Structural organization and regulation of the Small Proline-rich family of cornified envelope precursors suggest a role in adaptive barrier function.

Adriana Cabral, Patrick Voskamp, Anne-Marie Cleton-Jansen\(^1\), Andrew South\(^2\), Dean Nizetic\(^2\), Claude Backendorf\(^*\)

Department of Molecular Genetics, Leiden Institute of Chemistry, PO Box 9502, 2300 RA
Leiden, The Netherlands

\(^1\)Department of Pathology, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden,
The Netherlands

\(^2\)Center for Applied Molecular Biology, School of Pharmacy, University of London, 29/39 Brunswick Square, London WC1N 1AX, U.K.

\(^*\)Corresponding author
Laboratory of Molecular Genetics
Gorlaeus Laboratories
P.O. Box 9502
2300 RA Leiden
The Netherlands
Tel. (31) 71 527 4409
Fax (31) 71 527 4537
E-mail: Backendo@chem.leidenuniv.nl

Running Title: Functional mapping of the human SPRR gene cluster at 1q21
SUMMARY

The protective barrier provided by stratified squamous epithelia relies on the cornified cell envelope (CE), a structure synthesized at late stages of keratinocyte differentiation. It is composed of structural proteins, including involucrin, loricrin and the small proline-rich (SPRR) proteins, all encoded by genes localized at human chromosome 1q21. The genetic characterization of the SPRR locus reveals that the various members of this multigene family can be classified into 2 distinct groups with separate evolutionary histories. Whereas group 1 genes have diverged in protein structure and are composed of 3 different classes (SPRR1 (2x), SPRR3 and SPRR4), an active process of gene conversion has counteracted diversification of the protein sequences of group 2 genes (SPRR2 class, 7 genes). Contrasting with this homogenization process, all individual members of the SPRR gene family show specific in vivo and in vitro expression patterns and react selectively to UV irradiation. Apparently, creation of regulatory rather than structural diversity has been the driving force behind the evolution of the SPRR gene family. Differential regulation of highly homologous genes underlines the importance of SPRR protein-dosage in providing optimal barrier function to different epithelia, while allowing adaptation to diverse external insults.
INTRODUCTION

An essential function of stratified squamous epithelia is to provide a protective barrier for the organism against extracellular and environmental factors. The cornified cell envelope (CE), a specialized structure formed beneath the plasma membrane of differentiated cells, is a major component responsible for this protective function (1-3). The CE is an insoluble ≈15-nm-thick layer, which is the result of extensive cross-linking of several proteins by both disulfide and NE(γ-glutamyl)lysine isopeptide bonds catalyzed mainly by transglutaminases 1 and 3 (4,5). The assembly of the CE starts with the formation of a scaffold constituted of involucrin and envoplakin near the desmosomes. Subsequently, other reinforcing proteins, such as cystatin α, elafin, loricrin and SPRRs (6-9) are added to complete the CE structure, which serves as an attachment platform for specific lipids (10). Biochemical evidence has suggested that the characteristics of the CE related to toughness, strength and flexibility, exhibited by different stratified squamous epithelia, are dictated by SPRR proteins (11-15).

Human SPRRs consist of a multigene family (16-18), clustered in a 170kb region within the Epidermal Differentiation Complex (EDC) on human chromosome 1q21 (16,19-22). Orthologues of these genes have also been described in other mammalian species (reviewed in (23). The SPRR proteins have an identical structure consisting of head (amino-) and tail (carboxy-terminal) domains, comprising several glutamine and lysine residues, and a proline-rich central repetitive domain. Whereas the head and tail domains show a high degree of homology with other CE precursors (e.g. involucrin and loricrin) (24), the internal repeats, which vary in both number and consensus sequence, distinguish the various members of this gene family, allowing their classification into several SPRR classes. Based on the specific sequences
of these internal domains, secondary structure algorithms have predicted various degrees of flexibility for different classes (SPRR2<SPRR1<SPRR3) (25).

The SPRR classes are differentially regulated in various types of epithelia and their expression is modulated in response to environmental insult (UV irradiation), aging, diseased states and following carcinogenic transformation (18,23,26-37). Among all cornified envelope precursor proteins identified to date, SPRRs are the only ones that are encoded by a multigene family. Important questions concerning the reason for this complexity remain to be addressed. In order to provide novel insights to these questions, we have characterized the whole human SPRR locus, including the identification of several new members, the refinement of the physical and transcriptional maps, the comparison of gene and deduced protein structures and the establishment of in vivo and in vitro expression patterns at single gene level and after UV irradiation.
EXPERIMENTAL PROCEDURES

