MOLECULAR AND CELLULAR BIOLOGY, May 1988, p. 2195–2203
0270-7306/88/052195-09$02.00/0

Vol. 8, No. 5

Isolation, Characterization, and UV-Stimulated Expression of Two Families of Genes Encoding Polypeptides of Related Structure in Human Epidermal Keratinocytes

TONJA KARTASOVA†* AND PIETER VAN DE PUTTE

Laboratory of Molecular Genetics, State University of Leiden, Wassenaarseweg 64, 2333 AA Leiden, The Netherlands

Received 16 September 1987/Accepted 12 January 1988

By screening of a cDNA library made on mRNA isolated from UV-irradiated human epidermal keratinocytes for sequences whose relative concentration increases in the cytoplasm after irradiation, we have isolated 40 cDNA clones (T. Kartasova, B. J. C. Cornelissen, P. Belt, and P. van de Putte, Nucleic Acids Res. 15:5945–5962, 1987). Here we describe two distinct groups of cDNA clones which do not cross-hybridize to each other but nevertheless encode proteins of very similar primary structure. These polypeptides are small (8 to 10 kilodaltons) and exceptionally rich in proline, cysteine, and glutamine and have similar repeating elements not found elsewhere. The new proteins were designated sprI and sprII (small, proline rich). The presence of prolines and cysteines suggests that they may be either structural proteins with a strong secondary structure or metal-binding proteins such as metallothioneins. Southern blot and sequence analyses of the cDNAs indicate that at least the sprII group of clones represents a family of related genes. The nucleotide sequence of both groups seems to be conserved upon evolution. The level of mRNAs corresponding to the two groups of cDNAs is increased in the cytoplasm of human epidermal keratinocytes after both UV irradiation and treatment with 4-nitroquinoline 1-oxide or 12-*O*-tetradecanoylphorbol 13-acetate.

UV light is well known for its carcinogenic effect on human skin. Photoproducts which appear in DNA after UV exposure may be mutagenic and can lead to skin malignancy via mutation induction. To gain a better understanding of the interference of UV light with normal cellular functions and of the cellular response to a challenge with UV light, we have initiated a study of the UV-inducible regulation of gene expression in cultured human epidermal keratinocytes by cloning genes which are induced upon UV irradiation. Keratinocytes derived from the skin epidermis are functionally specialized to protect the organism against the damaging effects of different external agents, including UV light. Therefore, they represent an attractive system to study the influence of UV irradiation on gene expression.

A cDNA library was constructed from poly(A)$^+$ RNA isolated after UV irradiation. This library was differentially screened with cDNA probes synthesized on poly(A)$^+$ RNA isolated from either UV-irradiated or nonirradiated cells (9). Here we describe two groups of clones of which the deduced proteins show a striking homology at the level of both their amino acid composition and structural organization. For both groups a relative increase in the cytoplasmic mRNA level was detected on Northern (RNA) blots after UV irradiation and treatment with either 4-nitroquinoline 1-oxide (4-NQO) or the tumor promoter 12-*O*-tetradecanoylphorbol 13-acetate (TPA).

MATERIALS AND METHODS

Cell culture and treatment with UV, TPA, and 4-NQO. A primary culture of epidermal keratinocytes was established from human foreskin by the method of Rheinwald and Green (21) with some modifications described by Ponec et al. (20). Keratinocytes were grown in the presence of lethally irradiated 3T3 cells used as a feeder layer. The culture medium was a 3:1 mixture of Dulbecco-Vogt modified Eagle medium and Ham F12 medium supplemented with 5% fetal calf serum, 0.4 μg of hydrocortisone per ml, 10^{-10} M cholera toxin, and 10 ng of epidermal growth factor per ml. For UV irradiation shortwave (254-nm) UV light was used at 1 J/m^2 per s. Dose rates were determined with a UVX radiometer (Ultraviolet Products, Inc., San Gabriel, Calif.). Before irradiation cells were washed twice with prewarmed phosphate-buffered saline; then the phosphate-buffered saline solution was removed, and the cells were irradiated with the doses indicated below. TPA treatment was for 12 h at a concentration of 20 ng/ml. 4-NQO was used at a concentration of 0.5 mM. Cell incubation with 4-NQO was for 4 h at 37°C; then the cells were washed, and fresh medium was added. RNA was isolated 12 h later.

RNA isolation and Northern blot analysis. Total cytoplasmic RNA was isolated, and poly(A)$^+$ RNA was purified on poly(U)-Sepharose as described previously (9). The amount of poly(A)$^+$ RNA for each RNA preparation was quantitated by hybridization to ^{3}H-labeled poly(rU) (4). Equal amounts of poly(A)$^+$ RNA, ranging from 0.1 to 0.5 μg per slot for each experiment, were treated with glyoxal, electrophoresed in 1.5% agarose gel, blotted onto Gene-Screen filters (New England Nuclear Corp., Boston, Mass.), and hybridized to radiolabeled cDNA probes (22) as described by Sarachu et al. (24).

Primer extension. A 20-nucleotide-long primer was chemically synthesized, corresponding to nucleotides 35 through 54 of clone 128. The primer was ^{32}P labeled at its 5′ end by the kinase reaction described by Maniatis et al. (15). Eight micrograms of poly(A)$^+$ RNA isolated from TPA-treated keratinocytes was used as a template for cDNA synthesis (15) in the presence of 7.5 ng of primer. The product of the reverse transcriptase reaction was purified by electrophoresis on a 10% polyacrylamide–8 M urea gel. The radioactive

* Corresponding author.
† Present address: Building 37, Room 3B25, Laboratory of Cellular Carcinogenesis and Tumor Promotion, National Cancer Institute, Bethesda, MD 20892.

bands were localized by autoradiography and eluted from the gel as described by Maniatis et al. (15).

