*Proc. Natl. Acad. Sci. USA*
Vol. 83, pp. 5048–5052, July 1986
Biochemistry

# Clustered arrangement of keratin intermediate filament genes

**(separate gene clusters/variable transcriptional polarity/orphan exon)**

BARRY C. POWELL, GRAHAM R. CAM, MICHAEL J. FIETZ, AND GEORGE E. ROGERS

Adelaide Centre for Gene Technology, Department of Biochemistry, University of Adelaide, Adelaide, South Australia, 5000, Australia

*Communicated by Howard A. Bern, March 24, 1986*

ABSTRACT **We report here that component members of the keratin intermediate filament (IF) type I and type II gene families of sheep are closely linked but apparently the two families are not. Nine genes, accounting for up to half of the keratin IF gene repertoire, were mapped in four cosmid clones and the linkage between the genes ranged from several kilobases to 20 kilobases. In one cosmid, three tandem type I genes had the same transcriptional arrangement and were regularly spaced. In another cosmid, tandem genes encoding type II keratins were identified and, surprisingly, a solitary exon was discovered in the intergene region between the two type II genes. In a normal gene this exon encodes one of the most conserved amino acid regions of IF proteins, the C-terminal end of the α-helical core. Homologous C-terminal protein subdomains were encoded by two wool keratin type II genes and we suggest that this arrangement may also exist in the other wool keratin type II genes.**

Intermediate filaments (IF) are components of the cell cytoskeleton found in most eukaryotic cells and for the keratin IF of epithelial cells the fundamental building block is composed of two protein classes called type I and type II, aligned by interaction of their α-helical regions (1–4). The two keratin IF protein families are quite large, each including 10–20 proteins, which are encoded by a similar number of genes (5–7). Yet, from this large repertoire of type I and type II genes only a few, typically one coordinately expressed pair, seem to be expressed in any given epithelial phenotype (6, 8). This selective coexpression must reflect the different function of the IF in different keratinocytes. A comparison of coordinately expressed genes may indicate regulatory sequences that are involved in this coupled expression and it is possible that these coordinately expressed genes are linked in the genome. Recently, individual type I and type II keratin genes were isolated from human (9–11), mouse (12), and bovine (13) genomic libraries, but it is not known whether the type I and type II genes are linked or, indeed, if there is any linkage within the two gene families.

In the sheep wool keratinocyte, four type I and four type II keratin proteins have been identified (14, 15) and one representative cDNA clone of each type has been isolated (40). A comparison of these cDNA sequences with their human counterparts (9, 10, 16), derived from epidermal cells that express different subsets of keratin genes, suggests that the sheep cDNA probes would cross-hybridize with all keratins of the same type. Using these two probes, we report here that component members of the sheep keratin IF type I and type II gene families are closely linked, but the gene families are not.

## EXPERIMENTAL PROCEDURES

**Preparation of a Sheep Genomic Cosmid Library, Screening, and Isolation of Recombinant Clones Specific for Type I and Type II Keratin Genes.** High molecular weight DNA [>150 kilobases (kb)] was prepared from sheep liver and was partially digested with restriction endonuclease *Sau*3A under conditions that optimize the representation of DNA sequences in recombinant DNA libraries (17). Size-fractionated DNA was then cloned into cosmid vector pHC79 DNA (5:1 molar ratio) prepared by a modification of the method described by Hohn and Collins (18). Recombinant cosmid DNA was packaged with an efficiency of *ca.* $5.5 \times 10^5$ colonies per μg of insert DNA. Bacterial colonies containing recombinant plasmids were screened by the high-colony-density procedure (50,000 per 150-mm filter) described by Hanahan and Meselson (19). Filters were hybridized to a combination of nick-translated (20) type I and type II keratin probes (see Fig. 1) with $2 \times 10^6$ cpm per filter and then were washed and autoradiographed according to standard procedures (21). Positive clones were selected from master plates and duplicate filters were reprobed at low-density plating with the individual type I and type II keratin probes to allow the isolation of pure clones. Recombinant cosmid DNA was prepared by a modification of the procedure of Birnboim and Doly (22).

