tissues (hair follicle, nail bed and matrix, lingual papilla, thymic reticulum) Differentiation 37:215–230, 1988
3. Wu Y-J, Rheinwald JG: A new small (40kd) keratin filament protein made by some cultured human squamous cell carcinomas. Cell 25:627–635, 1981
4. Sun T-T, Eichner R, Schermer A, Cooper D, Nelson WG, Weiss RA: Classification, expression, and possible mechanisms of evolution of mammalian keratins: A unifying model. In: Levine AJ, Vande Woude GF, Topp WC, Watson JD (eds). Cancer Cells I, The Transformed Phenotype. Cold Spring Harbor Laboratory, Cold Spring Harbor, 1984, pp 169–176
5. Quinlan RA, Schiller DL, Hatzfeld M, Achtstätter T, Moll R, Jorcano JL, Magin TM, Franke WW: Patterns of expression and organization of cytokeratin intermediate filaments. Ann NY Acad Sci 455:282–306, 1985
6. Bartek J, Durban EM, Hallowes RC, Taylor RC, Papadimitrou J: A subclass of luminal epithelial cells in the mammary gland defined by antibodies to cytokeratins. J Cell Sci 75:17–33, 1985
7. Lane EB, Bartek J, Purkis PE, Leigh IM: Keratin antigens in differentiating skin. Ann NY Acad Sci 445:241–258, 1985
8. Bartek J, Taylor-Papadimitriou J, Miller N, Millis R: Patterns of ex-

pression of keratin 19 as detected with monoclonal antibodies in human breast tissues and tumors. Int J Cancer 36:299–306, 1985