Source of DNA and Southern blot analysis. High-molecular-weight DNA was isolated from cultured human epidermal keratinocytes and a rat kangaroo cell line, Pot3 (generous gift of J. de Wit, Department of Cell Biology, University of Rotterdam), by the method of Maniatis et al. (15). DNA preparations isolated from hamster and mouse cell lines and from a monkey kidney cell line (CV-1) were generous gifts of E. Zwarthoff (Department of Pathology, University of Rotterdam) and J. van der Lubbe (Sylvius Laboratory, University of Leiden), respectively. Equal amounts of DNA (10 μg per slot) were electrophoresed in 0.8% agarose gel. Southern blot and hybridization were performed by the method of Southern (25).

Sequence analysis. cDNA inserts were cloned in M13 phage vectors and sequenced by the dideoxy-chain termination method of Sanger et al. (23). The method of Maxam and Gilbert was used for the sequencing of the cDNA product after primer extension of mRNA with reverse transcriptase (16).

RESULTS

A library of 180,000 cDNA clones was prepared by the method of Okayama and Berg (19) by using a mammalian expression vector, a derivative of pSV2-glo (17). Poly(A)$^+$ RNA isolated from UV-irradiated human keratinocytes was used as a template for the cDNA synthesis. The cDNA library was differentially screened with cDNA probes synthesized on poly(A)$^+$ RNA isolated from UV-irradiated and nonirradiated keratinocytes (9). Among the 40 cDNA clones isolated and characterized, two groups of clones were selected because of the remarkable primary structure of the proteins deduced from their nucleotide sequences.

Primary structure and deduced amino acid sequence of group 1 cDNA inserts. The first group of clones (group 1) consisted of five related cDNAs, notably 128, 12F, 15B, 41D, and 109, which cross-hybridized to one other. Clone 128 contained the longest insert (622 nucleotides; Fig. 1A) and hybridized on a Northern blot to mRNA of a similar length (600 to 700 nucleotides), suggesting that the cDNA insert of clone 128 represents a full-length cDNA. To confirm this, a primer extension experiment was performed on cytoplasmic poly(A)$^+$ RNA isolated from keratinocytes. The sequence of the 20-nucleotide-long primer was complementary to nucleotides 35 to 54 from clone 128 (Fig. 1A). The products of the extension reaction were separated on a polyacrylamide gel. Two specific bands migrating close to one other were detected, and DNA purified from these bands was sequenced (data not shown). The primary structure of the slower-migrating DNA appeared to be identical to the 5′-end sequence of the insert of clone 128. The sequence of the faster-migrating band was four nucleotides shorter. Thus, clone 128 indeed represents a full-length DNA copy of the corresponding mRNA.

The first ATG at positions 69 through 71 started an open reading frame of 89 amino acids which ended with a TAA stop codon at positions 336 through 338. The sequence AGCATGA around the ATG codon was in good agreement with the consensus sequence PuNNATGPu determined by Kozak (11), which favors the optimal translation initiation on eucaryotic mRNAs. A possible polyadenylation signal (26) was 14 nucleotides upstream from the poly(A) tail. However, the sequence CAYTG, which is found adjacent to the polyadenylation site in many eucaryotic mRNAs (3), was not found in the 3′ part of the cDNA insert of clone 128.

The protein encoded by the open reading frame of clone 128 is small (M_r 9,888), basic, and hydrophilic and has an exceptionally high content of proline (29%), glutamine (18%), and cysteine (9%) (Fig. 1C). Comparison of the codon usage of the clone 128 protein with the codon usage of other eucaryotic proteins (5) indicates that the clone 128 codons in general belong to the frequently used ones (Table 1). Eight cysteines were distributed along the amino acid sequence in the following order: $N_9CN_{14}CN_7CN_7CN_7CN_7CN_7CN_7CN_{16}$. In the amino acid sequence of the clone 128 protein, two kinds of repeats could be identified. The sequence QQQKQPCIPPPQ (amino acids 4 through 15) was repeated with only one insertion and one substitution at positions 18 through 30, and the octamer PKVPEPCH was repeated tandemly with some variations six times in the clone 128 protein (Fig. 1C).

The cDNA insert of clone 12F was identical to that of clone 128, except that the cDNA insert of clone 12F was four nucleotides shorter. The differences in nucleotide and in deduced amino acid sequences between clones 128 and 12F on the one hand and clones 15B and 41D on the other hand are indicated in Fig. 1A and B, respectively. Clone 109 contained a cDNA insert of only 333 nucleotides, and its nucleotide sequence was identical to the 3′-end sequence of clones 128 and 12F.

Primary structure and deduced amino acid sequence of group 2 cDNA inserts. Initially, only one group 2 cDNA clone, clone 1740, was isolated by differential screening. Eight additional clones were obtained after rescreening of the library with the cDNA insert of clone 1740. Based on restriction enzyme and nucleotide sequence analyses, these

TABLE 1. Codon usage in clone 128 and 930 proteins

Amino acid	Codon	No. of codons used		Avg %[a]	Amino acid	Codon	No. of codons used		Avg %
		128	930				128	930	
Glu	GAG	5	2	64.5	Cys	TGT	0	4	38.3
Glu	GAA	1	0	35.5	Cys	TGC	8	8	61.7
Val	GTG	4	2	51.2	Tyr	TAT	0	1	37.0
Val	GTA	0	0	7.8	Tyr	TAC	0	0	63.0
Val	GTT	1	0	13.1					
Val	GTC	1	0	27.9	Leu	TTG	0	0	12.0
					Leu	TTA	0	0	4.9
Ala	GCG	0	0	10.1					
Ala	GCA	1	0	17.6	Ser	TCG	0	0	6.1
Ala	GCT	0	0	26.5	Ser	TCA	1	0	10.7
Ala	GCC	1	0	45.8	Ser	TCT	0	1	17.0
					Ser	TCC	1	2	26.0
Ser	AGT	1	0	9.9					
Ser	AGC	0	1	30.2	Gln	CAG	16	10	75.4
					Gln	CAA	0	0	24.6
Lys	AAG	8	7	64.0					
Lys	AAA	3	1	36.0	His	CAT	0	0	36.2
					His	CAC	3	0	63.8
Met	ATG	1	1	100.0					
					Leu	CTG	0	0	46.8
Ile	ATA	1	0	11.7	Leu	CTA	0	0	5.5
Ile	ATT	0	0	28.3	Leu	CTT	1	0	8.9
Ile	ATC	2	0	60.0	Leu	CTC	0	0	21.9
Thr	ACG	0	1	11.3	Pro	CCG	0	1	12.6
Thr	ACA	0	1	22.1	Pro	CCA	7	13	19.3
Thr	ACT	1	0	20.7	Pro	CCT	8	7	28.3
Thr	ACC	2	0	45.8	Pro	CCC	11	7	39.1

[a] Percentage of codon usage in human proteins (5).