Cosmid library screening and SPRR contig assembly: The chromosome 1 cosmid library ICRFc112 (38) was screened with SPRR class specific probes (16). A total of 27 cosmids covering the SPRR locus were identified and used for preliminary contig assembly (21). A minimal tiling path comprising 7 cosmids was chosen for further analysis and overlap was verified by cosmid walking (39). Cosmids, linearized with NruI or CpoI (both are unique restriction sites in the vector), were partially digested with either BamHI, BglII, EcoRI, HindIII, KpnI, NcoI, PstI or XbaI and analyzed by pulsed-field gel electrophoresis as previously described (21), allowing the establishment of a contiguous restriction map. All EcoRI bands, identified by hybridization of human DNA with an SPRR2 probe (16,21), were detected on the various cosmids and sequence analysis established the presence of 7 SPRR2 genes. SPRR1A, SPRR1B and SPRR3 genes have been described (40-42). A probe for SPRR4 was derived from an EST library and obtained by RT-PCR of epidermal RNA. Superposition of the sequencing data with the contiguous restriction map provided the physical map of the SPRR locus, with the exact position and the transcriptional orientation of the different genes. Gene orientation and intergenic distances were verified by long-distance (LD) PCR (Fig. 1).

Cell culture: Normal human keratinocytes were cultured as previously described (43) and induced to terminally differentiate by using the stratification assay (43). Shortly, monolayers of basal cells (-Ca++ conditions) were induced to stratify for 48 hours by adding DMEM containing 5% serum, without growth factors (+Ca++ conditions). UV-C irradiation (30 J/m²) was applied to the monolayers before the addition of +Ca++ medium. RNA was isolated
according to (44).

Expression studies: Frozen tissues, obtained from the Department of Pathology (Leiden University Medical Center), were homogenized and total RNA was isolated by the Trizol method (Gibco BRL). RNA from skin, esophagus and uterus were also purchased from Invitrogen. Total RNA (0.4 µg) was reverse transcribed using Super RT (SphaeroQ) and random hexamer primers (Pharmacia). PCR was performed for 25, 30 and 35 cycles with 20 pmol of gene specific primers and 0.4U AmpliTaq DNA polymerase (Perkin Elmer). The specificity of the various primer mixes was determined by using the various cosmids or derived plasmids containing a single SPRR gene and verified by sequencing (results not shown). The absence of DNA contamination in RNA preparations could be easily controlled as several primer combinations bridged an intron. The positions of the respective primers in the SPRR sequences and the size of the PCR fragments are indicated in Table I.
RESULTS

The refined physical map and transcriptional orientation identify 2 distinct SPRR groups: Figure 1 shows the physical map of the complete SPRR locus, which was determined by analyzing a cosmid contig with 8 different restriction enzymes (see Materials & Methods). The cluster, previously localized to a 170 Kbase region on human chromosome 1q21 (16,21), comprises 11 genes: SPRR1A, SPRR1B, 7 SPRR2 genes (A to G), SPRR3 and the recently identified SPRR4 gene (manuscript in preparation). Long distance PCR (LD-PCR) was performed to confirm the relative position and orientation of the various genes (stippled lines). The transcriptional orientation of the various members (indicated by arrows) is not random and allows the splitting of the SPRR cluster into 2 groups. One group consists of SPRR1A, 1B, 3 and 4, which are placed in a proximal region and are transcriptionally oriented from centromere to telomere. The other group comprises the 7 SPRR2 genes, clustered in a 100kb region, all oriented in the same direction, but opposite to group 1 genes.

Group specific differences in the diversity of SPRR protein structures: The subdivision of SPRR genes into 2 groups is also justified by the comparison of the predicted aminoacid sequences (Fig. 2A). Group 1 genes (SPRR1A, SPRR1B, SPRR3 and SPRR4) are characterized by a long N-terminus, an 8 aminoacid repeat motif and a short C-terminus, while group 2 genes (all SPRR2s) have a short N-terminus, a 9 aminoacid repeat motif and a more extended C-terminus. In the central repeats of the various proteins more diversity exists among group 1 proteins, which show differences in both number of repeats and consensus sequence of the repetitive unit. These differences justify the classification of the 4 group 1 genes into 3 classes (SPRR1, 3 and 4). On the contrary, the 7 group 2 proteins are characterized by a much higher homogeneity, as each member contains 3 repeats of the same nonamer consensus. Hence, all group 2 genes belong to a single class, SPRR2. In Fig. 2B the central repetitive domains of SPRR2 genes have been aligned. Although repeats 1, 2 and 3 of a single gene have different consensi at the nucleotide level (mainly due to variations in the wobble position), each of the 3 repeats is highly conserved among
the various members. This indicates that during evolution, repeat duplication has preceded gene duplication and was maintained hereafter in each gene. The seven group 2 proteins (Fig. 2C) are highly homologous. For instance, 2B differs from 2A by 1 aminoacid and from 2D and 2E by 2 residues. Notably, all aminoacids previously identified as being involved in transglutaminase-mediated crosslinking during CE formation (11,14) are conserved in all SPRR2 proteins (red residues).

