DOCTOR OF PHILOSOPHY

Skin barrier dysfunction in common genetic disorders

Chen, Huijia

Award date:
2011

DOCTOR OF PHILOSOPHY

Skin barrier dysfunction in common genetic disorders

Huijia Chen

2011

University of Dundee

CHAPTER 5

THE EPIDERMAL DIFFERENTIATION COMPLEX (EDC) CONTAINS PSORIASIS SUSCEPTIBILITY GENES

5.1 INTRODUCTION

5.1.1 Psoriasis is caused by genetic changes in the immune system and skin

Psoriasis (PS; OMIM#177900) is a common inflammatory skin disorder affecting approximately 2% of the northern European population (Nevitt and Hutchinson, 1996), 0.123% of the Han Chinese (Shao, 1996), but it is very rare in Africans (Bhalerao and Bowcock, 1998). Despite the common occurrence of this disorder, and a clear genetic component to its incidence, the full picture of the mechanisms responsible for psoriasis has not been uncovered.

Affected people suffer from chronic skin inflammation, manifested by erythematous scaly lesions that can be distributed over the scalp, palms, soles and extensor surfaces of the limbs (Schon and Boehncke, 2005). The nails are typically pitted and show onycholysis. A sero-negative spondyloarthropathy coexists in 40% of cases (Gladman, 1994) and can lead to progressive joint destruction (Nograles *et al.*, 2009); psoriasis patients also have increased risk of stroke and myocardial disease (Gelfand *et al.*, 2006; Gelfand *et al.*, 2009) At the microscopic level, psoriatic skin demonstrates features of abnormal differentiation (hyperkeratosis, parakeratosis) in the upper epidermal layers, and the hyper-proliferative basal layer projects finger-like papillae into the dermis. The proteins that are usually regulated for expression in the basal and

granular layers are abnormally expressed in the thickened spinous layer of psoriatic skin; premature expression of granular layer proteins like involucrin, loricrin, filaggrin and transglutaminases could lead to broad defects in the formation of the stratum corneum and result in parakeratosis (Guttman-Yassky *et al.*, 2009). S100A8 and S100A9 proteins, injury-induced K6, K16 and K17 are also upregulated followed by the triggering of leukocyte infiltration and dermal vessel dilation, causing an amplified inflammatory response (Lowes *et al.*, 2004; Schon and Boehncke, 2005). Thus psoriatic manifestations appear to result from the complex interplay of the immune system, the epidermal barrier and environmental factors. Currently, no curative therapy is available; all treatments that are effective in suppressing symptoms produce significant side effects (Lowes *et al.*, 2007).

The aetiology of psoriasis is incompletely understood (Barker, 1998). While environmental influences, such as streptococcal infection, medications and trauma, play a part in triggering psoriasis (Krueger and Duvic, 1994; Fry and Baker, 2007), greater attention has recently been focused on elucidating genetic factors that influence susceptibility to psoriasis (Bowcock and Cookson, 2004; Bowcock, 2005). Family-based and twin studies indicate a psoriasis heritability rate as high as 90% but identification of the causative gene have proven challenging (Brandrup *et al.*, 1978; Brandrup *et al.*, 1982). The epidemiology of psoriasis is consistent with it being a polygenic disease where multiple genes with small effects interact with one another and with the environment to determine an individual's propensity to develop psoriasis (Bhalerao and Bowcock, 1998). Geographic variations in psoriasis prevalence

are compatible with both genetic explanations that emphasize inter-ethnic differences and environmental variations.

5.1.2 Genome-wide association studies identify psoriasis susceptibility genes in the EDC

More than 10 genome-wide linkage studies have been conducted and have led to the identification of over 20 possible linked regions (Roberson and Bowcock, 2010). Independent studies of the major histocompatibility complex (MHC) in different ethnic groups consistently show an association between psoriasis and the *HLA-Cw6* allele (Tiilikainen *et al.*, 1980; Nair *et al.*, 1997; Trembath *et al.*, 1997; Enlund *et al.*, 1999a; Zhang *et al.*, 2002; Sagoo *et al.*, 2004; Lesueur *et al.*, 2007a). This locus at chromosome 6p21 has been dubbed psoriasis susceptibility locus 1, or *PSORS*1 (Elder *et al.*, 1994a; Elder *et al.*, 1994b; Trembath *et al.*, 1997). Although *PSORS1* is a predominant locus, it seems to be involved in only 30 to 50% of psoriasis cases (Asumalahti *et al.*, 2003). Linkage studies on small groups of individuals have identified other *PSORS* loci, including: *PSORS2* at 17q25 (Tomfohrde *et al.*, 1994), *PSORS3* at 4qter (Matthews *et al.*, 1996), *PSORS4* at 1q21 (Capon *et al.*, 1999), *PSORS5* at 3q21 (Enlund *et al.*, 1999b), *PSORS6* at 19p13 (Lee *et al.*, 2000), *PSORS7* at 1p (Veal *et al.*, 2001), *PSORS8* at 16q (Nair *et al.*, 1997) and *PSORS9* at 4q31-34. Another 12 susceptibility loci have also been postulated (International Psoriasis Genetics Consortium, 2003; Lesueur *et al.*, 2007b). Copy number variation studies have also revealed the importance of the β-defensin genomic copy number (2-12 copies) towards psoriasis