9. Moll R, Levy R, Czernobilsky B, Hohlweg-Mejert P, Dallenbach-Hellweg G, Franke WW: Cytokeratins of normal epithelia and some neoplasms of the female genital tract. Lab Invest 49:599–610, 1983
10. van Muijen GNP, Warnaar SO, Ponec M: Differentiation-related changes of cytokeratin expression in cultured keratinocytes and in fetal, newborn and adult epidermis. Exper Cell Res 171:334–345, 1987
11. Nagle RB, Moll R, Weidauer H, Nemetscheck H, Franke WW: Different patterns of cytokeratin expression in the normal epithelia of the upper respiratory tract. Differentiation 30:130–140, 1985
12. Bartek J, Bartkova J, Taylor-Papadimitriou J, Reijthar A, Kovarik J, Lukas Z, Vojtesek B: Differential expression of keratin 19 in normal human epithelial tissues revealed by monospecific monoclonal antibodies. Histochem J 18:565–575, 1986
13. Bader BL, Magin TM, Hatzfeld M, Franke WW: Amino acid sequence and gene organization of cytokeratin no. 19, an exceptional tail-less intermediate filament protein. EMBO J 5:1865–1875, 1986
14. Stasiak PC, Lane EB: Sequence of cDNA coding for human keratin 19. Nucl Acids Res 15:10058, 1987
15. Eckert RL: Sequence of the human 40-kDa keratin reveals an unusual structure with very high sequence identity to the corresponding bovine keratin. Proc Natl Acad Sci USA 85:1114–1118, 1988
16. Moll R, Franke WW, Volc-Platzer B, Krepler R: Different keratin polypeptides in epidermis and other epithelia of human skin: a specific cytokeratin of molecular weight 46,000 in epithelia of the pilosebaceous tract and basal cell epitheliomas. J Cell Biol 95:285–295, 1982
17. Moll R, Moll I, Wiest W: Changes in the pattern of cytokeratin polypeptides in epidermis and hair follicles during skin development in human fetuses. Differentiation 23:170–178, 1982
18. Lane EB: Monoclonal antibodies provide specific intramolecular markers for the study of epithelial tonofilament organisation. J Cell Biol 92:665–673, 1982
19. Taylor-Papadimitriou J, Purkis P, Lane EB, McKay IA, Chang SE. Effects of SV40 transformation on the cytoskeleton and behavioural properties of human keratinocytes. Cell Diff 11:169–180, 1982
20. Lane EB, Lane DP: A rapid antibody assay system for screening hybridoma cultures. J Immunol Meth 47:303–307, 1981
21. Karsten U, Papsdorf G, Roloff G, Stolley P, Abel H, Walther I, Weiss H: Monoclonal anticytokeratin antibody from a hybridoma clone generated by electrofusion. Eur J Cancer Clin Oncol 21:733–740, 1985
22. Gigi-Leitner O, Geiger B: Antigenic interrelationship between the 40-kilodalton cytokeratin polypeptide and desmoplakins. Cell Motil Cytosk 6:628–639, 1986
23. Nagle RB, Lucas DO, McDaniel KM, Clark VA, Schmalzel GM: New evidence linking mammary and extramammary Paget cells to a common cell phenotype. Am J Clin Path 83:431–438, 1985
24. Woodcock-Mitchell J, Eichner R, Nelson WG, Sun T-T: Immunolocalization of keratin polypeptides in human epidermis using monoclonal antibodies. J Cell Biol 95:580–588, 1982
25. Pruss RM, Mirsky R, Raff MC, Thorpe R, Dowding AJ, Anderton BH: All classes of intermediate filaments share a common antigenic determinant defined bt a monoclonal antibody. Cell 27:419–428, 1981
26. O'Farrell PH: High resolution two-dimensional electrophoresis of proteins. J Biol Chem 250:4007–4021, 1975
27. O'Farrell PZ, Goodman HM, O'Farrell PH: High resolution two-dimensional electrophoresis of basic as well as acidic proteins. Cell 12:1133–1142, 1977
28. Towbin H, Staehelin T, Gordon J: Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets. Proc Natl Acad Sci USA 76:4350–4354. 1979
29. Hancock K, Tsang VCW: India ink staining on nitrocellulose paper. Anal Biochem 133:157–162, 1983
30. Young RA, Davis RW: Efficient isolation of genes by using antibody probes. Proc Natl Acad Sci USA 80:1194–1198, 1983
31. Ruther U, Müller-Hill B: Easy identification of cDNA clones. EMBO J 2:1791–1794, 1983
32. Messing J, Vieira J: The pUC plasmids, an M13 mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene 19:259–268, 1982
33. Norrander J, Kempe T, Messing J: Construction of improved M13 vectors using oligonucleotide-directed mutagenesis. Gene 26:101–106, 1983
34. Sanger F, Nicklen S, Coulson AR: DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci 75:5463–5467, 1977
35. Staden R: Automation of the computor handling of gel reading data produced by the shotgun method of DNA sequencing. Nucl Acids Res 10:4731–4751, 1982
36. Garnier J, Osguthorpe DJ, Robson B: Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J Mol Biol 120:97–120, 1978
37. Novotny J, Auffray C: A program for prediction of protein secondary structure from nucleotide sequence data: application to histocompatibility antigens. Nucl Acids Res 12:243–255, 1984
38. Moll R, Moll I, Franke WW: Identification of Merkel cells in human skin by specific cytokeratin antibodies: changes of cell density and distribution in fetal and adult plantar epidermis. Differentiation 28:136–154, 1984
39. Gigi-Leitner O, Geiger B, Levy R, Czernobilsky B: Cytokeratin expression in sqamous metaplasia of the human uterine cervix. Differentiation 31:191–205, 1986
40. Morgan PR, Leigh IM, Purkis PE, Gardner ID, van Muijen GNP, Lane EB: Site variation in keratin expression in human oral epithelia—an immunocytochemical study of individual keratins. Epithelia 1:31–43, 1987
41. Morgan PR, Johnson NW, Leigh IM, Lane EB: Structure of gingival epithelium as revealed by monoclonal antibodies to keratins. In: Lehner T, Cimasoni G (eds.). Borderland between Caries and Periodontal disease III. Medicine and Hygiene, Geneva, 1986, pp 46–59
42. Chaudary MA, Millis RR, Lane EB, Miller NA: Paget's disease of the nipple: a ten year review including clinical, pathological and immunohistochemical findings. Breast Canc Res Treatm 8:139–146, 1986
43. Taylor-Papadimitriou J, Lane EB: Keratin expression in the mammary gland. In: Neville MC, Daniel CW (eds.). The Mammary Gland. Plenum, New York, 1987, pp 181–215
44. Kozak M: Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell 44:283–292, 1986
45. Kozak M: Comparison of initiation of protein synthesis in procaryotes, eukaryotes, and organelles. Microbiol Rev 47:1–45, 1983
46. Eckert RL, Green H: Cloning of cDNAs specifying vitamin A-responsive human keratins. Proc Natl Acad Sci USA 81:4321–4325, 1984
47. Crick, FH: The packing of α-helices: simple coiled-coils. Acta Crystal 6:687–688
48. Steinert PM, Parry DAD: Intermediate filaments: conformity and diversity of expression and structure. Ann Rev Cell Biol 1:41–65, 1985
49. Yaffe D, Nudel U, Mater Y, Neuman S: Highly conserved sequences in the 3′ untranslated region of mRNAs coding for homologous proteins in distantly related species. Nucl Acids Res 13:3723–3737, 1985
50. Steinert PM, Parry DAD, Idler WW, Johnson LD, Steven AC, Roop DR: Amino acid sequences of mouse and human epidermal type II keratins of Mr 67,000 provide a systematic basis for the structural and functional diversity of the end domains of keratin intermediate filament subunits. J Biol Chem 260:7142–7149, 1985
51. Klinge EM, Sylvestre YR, Freedberg IM, Blumenberg M: Evolution of keratin genes: different protein domains evolve by different pathways. J Mol Evol 24:319–329, 1987
52. Hanukoglu I, Fuchs E: The cDNA sequence of a human epidermal keratin: Divergence of sequence but conservation of structure among intermediate filament proteins. Cell 31:243–252, 1982
53. Steinert PM, Rice RH, Roop DR, Trus BL, Steven AC: Complete amino acid sequence of a mouse epidermal keratin subunit and implications for the structure of intermediate filaments. Nature 302:794–800, 1983
54. Miyatani S, Winkles JA, Sargent TD, Dawid IB: Stage-specific keratins in *Xenopus laevis* embryos and tadpoles: The XK81 gene family. J Cell Biol 103:1957–1965, 1986