Figure 3 provides a global view of the sequence conservation among group 2 genes (black graph). The highest similarity is found in exon 2 (94%) and corresponds to the N-terminus of the protein (from position 1250 to 1400). Nevertheless, high sequence conservation is not restricted to the coding sequence, since in SPRR2B and 2E (red graph) a 550 basepair region, with 100% identity, extends from the intron to the coding sequence (position 850-1400). The various promoters revealed an average homology of approximately 70%. The major differences are between positions 200 and 300 bases and are due to a deletion in SPRR2B.

Differential expression among single SPRR genes: The lower sequence conservation within the promoters of both group 1 (42) and group 2 prompted us to monitor the specific expression pattern for each gene. Initially, we analyzed RNA from various human tissues by hybridization with class specific probes (results not shown). Besides the expected expression in various stratified squamous epithelia (27,35), some SPRRs were also detected in tissues that (normally) do not contain these epithelia (uterus, bladder, ovary and trachea). Uterus, ovary and 3 stratified squamous epithelia, namely skin, esophagus and cervix, were chosen for single gene analysis. Likewise, expression in a well established in vitro system, which permits the study of keratinocyte terminal differentiation (45), was also analyzed.

Due to the high homology within the SPRR family, gene specific semi-quantitative RT-PCR was carried out to characterize the relative expression patterns for individual genes (Fig. 4). The analysis of calcium mediated in vitro keratinocyte differentiation (−Ca++ and +Ca++) revealed that all SPRRs are induced during this process, except for SPRR2F. In stratified squamous epithelia distinctions in gene expression between the different tissues were observed. Only SPRR2G and SPRR4 are preferentially expressed in skin. All other SPRRs show higher expression
levels in mucosal stratified squamous epithelia, but with tissue-specific modulation (e.g. compare the relative expression of SPRR2B, 2C and 2D in esophagus and cervix). SPRR1A and 3 are present in ovary, SPRR2D is found in uterus, and 2B, 2E and 2F in both uterus and ovary. Especially, the high expression of SPRR2F in ovary is remarkable.

*Individual SPRR genes respond selectively to UV irradiation:* In order to analyze the response of the SPRR gene family to external damaging insults, we have treated human keratinocyte cultures with UV light and measured the expression of individual SPRRs (Fig. 5). In 2 independent experiments, various members reacted selectively to this DNA damaging agent. Whereas SPRR4, 2C and 2G are consistently induced, a certain degree of variability is observed between individual experiments in the case of 2B, 2D and 2F. SPRR1A, 1B, 3, 2A and 2E do not respond to UV irradiation. The variability in SPRR2B, 2D and 2F induction is likely due to small differences in cell density, which are difficult to control at the start of the experiment, but might affect gene expression (46). UV irradiation did not affect the expression of involucrin. Overall these results indicate that individual SPRR genes are differentially expressed, although only a limited amount of biological samples (5 different human tissues and *in vitro* cultured keratinocytes) were analyzed and the effect of a single external agent (UV) was studied.
DISCUSSION

The cornified envelope (CE) has a vital role in the barrier function of stratified squamous epithelia. Recent biochemical studies suggest that SPRR proteins are the major modulators of the biomechanical properties of cornified envelopes (12-15). Among all CE precursor proteins identified to date, SPRRs are the only ones that are encoded by a gene family. Although all genes have a common ancestor (16), the present analysis indicates that the family can be divided into two distinct subgroups with separate evolutionary histories. Whereas group 1 genes have clearly diverged in protein structure, group 2 genes are characterized by a highly conserved coding sequence. Darwinian selection (recently reviewed in (47)) is not likely to be the driving force behind this conservation, as most wobble positions, including those which are specific for each of the 3 repeats (Fig. 2B), have been strongly preserved among all group 2 genes. The nature of the process responsible for this high similarity is revealed by comparing SPRR2B and 2E (Fig. 3). An identical 550 basepair-long region, flanked by non-identical DNA, points to gene conversion as the implicated mechanism. Gene conversion is a process of homologous recombination, which can be defined as a non-reciprocal transfer of information between two sequences. As one sequence can be converted into the other one this process can result in the homogenization of gene families (reviewed in (47)).

Whereas the chromosomai organization and the protein structures of SPRRs clearly distinguish group 1 and group 2 genes, such subdivision is not evident when examining the expression patterns of individual genes. In fact, the major finding of this work is that all human SPRR genes, irrespective of the group or class they belong to, are under the control of specific and selective regulatory processes. Apparently, during the evolution of the SPRR gene family,
creation of regulatory diversity was more important than diversification in protein structure. This implies that the control of protein-dosage must be of major importance for the function of these genes.