susceptibility (Hollox *et al.*, 2008). An increase in β-defensin copy number has been hypothesized to stimulate keratinocyte proliferation and increase cytokine production by immune cells (Huh *et al.*, 2002). Recently, large GWAS studies have been implemented to identify new susceptibility loci in psoriasis – the study by Stuart et al. consisting of 1,831 cases and 2,546 controls from Europe uncovered three previously unreported genomic regions containing *NOS2*, *FBXL19*, *PSMA6* and *NFKB1A* (Stuart *et al.*, 2010), while another study of 2,622 cases and 5,667 controls conducted by the Genetic Analysis of Psoriasis Consortium and the Wellcome Trust Case Control Consortium reported eight previously unreported genomic loci, of which seven loci contained genes with immune functions (*IL28RA*, *REL*, *IFIH1*, *ERAP1*, *TRAF3IP2*, *NFKBIA* and *TYK2*). Interestingly, this report also provided strong evidence for the interaction between *HLA-C* and *ERAP1* loci, with the *ERAP1* polymorphism exerting its effect on psoriasis risk only in individuals who also carried the *HLA-C* risk allele (Strange *et al.*, 2010).

The results of these studies have proved difficult to replicate, possibly because different susceptibility genes may be significant in different populations. A parallel GWAS study of 8,312 Han Chinese psoriasis cases and 12,919 ethnically matched controls identified six new susceptibility loci (*ERAP1*, *PPTG1*, *CSMD1*, *GJB2*, *SERPINB8* and *ZNF816A*; Sun *et al.*, 2010); presently, a limited number of susceptibility loci have been replicated in all ethnic groups. Comparison of the European and Han Chinese populations only showed the replication of the *PSORS1*, *PSORS9* (Zhang *et al.*, 2002), 9q33-34 (Sun *et al.*, 2007) and *ERAP1* regions (Sun *et al.*, 2010). Moreover,

the existence of more than one major susceptibility gene is likely to decrease the ability to detect linkage (Capon *et al.*, 2000). The actual identities of the susceptibility genes remain tenuous and further fine-mapping and functional analysis is required to identify the causative genes that contribute to the pathogenesis of psoriasis.

PSORS4, localized within chromosome 1q21, first emerged as a potential psoriasis susceptibility locus from American and Italian family studies (Bhalerao and Bowcock, 1998; Capon *et al.*, 1999). Fine mapping with densely spaced markers in 22 families of Italian descent further narrowed the region to a 350 kb interval in the distal portion of the epidermal differentiation complex (EDC) (Capon *et al.*, 2001). The EDC spans 2 Mb in 1q21 and contains at least 45 genes involved in cornification (Mischke *et al.*, 1996; Hoffjan and Stemmler, 2007), a differentiation process that leads to formation of a functional epidermal barrier. These genes can be grouped into 3 families: the cornified envelope precursor proteins, the keratin filament-associated proteins and the S100 calcium binding proteins. The first family includes loricrin (*LOR*), involucrin (*IVL*) and a family of small proline-rich proteins (*SPR*s). During cornification, these proteins are cross-linked by transglutaminases to form a strong, insoluble cornified cell envelope. The second family consists of filaggrin (*FLG*), trichohyalin (*TCHH*), filaggrin2 (*FLG2*), repetin (*RPTN*) (Krieg *et al.*, 1997), hornerin (*HRNR*) (Makino *et al.*, 2001) and cornulin (*CRNN*) (Contzler *et al.*, 2005). These proteins play an important role in bundling keratins to attain flattened squames during cornification. In addition to their structural role, it has been postulated that

many of these proteins also participate in calcium-dependent processes because they possess two EF-hand type calcium-binding domains (Huber *et al.*, 2005). The last family of genes in the EDC belong to the S100 family that include small, calcium binding proteins involved in signaling and cell cycle progression (Kligman and Hilt, 1988; Eckert *et al.*, 2004).

Many disorders that feature abnormal cornification and a breach of the skin barrier have been traced to dysfunctional genes in the EDC. A delayed stop codon in *LOR* was found to cause a variant form of Vohwinkel's syndrome (OMIM #604117), a condition characterised by mutilating keratoderma and ichthyosis (Korge and Krieg, 1996; Maestrini *et al.*, 1996). This was the first EDC gene associated with a skin disease. Recently, *FLG* has been shown to be the major gene responsible for ichthyosis vulgaris (Smith *et al.*, 2006); *FLG* also contributes significantly to the onset of atopic dermatitis (AD) and other atopic diseases in AD patients.