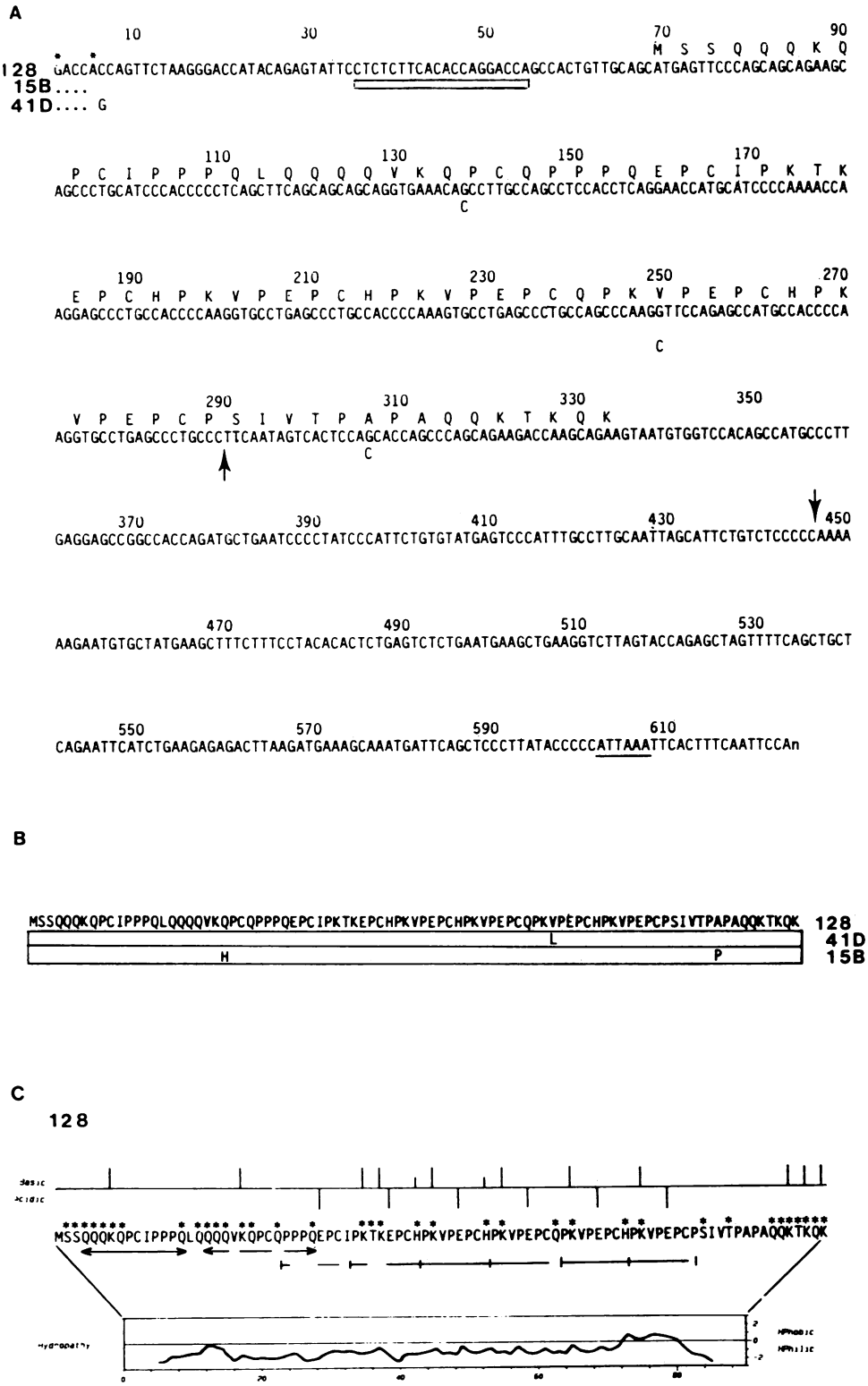

FIG. 1. (A) Nucleotide sequences of clones 128, 15B, 41D, and 109. Symbols: *, starts of mRNAs as determined by primer extension; ▭, sequence used as a primer; ↑, start of clone 109; ↓, end of clone 41D. The poly(A) signal is underlined. (B) Amino acid sequences of clones 128 (12F), 41D, and 15B. (C) Amino acid sequence of clone 128. Symbols: ⟷ and ⟼, repeats; *, amino acids that can interact with nucleic acids. The hydropathy plot was done by the method of Kyte and Doolittle (12).

cDNA clones could be divided into three subgroups: the first subgroup was represented by clones 930 and 1137, which had an internal *Eco*RI site; the second group was represented by clones 174, 1480, and 375, which had an internal *Pvu*II site; the third group included clones 1292, 60, 1546, and 1740, which had both internal *Eco*RI and *Pvu*II sites. Each subgroup had 90% homology with the two other subgroups on the DNA level. In the coding region the homology reached 94 to 98%.

The sequence of the cDNA insert of clone 930 of the first subgroup was 682 nucleotides long (Fig. 2A). On a Northern blot, mRNA of 700 to 800 nucleotides in length was detected after hybridization with this clone. Thus, the cDNA insert of clone 930 might represent a full-length copy of the corresponding mRNA. The first ATG was at positions 65 through 67 and was preceded by two in-frame TGA stop codons at positions 44 through 46 and 48 through 50, supporting the notion that this ATG is the start of the putative encoded protein. The sequence AGGATGT around the initiation codon had quite a good homology to the consensus PuNNATGPu determined by Kozak (11). The open reading frame of 72 amino acids ended at positions 280 through 282 at a TAA stop codon. There was no other significant open reading frame encoded by this sequence (Fig. 3). A possible polyadenylation signal (26) was 15 nucleotides upstream from the poly(A) tail. The sequence CAYTG (3) was not found in the 3′ part of the cDNA insert of clone 930.