Our RT-PCR analysis corroborates and extends earlier studies using class-specific DNA/RNA probes, antibodies or CE peptide sequencing, which have not allowed the detection of gene-specific differences within one class. The high expression of SPRR3 in esophagus and its absence from epidermis (27), as well as the elevated expression of SPRR1 in internal epithelia (48) have previously been observed. An interesting novel observation is the preferential expression of SPRR2G and SPRR4 in skin. Generally, it appears that genes that are well expressed in external “dry” epithelia (skin) are lower in internal “wet” epithelia and vice versa. Especially the preferential expression of SPRR2F in ovary is noteworthy. Expression of specific SPRR2 genes in murine uterus (23) and the presence of an SPRR homologue in cultured Chinese hamster ovary (CHO) cells (49) have previously been reported. At present, there is no satisfactory explanation for SPRR expression in these organs, which do not contain stratified squamous epithelia. It has been suggested that the presence of SPRRs in non-squamous epithelia might reflect a predisposition to undergo squamous metaplasia (23,50). Alternatively, SPRR genes could be involved in other forms of programmed cell death (apoptosis), which is known to occur in these tissues (51,52). A recent inspection of bladder epithelium with specific antibodies revealed SPRR1 and SPRR3 expression in the most superficial (umbrella) cells (our unpublished observation). Hence, a more thorough investigation, which is beyond the focus of this paper, will be imperative to assess the relevance of SPRR expression in these tissues.

Whereas differential expression of individual SPRR genes is likely to reflect the specific barrier
requirements of different epithelia, the UV experiment underlines the importance of barrier adaptation following external insults. UV responsiveness of SPRRs is not a novel finding, since they were originally isolated in our laboratory as UV inducible genes (17). The novelty resides in the fact that specific members of this gene family are selectively induced by UV light. Consequently, induction of SPRR4, 2C and 2G is not due to a global effect of UV irradiation on the process of terminal differentiation, during which most SPRRs are induced (Fig. 4). This view is also supported by the finding that involucrin expression is not modulated after UV irradiation (Fig. 5). These results indicate that, besides providing resistance and flexibility to very specialized tissues, SPRRs might fulfill a major role in the adaptation of epithelial barriers to a large variety of external and endogenous stimuli.

Recent evidence has indeed linked SPRR expression with barrier formation during mouse development (53). Within the cornified cell envelope, which constitutes a major determinant of the protective barrier, SPRRs have a specialized role as they function as crossbridging agents, which either interconnect or adjoin other CE precursor proteins. Both the structure and the concentration of the various SPRR proteins are believed to affect the biomechanical properties of the CE (25). It is possible that even small changes in aminoacid composition can influence these parameters. Whereas the use of one specific class is probably dictated by tissue specific requirements, adaptation to external signals is likely to be more efficient by varying the concentration of a given SPRR protein. Both mechanisms are however by no means exclusive. Indeed, as various epithelia are exposed to specific insults, some correlation between tissue expression and responsiveness to a given agent can be expected. As such, the finding that SPRR2G and SPRR4, which are preferentially expressed in the epidermis, are also responsive to
UV irradiation is not surprising.

External insults can be numerous and can differ largely between different epithelia (e.g. UV irradiation for the epidermis, tobacco smoke or food derived chemicals for oral epithelia, acid reflux for esophagus). By taking into account this large diversity of external insults, which might request barrier function adaptation, it is unlikely that all these signals are channeled to a single regulatory promoter region. A gene family, coding for highly homologous proteins, regulated by specialized promoters, responding to both inducing and repressing signals, is likely to allow fine-tuning of the barrier, in order to guarantee optimal protection to the organism.

The identification of 2 groups of UV inducible genes (dependent/independent on the culture conditions) within the SPRR2 class indicates that at least two different UV responsive signaling pathways selectively target specific members of the gene family. Other signal transduction cascades, initiated by other external or endogenous agents, are likely to regulate other family members. Our previous finding that the SPRR2A promoter, which is not affected by UV light (Fig. 5), is under the control of an interferon stimulated response element (ISRE) (43), not present in other SPRR2 genes (Fig. 6), supports such a view.

Previous work from our laboratory has focused on the promoter regions of specific members of the SPRR1, 2 and 3 classes and has revealed that integration of signals transmitted via various signaling pathways plays an essential role in the regulation of these genes (43). Such a strict regulation is a prerequisite for efficient barrier function adaptation. Indeed, ablation of the Klf4 transcription factor, one of the regulators of SPRR2 expression, results in severe barrier deficiency in the mouse (54). Differential regulation of SPRR promoters relies on variations in the precise position of specific cis-elements within the global promoter context. This variation
was recognized as a major factor in determining stimulus specific expression (40,42). As a matter of fact, diversification of control elements is also seen in the promoter regions of the highly homologous SPRR2 genes (Fig. 6), in concert with their differential regulation. Differences include the deletion of an element (AP-1 site in 2B; ETS site in 2F; octamer and ZNF site in 2G), the replacement of one element by another one (ISRE/ETS in 2A; TATA/initiator in 2C) and single point mutations in binding sites (AP-1 sites in 2D and 2G). This diversification in regulatory elements is likely to affect both the binding of specific transcription factors and their mutual cooperativity (43). For instance, the absence of ETS sites in SPRR2F might explain the loss of regulation of this gene during in vitro keratinocyte differentiation (Fig. 5) (43). Whether the same change is also responsible for the unexpected high expression of this gene in non-squamous epithelia of the uterus and ovary is not yet known.