55. Glass C, Kim KH, Fuchs E: Sequence and expression of a human type II mesothelial keratin. J Cell Biol 101:2366-2373, 1985

56. Leube RE, Bosch FX, Romano V, Zimbelmann R, Höfler H, Franke WW: Cytokeratin expression in simple epithelia III. Detection of mRNAs encoding human nos. 8 and 18 in normal and tumour cells by hybridization with cDNA sequences in vitro and in situ. Differentiation 33:69-85, 1986

57. Magin TM, Jorcano JL, Franke WW: Cytokeratin expression in simple epithelia II. cDNA cloning and sequence characteristics of bovine cytokeratin A (no. 8). Differentiation 30:254-264, 1986

58. Oshima RG, Milan JL, Cecena G: Comparison of mouse and human keratin 18: A component of intermediate filaments expressed prior to implantation. Differentiation 33:61-68, 1986

59. Romano V, Hatzfeld M, Magin TM, Zimbelmann R, Franke WW, Maier G, Ponting H: Cytokeratin expression in simple epithelia. I. Identification of mRNA coding for human cytokeratin no. 18 by a cDNA clone. Differentiation 30:244-253, 1986

60. Singer PA, Trevor K, Oshima RG: Molecular cloning and characterization of the Endo B cytokeratin expressed in preimplantation mouse embryos. J Biol Chem 261:538-547, 1986

61. Geisler N, Weber K: Amino acid sequence data on glial fibrillary acidic protein (GFA); implications for the subdivision of intermediate filaments into epithelial and non-epithelial members. EMBO J 2:2059-2063, 1983

62. Jorcano JL, Rieger M, Franz JK, Schiller DL, Moll R, Franke WW: Identification of two types of keratin polypeptides within the acidic cytokeratin subfamily I. J Mol Biol 179:257-281, 1984