The protein deduced from the nucleotide sequence of clone 930 was small (M_r, 7,830), basic, and hydrophilic (Fig. 2C) and was extremely rich in proline (39%), glutamine (17%), and cysteine (17%). Comparison of the codon usage of the clone 930 protein with the codon usage of other eucaryotic proteins (5) indicates that the clone 930 codons in general belong to the frequently used ones (Table 1). Twelve cysteines were distributed along the amino acid sequence in the following order: $N_7CN_3CN_5CN_4CN_3CN_4CN_3CN_4CN_8CN_3CN_8CN_3CN_5$. Furthermore, in the predicted protein sequence a nonamer, PKCPE(Q)PCPP, was repeated tandemly with some alterations four times.

The differences in nucleotide sequence and in deduced amino acid sequence between three subgroups of the cDNA clones and within the second subgroup (clones 174N, 1480, and 375) are indicated in Fig. 2A and B, respectively.

Comparison of the two groups of cDNA clones. When the nucleotide sequences of clone 128 (group 1) and clone 930 (group 2) were compared by computer analyses with UWGCG programs, a homology of 49% was found in the coding region (Fig. 4). No significant homology was found outside of this region.

The proteins deduced from the nucleotide sequence of group 1 and group 2 cDNA clones appeared to be very similar in their amino acid composition (Table 2). The predominant amino acids in both sequences were proline, glutamine, cysteine, and lysine, which together comprised 68.5 and 83.3% of the deduced protein sequences of clones 128 (group 1) and 930 (group 2), respectively. All other amino acids were represented poorly or not at all, so that the clone 128 protein consisted of only 13 different amino acids and the clone 930 protein consisted of only 10 (Table 2).

The most prominent feature of group 1 and group 2 cDNA clones was the presence of a striking similarity in the primary structures of their predicted polypeptides: they contained repeating elements (Fig. 1C and 2C). Moreover, when the two amino acid sequences were aligned, it was evident that the first 16 amino acids of the group 2 protein(s) were conserved in the N-terminal region of the group 1 protein(s) (Fig. 5). In addition, the sequence PK*PEPC was present both in the octamer PKVPEPCH, found six times in group 1, and in the nonamer PKCPE(Q)PCPP, occurring four times in group 2. Some homology was also found beyond the repeating elements at the C terminus of group 1 and group 2 protein(s) (Fig. 5).

The deduced group 1 and group 2 amino acid sequence as well as the sequence PK*PEPC were compared with the amino acid sequences of the NBRF bank by computer analysis with UWGCG programs. No significant homology was found, except for a homology of the PK*PEPC sequence with the N-terminal region of the chicken cardiac myosin L1 catalytic chain (192 amino acids): PKKPEPKKAPEPKKEEPKPAPKPAEPEPKK, which has a high content of proline, lysine, and glutamic acid (13).

Effects of UV, 4-NQO, and TPA. Figure 6 shows a Northern blot analysis of a gel run with equal amounts of cytoplasmic poly(A)$^+$ RNA from UV-irradiated and nonirradiated cells. Group 1 mRNAs were more abundant in nontreated cells than group 2 mRNAs, since a clear hybridization signal was obtained with the probe of group 1 (clone 15B) on Northern blots loaded with 0.1 to 1.0 μg of poly(A)$^+$ RNA. For Northern blot hybridizations, probes were used which cross-hybridized with other members of the same group of cDNA clones. Hybridizations were carried out under stringent conditions. On Northern blots each probe hybridized to a single band, indicating that mRNAs within each group are approximately the same size.

A clear relative increase in the cytoplasmic level of mRNA corresponding to both groups of cDNA clones was detected on Northern blots 12 h after UV irradiation with 50 J/m^2 (Fig. 6, lane a) or 6 h after irradiation with 35 J/m^2 (lane d). The level of mRNA corresponding to the elongation factor 1-alpha (Fig. 6) as well as of mRNA corresponding to the gamma-actin (data not shown) decreased after this treatment. The induction ranged between 3 to 5 and 5 to 10 times for group 1 and group 2 cDNA clones, respectively. Treatment of keratinocytes with a potent carcinogen, 4-NQO, which mimics the effects of UV light on living cells (7), led to a 5- to 10-fold increase in the level of the group 2 mRNA and to a somewhat smaller increase in the level of group 1 mRNA (lane c). Incubation of keratinocytes with the tumor promoter TPA affected the cytoplasmic level of the mRNAs for

TABLE 2. Amino acid composition of clone 128 and 930 proteins[a]

Amino acid	No. of amino acid in protein from clone:	
	128	930
Pro	26	28
Gln	16	12
Cys	8	12
Lys	11	8
Glu	6	2
Val	6	2
Ser	3	4
Thr	3	2
Met	1	1
Ile	3	
His	3	
Ala	2	
Leu	1	
Tyr		1

[a] The clone 128 protein contained 13 different amino acids; the clone 930 protein contained 10 different amino acids.