**M13 Subcloning, Sequencing, and Preparation of Strand-Specific Probes.** Appropriate restriction fragments were isolated from acrylamide gels (23) or were ligated directly from low-melting temperature agarose gels (24) into either M13 mp9 (25) or M13 mp18 (26) vectors. Specific 5′ and 3′ gene probes were prepared by subcloning *Pst* I-digested wool keratin type I and type II cDNA clones (see Fig. 1) into M13, and radioactively labeled strand-specific probes were prepared as described by Hu and Messing (27). DNA sequences were determined by the dideoxy chain-termination technique of Sanger *et al.* (28).

## RESULTS

**Keratin Type I and Type II Gene Families Are Not Closely Linked in the Genome.** We screened three to four genome equivalents of a sheep cosmid library with sheep wool keratin type I and type II cDNA clone probes under low-stringency conditions and detected several hundred positive clones of varying intensity. About 100 were selected for further screening and duplicate filters were hybridized with either the type I or the type II probe, but no clones were observed to hybridize to both probes. Several clones were subsequently purified and 2 type I and 2 type II positive clones have been characterized in detail.

**Tandem Type I Keratin Genes.** Our type I cDNA clone probe encoded 285 amino acids, about two-thirds of a type I protein, and 44 nucleotides from the 3′ noncoding region. We constructed specific 5′ and 3′ probes, each encoding about 140 amino acids from this cDNA clone (Fig. 1*A*). Available keratin type I gene sequence data (refs. 9 and 12; K. Ward, personal communication) indicate that in the α-helical region

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: IF, intermediate filament(s); bp, base pair(s); kb, kilobase(s).

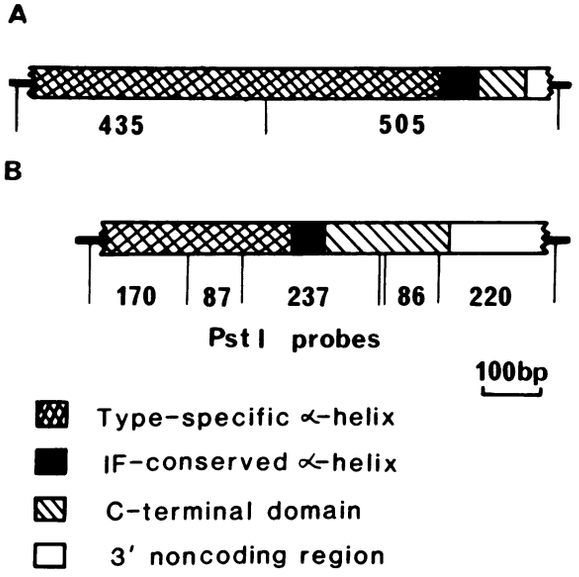


FIG. 1. Schematic representation of the wool keratin type I and type II cDNA clones. (*A*) This type I cDNA clone contains an insert into the *Pst* I site of pBR322 encoding 285 amino acids and 44 base pairs (bp) of a truncated 3′ noncoding region and is equivalent to the wool 8C1 low-sulfur keratin (40). The single *Pst* I site in the clone is indicated along with the size of the relevant fragments that were subcloned into M13 mp9 for use as 5′ (435 bp) and 3′ (505 bp) gene probes. (*B*) This type II cDNA clone encodes 187 amino acids and 180 bp of 3′ noncoding sequence and is equivalent to the wool 7C low-sulfur keratin (40). *Pst* I probes specifying α-helical domain (170, 87, and 237 bp), C-terminal domain (86 bp), and gene 3′ noncoding region were prepared in M13 mp9.