63. Knapp B, Rentrop M, Schweizer J, Winter H: Three cDNA sequences of mouse type I keratins. Cellular localization of the mRNAs in normal and hyperproliferative tissues. J Biol Chem 262:938-945, 1987

64. Knapp B, Rentrop M, Schweizer J Winter H: cDNA sequences and in situ localization of the mRNAs. Nucl Acids Res 14:751-763, 1986

65. Leube RE, Bader BL, Bosch FX, Zimbelmann R, Achtstätter T, Franke WW: Molecular characterization and expression of the stratification-related cytokeratins 4 and 15. J Cell Biol 106:1249-1261, 1988

66. Rosenberg M, RayChaudury A, Snows TB, LeBeau M, Fuchs E: A group of type I keratin genes on human chromosome 17: Characterization and expression. Mol Cell Biol 8:722-736, 1988

67. Darmon MY, Sémat A, Darmon MC, Vasseur M: Sequence of a cDNA encoding human keratin No. 10 selected according to structural homologies of keratins and their tissue-specific expression. Mol Biol Rep 12:277-283, 1987

68. Rieger M, Jorcano JL, Franke WW: Complete sequence of a bovine type I cytokeratin gene: conserved and variable positions in genes of polypeptides of the same cytokeratin subfamily. EMBO J 4:2261-2267, 1985

69. Winkles JA, Sargent TD, Parry DAD, Jonas E, Dawid IB: Developmentally regulated cytokeratin gene in *Xenopus Laevis*. Mol Cell Biol 5:2575-2581, 1985

70. Hoffman W, Franz JK: Amino acid sequence of the carboxy-terminal part of an acidic type I cytokeratin of molecular weight 51 000 from *Xenopus Laevis* epidermis as predicted from the cDNA sequence. EMBO J 6:1301-1306, 1984

71. Dowling LM, Crewther WG, Inglis AS: The primary structure of component 8c-1, a subunit protein of intermediate filaments in wool keratin. Relationships with proteins from other intermediate filaments. Biochem J 236:695-703, 1986

72. Kaufmann E, Weber K, Geisler N: Intermediate filament forming ability of desmin derivatives lacking either the amino-terminal 67 or the carboxy-terminal 27 residues. J Mol Biol 185:733-742, 1985

73. Geisler N, Fischer S, Vandekerckhove J, Plessmann U, Weber K: Hybrid character of a large neurofilament protein (NF-M): intermediate filament type sequence followed by a long and acidic carboxy-terminal extension. EMBO J 3:2701-2706, 1984

74. Roberts GP, Brunt J: Biosynthesis of a glycosylated keratin by human keratinocytes. Biochem Biophys Acta 883:413-419, 1986

75. Vidrich A, Gilmartin M, Ryan K, Freedberg IM: Localization of phosphate and glycosamine on keratins of ME180 cells (abstr). J Invest Dermatol 82:403, 1984

76. Dale BA, Holbrook KA, Kimball JR, Hoff M, Sun T-T: Expression of epidermal keratins and filaggrin during human fetal skin development. J Cell Biol 101:1257-1269, 1985

77. Al-Barwari SE, Potten CS: Regeneration and dose-response characteristics of irradiated mouse dorsal epidermal cells. Int J Radiat Biol 30:201-216, 1976

78. Winter GD: Epidermal regeneration studied in the domestic pig. In: Maibach HI, Rovee DT (eds.). Epidermal Wound Healing. Year Book Medical Publishers, Chicago, 1972, pp 71-112

79. Hunter JAA, Mcittie E, Comaish JS: Light and electron microscopic studies of physical injury to the skin I: Suction. Br J Dermatol 90:481-490, 1974

80. Schermer A, Galvin S, Sun T-T: Differentiation-related expression of a major 64K corneal keratin in vivo and in culture suggests limbal location of corneal epithelial stem cells. J Cell Biol 103:49-62, 1986

81. Roop DR, Huitfeldt H, Kilkenny A, Yuspa SH: Regulated expression of differentiation-associated keratins in cultured epidermal cells detected by monospecific antibodies to unique peptides of mouse epidermal keratins. Differentiation 35:143-150, 1987