See related article on pg 1644

COMMENTARY

SNPing at the Epidermal Barrier

David P. Kelsell[1] and Carolyn Byrne[1]

Filaggrin variants are well-established risk factors for atopic eczema (AE). Recent studies suggest additional epidermal differentiation complex (EDC) gene associations with AE. In this issue, Marenholz and colleagues confirm this prediction and show that a small proline-rich protein 3 (SPRR3) variant confers susceptibility to AE. This finding suggests that further genetic and functional characterization of SPRR3 should be performed in patients with AE.

Journal of Investigative Dermatology (2011) **131**, 1593–1595. doi:10.1038/jid.2011.92

The chromosomal region 1q21.3 (the epidermal differentiation complex, or EDC) has been widely studied in epidermal biology and disease because it harbors a large number of genes (around 60) expressed in late keratinocyte differentiation (Mischke *et al.*, 1996). These include genes encoding proteins such as filaggrin (FLG), loricin, and gene family clusters, including the late cornified envelope genes, S100s, and the small proline-rich proteins (SPRRs). They are involved in various aspects of stratum corneum formation, including keratin filament aggregation (FLG) and cornified envelope formation (loricin), or are implicated in host defense (the S100s). Mutations in FLG and loricin are associated with ichthyosis vulgaris (Smith *et al.*, 2006) and a variant of Vohwinkel's syndrome (Maestrini *et al.*, 1996), respectively. However, the most striking finding has been that loss-of-function mutations in FLG are strongly associated with atopic eczema (AE), implicating an impaired epidermal barrier in the development of this chronic inflammatory skin disease that affects up to 20% of children in some countries (Palmer *et al.*, 2006).

It is thus intriguing that FLG appears to be the main culprit in the EDC for AE susceptibility, although other EDC genes could be predicted to have similar deleterious effects on epidermal barrier function. Three recent areas of genetic investigation suggest that FLG may not be solely responsible. First, genetic studies (Cascella *et al.*, 2011; Sinclair *et al.*, 2009) show that, although present, FLG mutations are not strongly implicated in AE susceptibility in Italian and Bangladeshi populations. Second, expression studies show dysregulation of a number of EDC proteins in AE skin (Sugiura *et al.*, 2005; Saaf *et al.*, 2008; Guttman-Yassky *et al.*, 2009). Third, a genome-wide association study supports another locus (or loci) associated with AE within the EDC (Esparza-Gordillo *et al.*, 2009). The latter, in particular, led Marenholz *et al.* (2011, this issue) to investigate whether coding sequence variants in other EDC genes were associated with AE.

The investigators examined the National Center for Biotechnology Information (NCBI) database to identify likely deleterious coding sequence variants in other EDC genes, primarily insertion/deletion, frameshift, and stop codon mutations. They selected 20 such variants and found that rs28989168 (variant allele is an in-frame 24-bp insertion in SPRR3) was associated with eczema after replication studies (odds ratio of 1.30). These data demonstrate two distinct proteins within the EDC that are linked to the impaired epidermal barrier function associated with AE susceptibility.

SPRR family members are thought to have a structural role in the cornified envelope as linkers for transglutaminase-mediated protein cross-bridging (Steinert *et al.*, 1998), and they have been proposed as reactive oxygen species quenchers, with an important function at the edges of wounds (Vermeij and Backendorf, 2010). SPRR proteins contain glutamine and lysine-rich N- and C-termini, which provide transglutaminase substrates. The termini flank 8- to 9-amino-acid proline-rich internal repeats, varying in sequence and number among family members, and act as flexible, cross-bridging spacers (Steinert *et al.*, 1998). Repeat variability is thought to confer differences in flexibility, strength, and toughness to the stratum corneum. Accordingly, Marenholz *et al.* (2011) speculate that the eczema-associated SPRR3 variant, with an additional 8-amino-acid repeat, influences barrier function by altering the physical properties of the stratum corneum.

SPRR3 is a surprise candidate for a functionally important skin barrier gene in that it is abundantly expressed in internal epithelia (for example, the tongue, esophagus, and tonsil) and undetectable in normal skin (Cabral *et al.*, 2001; BioGPS, http://biogps.gnf.org). Its identification as a skin barrier component suggests that minor stratum corneum protein components can substantially affect barrier function. Alternatively, the SPRR3 variant may influence skin barrier activity as part of a skin stress response because the expression of SPRR proteins increases in response to diverse influences including UVR, aging, cancer, and wound healing (Vermeij and Backendorf, 2010; Cabral *et al.*, 2001, and references therein). A role for SPRR3 in events triggering eczema outbreaks (for example, repair of barrier microlesions after scratching or in response to infection) is plausible,

[1]*Centre for Cutaneous Research, The Blizard Institute, Barts & The London School of Medicine and Dentistry, Queen Mary, University of London, London, UK*

Correspondence: David P. Kelsell, Centre for Cutaneous Research, The Blizard Institute, Barts & The London School of Medicine and Dentistry, Queen Mary, University of London, 4 Newark Street, London E1 2AT, UK. E-mail: d.p.kelsell@qmul.ac.uk

Table 1. Nonsynonymous sequence variation at the SPRR3 locus

Variation ID	Variation type	Effect on protein
rs67156933	Deletion	In-frame 8-amino-acid insertion/deletion
rs28989168	**Deletion**	**In-frame 8-amino-acid insertion/deletion**
rs66616552	Deletion	In-frame 8-amino-acid insertion/deletion
rs17845559	Nonsynonymous coding	Amino acid substitution (valine or leucine)
rs1055935	Nonsynonymous coding	Amino acid substitution (leucine or valine)
rs2075740	Nonsynonymous coding	Amino acid substitution (threonine or methionine)

rs28989168 (in bold) is the insertion polymorphism associated with atopic eczema in the study by Marenholz *et al.* (2011). The other five coding polymorphisms were not genotyped.

and its function in skin requires further elucidation.

Another surprising finding reported by Marenholz *et al.* (2011) is the large number of sequence variants annotated in the NCBI database that they could not find in their patient samples, with only 4 of 20 being experimentally validated. The investigators suggested this may be attributable to the rarity of these alleles (although they tested 94 individuals), their population specificity, or their being sequence artifacts. We have revisited these original 20 sequence variants in a newer build of the human genome sequence; we also looked for their presence in the 1,000-genome database accessed via Ensembl, and we suggest that the 16 unvalidated variants were likely to be artifacts. For example, of the four FLG2 single-nucleotide polymorphisms (SNPs) genotyped by Marenholz *et al.*, only rs12568784 (the only FLG2 SNP they experimentally validated by PCR) was found in the 1,000-genome dataset. In light of this, we revisited SPRR3 to determine what other potentially deleterious sequence variants or SNPs have been annotated in the current sequence build. Table 1 shows that further SPRR3 coding variants warrant investigation as risk factors linked to AE susceptibility. These include two 8-amino-acid insertion/deletion sequence variants (rs67156933 and rs66616552) in addition to the AE-associated rs28989168, suggesting that these should also be assessed for association with AE.

The study by Marenholz *et al.* (2011) highlights how complex the genetic makeup of AE susceptibility is likely to be. On the other hand, it emphasizes that the epidermal barrier is a major factor in AE susceptibility. From a mechanistic point of view, developing topical applications that improve the stratum corneum should therefore result in improvement for many patients with AE despite a distinct genetic contribution to their etiology. In addition, the current and future impact of high-throughput sequencing in understanding genetic disease cannot be ignored. Since the completion of Marenholz and colleagues' study (2011), the 1,000-genome project is revealing new sequence variants and providing allele frequency data. The decreasing costs of high-throughput sequencing, including exome sequencing, allow geneticists to return to family-based studies (as opposed to the recent focus on identifying "low"-risk variants through large-scale genome-wide association studies). The exome sequencing of AE families is likely to reveal "private" genetic variation linked to AE susceptibility, which may be family specific—akin to that revealed for many monogenic skin diseases. The question is whether the majority will still be found within the EDC.

Clinical Implications

- Nonfilaggrin epidermal differentiation gene variants are risk factors for atopic eczema.
- A small proline-rich protein gene variant contributes to atopic eczema susceptibility.
- Technical advances and the human genomic knowledge explosion suggest imminent identification of additional genetic contributors to atopic eczema.

CONFLICT OF INTEREST

The authors state no conflict of interest.

REFERENCES

Cabral A, Voskamp P, Cleton-Jansen AM *et al.* (2001) Structural organization and regulation of the small proline-rich family of cornified envelope precursors suggest a role in adaptive barrier function. *J Biol Chem* 276:19231–7

Cascella R, Cuzzola VF, Lepre T *et al.* (2011) Full sequencing of the *FLG* gene in Italian patients with atopic eczema: evidence of new mutations, but lack of an association. *J Invest Dermatol* 131:982–4

Esparza-Gordillo J, Weidinger S, Folster-Holst R *et al.* (2009) A common variant on chromosome 11q13 is associated with atopic dermatitis. *Nat Genet* 41:596–601

Guttman-Yassky E, Suarez-Farinas M, Chiricozzi A *et al.* (2009) Broad defects in epidermal cornification in atopic dermatitis identified through genomic analysis. *J Allergy Clin Immunol* 124:1235–44

Maestrini E, Monaco AP, McGrath JA *et al.* (1996) A molecular defect in loricrin, the major component of the cornified cell envelope, underlies Vohwinkel's syndrome. *Nat Genet* 13:70–7

Marenholz I, Gimenez Rivera VA, Esparza-Gordillo J *et al.* (2011) Association screening in the epidermal differentiation complex (EDC) identifies an *SPRR3* repeat number variant as a risk factor for eczema. *J Invest Dermatol* 131:1644–9

Mischke D, Korge BP, Marenholz I *et al.* (1996) Genes encoding structural proteins of epidermal cornification and S100 calcium-binding proteins form a gene complex ("epidermal differentiation complex") on human chromosome 1q21. *J Invest Dermatol* 106:989–92

Palmer CN, Irvine AD, Terron-Kwiatkowski A

et al. (2006) Common loss-of-function variants of the epidermal barrier protein filaggrin are a major predisposing factor for atopic dermatitis. *Nat Genet* 38:441–6

Saaf AM, Tengvall-Linder M, Chang HY *et al*. (2008) Global expression profiling in atopic eczema reveals reciprocal expression of inflammatory and lipid genes. *PLoS One* 3:24

Sinclair C, O'Toole EA, Paige D *et al*. (2009) Filaggrin mutations are associated with ichthyosis vulgaris in the Bangladeshi population. *Br J Dermatol* 160:1113–5

Smith FJ, Irvine AD, Terron-Kwiatkowski A *et al*. (2006) Loss-of-function mutations in the gene encoding filaggrin cause ichthyosis vulgaris. *Nat Genet* 38:337–42s

Steinert PM, Candi E, Kartasova T *et al*. (1998) Small proline-rich proteins are cross-bridging proteins in the cornified cell envelopes of stratified squamous epithelia. *J Struct Biol* 122:76–85

Sugiura H, Ebise H, Tazawa T *et al*. (2005) Large-scale DNA microarray analysis of atopic skin lesions shows overexpression of an epidermal differentiation gene cluster in the alternative pathway and lack of protective gene expression in the cornified envelope. *Br J Dermatol* 152:146–9

Vermeij WP, Backendorf C (2010) Skin cornification proteins provide global link between ROS detoxification and cell migration during wound healing. *PLoS One* 5:e11957

See related article on pg 1745

Cutaneous Squamous Cell Carcinoma: A Smoking Gun but Still No Suspects

John T. Schiller[1] and Christopher B. Buck[1]

A viral etiology for cutaneous squamous cell carcinoma (cuSCC) has long been suspected, primarily on the basis of its dramatically increased incidence in immunocompromised individuals. In this issue, Arron and colleagues report a comprehensive hunt for viral gene transcription in cuSCC. Their findings show that it is very unlikely that any currently known virus is commonly responsible for the maintenance of this cancer.

Journal of Investigative Dermatology (2011) **131**, 1595–1596. doi:10.1038/jid.2011.151

Identification of a viral cause of a human cancer is desirable because it provides unique opportunities for intervention, particularly prevention. This is well illustrated by the recently introduced prophylactic vaccine to prevent infection by the sexually transmitted human papillomavirus (HPV) types most frequently associated with cervical and other anogenital and oral cancers. A common "smoking gun" trait of cancers with a well-documented viral etiology is their increased incidence in immunocompromised individuals (Grulich *et al*., 2007). This is presumably attributable to impaired immunological control over what might otherwise be a harmless infection. Risk of nonmelanoma skin cancer, including cutaneous squamous cell carcinoma (cuSCC), is dramatically increased after immunosuppression, with a more than 50-fold increase in incidence among recipients of solid-organ transplants. It has therefore long been suspected that cuSCC may be caused by an infectious agent. Significant attention has been focused on cutaneous HPVs, particularly those from papillomavirus genus-β (β-HPV), because they ubiquitously and chronically infect the skin and are clearly implicated in skin carcinogenesis in a rare genodermatosis known as epidermodysplasia verruciformis (Feltkamp *et al*., 2008).

The conclusion that the association of mucosatropic HPVs with anogenital cancers is causal is supported by many lines of laboratory and epidemiological evidence. The consistent maintenance of the viral genome, as well as ongoing transcription of the viral oncogenes in essentially every cell of the tumor, is a central component of this causal inference. To date, the evidence suggesting a causal role for cutaneous HPVs in cuSCC in the general population has been mixed, at best (Nindl *et al*., 2007). β-HPV DNA is frequently detected in cuSCC, and some studies have detected higher serological responses to β-HPV virions in cases than in controls. However, viral DNA is also frequently detected in normal skin and, in studies that have evaluated viral genome copy number in cuSCC tumor specimens, this number was usually much less than one viral genome per tumor cell. Because these studies used HPV DNA and serological assays designed to evaluate only a few of the many dozens of known HPV species, they left open the possibility that an unstudied subset of HPV types not targeted by the assays was more strongly associated with cuSCC.

In this issue, Arron and colleagues report their use of the powerful new technology of high-throughput mRNA sequencing to provide an unbiased and comprehensive assessment of the causal role of HPV gene expression in cuSCC. Using random primers, they amplified and sequenced a representative fraction of the total transcriptosome from 31 cuSCC and 8 patient-paired normal skin biopsies, with an impressive mean count of 3.5 million reads per sample. The sequence reads, averaging 54 nucleotides in length, were compared with a database of all known viruses. Only

[1]*Laboratory of Cellular Oncology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA*

Correspondence: John T. Schiller, Laboratory of Cellular Oncology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA. E-mail: schillej@mail.nih.gov