As a conclusion, we propose that the 2 structurally different groups of human SPRR genes provide on one hand specific resistance to very specialized tissues, while allowing on the other hand adaptation to a plethora of variable physiological and environmental insults. On this basis, the structural organization of the SPRR gene family reflects the functional duality with which epithelial barriers are confronted in order to guarantee optimal protection to the organism.

**Acknowledgement:** We thank Drs. D. Hohl (Lausanne) and T. Kartasova (Bethesda) for critically reading the manuscript and Dr. V.T.H.B.M. Smit (Dept. Pathology, LUMC) for providing human tissue. Drs. J. Brouwer and P. van de Putte are acknowledged for stimulating discussions, and A-M Borgstein for technical assistance. The hospitality of Dr. E. Bakker (Dept. Clinical Genetics, LUMC) was appreciated. This research was supported by grant BMH4-CT96-0319 of the EEC.
and by the J.A. Cohen Institute.
REFERENCES

1. Matoltsy, A. G., and Matoltsy, M. N. (1966) *J Invest Dermatol* 46(1), 127-9

2. Rice, R. H., and Green, H. (1977) *Cell* 11(2), 417-22

3. Hohl, D. (1990) *Dermatologica* 180(4), 201-11

4. Martinet, N., Kim, H. C., Girard, J. E., Nigra, T. P., Strong, D. H., Chung, S. I., and Folk, J. E. (1988) *J Biol Chem* 263(9), 4236-41

5. Candi, E., Melino, G., Mei, G., Tarcsa, E., Chung, S. I., Marekov, L. N., and Steinert, P. M. (1995) *J Biol Chem* 270(44), 26382-90

6. Steinert, P. M., and Marekov, L. N. (1995) *J Biol Chem* 270(30), 17702-11

7. Steinert, P. M., and Marekov, L. N. (1997) *J Biol Chem* 272(3), 2021-30

8. Steinert, P. M., and Marekov, L. N. (1999) *Mol Biol Cell* 10(12), 4247-61

9. Eckert, R. L., Yaffe, M. B., Crish, J. F., Murthy, S., Rorke, E. A., and Welter, J. F. (1993) *J Invest Dermatol* 100(5), 613-7

10. Nemes, Z., Marekov, L. N., Fesus, L., and Steinert, P. M. (1999) *Proc Natl Acad Sci U S A* 96(15), 8402-7

11. Candi, E., Melino, G., Sette, M., Oddi, S., Guerrieri, P., and Paci, M. (1999) *Biochem Biophys Res Commun* 262(2), 395-400

12. Candi, E., Tarcsa, E., Idler, W. W., Kartasova, T., Marekov, L. N., and Steinert, P. M. (1999) *J Biol Chem* 274(11), 7226-37

13. Steinert, P. M., Candi, E., Kartasova, T., and Marekov, L. (1998) *J Struct Biol* 122(1-2), 76-85
14. Steinert, P. M., Kartasova, T., and Marekov, L. N. (1998) *J Biol Chem* 273(19), 11758-69

15. Tarcsa, E., Candi, E., Kartasova, T., Idler, W. W., Marekov, L. N., and Steinert, P. M. (1998) *J Biol Chem* 273(36), 23297-303

16. Gibbs, S., Fijneman, R., Wiegart, J., van Kessel, A. G., van De Putte, P., and Backendorf, C. (1993) *Genomics* 16(3), 630-7

17. Kartasova, T., van Muijen, G. N., van Pelt-Heerschap, H., and van de Putte, P. (1988) *Mol Cell Biol* 8(5), 2204-10

18. Kartasova, T., and van de Putte, P. (1988) *Mol Cell Biol* 8(5), 2195-203

19. Marenholz, I., Volz, A., Ziegler, A., Davies, A., Ragoussis, I., Korge, B. P., and Mischke, D. (1996) *Genomics* 37(3), 295-302

20. Mischke, D., Korge, B. P., Marenholz, I., Volz, A., and Ziegler, A. (1996) *J Invest Dermatol* 106(5), 989-92

21. South, A. P., Cabral, A., Ives, J. H., James, C. H., Mirza, G., Marenholz, I., Mischke, D., Backendorf, C., Ragoussis, J., and Nizetic, D. (1999) *J Invest Dermatol* 112(6), 910-8