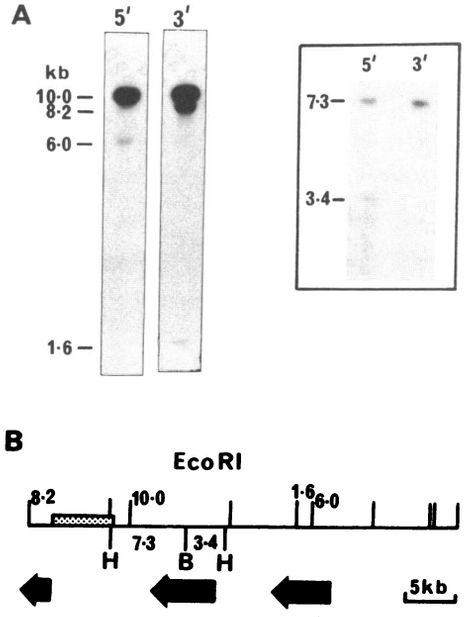


FIG. 2. Hybridization of wool keratin type I 5′ and 3′ gene probes to sheep cosmid 110. (*A*) Hybridization of M13 clones containing specific 5′ and 3′ gene probes prepared from a wool keratin type 1 cDNA clone (see Fig. 1*A*) to Southern blots (29) of *Eco*RI-digested cosmid 110. The blots were washed under low-stringency conditions (0.36 M NaCl/20 mM sodium phosphate, pH 7.4/2 mM EDTA/0.1% $NaDodSO_4$, 65°C) to detect the weak hybridization signals to the 6.0- and 1.6-kb fragments. (*Inset*) Hybridization of 5′ and 3′ gene probes to the middle gene of cosmid 110. A 10.7-kb *Hin*dIII fragment containing the gene was isolated, digested with *Bam*HI, transferred bidirectionally to nitrocellulose (29), and then hybridized with the type I 5′ and 3′ gene probes. The blots were washed under low-stringency conditions. The 5′ probe hybridized to both *Hin*dIII/*Bam*HI fragments and so spanned the *Bam*HI site. Since the 3′ probe hybridized to only the 7.3-kb *Hin*dIII/*Bam*HI fragment, the direction of transcription of the gene must be from right to left, from the 3.4-kb fragment to the 7.3-kb fragment. (*B*) *Eco*RI map and schematic interpretation of the hybridization results. The sizes of the hybridizing fragments are given in kilobases and the location of the vector sequence (▨), the genes (⬟), and the direction of transcription of all three genes are depicted. B, *Bam*HI; H, *Hin*dIII.

of type I genes the locations of the introns are conserved, and comparisons of our 5′ and 3′ type I probes with these data suggested that each probe would span at least two introns within type I keratin genes and therefore could detect a considerable length of DNA. The two probes hybridized to four different *Eco*RI fragments of cosmid 110, which were mapped to three regions separated by several kilobases of DNA (Fig. 2), suggesting the presence of three genes. The pattern of hybridization of the probes confirmed this as the middle and right gene regions hybridized to both of the specific 5′ and 3′ probes. The 3′ probe hybridized to *Eco*RI fragments from each region—namely, 10.0-, 8.2-, and 1.6-kb fragments—whereas the 5′ probe hybridized to the 10.0-kb fragment already detected by the 3′ probe and to a 6.0-kb fragment not detected previously. The 6.0- and 1.6-kb fragments that hybridized to the 5′ and 3′ probes, respectively, are adjacent in the cosmid and indicate the orientation of the right gene. Only the 3′ probe hybridized to the 8.2-kb fragment and since the 5′ probe did not hybridize to the contiguous fragment the gene must be only partly encompassed by the cosmid and must have the same transcriptional orientation as the gene at the extreme right (Fig. 2). The orientation of the gene within the 10.0-kb *Eco*RI fragment is the same as the other two and was determined by further blot analysis on an isolated and redigested fragment (see Fig. 2 *Inset*). These results demonstrate a regularly spaced linkage of three type I keratin genes, all with the same transcriptional polarity. Another, different cosmid that we have characterized contains one additional type I keratin gene (data not shown).

**Clustered Type II Keratin Genes: Variable Gene Orientation.** The type II cDNA clone used to screen the cosmid library is presented schematically in Fig. 1*B*. One of the cosmids selected with this probe, cosmid 151, contained 39 kb of insert DNA, and the *Eco*RI restriction enzyme fragments of this cosmid that hybridized with the total probe are shown in Fig. 3*A*. Of the three that hybridize, two are adjacent in the cosmid (4.1 and 3.5 kb), whereas the third, which hybridizes weakly by comparison, is 20 kb distant (Fig. 3*B*). To clarify the gene organization within this cosmid, short 5′ and 3′ probes were prepared from the wool keratin cDNA clone (Fig. 1*B*). The 170-, 87-, and 237-bp probes encode contiguous sections of the conserved α-helical region and on the basis of sequence comparison should cross-hybridize with all keratin type II genes, whereas the two more 3′ probes, the 86- and 220-bp probes, encode the more variant C-terminal domain of the protein and the 3′ noncoding region of the gene, respectively, and may both be gene specific. Only one probe, the 170-bp 5′-specific probe, hybridized to the 4.6-kb *Eco*RI fragment (Fig. 3*A*). As none of the more 3′ probes hybridized to this fragment or any other in the vicinity at low stringency (data not shown), not even the 237-bp probe that covers the highly conserved α-helical region abutting the C-terminal domain (Fig. 1*B*; refs. 30 and 31), then the 3′ end of that gene must lie on DNA not encompassed by this cosmid. The other type II keratin gene in this cosmid, located in adjacent 4.1- and 3.5-kb *Eco*RI fragments, hybridized to all probes under stringent conditions (Fig. 3*A*), and the strong hybridization of the 220-bp 3′ noncoding probe suggests that this gene is identical to the wool keratin clone probe. The hybridization pattern of the 5′ and 3′ probes to this cosmid

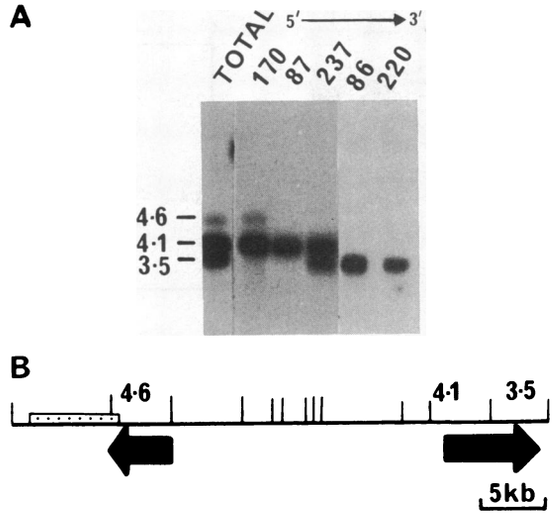


FIG. 3. Hybridization of wool keratin type II gene probes to sheep cosmid 151. (*A*) The left track shows hybridization of total nick-translated (20) type II cDNA clone insert to a *Bam*HI digest of cosmid 151. Then, proceeding from left to right are shown hybridization of 5′ through to 3′ type II gene probes (see Fig. 1*B*). All blots were washed at high stringency (18 mM NaCl/1 mM sodium phosphate, pH 7.4/0.1 mM EDTA/0.1% $NaDodSO_4$, 65°C). The sizes of hybridizing fragments are shown in kilobases. (*B*) The cosmid 151 map in which the location of the vector sequence (▨) and the approximate position and direction of transcription of the two type II genes are depicted (➡).

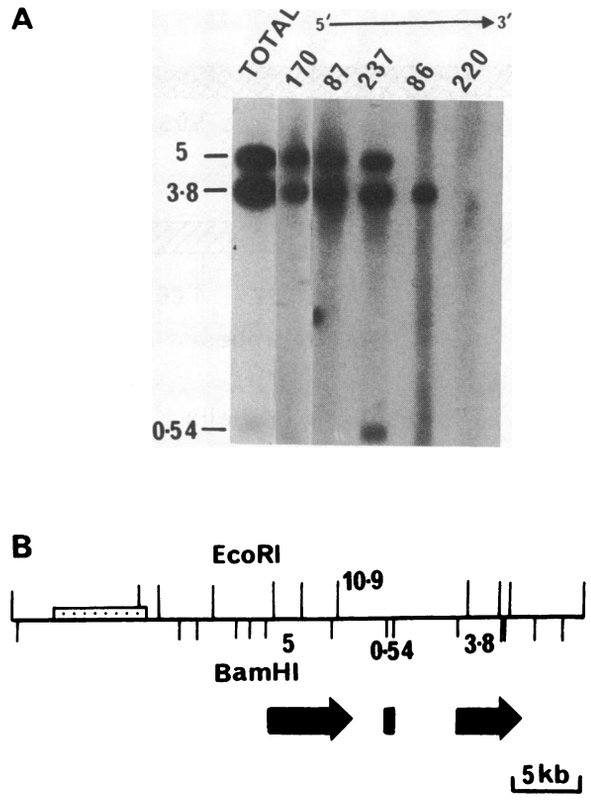


FIG. 4. Hybridization of wool keratin type II gene probes to sheep cosmid 150. (*A*) The left track shows hybridization of total nick-translated (20) type II cDNA clone insert to a *Bam*HI digest of cosmid 150. Then, proceeding from left to right are shown hybridization of 5′ through to 3′ type II gene probes (see Fig. 1*B*). Note that the "TOTAL" track was blotted onto nitrocellulose and the other five probe tracks were blotted onto Zeta-Probe membrane, and the variable retention of small DNA fragments by nitrocellulose resulted in a relatively weak hybridization signal by the 0.54-kb *Bam*HI fragment. All blots were washed at low stringency (0.36 M NaCl/20 mM sodium phosphate, pH 7.4/2 mM EDTA/0.1% $NaDodSO_4$, 65°C). (*B*) *Eco*RI and *Bam*HI restriction maps of cosmid 150 and the locations of the vector sequence (▨) and type II probe-positive fragments are depicted (▮). The direction of transcription of the two genes is also shown (➡).

demonstrates that these two linked type II keratin genes must be divergently transcribed (Fig. 3).

Hybridization of the type II cDNA clone to another, nonoverlapping cosmid, cosmid 150, detected three *Bam*HI fragments distributed over about 18 kb of DNA (Fig. 4). The 5′ and 3′ probe results show that there are two genes with sequences homologous to the probe separated by several kilobases of DNA and centered around the 3.8- and 5-kb *Bam*HI fragments (Fig. 4). No adjacent *Bam*HI fragments hybridized to any of the probes. Both genes hybridized to the three consecutive probes (170, 87, and 237 bp) from the α-helical region of the wool keratin type II clone, only one gene hybridized to the C-terminal domain probe, and neither hybridized to the 3′ noncoding probe. Thus, one gene is closely related but not identical to the wool keratin clone probe, whereas the other type II gene is not. Although the present 5′ and 3′ probe results do not reveal the orientation of the two type II keratin genes in this cosmid, further blot analyses from the appropriate *Bam*HI and *Eco*RI fragments have shown that both genes are transcribed in the same direction (data not shown).

**An Orphan Type II Exon Is Situated Between Two Type II Keratin Genes.** The 540-bp *Bam*HI fragment of cosmid 150, which is located between two type II genes, hybridized to the 237-bp probe, yet the adjacent *Bam*HI fragments in the cosmid did not hybridize to any other probe (Fig. 4). To elucidate the nature of this isolated hybridization signal we sequenced this small *Bam*HI fragment. It apparently encodes an IF type II exon of 73 amino acids (Fig. 5) bounded by sequences in excellent agreement with consensus splice junction sequences (32) and the postulated intron/exon junctions occur in the same positions relative to the protein sequence as determined for nine IF genes (9–13, 34).

When the 540-bp *Bam*HI fragment was probed to an *Eco*RI genomic blot, a 10.9-kb *Eco*RI fragment produced the strongest signal under stringent conditions (Fig. 6). This strongly suggests that the assigned location of the 540-bp *Bam*HI fragment within the 10.9-kb *Eco*RI fragment of cosmid 150 is correct (Fig. 4*B*) and does not result from an artificial rearrangement of sequences adjacent to the *Bam*HI fragment. The faintly visible bands are due to cross-reaction with other keratin type II genes, and the full complement of the type II family, detected under low-stringency conditions, is shown in the adjacent track (Fig. 6). Thus, it is unlikely that the exon encoded by the 540-bp *Bam*HI fragment is part of a normal gene but is rather an orphan exon.

## DISCUSSION

**Type II Orphan Exon.** The amino acid sequence of the orphan exon shows 31 differences from the equivalent sequence in the wool keratin type II cDNA clone, 21 in the 5′ half of the exon and 12 in the 3′ half (Fig. 5). Of these differences, 8 in the 5′ half involve charge changes, whereas only 1 in the 3′ half does. In normal IF proteins the 3′ half of this sequence demarcates the C-terminal end of the α-helical core section and includes one of the most conserved amino acid regions among all of the IF proteins (30). A comparison of the 5′ half of the orphan exon with known IF amino acid sequences (4, 9–12, 16, 30) suggests that this exon was derived from a keratin type II gene, but whether it was a wool or epidermal gene cannot be distinguished because the wool

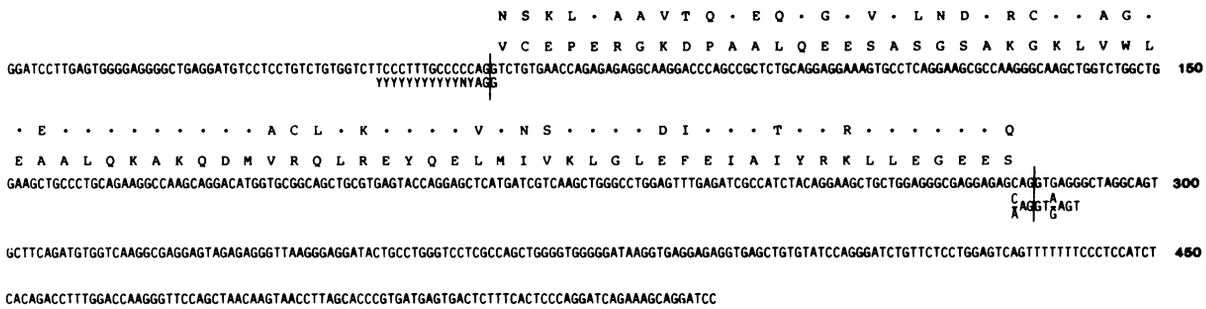


FIG. 5. Sequence of the 540-bp *Bam*HI fragment of cosmid 150. Consensus splice junction sequences (32) are shown below the DNA sequence and the intron/exon junctions are indicated by vertical lines. The conceptual amino acid sequence of the proposed exon is given immediately above the nucleotide sequence, using the single-letter code (33). The equivalent amino acid sequence from the wool keratin type II cDNA clone (K. Ward, personal communication) is presented above that encoded by the exon, and identical amino acids are represented by dots.

keratin (4) and human epidermal keratin (4, 10, 16) sequences both show a similar number of differences from this orphan exon. Interestingly, mapping and sequence data from a λ Charon 4A sheep genomic clone suggest that an equivalent exon is present as an orphan in another part of the sheep genome (unpublished). Unlike the sequence presented here, the other orphan exon is nearly identical to the equivalent sequence encoded by the wool keratin type II cDNA, differing in only three amino acids and at 15 nucleotide positions (93% nucleotide sequence homology; unpublished), indicating that it was derived from a wool keratin type II gene.

These orphan exons probably arose through unusual recombination events either during gene duplication or replication, but whether they were produced from the same gene is not known. If they were generated during the gene duplications that led to the current repertoire of mammalian keratin IF genes then they should be detectable in related genomes (35). As far as we are aware, an instance of solitary extragenic exons has not been reported previously, yet our observations are not completely without precedent, as solitary histone genes displaced from their regular clusters have been reported (36, 37).

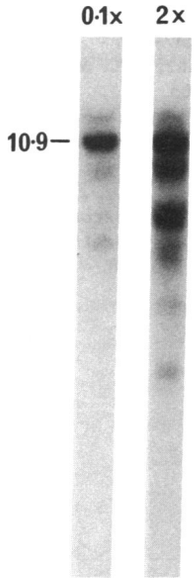


FIG. 6. Hybridization of the 540-bp *Bam*HI fragment of cosmid 150 to a sheep *Eco*RI genomic blot. Sheep DNA (10 μg per track), from the same DNA preparation that was used to construct the cosmid library, was digested to completion with *Eco*RI and transferred to nitrocellulose (29). An *Eco*RI digest of cosmid 150, electrophoresed and transferred in an adjacent track, provided a specific marker for the 10.9-kb fragment (data not shown). The 540-bp *Bam*HI fragment was nick-translated (20) and hybridized to the blots, which were then washed at low [0.36 M NaCl/20 mM sodium phosphate pH 7.4/2 mM EDTA (2×)/0.1% $NaDodSO_4$, 65°C] and high [18 mM NaCl/1 mM sodium phosphate, pH 7.4/0.1 mM EDTA (0.1×)/0.1% $NaDodSO_4$, 65°C] stringency.

**Independent Sheep Keratin Type I and Type II Gene Clusters.** The two sheep wool type I and type II keratin gene probes detected several hundred recombinant clones from our cosmid library and of those we rescreened about 100 for close linkage of the two gene families. However, none of the cosmids hybridized to both probes, suggesting that within the sheep genome these two keratin gene families are not closely linked and are separated by at least 35–40 kb, the average cosmid insert size. As both gene probes covered the highly conserved α-helical region abutting the C-terminal domain (Fig. 1; ref. 30), we believe that each probe should have cross-hybridized with all genes of the same type under our low stringency screening conditions.

There are probably 10–20 genes in each keratin type I and type II gene family (1, 6, 7, 35), and the detection of 10–15 sheep *Eco*RI fragments with a type II gene probe (Fig. 6) agrees with the estimates for that family in other genomes (1, 6, 7, 35). Thus, from a combined total of 20–40 type I and type II genes we have detected 9 of them within four cosmid clones and, although it is not yet known whether any of these cosmids are linked together into a larger array, it is apparent that a fair proportion of the genes are arranged in clusters of at least two or three genes in the sheep genome.

**Hair (Wool) Keratin Type II Proteins May Contain Conserved C-Terminal Domains Different from Those of the Epidermis.** Of the 10–20 type II proteins typically encoded in mammalian genomes (1, 6, 7, 35), only 4 appear to be synthesized in the wool follicle (14, 15) and the number of wool keratin IF genes may be similarly small. Comparison of amino acid sequence data for the one known wool type II protein (K. Ward, personal communication) with that of several epidermal type II proteins (10, 11, 16, 31, 38, 39) reveals that the wool keratin C-terminal domain is rich in cysteine and devoid of the glycine and serine tracts found in the epidermal keratins. Accordingly, the known wool and epidermal keratins form two subsets recognizable by their C-terminal domains. Further, the amino acid sequence data for the epidermal type II keratins suggest that their C-terminal domains contain three distinct subdomains (1, 39).

One of the short type II gene probes used in this work, the 86-bp probe, was derived from a wool keratin IF type II cDNA clone and covered about half of the C-terminal domain (Fig. 1*B*). The corresponding region in the epidermal type II genes is a variable protein subdomain (1, 39) and if the same subdomain organization exists for both, then the 86-bp probe should be gene specific. However, the 86-bp probe hybridized to two type II IF genes in separate cosmids (Figs. 3 and 4). Since one of those genes also was detected by the 3′

noncoding probe, it is likely that one gene is equivalent to the wool keratin IF type II cDNA clone and the other is a second wool keratin type II gene. Further, it is possible that all four wool keratin IF type II proteins possess these conserved C-terminal domain sequences, which are different from those characteristic of epidermal keratins, reflecting their different functions in the hair and epidermis.

We thank Dr. K. Ward (Commonwealth Scientific and Industrial Research Organisation, Division of Animal Production, Prospect, N.S.W., Australia) for the gift of the sheep keratin type I and type II cDNA clones and communication of his nucleotide sequences, Lesley Crocker and Imbi Semenov for excellent technical assistance, and Leanne Goodwin for typing the manuscript. Recombinant DNA work was undertaken at the appropriate level of containment in accordance with the guidelines of the Australian Scientific Committee on Recombinant DNA. This work was supported by a grant from the Wool Research Trust Fund to G.E.R. and B.C.P. on the recommendation of the Australian Wool Corporation.

1. Steinert, P. M., Steven, A. C. & Roop, D. R. (1985) *Cell* **42,** 411–419.
2. Parry, D. A. D., Steven, A. C. & Steinert, P. M. (1985) *Biochem. Biophys. Res. Commun.* **127,** 1012–1018.
3. Woods, E. F. & Inglis, A. S. (1984) *Int. J. Biol. Macromol.* **6,** 277–283.
4. Crewther, W. G., Dowling, L. M., Steinert, P. M. & Parry, D. A. D. (1983) *Int. J. Biol. Macromol.* **5,** 267–274.
5. Lazarides, E. (1982) *Annu. Rev. Biochem.* **51,** 219–250.
6. Moll, R., Franke, W. W., Schiller, D., Geiger, B. & Krepler, R. (1982) *Cell* **31,** 11–24.
7. Fuchs, E. V., Coppock, S. M., Green, H. & Cleveland, D. W. (1981) *Cell* **27,** 75–84.
8. Eichner, R., Bonitz, P. & Sun, T.-T. (1984) *J. Cell Biol.* **98,** 1388–1396.
9. Marchuk, D., McCrohon, S. & Fuchs, E. V. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 1609–1613.
10. Johnson, L. D., Idler, W. W., Zhou, X.-M., Roop, D. R. & Steinert, P. M. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 1896–1900.
11. Tyner, A. L., Eichman, M. J. & Fuchs, E. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 4683–4687.
12. Krieg, T. M., Schafer, M. P., Cheng, C. K., Filpula, D., Flaherty, P., Steinert, P. M. & Roop, D. R. (1985) *J. Biol. Chem.* **260,** 5867–5870.
13. Lehnert, M. E., Jorcano, J. L., Zentgraf, H., Blessing, M., Franz, J. K. & Franke, W. W. (1984) *EMBO J.* **3,** 3279–3287.
14. Crewther, W. G., Dowling, L. M., Gough, K. H., Marshall, R. C. & Sparrow, L. G. (1980) in *Fibrous Proteins—Scientific, Industrial and Medical Aspects*, eds. Parry, D. A. D. & Creamer, L. K. (Academic, London), Vol. 2, pp. 151–159.
15. Powell, B. C. & Rogers, G. E. (1985) in *Biology of the Integument*, eds. Bereiter-Hahn, J., Matoltsy, A. G. & Richards, K. S. (Springer, New York), Vol. 2, pp. 695–721.
16. Hanukoglu, I. & Fuchs, E. (1983) *Cell* **33,** 915–924.
17. Seed, B., Parker, R. C. & Davidson, N. (1982) *Gene* **19,** 201–209.
18. Hohn, B. & Collins, J. (1980) *Gene* **11,** 291–298.
19. Hanahan, D. & Meselson, M. (1980) *Gene* **10,** 63–67.
20. Roop, D. R., Tsai, M.-J. & O'Malley, B. W. (1980) *Cell* **19,** 63–68.
21. Grunstein, M. & Hogness, D. S. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 3961–3965.
22. Birnboim, H. C. & Doly, J. (1979) *Nucleic Acids Res.* **7,** 1513–1523.
23. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–560.
24. Crouse, G. F., Frischauf, A. M. & Lehrach, H. (1983) *Methods Enzymol.* **101,** 78–89.
25. Messing, J. & Vieira, J. (1982) *Gene* **19,** 269–276.
26. Norrander, J., Kempe, T. & Messing, J. (1983) *Gene* **26,** 101–106.
27. Hu, N.-T. & Messing, J. (1982) *Gene* **17,** 271–277.
28. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
29. Smith, G. E. & Summers, M. D. (1980) *Anal. Biochem.* **109,** 123–129.
30. Geisler, M., Kaufmann, E., Fischer, S., Plessman, U. & Weber, K. (1983) *EMBO J.* **2,** 1295–1302.
31. Steinert, P. M., Parry, D. A. D., Racoosin, E. L., Idler, W. W., Steven, A. C., Trus, B. L. & Roop, D. R. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 5709–5713.
32. Mount, S. M. (1982) *Nucleic Acids Res.* **10,** 459–472.
33. IUPAC-IUB Commission on Biochemical Nomenclature (1968) *Eur. J. Biochem.* **5,** 151–153.
34. Quax, W., Egberts, W. V., Hendriks, W., Quax-Jeuken, Y. & Bloemendahl, H. (1983) *Cell* **35,** 215–223.
35. Fuchs, E. & Marchuk, D. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 5857–5861.
36. Childs, G., Maxson, R., Cohn, R. H. & Kedes, L. (1981) *Cell* **23,** 651–663.
37. Maxson, R., Cohn, R., Kedes, L. & Mohun, T. (1983) *Annu. Rev. Genetics* **17,** 239–277.
38. Jorcano, J. L., Franz, J. K. & Franke, W. W. (1984) *Differentiation* **28,** 155–163.
39. Steinert, P. M., Parry, D. A. D., Idler, W. W., Johnson, L. D., Steven, A. C. & Roop, D. R. (1985) *J. Biol. Chem.* **260,** 7142–7149.
40. Ward, K. A., Sleigh, M. J., Powell, B. C. & Rogers, G. E. (1982) *Proceedings of the Second World Congress on Genetics Applied to Livestock Production* (Editorial Garsi, Madrid, Spain), Vol. 6, pp. 146–156.