As psoriasis also displays abnormal epidermal differentiation, the genes in the EDC cluster are biologically plausible candidates for psoriasis susceptibility. In psoriatic patients, upregulation of the *S100A7, S100A8* and *S100A9* genes have been observed in the epidermis (Saintigny *et al.*, 1992; Hoffmann *et al.*, 1994) and *LOR* is downregulated (Giardina *et al.*, 2004). *HRNR* was reported to be expressed only in psoriatic and regenerating skin (Takaishi *et al.*, 2005).

5.1.3 Aims of this chapter

- Identify gene(s) within EDC that confer susceptibility to psoriasis in Singaporean Chinese psoriatic patients, using high-density single-nucleotide polymorphism (SNP) genotyping within the 2 Mb EDC region.

5.2 METHODS

5.2.1 Study population

Singaporean Chinese patients with psoriasis were clinically ascertained and recruited through the Psoriasis Clinic of the National Skin Centre, Singapore, and unaffected Singaporean Chinese with a similar age and gender distribution to the cases were selected as controls from stored samples in the Singapore Bio Bank (SBB). Informed consent was obtained under protocols approved by the Institutional Review Boards of the participating institutions.

DNA extraction was performed within the SBB facilities using the Gentra PureGene Cell Kit (QIAGEN, Hilden, Germany, www.qiagen.com).

The first analysis with the Sequenom® platform was performed with DNA extracted from 405 patients and 431 controls, while the second experiment (Infinium™ platform) was carried out with the same set of samples and additional DNA, collected from a total of 494 cases and 588 controls.

5.2.2 Selection of EDC candidate genes and SNPs for genotyping

For Sequenom® analysis, eleven candidate genes were chosen by virtue of their function in skin differentiation and localization to the distal 350 kb of the EDC (**Table 5.1**). All SNPs within these candidates and in 100 kb regions

upstream and downstream of them were identified through the Genome Institute of Singapore (GIS) SNP collective in-house database. A subset of bi-allelic SNPs spaced at approximately 1.5 kb intervals in the region of these candidate genes was selected, with priority given to those which had previously been validated and shown to be present in Han Chinese with minor allele frequency (MAF) ≥ 20% wherever possible. These were then screened with the RealSNP™ tool (Sequenom®, www.realsnp.com) to remove SNPs with ambiguous or highly repetitive flanking sequences.

In order to further analyze the initial findings, we decided to examine the entire EDC with 267 SNPs from the Human HapMap 550v3 panel (Illumina®) that were located in an interval spanning positions 150000000 to 152000000 in chromosome 1 based on NCBI build 36 and dbSNP build 126 data (**Appendix Table 2**).

Table 5.1 List of candidate genes in the epidermal differentiation complex (EDC) chosen for SNP genotyping

Candidate genes/ genomic regions	Chromosomal start	Chromosomal end	Length (kb)	No. of SNPs genotyped
TCHH	150345417	150353180	7.76	5
RPTN	150392695	150398328	5.63	3
Region between *RPTN* and *HRNR*	150398329	150451175	52.8	35
HRNR	150451176	150463296	12.1	10
FLG	150541275	150564303	27.9	16
FLG2	150587837	150599106	11.3	8
CRNN	150648343	150653363	5.02	3
Region between *CRNN* and *IVL*	150653364	151147644	494	17
IVL	151147663	151150985	3.32	4
Region between *IVL* and *SPRR1A*	151150986	151223180	72.2	8
SPRR1A	151223181	151224913	1.73	2
SPRR3	151240849	151242956	2.11	2
SPRR2A	151295220	151296612	1.39	2
LOR	151498803	151501224	2.42	1
			Total	**116**

5.2.3 Genotyping with Sequenom® and Infinium™ platforms

Two sets of assays were performed using different genotyping technologies. SNPs selected from the GIS SNP database were genotyped on 405 patients and 431 controls using Sequenom® MassARRAY technology (Sequenom®, San Diego, CA, USA, www.sequenom.com). The iPLEX™ assay was performed according to manufacturers' instructions (www.sequenom.com). Briefly, PCR and extension reactions were designed by the MassARRAY design software, and were carried out using 5 ng of template DNA. The unincorporated nucleotides in the PCR product were deactivated with shrimp alkaline phosphatase. SNP sites were amplified using MassExtend primers in a termination master mix with differentially labeled ddNTPs. The primer extension products were desalted with the Clean Resin, spotted onto a SpectroCHIP® bioarray and MALDI-TOF MS analysis was performed with MassARRAY Workstation Version 3.3 software. Resulting spectra were analyzed and genotypes were determined using the Sequenom TYPER software.