22. Volz, A., Korge, B. P., Compton, J. G., Ziegler, A., Steinert, P. M., and Mischke, D. (1993) *Genomics* 18(1), 92-9

23. Song, H. J., Poy, G., Darwiche, N., Lichti, U., Kuroki, T., Steinert, P. M., and Kartasova, T. (1999) *Genomics* 55(1), 28-42

24. Backendorf, C., and Hohl, D. (1992) *Nat Genet* 2(2), 91

25. Kartasova, T., Parry, D. A. D., and Steinert, P. M. (1995) *J. Invest. Dermatol.* 104, 611

26. Abraham, J. M., Wang, S., Suzuki, H., Jiang, H. Y., Rosenblum-Vos, L. S., Yin, J., and
27. Hohl, D., de Viragh, P. A., Amiguet-Barras, F., Gibbs, S., Backendorf, C., and Huber, M. (1995) *J Invest Dermatol* 104(6), 902-9

28. Jarnik, M., Kartasova, T., Steinert, P. M., Lichti, U., and Steven, A. C. (1996) *J Cell Sci* 109(Pt 6), 1381-91

29. Kartasova, T., Darwiche, N., Kohno, Y., Koizumi, H., Osada, S., Huh, N., Lichti, U., Steinert, P. M., and Kuroki, T. (1996) *J Invest Dermatol* 106(2), 294-304

30. Lohman, F. P., Medema, J. K., Gibbs, S., Ponec, M., van de Putte, P., and Backendorf, C. (1997) *Exp Cell Res* 231(1), 141-8

31. Gibbs, S., Lohman, F., Teubel, W., van de Putte, P., and Backendorf, C. (1990) *Nucleic Acids Res* 18(15), 4401-7

32. Garmyn, M., Yaar, M., Boileau, N., Backendorf, C., and Gilchrest, B. A. (1992) *J Invest Dermatol* 99(6), 743-8

33. Fujimoto, W., Marvin, K. W., George, M. D., Celli, G., Darwiche, N., De Luca, L. M., and Jetten, A. M. (1993) *J Invest Dermatol* 101(3), 268-74

34. Ishida-Yamamoto, A., Iizuka, H., Manabe, M., O'Guin, W. M., Hohl, D., Kartasova, T., Kuroki, T., Roop, D. R., and Eady, R. A. (1995) *Arch Dermatol Res* 287(8), 705-11

35. Xu, X. C., Mitchell, M. F., Silva, E., Jetten, A., and Lotan, R. (1999) *Clin Cancer Res* 5(6), 1503-8

36. Yaar, M., Eller, M. S., Bhawan, J., Harkness, D. D., DiBenedetto, P. J., and Gilchrest, B. A. (1995) *Exp Cell Res* 217(2), 217-26

37. Saunders, N. A., Smith, R. J., and Jetten, A. M. (1993) *Biochem Biophys Res Commun*
38. Nizetic, D., Monard, S., Young, B., Cotter, F., Zehetner, G., and Lehrach, H. (1994) *Mamm Genome* 5(12), 801-2

39. Ivens, A., and Little, P. (1995) in *DNA Cloning 3 A Practical Approach* (Glover, D., and Hames, B., eds), IRL Press, Oxford

40. Sark, M. W., Fischer, D. F., de Meijer, E., van de Putte, P., and Backendorf, C. (1998) *J Biol Chem* 273(38), 24683-92

41. An, G., Tesfaigzi, J., Chuu, Y. J., and Wu, R. (1993) *J Biol Chem* 268(15), 10977-82

42. Fischer, D. F., Sark, M. W., Lehtola, M. M., Gibbs, S., van de Putte, P., and Backendorf, C. (1999) *Genomics* 55(1), 88-99

43. Fischer, D. F., Gibbs, S., van De Putte, P., and Backendorf, C. (1996) *Mol Cell Biol* 16(10), 5365-74

44. Belt, P. B., Groeneveld, H., Teubel, W. J., van de Putte, P., and Backendorf, C. (1989) *Gene* 84(2), 407-17

45. Rheinwald, J. G., and Green, H. (1977) *Nature* 265(5593), 421-4

46. Lee, Y. S., Yuspa, S. H., and Dlugosz, A. A. (1998) *J Invest Dermatol* 111(5), 762-6

47. Li, W.-H. (1997) *Molecular Evolution*, Sinauer Associates, Inc.

48. Lee, C. H., Marekov, L. N., Kim, S., Brahim, J. S., Park, M. H., and Steinert, P. M. (2000) *FEBS Lett* 477(3), 268-72

49. Tesfaigzi, J., and Carlson, D. M. (1996) *Exp Cell Res* 228(2), 277-82

50. Jetten, A. M., De Luca, L. M., Nelson, K., Schroeder, W., Burlingame, S., and Fujimoto, W. (1996) *Mol Cell Endocrinol* 123(1), 7-15
51. Gosden, R., and Spears, N. (1997) Br Med Bull 53(3), 644-61