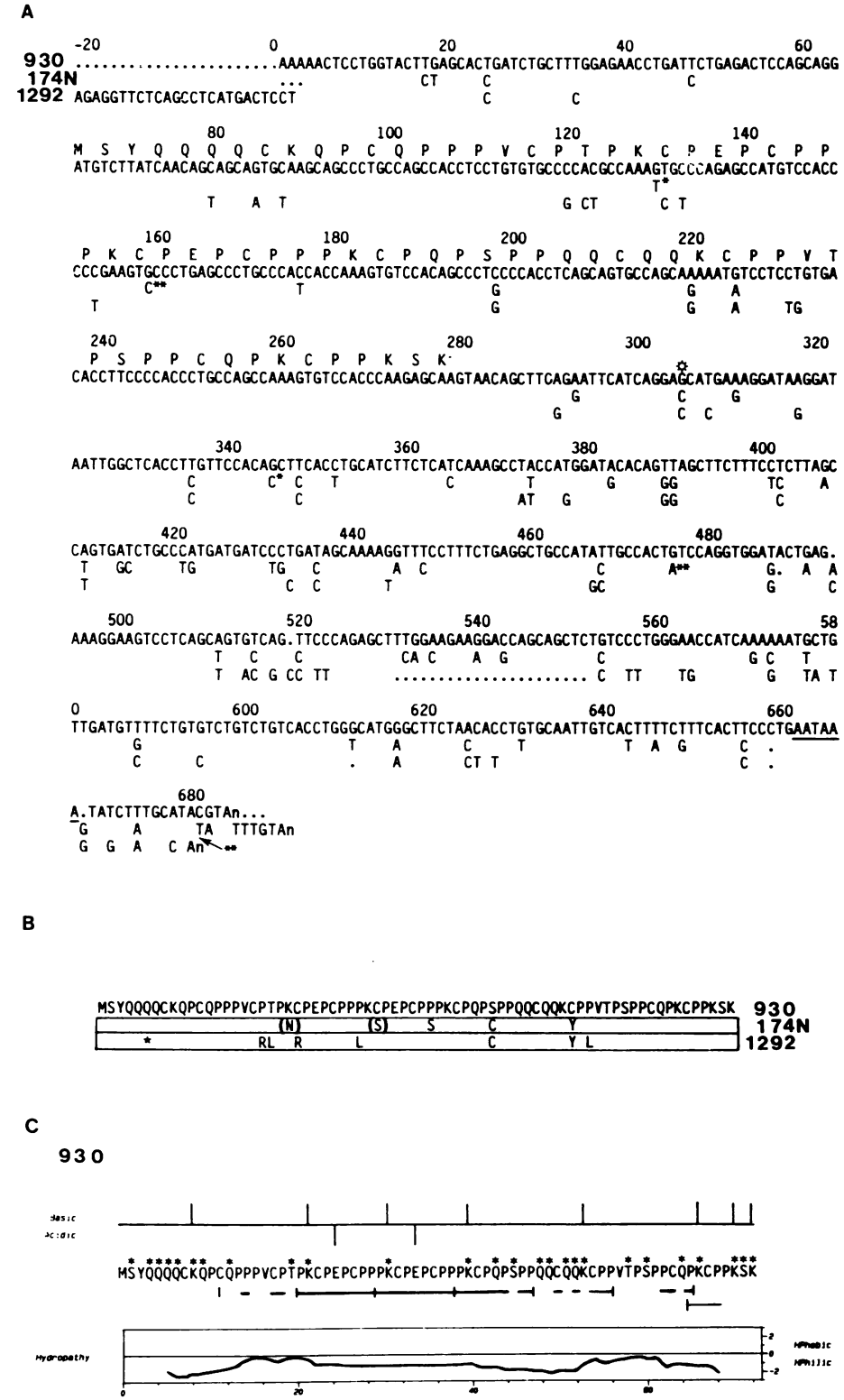

FIG. 2. (A) Nucleotide sequences of clones 930 (1137), 174N (375, 1480), and 1291 (60, 1548, 1740). Symbols: *, base substitutions in clone 1480; **, base substitutions in clone 375; ⟷, substitution in 1480 and 375. The 3′ end of clone 375 is indicated by an arrow. The poly(A) signal is underlined. (B) Amino acid sequence of clone 930 (1137) compared with those of clones 174N (375, 1480) and 1292 (60, 1548, 1740). (N) in clone 1480 and (S) in clone 375. Symbol: *, nonsense codon in clone 1292. (C) Amino acid sequence of clone 930. Symbols: ⟷, repeats; *, amino acids that can interact with nucleic acids. The hydropathy plot was done by the method of Kyte and DooLittle (12).

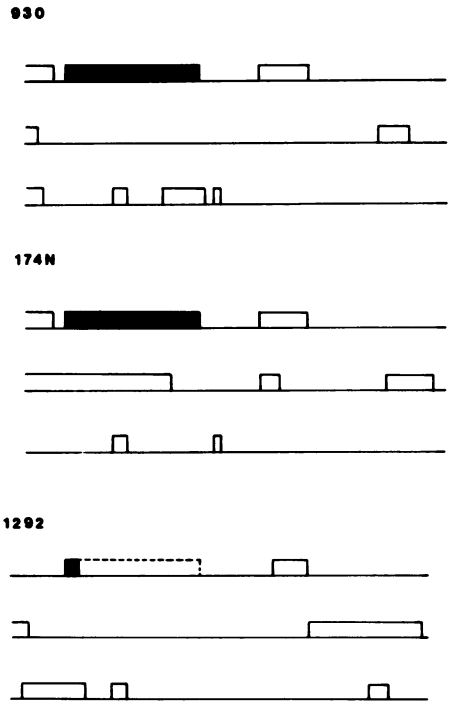

FIG. 3. Possible open reading frames in three subgroups of group 2 cDNA clones. Symbol: ▬, open reading frame of 72 amino acids.

both groups (lane e). The relative increase in the mRNA level in the cytoplasm after treatment with TPA was between 3 to 5 and 10 to 20 times for group 1 and group 2 cDNA clones, respectively.

Southern blot analysis. The genomic organization of the genes corresponding to group 1 and 2 cDNA clones was studied by using Southern blot analysis (Fig. 7 and 8). After digestion with restriction enzymes *Kpn*I (Fig. 7A, lane g), *Bgl*II, and *Bam*HI (data not shown), which do not cut within the group 1 cDNA probe, only one band was detected on Southern blots. Further restriction analysis by *Eco*RI, *Hin*dIII, and *Pst*I enzymes and by *Eco*RI-*Pst*I double digestions indicated the presence of at least one intervening sequence within the group I genomic sequence (Fig. 7A, lanes a through f). Weak cross-hybridizing bands were also detected. The group 2 cDNA probe was hybridized to genomic DNA cut with *Pst*I, *Hin*dIII, and *Bam*HI (Fig. 8A, lanes a through c). These restriction enzymes did not cut within the group 2 cDNA inserts. On Southern blot numerous bands appeared when the genomic DNA was cut with *Pst*I and *Hin*dIII. The bands generated by *Bam*HI digestion were concentrated in the region of 15 to 25 kilobases. This made it difficult to estimate the exact gene copy number. The analysis with *Pst*I and *Hin*dIII indicates the presence of several copies of group 2-like genes in the human genome and/or the presence of intervening sequences within the coding region.

We have examined the genomes of other eucaryotic species for the presence of sequences similar to group 1 and group 2 genes. Monkey, mouse, hamster, and rat kangaroo DNAs were digested with *Eco*RI and run on agarose gels. After hybridization to a group 1 cDNA probe, two bands could be detected on Southern blots with monkey, mouse, and hamster DNAs and three bands were detected after hybridization to rat kangaroo DNA (Fig. 7B, lanes a through d). After hybridization to a group 2 cDNA probe, several bands appeared on a Southern blot with monkey and mouse DNAs and at least two bands appeared with rat kangaroo DNA (Fig. 8A, lanes d through f). The hybridization of group 1 and 2 cDNA probes to the rat kangaroo DNA was much weaker than to the DNA from other higher mammals. This might be the consequence of a high sequence divergency between human and kangaroo group 1 and group 2 genes.

DISCUSSION

Two groups of cDNA clones, group 1 and 2, were isolated by screening of a cDNA library for sequences whose levels increase in the cytoplasm of human epidermal keratinocytes after UV irradiation. Sequence analysis of these clones demonstrates that although the homology at the DNA level is limited, the primary structures of the deduced proteins are very similar. Analysis of the amino acid sequence predicted from the open reading frames of group 1 and group 2 cDNA clones revealed unusual amino acid composition, repeating elements, and overall homology in the structures of these proteins. We term the group 1 and group 2 deduced proteins sprI and sprII (small, proline rich).

Comparison of the codon usage of the sprI and sprII proteins with the codon usage of other eucaryotic proteins suggests that the translation of the sprI and sprII mRNAs is not hampered by an unusual codon usage. The high proline content largely excluded alpha-helical or beta-sheet conformations in the secondary structure of the proteins. Only in the 23 N-terminal and 8 C-terminal amino acids of the sprI protein(s) and in the 10 N-terminal amino acids of the sprII protein(s) some alpha-helix or beta-sheet configuration was found (6). The repeating unit of the sprI protein(s) contains one cysteine per unit, and the repeating unit of the sprII protein(s) harbors two. The cysteines can play a role in intra- and intermolecular formation of S-S bonds and protein dimerization or polymerization or form potential metal-binding sequences (2). The high proportion of amino acids such as glutamine, lysine, serine, and threonine, which were shown to mediate the DNA-protein interaction in repressor proteins (18), indicates that the sprI and II proteins might be

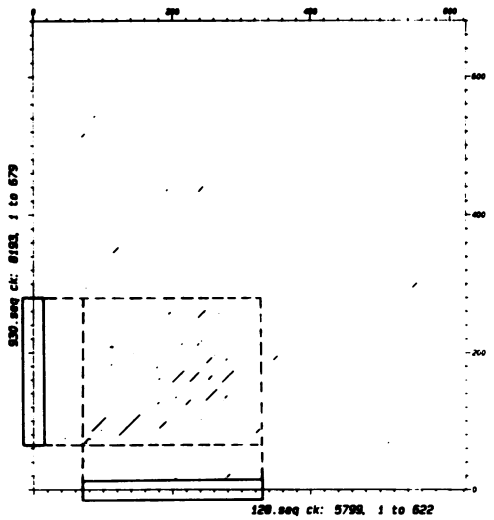

FIG. 4. Comparison of the nucleotide sequence of clones 128 (group 1) and 930 (group 2). Comparison was done by the method of Maizel and Lenk (14), with window 21 and stringency 14, Symbol: ▭, open reading frames.

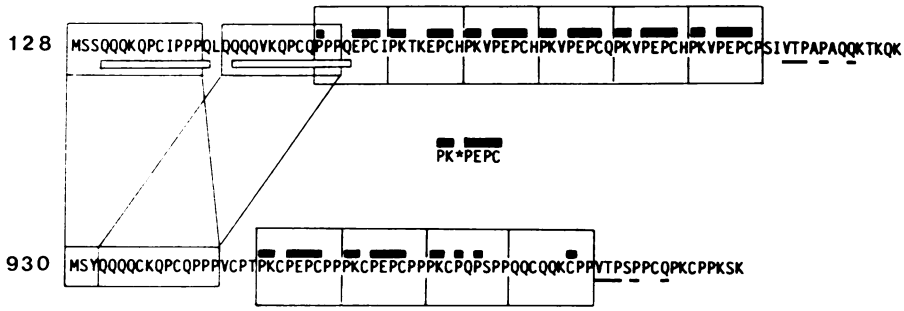

FIG. 5. Comparison of the amino acid sequences of clone 128 and 930 proteins. Symbols: ▭, repeats; ■ ■, homology with the sequence PK*PEPC; – – –, homology in the C terminus.

active in binding nucleic acids. On the other hand, the homology of the N-terminal region of myosin L1 chain (13) suggests a structural function for these proteins.

Northern blot analysis showed that the expression of group 1 and 2 genes is stimulated by UV light, 4-NQO, and TPA. It would be interesting to find out whether the expression of these genes is also affected by other stress treatments like heat shock or anoxia.