The Human HapMap SNPs were genotyped with the Infinium™ assay protocol (Illumina®, San Diego, CA, USA, www.illumina.com) on 750 ng of genomic DNA each from 494 cases and 588 controls (including all individuals from the first set) using the standard reagent kit provided. Genotypes were assigned with BeadStudio 3.0 software (Illumina®).

5.2.4 Statistical and stratification analyses

Genotype data derived from the Sequenom® platform was analyzed with HelixTree 2.0 software (Golden Helix, Bozeman, MT, USA, www.goldenhelix.com). Allele-based association studies were performed with missing alleles excluded from analysis. Analysis for conformation with the Hardy-Weinberg equilibrium was performed with a χ^2 test to identify potential errors in genotype calls, and SNPs that yielded *P* values > 0.05 were discarded.

The genotype data from the Infinium™ platform was initially subjected to a quality check with BeadStudio 3.0 and SNPs with call frequencies < 88% and samples with call rates < 96% were discarded. Data from the remaining 469 cases and 585 controls was analyzed with PLINK, a toolset for whole genome association analysis (http://pngu.mgh.harvard.edu/~purcell/plink). 3 filters were applied to the data:

i. a gender concordance check
ii. detection of ethnic outliers using principal components analysis (PCA) as implemented in the program EIGENSTRAT (Price *et al.*, 2006) with the following parameters: MAXPOPS 3, BURNIN 10000, NUMREPS 10000, INFERLAMBDA 0, FREQSCORR 1.
iii. detection of 1st degree relatives (one of each pair was dropped from subsequent analysis for association)

The Cochran-Armitage TREND test was then used for association analysis on the filtered data (derived from 450 cases and 570 controls) while conformation with the Hardy-Weinberg equilibrium was checked using the exact test (Wigginton *et al.*, 2005).

The Infinium™ genotype data from the controls was analyzed for linkage disequilibrium using Haploview 4.0 (http://www.broad.mit.edu/mpg/haploview). Standard filtering procedures were implemented (Hardy-Weinberg, p-value < 0.001, percentage of non-missing genotypes < 75%, number of Mendelian errors > 0, minor allele frequency < 0.001.

I prepared all statistical data with the advice and guidance from Dr Terry Toh, while stratification analysis was conducted with assistance from Dr Terry Toh and Mr Rick Ong.

5.3 RESULTS

5.3.1 Fine mapping of the EDC reveal significant association of psoriasis with the involucrin gene (*IVL*) region

The majority of the psoriasis patients in the study had common plaque psoriasis (471); there were 7 with palmoplantar psoriasis, 9 with pustular psoriasis and 4 with erythrodermic psoriasis. There were 3 individuals with overlapping features (1 instance each of plaque and palmoplantar, plaque and pustular, plaque and erythrodermic). A total of 494 Singaporean Chinese psoriatic individuals were eventually studied; 360 had an age of onset ≤ 40 years while 129 were diagnosed after the age of 40 (there was no clinical data available for 5 individuals).

The EDC locus maps to a 2.0 Mb interval in chromosome 1q21 and contains at least 45 genes whose products are involved in the process of skin cornification (Mischke *et al.*, 1996; Hoffjan and Stemmler, 2007) We selected 11 candidate genes (**Figure 5.1, Table 5.1**) from this region that was previously associated with psoriasis susceptibility (Capon *et al.*, 1999; Sun *et al.*, 2006). After SNP selection and screening, we had an initial set of 94 SNPs extending from positions 150346613 to 151500919 on chromosome 1 (NCBI 36). Once the analysis of this set was completed, a second set of 22 SNPs, clustered in a region that showed statistically significant association with psoriasis, was chosen from the GIS SNP database and genotyped together with

Figure 5.1 SNP-based association analysis of candidate genes within chromosome 1q21

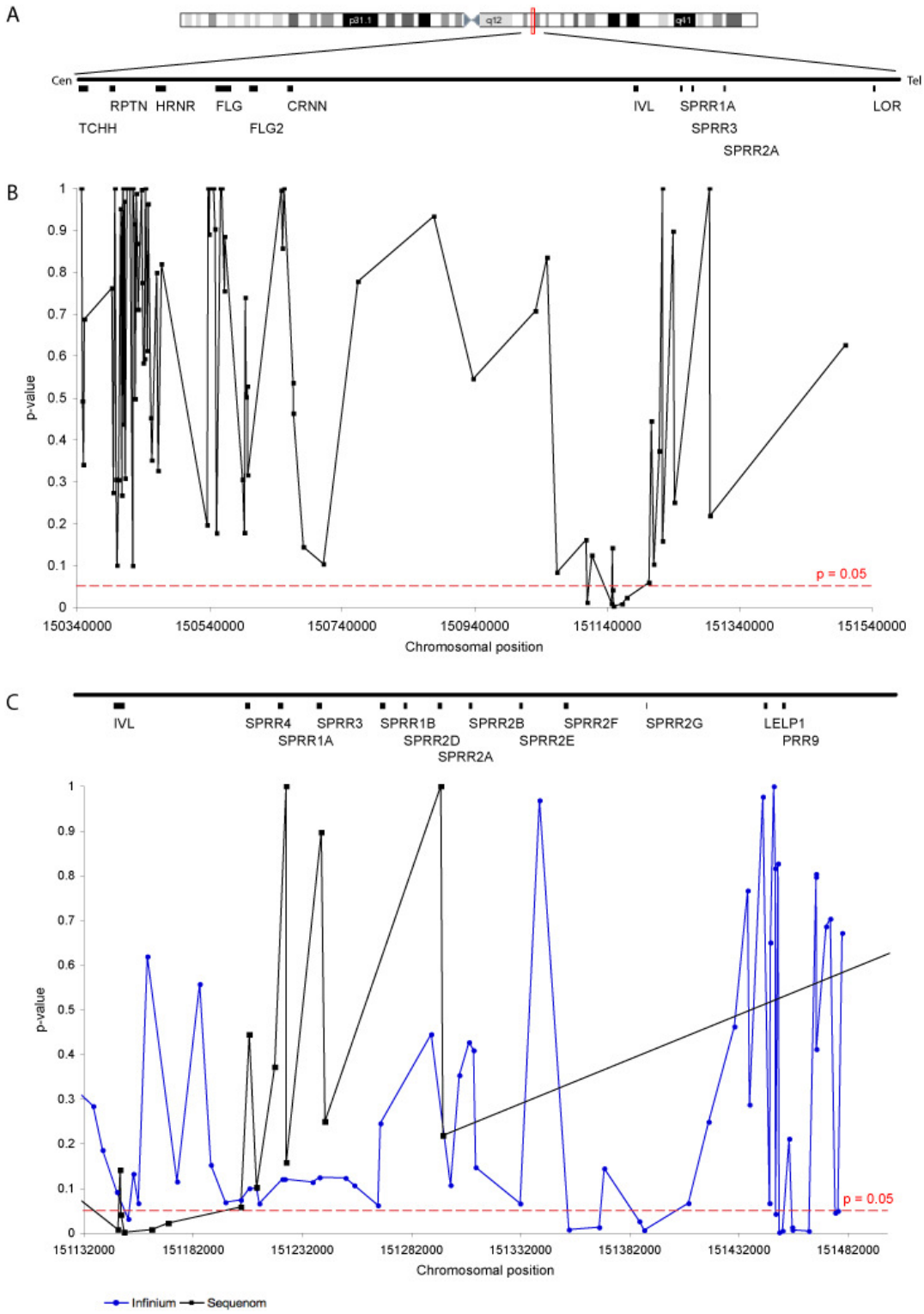


Figure 5.1 SNP-based association analysis of candidate genes within chromosome 1q21

Figure legend

A**.** Schematic diagram of candidate genes chosen for Sequenom® genotyping.

B. *P* value plot of 94 SNPs extending from positions 150346613 to 151500919 on chromosome 1 (NCBI build 36) genotyped by Sequenom®; distribution of all these SNPs conforms to the Hardy-Weinberg equilibrium. The red dotted line indicates the threshold of statistical significance ($P < 0.05$). A cluster of 5 SNPs had *P* values < 0.05 and were mapped either within the 3.3 kb involucrin gene (*IVL*) or in regions flanking it.

C. The Infinium™ system was used to expand the genoptyping of the entire EDC (2 Mb) with 267 SNPs. This figure shows a high resolution map of the region flanked by *IVL* and loricrin (*LOR*) genes and the corresponding *P* value plot of SNPs genotyped with Infinium™ (blue) and Sequenom® (black) platforms showing the close proximity of possible PS susceptibility locus. In addition to the cluster in the *IVL* region, this expanded approach yielded a cluster of SNPs further downstream of *IVL* and positioned within a 120 kb segment of EDC, between the small proline-rich protein 2A gene (*SPRR2A*) and *LOR*.

9 SNPs from the first set. Ultimately, 116 SNPs were genotyped (**Appendix Table 1**).

After discarding 22 SNPs whose genotype data did not conform to Hardy-Weinberg equilibrium, significant association ($P < 0.05$) was detected with 6 SNPs (**Table 5.2**). Strikingly, 5 of these were mapped either within the 3.3 kb long involucrin gene (*IVL*) or in regions flanking it. The lowest *P*-value (0.00157) was obtained with rs913996 located in the 3' UTR of *IVL*, but others were found in the single intron and in exon 2 of the *IVL*.

5.3.2 Expanded analysis further reveal significant association of psoriasis with an additional EDC gene region

When the survey was expanded to the entire EDC, 15 cases and 10 controls that were used in the study with the Sequenom® platform were filtered out as either ethnic outliers or members of 1stdegree relative pairs. Data from 267 SNPs distributed in a 2 Mb interval from positions 150000000 to 152000000 of chromosome 1 (NCBI build 36) were analyzed (**Appendix Table 2**). Statistically significant association ($P < 0.05$) was still obtained with a SNP (rs2879485) mapped to the vicinity of the *IVL* gene. The expanded approach yielded a cluster of SNPs with statistical evidence for association with psoriasis (**Table 5.3**). These SNPs are further downstream of *IVL* and positioned within a 120 kb segment of EDC, between the small proline-rich protein 2A (*SPRR2A*) and loricrin (*LOR*) genes, which were not previously selected for analysis with the Sequenom® platform (**Table 5.1**). One of these

Table 5.2 SNPs showing significant *P* values in Sequenom® genotyping

SNP	Chromosomal location	Gene annotation	Remarks	Major/minor allele	p-value
rs11205128	151110766	-	NA	G/A	0.0109
rs16834751	151147835	*IVL* intron	NA	A/C	0.0075
rs2229496	151149234	*IVL* exon 2	non-synonymous	G/A	0.0408
rs913996	151150735	*IVL* 3' UTR	NA	C/T	0.0016
rs6661932	151163358	downstream of *IVL*	12.3 kb	C/T	0.0071
rs4845497	151170786	downstream of *IVL*	19.8 kb	C/T	0.0220

The chromosomal location is based on NCBI build 36 and dbSNP 126.

NA – Not applicable

Table 5.3 Human HapMap 550v3 SNPs along the EDC within chromosome 1q21 showing significant *P* values in Infinium™ genotyping

SNP	Chromosomal Location	Gene Annotation	Remarks	Major/minor allele	*P* value
rs1011297	150864572	upstream of LCE3A	2639 bp	T/C	0.0458
rs2879485	151152537	downstream of *IVL*	1551 bp	G/A	0.0309
rs1500941	151354285	downstream of *SPRR2F*	1672 bp	A/G	0.00758
rs428913	151368112	upstream of *SPRR2G*	11479 bp	C/T	0.01282
rs533437	151386593	*SPRR2C** intron	NA	A/G	0.02588
rs10494291	151448915	downstream of *LELP1*	4695 bp	G/T	0.04256
rs10494292	151450469	downstream of LELP1/upstream of *PRR9*	6249 bp/6215 bp	T/G	0.00147
rs10788861	151452018	upstream of *PRR9*	4666 bp	C/A	0.00495
rs1410859	151456561	upstream of *PRR9***	123 bp	G/A	0.01235
rs1410860	151456602	upstream of *PRR9***	82 bp	A/G	0.00681
rs12127862	151464064	downstream of *PRR9*	5647 bp	G/A	0.00393
rs10888541	151476212	downstream of *PRR9*	17795 bp	A/G	0.04513
rs4845342	151477432	downstream of *PRR9*	19015 bp	C/A	0.04937

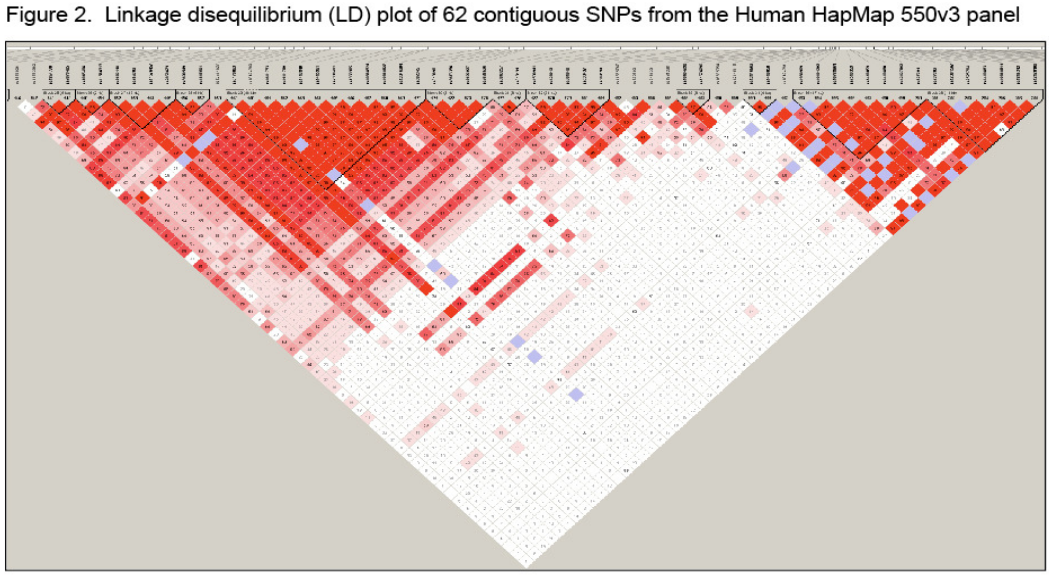
*Unprocessed pseudogene

**Possible promoter region of *PRR*

SNPs is located in the intron of *SPR2C* pseudogene, the rest are found in the neighbourhood of the small proline-rich protein 2F and 2G (*SPR2F*, *SPR2G*), late cornified envelope-like proline-rich protein 1 (*LELP1*) and proline-rich protein 9 (*PRR9*) genes (**Figure 5.1c**).

The linkage disequilibrium (LD) plot for a subset of the EDC SNPs that encompass all that are statistically associated with psoriasis is displayed in **Figure 5.2**. rs2879485 forms a small LD block of 5 kb (block 25) with rs1854779 (350 bp upstream of *IVL*). The remaining SNPs associated with psoriasis separate into 3 well-defined LD blocks (blocks 32, 35 and 36 that span 34, 17 and 11 kb, respectively.

Figure 5.2 Linkage disequilibrium (LD) plot of 62 contiguous SNPs from the Human HapMap 550v3 panel



Linkage disequilibrium (LD) plot of 62 contiguous SNPs from the Human HapMap 550v3 panel, including all that are significantly associated with psoriasis. If you take the SNP position along the top and draw lines parallel to the edge of the triangle to get an intersection, the square represents the probability that these 2 SNPs are inherited together. The squares are coloured red, light red, blue and white in decreasing probability of co-inheritance. Haplotype blocks are determined by computer algorithms. This subset of EDC SNPs is shown along the top of the triangle and extends from positions 151136387 to 151479292 (dbSNP 126). They are grouped into 12 haplotype blocks (blocks 25 to 36) and those that are significantly associated with psoriasis fit into 4 well defined haplotype blocks (blocks 25, 32, 35 and 35 that span 5, 34, 17 and 11 kb, respectively). SNPs that are excluded from the blocks were filtered using criteria specified in the text.

5.4 DISCUSSION AND CONCLUSION

5.4.1 The EDC is an important psoriasis susceptibility locus in Asia

For many years, psoriasis has been perceived and studied primarily as an epidermal condition (Lebwohl, 2003). However, the discovery of cyclosporin and other lymphocyte-selective toxins, as an effective treatment of psoriasis, coupled with the early infiltration of T-cells and macrophages into psoriatic lesions before epidermal changes, altered the pathogenenesis notion psoriasis into that of a chronic inflammatory disorder (Bjerke *et al.*, 1978; Gottlieb *et al.*, 1995). Changes in the amounts of inflammatory cells and cytokines in the dermis are thought to lead to epidermal hyperkeratosis and parakeratosis as a secondary phenomenon (Ghoreschi *et al.*, 2007; Nickoloff *et al.*, 2007). Recent advances in our understanding of atopic dermatitis (eczema) may necessitate a revision of this model. Mutations in filaggrin (*FLG*) have recently been shown to contribute to as many as 50% of moderate-severe atopic dermatitis cases (Sandilands *et al.*, 2007). It has been hypothesised that the abnormal cornification that results from the mutations disrupts the epidermal barrier, allowing percutaneous entry of antigens that trigger the inflammation observed in atopic dermatitis. A similar pathogenetic mechanism where inflammation occurs secondary to epidermal disarray (Albanesi *et al.*, 2007) may explain the association of psoriasis with skin trauma. The search for psoriasis susceptibility genes should therefore be

widened to include genomic regions that play a role in skin cornification (Tschachler, 2007).

The proteins expressed by the genes encoded by the epidermal differentiation complex (EDC), of which *FLG* is one, play a significant role in cornification. The EDC itself maps to chromosome 1q21, the same locus as *PSORS4*, which has previously been reported as a psoriasis susceptibility locus in Europeans (Capon *et al.*, 1999) and Americans (Bhalerao and Bowcock, 1998). However, at the time of our experiment, this locus has not previously been associated with psoriasis in any Asian population. Involucrin, filaggrin and loricrin are the major components of the cornified envelope. Previous analyses of mutations and SNPs in the *FLG* and *LOR* genes indicate that both of them are unlikely to be the gene(s) that confer susceptibility to psoriasis in *PSORS4* (Giardina *et al.*, 2004; Zhao *et al.*, 2007). Interestingly, SNPs from peptidoglycan recognition protein coding region, which lies between *LOR* and *S100A9* genes, have shown association with psoriasis in a family-based Caucasian patient cohort (Sun *et al.*, 2006).

5.4.2 Candidate genes for psoriasis susceptibility lie in close proximity with *IVL* or small proline rich repeat genes (*SPRRs*)

Our preliminary association study of 11 selected genes within the EDC revealed 6 SNPs in the vicinity of *IVL* with significant *P* values ($0.0015 < P <$

0.040) (**Table 5.2**). Involucrin was the first structural protein to be identified as a cornified envelope component and has been reported to be show abnormal expression in psoriatic skin (Ishida-Yamamoto and Iizuka, 1995; Pena-Penabad *et al*., 1999). Involucrin is an early component in the process of skin cornification; it becomes crosslinked to many structural proteins in the epidermal layer, serving as a scaffold to aid the subsequent maturation of the stratum corneum. Regulation of involucrin expression is highly complicated and involves spatially distinct regulatory elements (Crish *et al.*, 2002). The proximal regulatory region (PRR) containing AP1-1 and C/EBP transcription factor binding sites lies directly upstream of exon 1 whereas the distal regulatory region (DRR) between -2473/-1953 include ISE, AP1-5 and Sp1 binding sites (Crish *et al.*, 2006). Three out of 6 SNPs that scored as significant were within *IVL*, one is only 300 bp from DRR on the centromeric side, and two others are on the telomeric side. This region may also harbour transcription factor-binding sites that participate in the complex regulation of the involucrin gene previously alluded to.

Our subsequent expanded analysis of the EDC revealed further interesting insights. Borderline evidence (P = 0.030) was detected with a SNP in the flanking region of *IVL*, and a cluster of 11 SNPs significantly associated with psoriasis ($0.0014 < P < 0.049$) was identified within a 120 kb region of EDC 200 kb distal from *IVL* (**Table 5.3, Figure 5.1C**). This region includes the small proline-rich protein 2F, 2G and 2C (*SPRR2F SPRR2G* and *SPRR2C*), proline-rich protein 9 (*PRR9*) and late cornified envelope-like proline-rich protein 1 (*LELP1*) genes. A literature search revealed that these genes have

been less intensively studied than other EDC components. SPRRs, also known as cornifins, are precursors of the cornified envelope and are more conspicuously expressed in psoriatic skin (Iizuka *et al.*, 2004). *LELP1* has not previously been associated with psoriasis, although SNPs in its proximity are associated with elevated immunoglobulin E levels, a feature of atopic dermatitis (Sharma *et al.*, 2007). The majority of the SNPs are positioned in the surrounding region of *PRR9*, of which rs14100859 and rs1410860 are polymorphisms that could potentially affect promoter activity (**Table 5.3**). Since the majority of the patients were classified as early-onset of plaque psoriasis, further statistical sub-analysis was carried out on patients diagnosed before the age of 40 years old. Interestingly, this revealed significant association ($P < 0.05$) for all *IVL* SNPs, supporting our Sequenom® genotyping data and reaffirming the possible involvement of this region in psoriasis susceptibility (results not shown).

Our data are consistent with previous findings in Caucasians that polymorphisms in precursor proteins of cornified envelope genes are associated with susceptibility to psoriasis (Capon *et al.*, 2001; Sun *et al.*, 2006), thus confirming PSORS4 as a susceptibility locus in Singaporean Chinese as well. Despite the recent discovery of *FLG* as an important gene in the maintenance of an efficient epidermal barrier (Irvine and McLean, 2006), our data confirms, in line with Zhao et al. (2007), that *FLG* is probably not important in psoriasis (**Figure 5.1B**).

Recently, another cluster of late cornified envelope (*LCE*) genes within the EDC have been reported to be candidate genes for psoriasis susceptibility in both European and Chinese populations (de Cid *et al.*, 2009; Zhang *et al.*, 2009). The *LCE* genes span 340 kb are usually expressed in the stratum corneum very late during terminal differentiation and are grouped into three main families (*LCE1*, *LCE2* and *LCE3*). Some individuals with increased psoriasis susceptibility have a deletion of ~30 kb which leads to the lack of *LCE3B* and *LCE3C* gene expression (de Cid *et al.*, 2009). It has been hypothesised that the absence of the *LCE3B* and *LCE3C* genes could cause an impaired repair response in stressed psoriatic skin. In addition, the deletion of the *LCE3B* and *LCE3C* genes coincidentally disrupts a keratinocyte differentiation enhancer, which could contribute to a regulatory loss of skin barrier function and increase psoriasis susceptibility (de Guzman Strong *et al.*, 2010).

Clearly, further work is necessary to determine the interactions between involucrin and the proline-rich and the late cornified envelope proteins that lead to the formation of a functional stratum corneum, and the role that abnormal expression of these genes may play in the development of psoriasis. It will also be interesting to see if our findings are replicated in other populations, especially in neighbouring Asian regions.