52. Salamonsen, L. A., Kovacs, G. T., and Findlay, J. K. (1999) Baillieres Best Pract Res Clin Obstet Gynaecol 13(2), 161-79

53. Marshall, D., Hardman, M. J., and Byrne, C. (2000) J Invest Dermatol 114(5), 967-75

54. Segre, J. A., Bauer, C., and Fuchs, E. (1999) Nat Genet 22(4), 356-60
FOOTNOTES

The abbreviations used are: CE: cornified envelope; SPRR: small proline-rich protein; RT-PCR: reverse transcriptase polymerase chain reaction
Fig. 1. **Contiguous map of the entire SPRR locus identifying 2 groups and 4 sub-classes of genes.** A cosmid contig covering the whole SPRR locus was analyzed with 8 restriction enzymes. The SPRR 1, 2, 3 and 4 sub-classes are indicated by different ovals. Arrows define the transcriptional orientation of each gene. Fragments hybridizing with the different SPRR probes are indicated with a small open square beneath each restriction bar. Fragments amplified by long-distance PCR (LD-PCR) are indicated with stippled lines. Regions with more than 20kb could not be resolved by LD-PCR (SPRR3/1B and 2F/2C).

Fig. 2. **Structural features of SPRR group 1 and 2 proteins.** (A) Aminoacid sequence comparison of N- and C- terminal domains. The number of reiterated repeats and the number of aminoacids present in each repeat motif are represented. Aminoacids identical in both groups are indicated in red. Light- or dark-grey backgrounds indicate aminoacid identity within group 1 or group 2 respectively. Asterisks indicate stop codons. (B) Comparison of the nucleotide sequences of SPRR2 repetitive domains. Repeat specific nucleotide substitutions are indicated in red. Class specific differences are represented in bold. (C) Aminoacid conservation among SPRR2 proteins. The N termini, the 3 internal repeats (R1, R2 and R3) and the C termini of the 7 SPRR2 proteins are compared. Aminoacids involved in the transglutaminase crosslinking reaction are represented in red, whereas sequence differences between the various members are in bold.

Fig. 3. **Similarity analysis of human SPRR2 genomic sequences.** The genomic structure of SPRR2 genes is schematically indicated. Exon 2 comprises the whole coding sequence (CDS).
and the 3' untranslated region (UTR). The similarity plot of the 7 SPRR2 genes is represented in black, whereas the red plot compares SPRR2B and 2E. The analysis was carried out with the plotsimilarity program of the GCG (Wisconsin Package Version 10.0).

**FIG. 4.** *RT-PCR analysis of gene-specific SPRR expression profiles.* RT-PCR products were derived from RNA of the indicated sources and amplified during 25, 30 and 35 cycles. –Ca++ and +Ca++ indicates RNA from keratinocytes grown *in vitro* in the absence or presence of calcium, respectively.

**FIG. 5.** *Selective induction of individual SPRR genes after UV irradiation of keratinocyte cultures.* Stripped monolayers of undifferentiated keratinocytes were irradiated (30 J/m², UV-C) or mock-irradiated and induced to differentiate by the addition of +Ca++ medium (see Materials and Methods). RNA was isolated 48 hours later and analyzed by gene specific RT-PCR (35 PCR cycles). The results presented are from 2 independent experiments (A and B).

**FIG. 6.** *Regulatory elements in SPRR2 promoters.* Transcription factor binding sites previously identified in the SPRR2A promoter (43) are represented and were placed in the other promoters by sequence homology. Dotted lines indicate deletions. Single point mutations are represented in bold. TATA boxes and an initiator sequence (INR) in SPRR2C are indicated. ZNF: binding site for Krüppel-like zinc-finger factors. ETS: binding site for the Ets family of transcription factors. ISRE: interferon stimulated response element. Octamer: binding site for Oct transcriptional
regulators. AP-1: binding site for Jun/Fos factors.
### TABLE I
Primers used in RT-PCR