The exact number of genes corresponding to group 1 cDNA clones is difficult to estimate and needs further investigation. Preliminary analysis of genomic blots indicates the presence of at least one gene copy of approximately 12 kilobases in length, although a tandem organization of several copies cannot be excluded. A weak hybridization to multiple bands on Southern blots can be due to the presence of group 1 related sequences in the human genome or to the nonspecific hybridization to G+C-rich sequences, since the cDNA insert used for nick translation contained the G+C tail. The detection on Southern blots of fragments that were smaller than the DNA insert itself suggests the presence of at least one intervening sequence in the group 1 gene(s). Sequence variation between clones 128 and 12F on the one hand and clones 41D and 15B on the other hand (nucleotides 137, 249, and 306 in Fig. 1) can be explained by the expression of three slightly different alleles or might be the consequence of the introduction of wrong bases into the cDNA sequence during the cDNA synthesis. Group 2 cDNA clones represent a multigene family of highly homologous genes. The stop codon at position 79 in clones 1292, 60, 1548, and 1740 (Fig. 2A) is probably a nonsense mutation in one of the genes of this family which is still expressed at a relatively high level as compared with that of other group 2 sequences, since four independent cDNA clones carrying this mutation were isolated. Preliminary results on the presence of group 1 and 2 genes in different

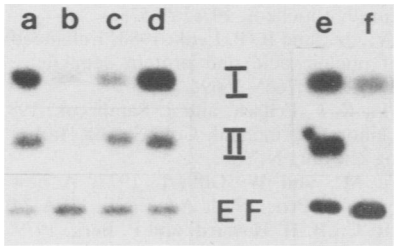

FIG. 6. Northern blot analysis of mRNA corresponding to group 1 and 2 cDNA clones. Equal amounts of poly(A)$^+$ RNA (0.1 μg), determined by hybridization to ^{3}H-labeled poly(U), were separated on 1.5% agarose gel after denaturation with glyoxal. Lanes: a and d, poly(A)$^+$ RNA from keratinocytes 12 h after irradiation with 50 J/m^2 and 6 h after irradiation with 35 J/m^2, respectively; b and f, poly(A)$^+$ RNA from nontreated cells; c and e, poly(A)$^+$ RNA from cells treated with 4-NQO and TPA, respectively. I, Group 1 (clone 15B); II, group 2 (clone 1740); EF, elongation factor 1-alpha (control). Only the parts of the gel containing the bands of RNA recognized by each particular probe are shown.

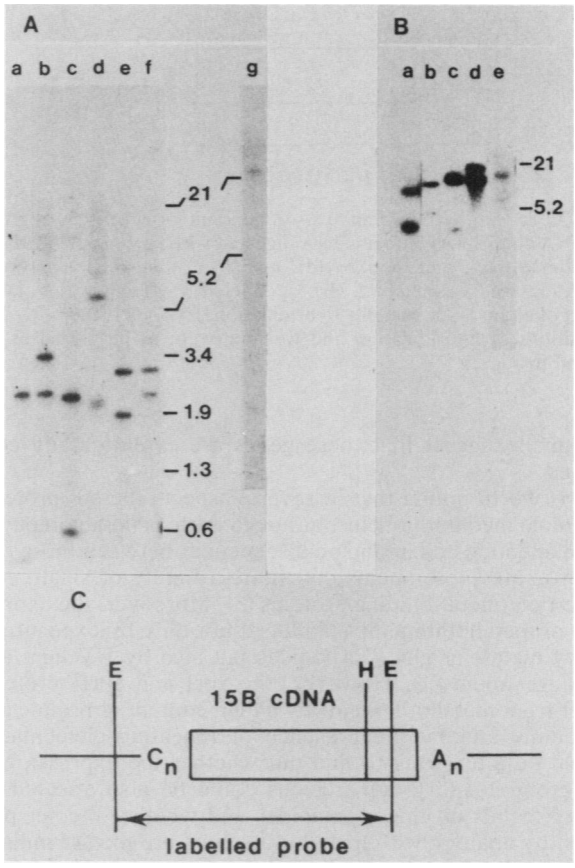

FIG. 7. Southern blot analysis of genomic organization of group 1 cDNA clones. (A) Human DNA digested with *Pst*I-*Eco*RI (a), *Pst*I (b), *Hin*dIII-*Pst*I (c), *Hin*dIII (d), *Eco*RI-*Hin*dIII (e), *Eco*RI (f), and *Bam*HI (g). (B) Monkey (a), mouse (b), hamster (c), rat (d), and rat kangaroo (e) DNA digested with *Eco*RI. (C) ^{32}P-labeled fragment of the cDNA insert of clone 15B used for hybridization. E, *Eco*RI; H, *Hin*dIII. Hybridization was performed under stringent conditions (15) at 42°C in 50% formamide and 5× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate); blots were washed at 60°C in 2× SSC and finally in 0.1× SSC.

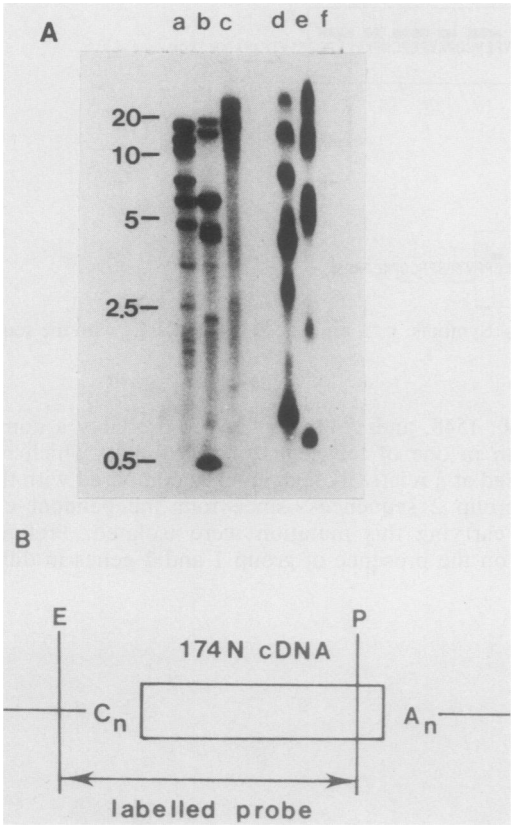

FIG. 8. Southern blot analysis of genomic organization of group 2 cDNA clones. (A) Human DNA digested with *Pst*I (a), *Hin*dIII (b), and *Bam*HI (c), and monkey (d), mouse (e), and rat kangaroo (f) DNA digested with *Eco*RI. (B) ^{32}P-labeled fragment of the cDNA insert of clone 174N used for hybridization. E, *Eco*RI; P, *Pvu*II. The conditions of hybridization and washing were as described in the legend to Fig. 7.

mammals suggest that these genes are evolutionarily conserved.

It is worth noting that in several aspects the spr proteins resemble metallothioneins, although there is no homology in the nucleotide and amino acid sequences between these two groups. Metallothioneins constitute a family of small, cysteine-rich, metal-binding proteins (8). Moreover, the expression of metallothioneins is induced not only by exposure to heavy metals or glucocorticoides but also by UV light and TPA treatment (1). However, the sprI and sprII proteins differ from metallothioneins by a high content of proline and glutamine and by the presence of repeating elements. It would be interesting to find out whether the expression of the group 1 and group 2 genes could be also affected by heavy metals or glucocorticoides and whether the spr proteins, by analogy with metallothioneins, are metal binding.

It is still to be proven whether the sprI and sprII polypeptides contain metal ions and whether they can bind nucleic acids. Further characterization of these proteins will reveal their possible role in UV-inducible response or in any other cellular event. Recent results obtained with the antibodies specific to the chemically synthesized polypeptide corresponding to one of the group I genes confirm the existence of the predicted protein in vivo and indicate the regulation of the expression of group 1 proteins during keratinocyte differentiation (10).

ACKNOWLEDGMENTS

We acknowledge J. Kempenaar (Department of Dermatology, University Hospital, Leiden) for the help with the cell culture, M. Hoffmann and F. Brederode for assistance in the experiments summarized in Fig. 7 and 8, and B. J. C. Cornelissen, I. B. Roninson, J. H. J. Hoeijmakers, and E. C. Zwarthoff for critical reading of the manuscript.

This work was supported in part by the Foundation of Medical Scientific Research in the Netherlands and by the J. A. Cohen Institute for Radiopathology and Radiation Protection.

LITERATURE CITED

1. **Angel, P., A. Poting, U. Mallick, H. J. Rahmsdorf, M. Schorpp, and P. Herrlich.** 1986. Induction of metallothionein and other mRNA species by carcinogens and tumor promoters in primary human skin fibroblasts. Mol. Cell. Biol. **6**:1760–1766.
2. **Berg, J.** 1986. Potential metal-binding domains in nucleic acid binding proteins. Science **232**:485–487.
3. **Berget, S. M.** 1984. Are U4 small nuclear ribonucleoproteins involved in polyadenylation? Nature (London) **309**:179–181.
4. **Bishop, J. O., M. Rosbash, and D. J. Evans.** 1974. Polynucleotide sequence in eukaryotic DNA and RNA that form ribonuclease-resistant complexes with polyuridilic acid. J. Mol. Biol. **85**:75–86.
5. **Chen, H. R., and W. C. Barker.** 1985. The protein identification resource and its applications. Trends Genet. **1**:221–223.
6. **Chou, P. Y., and G. D. Fasman.** 1978. Prediction of the secondary structure of proteins from their amino acid sequence. Adv. Enzymol. **47**:45–148.
7. **Ikeneda, M., H. Ichikawa-Ryo, and S. Kado.** 1975. The major cause of inactivation and mutation by 4-nitroquinoline 1-oxide in *Escherichia coli*: excisable 4-NQO-purine adducts. J. Mol. Biol. **92**:341–356.
8. **Kagi, J. H. R., and M. Nordberg (ed.).** 1979. Metallothionein. Birkhauser, Basel.
9. **Kartasova, T., B. J. C. Cornelissen, P. Belt, and P. van de Putte.** 1987. Effects of UV, 4-NQO and TPA on gene expression in cultured human epidermal keratinocytes. Nucleic Acids Res. **15**:5945–5962.
10. **Kartasova, T., G. N. P. van Muijen, H. van Pelt-Heerschap, and P. van de Putte.** 1988. Novel protein in human epidermal keratinocytes: regulation of expression during differentiation. Mol. Cell. Biol. **8**:2204–2210.
11. **Kozak, M.** 1986. Point mutations define a sequence flanking the initiator codon that modulates translation by eukaryotic ribosomes. Cell **44**:283–292.
12. **Kyte, J., and R. F. DooLittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157**:105–132.
13. **Maita, T., T. Umegane, Y. Kato, and G. Matsuda.** 1980. Amino-acid sequence of the L-1 light chain of chicken cardiac-muscle myosin. Eur. J. Biochem. **107**:565–575.
14. **Maizel, J. V., Jr., and R. P. Lenk.** 1980. Enhanced grafic matrix analysis of nucleic acid and protein sequences. Proc. Natl. Acad. Sci. USA **78**:7665–7669.
15. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
16. **Maxam, A. M., and W. Gilbert.** 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. USA **74**:560–564.
17. **Mulligan, R. C., B. H. Howard, and P. Berg.** 1979. Synthesis of rabbit beta-globin in cultured monkey kidney cells following infection with a SV40 beta-globin recombinant genome. Nature (London) **277**:108–114.
18. **Ohlendorf, D. H., and B. W. Mattews.** 1983. Structural studies of protein-nucleic acid interactions. Annu. Rev. Biophys. Bioeng. **12**:259–284.
19. **Okayama, H., and P. Berg.** 1982. High-efficiency cloning of full-length cDNAs. Mol. Cell. Biol. **2**:161–170.
20. **Ponec, M., J. A. Kempenaar, and E. R. de Kloet.** 1981. Corticoids and cultured human epidermal keratinocytes: specific intra-cellular binding and clinical efficacy. J. Invest. Dermatol. **76**:211–214.

21. **Reinwald, J. G., and H. Green.** 1975. Serial cultivation of strains of human epidermal keratinocytes: the formation of keratinizing colonies from single cells. Cell **6**:331–344.
22. **Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. **113**: 237–251.
23. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
24. **Sarachu, A. N., M. J. Huisman, L. van Vloting-Doting, and J. F. Bol.** 1985. Alfalfa mosaic virus temperature-sensitive mutants. I. Mutants defective in viral RNA and protein synthesis. Virology **141**:14–22.
25. **Southern, E.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**: 503–517.
26. **Wickens, M., and P. Stephenson.** 1984. Role of a conserved AAUAAA sequence: four AAUAAA point mutations prevent messenger 3′ end formation. Science **226**:1045–1051.