| GENE   | PRIMER      | FRAGMENT SIZE | ACCESSION No. |
|--------|-------------|---------------|---------------|
| SPRR1A | F: 862–881  | 454 bp        | L05187        |
|        | R: 2397–2377|               |               |
| SPRR1B | F: 967–987  | 505 bp        | M84757        |
|        | R: 2548–2528|               |               |
| SPRR2A | F: 1445–1467| 318 bp        | X53064        |
|        | R: 2478–2455|               |               |
| SPRR2B | F: 1437–1460| 199 bp        | AF333952      |
|        | R: 1635–1616|               |               |
| SPRR2C | F: 418–439  | 524 bp        | AF333953      |
|        | R: 1636–1616|               |               |
| SPRR2D | F: 1602–1624| 179 bp        | AF333954      |
|        | R: 1880–1757|               |               |
| SPRR2E | F: 1484–1505| 227 bp        | AF333955      |
|        | R: 1710–1684|               |               |
| SPRR2F | F: 1528–1548| 158 bp        | AF333956      |
|        | R: 1885–1665|               |               |
| SPRR2G | F: 376–398  | 481 bp        | AF333957      |
|        | R: 1576–1556|               |               |
| SPRR3  | F: 1689–1709| 622/702 bp    | AF077374      |
|        | R: 3519–3499|               |               |
| SPRR4  | F: 24–40    | 371 bp        | AF335109      |
|        | R: 394–37   |               |               |
| Involucrin | F: 101–120 | 373 bp        | M13903        |
|         | R: 473–453  |               |               |
### A

| N TERMINUS | number | consensus sequence | C TERMINUS |
|------------|--------|-------------------|------------|
| SPRR1A     | 6X     | PKVPEPC          | STVTPPAQQ  |
| SPRR1B     | 16X    | TKVPEPCG         | KVTPPAQQ   |
| SPRR3      | 4X     | PK.K..CA         | SKKQK*     |
| SPRR4      | 3X     | PKCPEPCPP        | SKKQK*     |

### B

**Repeat 1:**

| Consensus | 2A | 2B | 2C | 2D | 2E | 2F | 2G |
|-----------|----|----|----|----|----|----|----|
| PKCPQCPP  | CCAAGTGCGCGAGCCATGCTTCCACCC | CCAAGTGCGCGAGCCATGCTTCCACCC | CCAAGTGCGCGAGCTTCCACCC | CCAAGTGCGCGAGCTTCCACCC | CCAAGTGCGCGAGCTTCCACCC | CCAAGTGCGCGAGCTTCCACCC | CCAAGTGCGCGAGCTTCCACCC |

**Repeat 2:**

| Consensus | 2A | 2B | 2C | 2D | 2E | 2F | 2G |
|-----------|----|----|----|----|----|----|----|
| PKCPQCPP  | CGGAAGTGCGCGAGCTTCCACCC | CGGAAGTGCGCGAGCTTCCACCC | CGGAAGTGCGCGACCTTCCACCC | CGGAAGTGCGCGACCTTCCACCC | CGGAAGTGCGCGACCTTCCACCC | CGGAAGTGCGCGACCTTCCACCC | CGGAAGTGCGCGACCTTCCACCC |

**Consensus:**

| NUCLEOTIDE | AMINOACID |
|------------|-----------|
| R1: CCAAGTGCGCGAGCTTCCACCC | PKCPQCPP |
| R2: CGGAAGTGCGCGACCTTCCACCC | PKCPQCPP |
| R3: CGGAAGTGCGCGACCTTCCACCC | PKCPQCPP |

### C

| N TERMINUS | R1 | R2 | R3 | C TERMINUS |
|------------|----|----|----|------------|
| SPRR2A     | MSYQQQCKPCQCPPPVCPPT | PKCPEPCPP | PKCPEPCPP | PKCPQCPP | QQQKYYPPVTPSPCPPQSKYPPKSK |
| SPRR2B     | MSYQQQCKPCQCPPPVCPPT | PKCPEPCPP | PKCPEPCPP | PKCPQCPP | QQQKYYPPVTPSPCPPQSKYPPKSK |
| SPRR2C     | MSYQQ*CKPCQCPPPVCPPT | PKRPEPCPP | PKKPEPCPS | PKCPQCPP | QQQKYYPPVTPSPCPPQSKYPPKSK |
| SPRR2D     | MSYQQQCKPCQCPPPVCPPT | PKCPEPCPP | PKCPEPCPP | PKCPQCPP | QQQKYYPPVTPSPCPPQSKYPPKSK |
| SPRR2E     | MSYQQQCKPCQCPPPVCPPT | PKCPEPCPP | PKCPEPCPP | PKCPQCPP | QQQKYYPPVTPSPCPPQSKYPPKSK |
| SPRR2F     | MSYQQQCKPCQCPPPVCPPT | PKCPEPCPP | PKCPEPCPP | PKCPQCPP | QQQKYYPPVTPSPCPPQSKYPPKSK |
| SPRR2G     | MSYQQQCKPCQCPPPVCPPT | PKCPEPCPP | PKCPEPCPP | PKCPQCPP | QQQKYYPPVTPSPCPPQSKYPPKSK |
Structural organization and regulation of the small proline-rich family of cornified envelope precursors suggest a role in adaptive barrier function
Adriana Cabral, Patrick Voskamp, Anne-Marie Cleton-Jansen, Andrew South, Dean Nizetic and Claude Backendorf

*J. Biol. Chem.* published online March 15, 2001

Access the most updated version of this article at doi: 10.1074/jbc.M100336200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts