

Washington University in St. Louis

Washington University Open Scholarship

All Theses and Dissertations (ETDs)

January 2011

Identification of Genetic and Epigenetic Risk Factors for Psoriasis and Psoratic Arthritis

Ying Liu

Washington University in St. Louis

Follow this and additional works at: <https://openscholarship.wustl.edu/etd>

Recommended Citation

Liu, Ying, "Identification of Genetic and Epigenetic Risk Factors for Psoriasis and Psoratic Arthritis" (2011). *All Theses and Dissertations (ETDs)*. 206.
<https://openscholarship.wustl.edu/etd/206>

This Dissertation is brought to you for free and open access by Washington University Open Scholarship. It has been accepted for inclusion in All Theses and Dissertations (ETDs) by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.

WASHINGTON UNIVERSITY IN ST. LOUIS

Division of Biology and Biomedical Sciences

Program in Human and Statistical Genetics

Dissertation Examination Committee:

Anne M. Bowcock, Ph.D., Chair

Alison M. Goate, D.Phil.

Ted H. Hansen, Ph.D.

Michael Lovett, Ph.D.

John P. Rice, Ph.D.

Nancy L. Saccone, Ph.D.

**IDENTIFICATION OF GENETIC AND EPIGENETIC RISK FACTORS FOR
PSORIASIS AND PSORATIC ARTHRITIS**

by

Ying Liu

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2011

Saint Louis, Missouri

ABSTRACT OF THE DISSERTATION

Identification of Genetic and Epigenetic Risk Factors for Psoriasis and Psoriatic Arthritis

by

Ying Liu

Doctor of Philosophy in Biological and Biomedical Sciences

Human and Statistical Genetics Program

Washington University in Saint Louis, 2011

Dr. Anne M. Bowcock, Chair

Dr. Nancy L. Saccone, co-Chair

Psoriasis (PS) is a common incurable inflammatory skin disease affecting 2–3% of the European population. ~10–30% of patients develop psoriatic arthritis (PsA). Genetic variation in the major histocompatibility complex (MHC) increases risk of developing PS. However, only ~10% of individuals with this risk factor develop PS, indicating that other genetic effects and environmental triggers are important. In order to identify novel susceptibility genes of PS and PsA, I performed the first large scale genome wide association scan for psoriasis susceptibility loci using 233 cases and 519 controls. It revealed that genes of the immune system and of the barrier are associated with psoriasis. The MHC (psoriasis susceptibility 1 or PSORS1) conferred the strongest risk factor for PS and PsA. The study also confirmed recently identified associations with interleukin-23 receptor and interleukin-12B in both PS and PsA. Novel loci with modest effect were also identified, including a region on chromosome 4q27 that contains genes for interleukin 2 and interleukin 21 that has been implicated in other autoimmune

diseases, and seven additional regions that included chromosome 13q13 and 15q21. A follow-up study, aimed to identify potential functional SNPs in the PSORS1 region, implicated an allele-specific repressor role of SNP rs10456057 via binding to nuclear transcriptional factors. Further study with additional PSORS1 SNPs identified “enhancer” activity of the risk allele of SNP rs13191343 in differentiating keratinocytes, and the presence of the PSORS1 risk allele is correlated with CDSN (corneodesmosin) expression, which would affect skin barrier formation. Finally, this thesis also describes the first genome-wide study of altered CpG methylation in psoriatic skin. The study determined the methylation levels at 27,578 CpG sites in skin samples from individuals with psoriasis (12 involved, 8 uninvolved) and 10 unaffected individuals. Involved skin differed from normal skin at 1,108 CpG sites at adjusted p-value < 0.05. Twelve of those CpG sites mapped to the epidermal differentiation complex close to genes that are highly up-regulated in psoriasis. Hierarchical clustering of 50 of the top differentially methylated sites accurately separated all psoriatic skin samples (involved and uninvolved) from normal skin. Methylation at 12 CpG sites was significantly correlated with expression levels of a nearby gene. Taken together, the thesis reveals that the genetic and epigenetic risk factors of psoriasis lead to alterations in genes of skin barrier and immune system which act together to trigger the pathogenesis of the disease.

ACKNOWLEDGEMENTS

My first, and most earnest, acknowledgment must go to my advisor and chair of my thesis examination committee, Dr. Anne Bowcock. It is difficult to overstate my gratitude to Anne. It was such a great honor and pleasant to work with Anne and during the past 6 years, she has been giving me enthusiastic supervision in ensuring my academic, professional, financial, and moral well-being. With her enthusiasm, inspiration, and great efforts to explain things clearly and simply, she helped to make Ph.D. study to be a stimulating and fun place to learn and grow. Throughout my Ph.D. study and thesis-writing period, she provided encouragement, sound advice, good teaching, good company, and lots of inspiring advices. I also salute my co-advisor, Dr. Nancy Saccone and committee members Drs. John Rice, Alison Goate, Ted Hansen and Michael Lovett, for their kind intellectual advice and assistance in many different ways. Far too many people to mention individually have assisted in so many ways during my work at Dr. Bowcock Lab. They all have my sincere gratitude. In particular, I would like to thank Dr. Li Cao, Shenghui Duan, Cynthia Helms, Cailin Joyce, Catherine Jordan and Dr. Elisha Roberson. Dr. Robinson and I had numerous stimulating discussions on the methylation paper and he helped with the re-analyzing of the experimental data and general advices on the drafting the manuscript. I also owe a huge debt of gratitude to our project collaborators who provided skin samples to the study and several anonymous reviews for their help and comments that improved various published papers. My final, and most heartfelt, acknowledgment must go to my wife Su. Her support and encouragement has turned my journey through graduate school into a pleasure.

TABLE OF CONTENTS

	PAGE
ABSTRACT OF THE DISSERTATION	ii
ACKNOWLEDGEMENTS.....	iv
TABLE OF CONTENTS.....	v
LIST OF TABLES.....	viii
LIST OF FIGURES	x
LIST OF ABBREVIATIONS.....	xiii
CHAPTER 1: INTRODUCTION AND LITERATURE REVIEW	1
1.1 Introduction.....	2
1.2 Genetics of Psoriasis and Psoriatic Arthritis.....	4
1.2.1 MHC (PSORS1) and Psoriasis Susceptibility	4
1.2.2 Chromosome 17q25 (PSORS2) and Psoriasis Susceptibility	6
1.2.3 Other non-MHC loci and Psoriasis Susceptibility	7
1.2.4 Genetics of Psoriatic Arthritis.....	8
1.2.5 Current Challenges.....	9
1.3 Epigenetics of Common Diseases.....	10
1.3.1 Overview of Epigenetics.....	10
1.3.2 Epigenetics alterations and Common Diseases.....	10
1.4 Aims of the Dissertation	14
CHAPTER 2: GENOME-WIDE ASSOCIATION STUDY OF PSORIASIS AND PSORIATIC ARTHRITIS	18
2.1 Introduction.....	19
2.2 Materials and Methods.....	21
2.2.1 Subjects.....	21
2.2.2 Genotyping Methods.....	23
2.2.3 Quality Control	24
2.2.4 Statistical Analysis for Association	24
2.2.5 Immunohistochemistry	25

2.3 Results.....	27
2.3.1 Genome-Wide Association Scan	27
2.3.2 Association with MHC Region.....	30
2.3.3 Association with IL23R	32
2.3.4 Association with IL12B	33
2.3.5 Association with Novel Psoriasis Loci	34
2.3.6 Association with Chromosome 4q27	37
2.4 Conclusions.....	42

CHAPTER 3: IDENTIFICATION OF FUNCTIONAL GENETIC

VARIATIONS IN THE PSORS1 REGION.....	54
3.1 Introduction.....	55
3.1.1 Human genetic variations	55
3.1.2 Genetic Variants in the PSORS1 Region.....	56
3.1.3 Experimental assessment of regulatory SNPs.....	58
3.2 Materials and Methods.....	60
3.2.1 Plasmid Construction (Luciferase Assay).....	60
3.2.2 Transient Transfection and Measurement of Luciferase activity	60
3.2.3 Electrophoretic mobility assay.....	60
3.2.4 Luciferase Assay in Mouse Keratinocyte Cells	61
3.2.5 Bioinformatics Evaluation of PSORS1 variants	61
3.3 Results.....	63
3.3.1 Allele-specific repressor activity of SNP9.....	63
3.3.2 Oligos containing SNP9 binds to nuclear extra proteins	64
3.3.3 Systematic evaluation of enhancer/repressor activities of potential rSNPs in PSORS1	65
3.3.4 Genotypes of SNP rs13191343 are associated with CDSN expression.....	66
3.3.5 Bioinformatic Evaluation of PSORS1 Variants.....	67
3.4 Discussion.....	68

CHAPTER 4: GLOBAL METHYLATION CHANGES IN PSORIATIC

SKIN	80
4.1 Introduction.....	81
4.2 Materials and Methods.....	83
4.2.1 Skin biopsy samples.....	83
4.2.2 DNA methylation profiling using bead array	83
4.2.3 Differential Methylation Analysis	84
4.2.4 Selection of the top 50 group discriminating differentially methylated CpG sites	85
4.2.5 Correlation with gene expression.....	86
4.2.6 Pyrosequencing.....	86
4.3 Results.....	87
4.3.1 Differential CpG site methylation in psoriatic skin	87
4.3.2 Validation of differential methylation with pyrosequencing.....	89
4.3.3 Methylation levels correctly classify involved, uninvolved, and normal skin samples	90
4.3.4 Correlation of methylation with nearby gene expression	93
4.4 Discussion.....	96
CHAPTER 5: SUMMARY AND FUTURE DIRECTIONS.....	112
5.1 Summary.....	113
5.2 Future Perspectives on the Genetics of Psoriasis.....	120
5.3 Future Perspectives on the Epigenetic Study of Psoriasis	122
5.4 Integrated Genetic and Epigenetic Approach	126
References.....	131
APPENDIX I: SUPPLEMENTARY TABLES/FIGURES FOR CHAPTER 2	169
APPENDIX II: SUPPLEMENTARY TABLES/FIGURES FOR CHAPTER 4.....	172

LIST OF TABLES

		PAGE
Chapter 1:		
Table 1.1	Locations of psoriasis susceptibility (PSORS) loci and gene candidates in each interval	16
Chapter 2:		
Table 2.1	Summary of association with previously reported PS susceptibility loci (MHC, IL23R and IL12B) in U.S. PS cohort (810 cases, 1256 controls)	44
Table 2.2	Summary of association with previously reported PS susceptibility loci in U.K. PsA cohort (576 cases, 480 controls)	45
Table 2.3	Potential novel loci from GWA scan	46
Table 2.4	Potential Association results at chromosome 4q27 in PsA and PS cohorts (U.S.: 810 PS cases, 1256 U.S. controls; U.K.: 576 PsA cases, 480 controls)	47
Chapter 3:		
Table 3.1	Sequences of the plasmid constructs for luciferase assay in mouse keratinocyte cells	71
Chapter 4:		
Table 4.1	CPG sites exhibiting two-fold differences in methylation status in involved versus normal/uninvolved and involved/uninvolved versus normal skin	99
Table 4.2	CpG sites that correlated with expression levels	100
Table 4.3	Top functional pathways identified by MetaCore pathway analysis	101

LIST OF TABLES (Continues)

		PAGE
Chapter 5:		
Table 5.1	Psoriasis-associated genes identified by GWAS studies	128
Appendix I:		
Table A1.1	Summary of cases and controls used in discovery and replication stages	170
Appendix II:		
Table A2.1	Sample information	173
Table A2.2	Pyrosequencing assay information	174
Table A2.3	Top 50 sites that best discriminate involved from normal skin	176

LIST OF FIGURES

	PAGE
Chapter 1:	
Figure 1.1	Flow diagram illustrating the integrated approach used in this thesis to identify genetic/epigenetic risk factors for psoriasis 17
Chapter 2:	
Figure 2.1	Summary of genome-wide association scan results for all cases and the PsA subgroup 48
Figure 2.2	Association localization plots for the MHC following discovery and replication phases 49
Figure 2.3	Association localization plots for the ILI23R region on chromosome 1 50
Figure 2.4	Association localization plots for novel replicated region on chromosome 13 51
Figure 2.5	Immunostaining of normal, non-lesional and lesional skin for IL2, IL21, COG6 and SPPL2A proteins 52
Figure 2.6	Association localization plots for novel replicated region on chromosome 15 53
Chapter 3:	
Figure 3.1	Repressor activity of SNP9 (rs10456057) 72
Figure 3.2	Electrophoretic mobility shift and competition assays with Jurkat nuclear cell extract for allelic variants of SNP9 and SNP7 73
Figure 3.3	Scheme for candidate rSNPs and nearby genes in PSORS1 region 74
Figure 3.4	Reporter activity of PSORS1 alleles in keratinocytes 76
Figure 3.5	CDSN gene expression is associated with genotypes of PSORS1 SNPs 78

LIST OF FIGURES (Continues)

		PAGE
Chapter 4:		
Figure 4.1	Venn diagram of the CpG sites exhibiting differential methylation for each of three contrasts using a significance cutoff of 0.05 for the adjusted p-value	102
Figure 4.2	Differentially methylated CpGs that map to the epidermal differentiation complex (EDC)	103
Figure 4.3	Pyrosequencing data in PP, PN and NN skin biopsies at 9 CpG sites in the gene IFI27	104
Figure 4.4	Pyrosequencing data for C10orf99 in PP, PN and NN skin biopsies at 3 CpG sites	105
Figure 4.5	Pyrosequencing data for LGALS3BP in PP, PN and NN skin biopsies at 7 CpG sites	106
Figure 4.6	Pyrosequencing data for SERPINB4 in PP, PN and NN skin biopsies	107
Figure 4.7	Heatmap showing PP, PN, and NN samples clustered using the top 50 CpG sites that differentiate PP from NN skin	108
Figure 4.8	The top scored network identified by using the top 50 differential methylation genes	109
Figure 4.9	Boxplots of methylation levels in three sample groups	110
Figure 4.10	Correlation between methylation level of CpG sites and expression of C10orf99 (A) and KYNU (B) genes	111
Chapter 5:		
Figure 5.1	An integrated genetic and epigenetic approach to psoriasis and PsA	130

LIST OF FIGURES (Continues)

		PAGE
Appendix I:		
Figure A1.1	Quality control measures for discovery study	171
Appendix II:		
Figure A2.1	Hierarchical clustering analysis of psoriatic skin samples	178
Figure A2.2	Comparison of raw and normalized distributions for β - and M-values	179
Figure A2.3	Theoretical power for detecting methylation changes	180

LIST OF ABBREVIATIONS

<i>AIM</i>	<i>ancestry informative SNPs</i>
<i>ANOVA</i>	<i>analysis of variance</i>
<i>CDGE</i>	<i>common disease genetic and epigenetic (hypothesis)</i>
<i>CDSN</i>	<i>corneodesmosin</i>
<i>CEPH</i>	<i>Centre du Etude Polymorphisme Humain</i>
<i>CGIs</i>	<i>CpG islands</i>
<i>CI</i>	<i>confidence interval</i>
<i>CNV</i>	<i>copy number variation</i>
<i>dbSNP</i>	<i>NCBI's SNP database</i>
<i>DCs</i>	<i>dendritic cells</i>
<i>DNMT</i>	<i>DNA methyltransferase</i>
<i>EDC</i>	<i>epidermal differentiation complex</i>
<i>EMSA</i>	<i>electrophoretic mobility shift assay</i>
<i>FDR</i>	<i>false discovery rate</i>
<i>GAIN</i>	<i>genome wide association network</i>
<i>GRR</i>	<i>genotype relative risk</i>
<i>GWAS</i>	<i>genome-wide association studies</i>
<i>HDAC</i>	<i>histone deacetylase</i>
<i>HLA</i>	<i>human leukocyte antigen</i>
<i>HWE</i>	<i>Hardy-Weinberg equilibrium</i>
<i>KIR</i>	<i>killer cell immunoglobulin-like receptors</i>
<i>LD</i>	<i>linkage disequilibrium</i>
<i>LOD</i>	<i>logarithm of the odds</i>
<i>MHC</i>	<i>major histocompatibility complex</i>
<i>miRNA</i>	<i>micro RNA</i>
<i>nsSNP</i>	<i>non-synonymous coding SNP</i>
<i>OR</i>	<i>odds ratio</i>
<i>PASI</i>	<i>psoriasis area severity index</i>
<i>PS</i>	<i>psoriasis</i>
<i>PsA</i>	<i>psoriatic arthritis</i>
<i>PSORS1</i>	<i>psoriasis susceptibility 1</i>
<i>rSNP</i>	<i>regulatory SNP</i>
<i>SNP</i>	<i>single nucleotide polymorphism</i>
<i>TF</i>	<i>transcription factor</i>
<i>TNF</i>	<i>Tumor necrosis factor</i>
<i>UTR</i>	<i>un-translated region</i>

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

1.1 Introduction

Psoriasis (PS) is a common chronic inflammatory skin disease that affects 2–3% of the Caucasian population. It is less common in individuals of Asian descent (0.1% or less), and is exceedingly rare in Africa [1]. It frequently develops in early adulthood in individuals in their twenties, although individuals of all ages can be affected [2]. The disease is characterized by thickened, scaly skin patches or psoriatic plaques, caused by abnormal keratinocyte proliferation and infiltration of inflammatory cells into the dermis and epidermis. Patients have a natural history of outbreaks (flares) followed by temporary remissions. Approximately 10–30% of psoriatic patients also develop psoriasis arthritis (PsA), which affects joints and surrounding tissues, markedly impacts mobility and can cause irreversible joint destruction [3]. Some of the earliest changes in the pathogenesis of psoriasis are thought to be mediated by an array of environmental triggers, including HIV infection [4], use of drugs such as lithium, beta-blockers, or anti-malarials, and the withdrawal of corticosteroids [5]. Although these are diverse stimuli, all might be viewed as trigger factors that can activate cellular immunity, either through innate or acquired pathways. Psoriasis is believed to be a T-cell mediated ‘Type-1’ autoimmune disease [6]. Gene expression changes in psoriasis lesions have been well documented, and strongly support an important role for tumor necrosis factor and interferon gamma signal pathways in its pathogenesis [7,8]. One theory for the development of psoriasis is that T cells and dendritic cells elaborate inflammatory cytokines and chemokines to create an environment in the skin that stimulates proliferation of resident keratinocytes and endothelial cells in genetically susceptible individuals, producing a pattern of growth recognized as psoriasis. It is expected that

knowledge of predisposing variants in susceptible individuals and exposure to environmental triggers will eventually explain the numerous changes that exist in psoriatic skin. PS and PsA are serious but poorly understood diseases. There are no cures and they require sophisticated medical care and treatments. Moreover, having psoriasis increases the risk of heart disease and stroke [9,10].

1.2 Genetics of Psoriasis and Psoriatic Arthritis

Although psoriasis often appears sporadically, the existence of a genetic contribution of psoriasis is well accepted. It is a complex disease that includes familial components. Individuals with an affected family member are at increased risk of developing the disease, and siblings of an individual with psoriasis have a 4–6-fold increased risk of developing psoriasis compared to the general population [1]. Previous twin studies in psoriasis report that concordance of the disease in monozygotic twins is much higher than in dizygotic twins, being approximately 72% and 15–23%, respectively, for northern European individuals. These results indicate that genetic components play a role in predisposition to the disease, and it is estimated that the heritability for psoriasis is between 60% and 90% [11]. Recent developments in genetic analysis have provided better understanding of the fundamental biological pathways in disease susceptibility. Currently, few genes for psoriasis have been conclusively identified, although genome-wide linkage/association scans have revealed over 20 candidate loci during the past years [12] (**Table 1.1**).

1.2.1 MHC (PSORS1) and Psoriasis Susceptibility

Among the reported psoriasis susceptible loci, the locus in the major histocompatibility complex (MHC) class I region on chromosome 6p21.3, known as PSORS1 (psoriasis susceptibility 1) consistently identified in both linkage and association genome-wide scans, accounting for one-third to one-half of the genetic liability to psoriasis. Over 30 years ago, this region was found to harbor human leukocyte antigen (HLA) genes that associated with autoimmune diseases [13]. Psoriasis was found to be associated with

HLA-C and several HLA-B alleles; however, the association with HLA-B was later determined to be due to strong extended haplotypes and linkage disequilibrium with HLA-C [14]. This region was subsequently identified by linkage analysis in 1997 [15,16] and replicated in numerous populations. The strong association of HLA-Cw6 allele in this locus with psoriasis was first reported in the Finnish population over 26 years ago [13]. In these early studies, which were performed in Northern European populations, the frequency of HLA-Cw6 was ~46% in cases with psoriasis vulgaris and 7.4% in controls. However, the exact location of PSORS1 gene remains controversial owing to extensive linkage disequilibrium across this region. Two candidate genes lying just telomeric to HLA-C were appealing and have been intensively studied with respect to their role in psoriasis susceptibility. One is HCR (helix coiled coil rod homolog) [17]; the other is corneodesmosin (CDSN). CDSN is expressed in terminally differentiated keratinocytes and in the inner root sheath of hair follicles [18,19,20]. It localizes to the modified desmosomes of keratinocytes in the stratum granulosum and stratum corneum. The serine and glycine-rich terminal domains of CDSN that are essential for cell adhesion are sequentially cleaved during skin desquamation [21,22]. Apart from the skin, CDSN mRNA is only detected significantly in the placenta and in the thymus [23]. Several studies indicate that PSORS1 is likely to lie very close to HLA-C, and distinct from the region harboring CDSN and HCR. Two single-nucleotide polymorphisms (SNPs) (SNP9 rs10456057 and SNP7 rs12208888), lying ~5.7 and ~10 kb, respectively upstream from the start site of HLA-C exhibit stronger association with psoriasis than any other SNP in the region [24]. A study from our group looking at both SNPs and classical HLA alleles revealed that haplotypes harboring HLACw*0602 and HLA-Cw*1203 were over-

transmitted to affecteds, suggesting that a variant or combination of variants exclusive to these haplotypes were predisposing [25]. Interestingly, one of these variants leads to acquisition of an additional site for the transcription factor RUNX, within intron 7. This is highly correlated with SNP9 (rs10456057) and SNP7 (rs12208888) polymorphisms ($r^2=1$).

1.2.2 Chromosome 17q25 (PSORS2) and Psoriasis Susceptibility

Although PSORS1 is the only locus for psoriasis identified in all genetic studies to-date and is generally understood to confer the most risk for psoriasis, harboring HLA-Cw6 was not sufficient to develop disease, and the penetrance of this allele was estimated to be only 10% [11], indicating that other genetic/environmental factors may also contribute to the liability of the disease. Numerous non-MHC susceptibility loci also have been identified by linkage/association approaches. Approximately 10 genome-wide linkage scans, primarily with polymorphic microsatellites, have been conducted in psoriasis. This has led to the identification of over 20 possible linked regions [26]. **Table 1.1** lists the identified psoriasis susceptibility loci and gene candidates, along with the literature reference. PSORS2 on chromosome 17q25 (autosomal dominant) was the first identified as non-MHC locus that confers susceptibility to psoriasis [27]. Our group initially identified this region by genome-wide linkage scan on eight Caucasian affected families, among which PS1 family (19 affected and 12 unaffected members) showed strong linkage to D17S784 marker with a maximal two point logarithm of the odds (LOD) score of 5.33, with high penetrance [27]. Evidence from a recent linkage study on a single large-pedigree in Taiwan replicated our linkage findings, mapping this psoriasis

susceptibility locus to the distal end of chromosome 17q [28,29]. Subsequent sequencing of 78 genes within this region identified a c.-625A>C mutation in ZNF750, a putative C2H2 zinc-finger protein, resulted in a 42% reduction of the promoter activity [30]. Overall, the mutation accounted for 1.7% of the psoriasis in the Chinese population although recent evidence suggests that it is a rare variant that is unrelated to disease susceptibility. In addition, a third independent PSORS2-linked Israeli Jewish Moroccan family was also recently identified [31]. The disease in this family was linked to D17S928 (maximum multipoint LOD score of 8.79 at $\theta=0$) and segregates with a frame-shift mutation in ZNF750. Although mutations in ZNF750 could contribute to psoriasis susceptibility, the general applicability of their impact in the common form of psoriasis remains to be elucidated.

Finally, a five-marker variant in a region on chromosome 17q25 harboring SLC9A3R1 and NAT9 was identified with association mapping [32]. One psoriasis-associated allele from this five marker haplotype leads to loss of a putative site for the RUNX family of transcription factors. This is of interest as RUNX1 and RUNX3 play a major role in hematopoietic development and thymic selection, and alterations of RUNX binding sites have also been reported as susceptibility variants for systemic lupus erythematosus and rheumatoid arthritis [33,34].

1.2.3 Other non-MHC loci and Psoriasis Susceptibility

Other known potential non-MHC loci include SLC12A8, (solute carrier family 12 (bumetanide-sensitive Na-K-Cl co-transporters), member 8, known as PSORS3), epidermal differentiation complex (EDC, known as PSORS4) region on chromosome

1q21, as well as PSORS9 on chromosome 4q28-q31, and so on (**Table 1.1**). SLC12A8 was the first gene proposed to be associated with psoriasis susceptibility [35], and it has been replicated in a recent study [36]. This gene is a member of the cation-chloride-coupled cotransporter gene family [37], although the substances it transports are unknown. Finally, chromosome 1q21 region (PSORS4) harbors a cluster of genes, named as the epidermal differentiation complex (EDC), spanning over 2 Mb. Evidence of linkage to this region was first described in Italian families [38], although suggestive evidence for linkage was seen at the same time in our cohort of multiplex families from the US, and was replicated in our cohort of affected sibship pairs [1].

1.2.4 Genetics of Psoriatic Arthritis

Moll and Wright were the first to demonstrate familial aggregation of PsA, and estimated the recurrence risk ratio in first-degree relatives (λ_1) to be 55 [39], compared with estimates ranging from 5 to 10 in cutaneous psoriasis, implicating a strong genetic basis in PsA. A more recent study has estimated the λ_1 to be 47 in a British population [40]. HLA antigens were identified as prognostic factors in patients with PsA, and polymorphisms in the genes coded in the HLA region on chromosome 6p have been shown to be associated with PsA [41]. Karason *et al.* published their genome-wide linkage study in PsA, obtaining a LOD score of 2.17 on 16q, which is close to PSORS8 locus for psoriasis. Other associated loci for PsA outside the MHC region were also reported elsewhere [42,43,44], such as the IL-1 gene cluster on chromosome 2q and KIR (Killer cell Immunoglobulin-like Receptors) genes on 19q13.4. KIR genes encode a family of inhibitory and activating receptors expressed by most natural killer (NK) cells

and small subpopulations of T cells. There are two groups of KIR gene-family haplotype, known as A and B. Whereas the group A KIR haplotype mainly encodes inhibitory KIRs, the group B KIR haplotypes encode more activating KIRs. It has been recognized that various combinations of MHC class I molecules and KIRs are dominant in the regulation of human NK cells and thus correlate with susceptibility to autoimmune diseases [45].

1.2.5 Current Challenges

As with all complex diseases, linkages and associations with psoriasis have not always been replicated in other cohorts. This can be due to low effect size of individual genetic variation, ascertainment bias, gene–environment interactions and other confounding factors. Besides, the functional relationship between predisposing genetic variation and the mechanism of interaction between environmental trigger factors and genetic effects still remain unknown, which, however, provides further evidence for the complex basis of this disease. A more systematic approach, with the capability of integrating and analyzing different sources of biological information, i.e. phenotype/genotype data, gene expression profiling and epigenetic changes, is described in this dissertation as a means of attempting to understand the pathogenesis of the disease better (**Figure 1.1**).

1.3 Epigenetics of Common Diseases

1.3.1 Overview of Epigenetics

It has been widely recognized that classic genetics alone is not sufficient to explain the diversity of phenotypes within a population. Nor does classic genetics explain how, despite their identical DNA sequences, monozygotic twins or cloned animals can have different phenotypes and different susceptibilities to a disease [46,47]. The concept of epigenetics offers a partial explanation of these phenomena. It focuses on non-sequence mediated forms of gene regulation and heredity. The term epigenetics is derived from epi (above and beyond) and genetics, and is commonly used to describe heritable biological information that is not encoded in the DNA sequence. Epigenetic mechanisms have many layers of complexity, including DNA methylation by DNA methyltransferases (Dnmts), histone modifications such as methylation, acetylation, and phosphorylation, structural modifications of chromatin, and microRNAs as well as other noncoding regulatory RNA [48]. Epigenetic changes are crucial for the development and differentiation of the various cell types in an organism, as well as for normal cellular processes such as X-chromosome inactivation in female mammals [49]. However, epigenetic states can become disrupted by environmental influences or during aging, and the importance of epigenetic changes in the development of cancer and other common diseases is increasingly being appreciated.

1.3.2 Epigenetics alterations and Common Diseases

The best-known epigenetic marker is DNA methylation, the addition of a methyl group to DNA at the 5-carbon of the cytosine pyrimidine ring. It has critical roles in the control of

gene activity and the architecture of the nucleus of the cell. In humans, DNA methylation commonly occurs in cytosines that precede guanines; these are called CpG dinucleotides [50]. Approximately 70-80% of the CpG dinucleotides in the human genome are methylated, predominately in areas harboring repetitive sequences [51]. However, regions rich in CpGs, termed CpG islands (CGIs) are also found in promoters of more than 70% of annotated genes [52,53]. These islands are usually not methylated in normal cells [54,55]. The methylation of particular subgroups of promoter CpG islands can, however, be detected in normal tissues.

It has been well accepted that DNA methylation acts in concert with other epigenetic mechanisms, such as histone modification, to regulate normal gene expression and facilitate chromatin organization within cells. Such regulation may be part of normal developmental or differentiation processes but can also be triggered by environmental factors [56,57,58,59,60]. These and other demonstrations of how epigenetic changes can modify gene expression have led to human epigenome projects and epigenetic therapies [61,62,63]. Bell *et al.* recently reported that methylation of CpG dinucleotides is inversely correlated with gene expression in the human genome [64]. By comparing methylation levels to estimates of gene expression, the authors found a significant negative correlation (mean rank correlation $r = -0.454$) between methylation and gene expression levels across 11,657 genes in HapMap cell lines [64]. In addition, the drop in methylation levels near to the transcriptional start site (TSS) was only observed in highly expressed genes.

On the other hand, aberrant DNA methylation patterns are often observed in human common diseases [65], especially for cancer development [66,67]. Such events

are often accompanied by alterations in chromatin structure at gene regulatory regions. It is believed that during carcinogenesis, epigenetic switching and reprogramming result in the aberrant hypermethylation of CpG islands, reducing epigenetic plasticity of critical tumor suppressor genes, and rendering them unresponsive to normal stimuli. One of the first epigenetic alterations found in human cancer was the global hypomethylation of DNA in human tumors as compared with the level of DNA methylation in their normal-tissue counterparts [68]. A recent large-scale study of DNA methylation with the use of genomic microarrays has detected extensive hypomethylated genomic regions in gene-poor areas [60]. During the development of a neoplasm, the degree of hypomethylation of genomic DNA increases as the lesion progresses from a benign proliferation of cells to an invasive cancer [69]. Moreover, another major event in the origin of many cancers is hypermethylation of the CpG islands in the promoter regions of tumor-suppressor genes [70]. These changes occur at different stages in the development of cancer and can affect genes involved in the cell cycle, DNA repair, the metabolism of carcinogens, cell-to-cell interaction, apoptosis, and angiogenesis, all of which are involved in the development of cancer [67,71]. Recent studies on the profiles of microRNA (miRNA) expression suggested a third possible mechanism of epigenetic lesion in tumorigenesis. DNA hypermethylation in the miRNA 5' regulatory region may account for the inactivation of miRNA in tumors [72,73].

One new frontier in the study of the epigenetics of human diseases is to establish its potential role in common non-neoplastic human diseases, especially the autoimmune diseases [74]. Epigenetic mechanisms are essential for the function of immune system. Moreover, a failure to maintain epigenetic homeostasis in the immune response due to

factors including environmental influences, leads to aberrant gene expression, contributing to immune dysfunction and in some cases the development of autoimmunity in genetically predisposed individuals. For instance, aberrant hypomethylation is found in T cells of patients with systemic lupus erythematosus, including in genes such as lymphocyte function-associated antigen-1, which is overexpressed in lupus T cells [75]. Another study on epigenetic changes in the blood of systemic lupus erythematosus patients revealed altered methylation of several genes contributing to T-cell autoreactivity, B-cell overstimulation and macrophage killing [76]. Compared to many other autoimmune diseases, psoriasis is more tractable due to the accessibility of its target organ: the skin. There have been a few reports of altered methylation within promoters of single genes of psoriatic skin. For example, the *SHP-1* (PTPN6) promoter is reported to be demethylated in psoriatic skin but not in Atopic Dermatitis (AD) or normal skin [77]. However, global methylation changes in psoriasis have not been described.

1.4 Aims of the Dissertation

The aims of the present dissertation are threefold as shown in **Figure 1.1**. Aim 1 was to perform a genome-wide association study (GWAS) of psoriasis and psoriatic arthritis. This is described in Chapter 2 where the major findings of this aim are highlighted. Unlike many Mendelian disorders in which the disease alleles are rare and of catastrophic effect with high penetrance, the alleles underlying complex genetic disorders, like those of psoriasis, are relatively common and make low to modest contributions to disease risk, rendering them difficult to identify by linkage study [78]. In this setting, tests of association are much more powerful than tests of linkage, provided causal variants or proxies for them can be genotyped [79]. With the advent of HapMap project, millions of genetic markers, or single nuclear polymorphisms (SNPs), have been identified throughout genome [80]. By comparing the allele frequencies of each marker in cases and controls, association with a polymorphism can be established, allowing more precise localization of a risk variant for further sequence analysis [81]. This was being done for a number of common diseases, where linkage analysis had not been fruitful [82]. In this GWAS, we scanned more than 300,000 SNPs in the genomes of 223 psoriasis patients, including 91 who had psoriatic arthritis. We compared the DNA variations in people with psoriasis to those found in 519 healthy control patients, looking for specific differences that may be linked to the disease. The initial findings were further confirmed in an independent replication study with larger set of patients and healthy controls.

The Aim 2 of this thesis was to perform follow up functional analyses to identify the PSORS1 variant. Chapter 3 describes a systematic approach for screening the PSORS1 genetic variants for enhancer activities by transfecting the construct plasmid

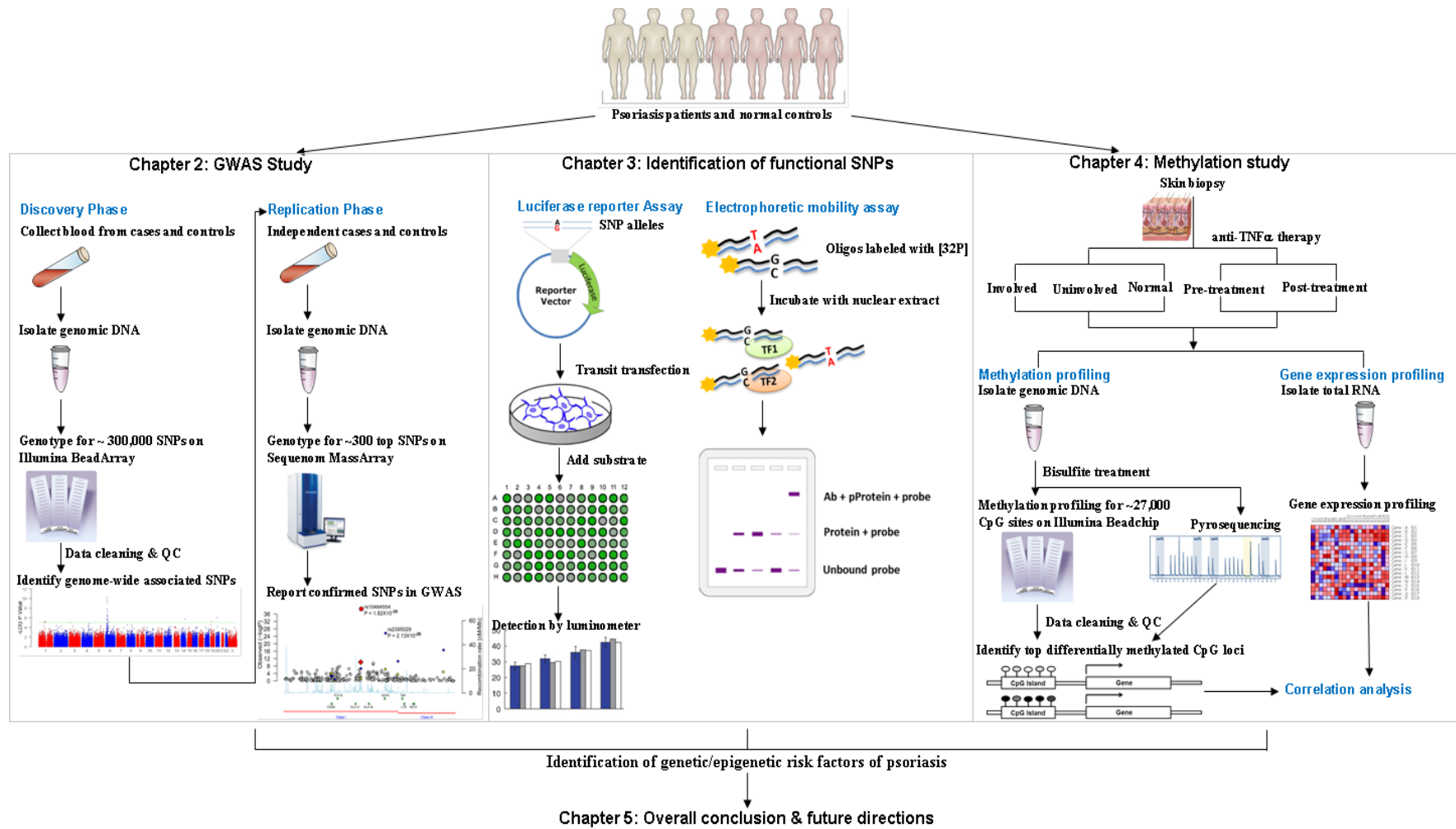
that contains the different alleles of these SNPs into keratinocytes. This approach allowed us to investigate the regulatory activity of PSORS1 risk alleles in both proliferating and differentiating keratinocytes (see Methods in Chapter 3). Since psoriasis is a defect of differentiation, we hypothesized that the PSORS1 allele would show a difference in differentiating keratinocytes. In addition, correlations of PSORS1 risk alleles with expression levels of nearby candidate genes were also analyzed and reported.

Finally, Aim 3 of this thesis was to perform a global study of altered CpG methylation in psoriatic skin. Since the genetic variants only account for about 60% of phenotypic variance of psoriasis [11], we hypothesized that both genetic and environmental can modify the epigenetic changes in psoriatic skin, which in turn are attributable to the pathogenesis of the disease. Chapter 4 describes profiling of global changes of methylation in involved psoriatic skin versus uninvolved and normal skin by querying 27,578 CpG sites with Illumina bead arrays. Significant methylation changes at certain CpG sites were identified and well characterized. Hierarchical clustering analyses of the top 50 differentiating loci were used to classify the lesion and normal samples. The correlations of methylation level at specific CpG loci with expression levels of a nearby gene were further examined, and finally, global changes in methylation and expression status as a consequence of treatment with a TNF-alpha inhibitor were assessed and reported in Chapter 4. In Chapter 5, I provide an overall conclusion of my findings and discuss future directions. The flow diagram showed in **Figure 1.1** outlines the integrative approach that was used in this dissertation to explore both genetic and environmental effects through epigenetic modifications of psoriasis.

Table 1.1 Locations of psoriasis susceptibility (PSORS) loci and gene candidates in each interval.

Locus	Chrom Location	Candidate Genes	References
<i>PSORS1</i>	6p21.3	<i>HLA-C, CDSN, HCR, HERV-K, HCG2, POU5F1, TCF19, CCHCR1, LMP, SEEK1, SPR1</i>	[16,83,84,85,86,87,88,89,90,91]
<i>PSORS2</i>	17q25	<i>RUNX1, RAPTOR, SLC9A3R1, NAT9, TBCD</i>	[27,32,92,93,94,95]
<i>PSORS3</i>	4q34	<i>IRF-2</i>	[96,97]
<i>PSORS4</i>	1q21	<i>S100 genes and late cornified envelope within EDC, Loricrin, Filaggrin, Pglyrp3,4</i>	[38,98,99,100,101,102,103,104]
<i>PSORS5</i>	3q21	<i>SLC12A8, cystatin A, zinc finger protein 148</i>	[35,36,105,106]
<i>PSORS6</i>	19p13	<i>JUNB</i>	[107,108,109,110]
<i>PSORS7</i>	1p13	<i>PTPN22</i>	[111,112]
	1p32.1-31.2	<i>IL23R</i>	
<i>PSORS8</i>	16q13	<i>CX3CL1/CX3R1, NOD2/CARD15</i>	[113,114]
<i>PSORS9</i>	4q31.2	<i>IL-15</i>	[115,116]
<i>PSORS10</i>	18p11		[117,118]

Figure 1.1 Flow diagram illustrating the integrated approach used in this thesis to identify genetic/epigenetic risk factors for psoriasis. Abbreviation: GWAS, genome-wide association study.



CHAPTER 2
GENOME-WIDE ASSOCIATION STUDY OF PSORIASIS
AND PSORIATIC ARTHRITIS

2.1 Introduction

Psoriasis (PS) is a chronic inflammatory disease of the skin affecting 2-3% of the population [119]. Approximately 25% of patients also develop psoriatic arthritis (PsA), a common, debilitating auto-immune disease belonging to the family of spondyloarthritides [120,121]. The recurrence risk (λ_s) of PsA is high, and estimates of 27-47 have been proposed [40,122]. This is much higher than the estimated λ_s of PS which is estimated to be between 4 and 11 [1].

PS and PsA are interrelated disorders, and the prevalence of PS is 19 times higher among first degree relatives of probands with PsA compared with the general population [123]. The pathogenesis of PS and PsA is complex, involving both genetic and environmental risk factors. Strong association of PS with the MHC class I region (PSORS1 or psoriasis susceptibility locus 1) was demonstrated in the 1970s [124] and has been confirmed in numerous subsequent studies [16,125,126]. However, the genetics of PsA is not as clear-cut and association with alleles of the HLA class I region is not reported to be as strong with PsA as with PS [123]. Hence, it has not been clear if PsA is a clinical phenotype that is distinct from PS without psoriatic arthritis and if is due to different predisposing genetic factors.

A number of regions in the genome have been reported to be associated with PS [8,32,35,119], and some have been convincingly replicated. This includes the 3'UTR of interleukin 12B (IL12B) [127,128] and two non-synonymous SNPs of interleukin 23 receptor (IL23R) [128]. One of these (R381Q) was also shown to be associated with Crohn's disease [129]. However, together with the PSORS1 locus, the combined effect of these loci is unable to account entirely for genetic susceptibility to PS.

In order to systematically search for other susceptibility loci, we undertook a genome wide association scan (GWAS) to identify genetic factors predisposing to PS and PsA. Besides detecting strong association with the HLA class I region in the combined and PsA cohort, and replicating the recently reported associations with IL23R and IL12B, we identified a number of novel associations. These include a region on chromosome 13q13 harboring LHFP and COG6, a region on chromosome 15q21 harboring USP8-SPPL2A-TNFAIP8L3, association with the LCE cluster of genes on chromosome 1q21 from the PSORS4 locus, and a region of chromosome 4q27 recently reported to be associated with several other autoimmune diseases and associated with PsA and potentially PS.

2.2 Materials and Methods

2.2.1 Subjects

The cohorts for the discovery and replication phases of this study are all of European descent and are described in **Table A1.1, Appendix I**. The discovery cohort consisted of 223 Caucasian individuals with PS or PsA from the US. Cases were ascertained through Texas Dermatology (Dallas, TX) and the dermatology clinics at the University of California, San Francisco (UCSF). 89 of the PsA cases had a first degree relative with psoriasis and were members of affected sib pair families, described elsewhere [130]. All except for two of these PsA cases also had PS. These cases were from affected sib pair families with psoriasis and both of these cases had several first degree relatives with PS.

Genotypes of 519 European controls obtained following hybridization to the Illumina HumanHap300 array were from the New York Cancer Project (NYCP) [131] and were downloaded from <http://inragen.c2b2.columbia.edu/>. These were random controls and there was no specific information about autoimmune/inflammatory disease. Recent large genome-wide association studies using controls of this type have been shown to be successful, leading to only a modest effect on power unless the event of misclassification bias is substantial [132]. Informed consent was obtained from all participants. Protocols were approved by the local institution review boards of all participating institutions. All subjects over 18 years of age gave written informed consent, filled out a clinical questionnaire and received a skin examination by the study dermatologist, who confirmed the diagnosis of plaque PS and graded PS severity. All adults with PsA satisfied the inclusion criteria of having both clinically documented

inflammatory synovitis and PS, confirmed by a rheumatologist and dermatologist respectively.

Blood samples were obtained by venipuncture for all subjects, and genomic DNA was isolated from whole blood by standard procedures.

Replication cohorts were from both the U.K. and the U.S. The U.K. cohort consisted of 576 PsA patients from the UK and are described elsewhere [133]. In brief, PsA patients under active follow up by hospital rheumatologists were recruited from throughout the UK although the majority came from the North-West region of England. All patients satisfied the inclusion criteria of having both clinically documented inflammatory synovitis and PS. Each patient was assessed by a trained research nurse, who undertook a standardized clinical history and examination. Detailed demographic and clinical information was obtained and whole blood was taken for DNA extraction and subsequent genetic analysis. Control samples (n = 480) were obtained from blood donors. All patients and controls were white and of UK descent. They were recruited with ethical committee approval (MREC 99/8/84) and provided written informed consent.

The replication cohort from the U.S. for cases consisted of 577 patients with PS (94 of these were also diagnosed with PsA), ascertained at the University of California, San Francisco, CA or at Texas Dermatology, Dallas, TX. The replication cohort for controls consisted of 479 unrelated Caucasian individuals from the University of California, San Francisco, ascertained as a set of healthy controls, for cardiovascular studies. A separate cohort of 258 controls was ascertained in Texas. The latter controls were all > 40 years of age and were ascertained on the basis of not having PS, PsA, or any other inflammatory or autoimmune disease. **Table A1.1 in Appendix I** also provides

information on how well the cases and controls were matched in terms of age and gender. It can be seen that the gender proportions and ages are similar in cases and controls, for both discovery and replication studies.

2.2.2 Genotyping Methods

DNA was normalized to a concentration of 100 ng/ μ l (diluted in 10mM Tris/1mM EDTA). Samples were quantitated with a Nanodrop Spectrophotometer (ND-1000). For the discovery phase, approximately 1 μ g of genomic DNA was used to genotype each sample on the Illumina HumanHap300v2A Genotyping BeadChip. This was performed at the Robert S. Boas Center for Genomics and Human Genetics at The Feinstein Institute for Medical Research, Manhasset, NY. This assay relies on allele specific primer extension and the use of a single fluorochrome. Samples were processed according to the standard Illumina Infinium II automated protocol. This involved whole genome amplification, fragmentation, precipitation, resuspension in hybridization buffer and hybridization to the Illumina Bead Chips for a minimum of 16 h at 48°C. After hybridization the BeadChips were processed for the single base extension reaction, followed by staining and imaging on an Illumina Bead Array Reader. Normalized bead intensity data were loaded into the Illumina Beadstudio 2.0 software which converted fluorescence intensities into SNP genotypes.

Genotyping for all the replication studies was performed with the Sequenom MassArray system (iPlex assay). This involves primer extension chemistry and mass spectrometric analysis described at our web site http://hg.wustl.edu/info/Sequenom_description.html.

2.2.3 Quality Control

Before analysis, we performed quality filtering of both samples and SNPs to ensure robust association tests. Based on previous criteria [134], we required that all samples used for the discovery phase pass a 93% genotyping call rate threshold, and that all SNPs pass a 95% call rate threshold.

In the case of the replication studies, 57 individuals from the total of 2370 individuals in the replication study were removed because of low genotyping (i.e. when over half of the genotypes for a sample were missing). SNPs with $< 75\%$ call rates were also excluded from analysis to obtain an average genotyping rate of 0.902. Genotypes were also evaluated for departure from HWE in the controls and SNPs with $P < 0.001$ were removed from further analysis. After pruning, 244 SNPs remained.

A total of 463 ancestry informative SNPs (AIM) present on the Illumina HumanHap300v2A Genotype BeadChip were used to check for possible confounding population substructure in the discovery sample with STRUCTURE software [135]. For this analysis, genotypes at these SNPs were analyzed for all 742 samples (223 PS cases and 519 controls).

To investigate other biases [136] that could be introduced with shared controls we assessed the potential effect of substructure with the genomic-control method [137] and with EIGENSTRAT [138].

2.2.4 Statistical Analysis for Association

The Cochran-Armitage Test for trend was conducted with Purcell's PLINK program (<http://pngu.mgh.harvard.edu/~purcell/plink>). However, several SNPs in the current study that exhibited significant differences in cases/controls, were also different when allele frequencies in controls were compared with those from European CEPH typed for SNPs in the HapMap project. NYCP participants are quite diverse with respect to European origin, and many SNPs are reported to show differences among European subgroups [139]. These were identified by a comparison of SNP allele frequencies in European CEPH individuals typed for the HapMap project and were not selected for follow-up studies.

Measures of linkage disequilibrium, D' and r^2 , and allele frequencies were based on pre-computed scores from the International HapMap website or were computed locally from HapMap genotypes or from our own case and control genotypes with Haploview 3.2 (<http://www.broad.mit.edu/mpg/haploview/>). Power calculations for association were calculated at: <http://pngu.mgh.harvard.edu/~purcell/gpc/>. Association localization plots were generated with an R code modified from snp.plotter (<http://cbdb.nimh.nih.gov/~kristin/snp.plotter.html>) and Regional Association Plot (<http://www.broad.mit.edu/diabetes/scandinav/figures.html>). Family based association tests on 271 nuclear families were performed with the Pedigree Transmission Disequilibrium Test [140] as described elsewhere [32,126,141].

2.2.5 Immunohistochemistry

Tissue sections were fixed with acetone and stained with 10ug/mL purified mouse anti-human monoclonal antibodies to IL-2 (R&D, clone 5334.21), IL-21 (R&D, J148-1134),

COG6 (Abnova, H00057511-M01) and SPPL2A (Abgent, AP6312a). Biotin labeled horse anti-mouse antibodies (Vector Laboratories) were amplified with avidin-biotin complex (Vector Laboratories) and developed with chromogen 3-amino-9-ethylcarbazole (Sigma Aldrich).

2.3 Results

2.3.1 Genome-Wide Association Scan

For our “discovery” phase, 223 PS cases (132 cases with PS without arthritis and 91 PS cases with arthritis (PsA) were typed on the Illumina HumanHap300 arrays. We compared case data to publicly available genotype data of 519 European controls from the New York Cancer Project [142] collected with the same platform. The number of cases used for this scan is smaller than that used in many recently described genome wide association scans. However, the 91 cases of PsA had at least one first degree relative with PS and were expected to be enriched for genetic factors. Power calculations based on 223 cases and 519 controls indicated that using a threshold of $P < 5 \times 10^{-5}$, we had 70% power of detecting a locus with a genotype relative risk (GRR) of 2.0, and over 99% power to detect a locus with a GRR of 3.0 such as the MHC (see below). However, many replicated associations have small GRRs [132] and we had only 10% power to detect a locus with a GRR of 1.5.

Following the genotyping of samples, stringent quality control measures were implemented. We required that all samples used for the discovery phase passed a 93% genotyping call rate threshold, and that all SNPs passed a 95% call rate threshold. Justification for this threshold is based on the evaluation of empirical distributions (**Appendix I, Figure A1.1**). With sample call rates $\leq 93\%$, there was an elevation in observed sample heterozygosity, i.e. deviation from Hardy-Weinberg equilibrium, suggesting possible genotyping errors (e.g., sample contamination or allele drop-out). Likewise, there was a significant discrepancy of missingness between case and control groups when the SNP success rate was $< 95\%$.

For the discovery phase, a total of 311,398 SNPs were pruned to 305,983 SNPs after filtering for low call rate, minor allele frequency < 0.01 and deviation from Hardy-Weinberg equilibrium ($P < 0.001$). Quality control also led to the removal of 29 samples, leaving 218 cases for further analysis. The average genotyping rate in the remaining individuals was 0.995.

To investigate other biases [136] that could be introduced with shared controls, we assessed the median distribution of test statistics using the genomic-control factor λ_{GC} [137]. With a set of 463 ancestry informative SNPs (AIMs), $\lambda_{GC} = 1$ indicating no inflation). We also performed analysis on the same set of AIM SNPs with STRUCTURE software [143]. Under the assumption of two population clusters, there was no association between the most likely inferred cluster and case/control status and the average allele frequency difference between clusters was less than 2.5%. These results showed that population substructure is unlikely to be confounding our results. However, analysis of all markers used in the discovery study yielded $\lambda_{GC}=1.101$ before correction. A similar value was obtained with EIGENSTRAT [138] where $\lambda=1.107$. Examining stratified subsets of cases (PS without arthritis or PsA) with all markers also yielded similar λ values (PS without arthritis: $\lambda_{GC}=1.07$; PsA: $\lambda_{GC}=1.05$). Following adjustment of P values with the genomic control method, $\lambda=1$. The discovery P values adjusted by the genomic control method as implemented in PLINK [144] are presented in the tables and figures.

To detect associations, we first performed a preliminary analysis with a Cochran-Armitage trend test. **Figure 2.1** illustrates negative logarithm of the P values obtained across the genome, considering all cases and all controls (**Figure 2.1A**) and considering

only the 91 PsA cases and all controls (**Figure 2.1B**). Results were then rank ordered on the basis of P values. 84 SNPs in 35 genomic regions were associated with $P < 5 \times 10^{-5}$; a level that we would informally expect to observe by chance roughly 15 times in this scan given the number of tests performed if all SNPs were independent. A subset of SNPs from 120 regions were investigated further. Criteria for selection included the strength of the discovery P value, particularly when several SNPs from a single region showed evidence for association, a possible biological role of a gene harboring a SNP with some evidence for association, or localization of SNPs with moderate evidence for association to a known psoriasis susceptibility locus (e.g. PSORS4). We also included the previously reported associated SNPs in IL23R and IL12B [128].

An independent cohort of 577 PS cases from the U.S. and 737 U.S. controls were used for the replication stage; 94 of these cases had also been diagnosed with PsA. To examine the potential role of variants upon PsA susceptibility specifically, 576 PsA cases from the UK and 480 controls from the UK were also employed. An alternative genotyping technology (iPlex; Sequenom) was used for the replication phase. The platforms used for the discovery and replication phases gave very similar results: Concordance rates on the basis of 116 samples and 301 SNPs typed with both platforms was 98.74%.

Our 100 top ranked SNPs with any cohort (PS, PS without arthritis, or PsA) are listed in **Table S2** of the article website (<http://www.plosgenetics.org/article/info:doi/10.1371/journal.pgen.1000041#s4>) to facilitate future attempts to replicate our findings. A total of 289 SNPs, including SNPs from the MHC, and two previously reported

associated SNPs within IL12B and IL23R [128] were genotyped in the replication analysis.

2.3.2 Association with MHC Region

The MHC, and in particular, the HLA class I region, is the only region that has been shown to be consistently associated with PS. The first nine top-ranking SNPs were from the MHC and seven were significant, even when adjusted with the Bonferroni correction for multiple tests. Overall, 32 SNPs from the MHC had adjusted P values $< 5 \times 10^{-5}$ (**Figure 2.2**). The most significant association was with rs10484554 (adjusted P = 7.8×10^{-11} , GWA scan; P = 5.61×10^{-28} , replication; P = 9.772×10^{-38} , combined) (**Figure 2.2, Table 2.1**). This SNP lies 34.7kb upstream from the transcriptional start site of HLA-C. Strongest association with this region is consistent with previous results from our group and others [16,125,126]. The rs10484554*T allele had frequencies of 0.325 in U.S. cases and 0.15 in U.S. controls (OR: 2.8 (95% CI: 2.4 – 3.3). To determine the relationship of this allele with the classical HLA-C allele strongly associated with psoriasis (HLA-Cw*0602), we investigated the transmission of this allele with classical HLA-C alleles in ~250 nuclear families with psoriasis that we have reported elsewhere [126]. The rs10484554*T allele was detected on nearly all haplotypes with HLA-Cw*0602 or HLA-Cw*1203 alleles (results not shown), and was also strongly correlated with the previously described highly associated PSORS1 SNP n9*G [24,126] (rs10456057*G) allele. We have previously shown that SNPs upstream from HLA-C are more strongly correlated with PS than HLA-Cw*0602 is, and that these risk alleles are also correlated with HLA-Cw*1203 [126]. Hence, rs10484554*T may be a good proxy for the PSORS1 variant.

In the case of the U.K. PsA replication samples, rs10484554 was again highly significant ($P = 6.86 \times 10^{-11}$) (**Table 2.2**), although the frequency of the rs10484554*T allele exhibited population differences when frequencies in the U.K. and U.S. were compared. In the U.K. the rs10484554*T allele was found at a lower frequency in cases and controls (0.19 and 0.07 respectively; OR: 2.4 (95% CI: 1.8 - 3.1)).

A second SNP from the HLA class I region lying between MICA and MICB (rs2395029) was highly associated with PS and PsA. This SNP results in the G2V polymorphism of the class I gene HCP5 (HLA complex P5) which encodes an endogenous retroviral element. For this SNP, PS was associated with a combined $P = 2.13 \times 10^{-26}$ in the U.S. cohort and 1.86×10^{-10} in the U.K. PsA cohort (**Table 2.1, 2.2**). The OR of the rs2395029*C allele with both PS and PsA was higher than with any other SNP tested (4.1 and 3.2 with PS and PsA respectively). This allele was found at a frequency of ~ 0.12 in cases and 0.04 in controls and did not exhibit the population frequency differences of rs10484554. The LD relationship between rs2395029 and rs10484554 is not strong ($r^2 = 0.33$ in European CEPH HapMap samples and $r^2 = 0.23$ in our U.S. case/control cohort). Conditioning upon rs10484554, the P value for rs2395029 was still significant ($P = 7 \times 10^{-10}$), hence effects from this SNP are likely to be independent.

HCP5 is expressed primarily in cells of the immune system such as spleen, blood and thymus (<http://smd-www.stanford.edu/>), consistent with a potential role in autoimmunity. This allele was recently shown to explain 9.6% of the total variation in viral set point following HIV-1 infection [145]. This is of interest, since psoriasis can be triggered by infection with HIV and other viruses. Hence, it is possible that HCP5-C

carriers mount a strong immune reaction to viral infection, but that in genetically susceptible individuals, this reaction leads to excessive inflammation in skin and joints. Overall, our observations indicate that MHC class I region SNPs are more highly associated with both PS and PsA than any other SNPs.

2.3.3 Association with IL23R

A recent global association scan using a set of pooled PS samples and controls against a set of 25,215 gene-centric SNPs confirmed a previously reported association with IL12B (rs3212227 in its 3' UTR) [127] and identified a second region of association 60kb upstream from its mRNA start site (rs6887695) [128]. An analysis of additional genes encoding components of the IL12B pathway lead to the identification of associations with IL23R (R381Q: rs11209026 and L310P: rs7530511) [128]. These SNPs were proposed to mark a common psoriasis-associated haplotype. Rs11209026 is also the SNP within IL23R reported to be associated with Crohn's disease [129].

In our discovery cohort, the most significant association in the IL23R interval was obtained with a different SNP (rs12131065) from that described previously as being associated with PS (rs11209026). This SNP rs12131065 has $P = 0.0039$ in the discovery cohort (**Table 2.1**) and has not previously been reported to be associated with PS. The LD relationship between rs12131065 and the previously associated rs11209026 SNP is low ($r^2 = 0.031$ in HapMap CEPH European samples; 0.009 in cases; 0.026 in controls). Conditioning upon rs11209026, the P value for rs12131065 was 0.013. Hence, effects from this SNP may be independent of rs11209026 and its association with PS should be investigated in other cohorts.

SNP rs12131065 lies downstream from IL23R (63kb from rs11209026) and 4.041kb upstream from the gene for interleukin 12 receptor B2 (IL12RB2) (**Figure 2.3**). IL12RB2 is involved in IL12 dependent signaling, is upregulated by gamma interferon in T_h1 cells and plays a role in T_h1 differentiation [146]. Association with a SNP closer to IL12RB2 than IL23R is of interest since animals where IL12RB2 is inactivated develop autoimmune disease [147].

Association with the previously reported IL23R associated SNP rs11209026 in the discovery cohort was not significant (adjusted P = 0.081). Genotyping of rs11209026 and rs12131065 in the U.S. replication cohort yielded combined P values of 1.4×10^{-4} and 0.001 respectively (**Table 2.1**) consistent with replication of this locus with respect to previous studies. In the case of these two SNPs, the protective T and A alleles were found at frequencies of ~0.04 and 0.2 in cases versus ~0.07 and ~0.24 of controls respectively. In the U.K. replication PsA cohort, association with rs11209026 was consistent with replication ($P = 8.3 \times 10^{-4}$), with the rs11209026*T allele being found at frequencies of ~0.04 in cases and ~0.08 in controls (**Table 2.2**).

2.3.3 Association with IL12B

Although the associated IL12B SNPs rs3212227 and rs6887695 were not interrogated by the Illumina screening panel of SNPs used here, typing of these SNPs in our replication U.S. case/control cohorts yielded P values of 0.021 and 5×10^{-5} and replicated previously reported associations (**Table 2.1**). In the U.K. PsA cohort, association with rs6887695 was also consistent with replication ($P = 0.0013$) (OR: 0.69; 95% CI: 0.56 - 0.85)) (**Table 2.2**).

2.3.4 Association with Novel Psoriasis Loci

In the discovery cohort, there were four SNPs from 13q13 where $P < 5 \times 10^{-5}$ (adjusted $P < 2 \times 10^{-4}$). These were: rs1186468, rs4514547, rs4569133 and rs7993214. These SNPs lie within a region on chromosome 13q13 that encodes the conserved oligomeric golgi complex component 6 (COG6) gene and a lipoma HMGIC fusion partner (LHFP) [148]. Three of the top associated SNPs were tested in the U.S. replication cohort, and all showed evidence of replication at $P < 0.05$ (**Table 2.3, Figure 2.4**). Results were most significant with rs7993214 (adjusted $P = 10^{-4}$, GWA scan; $P = 0.0033$, replication; $P = 2 \times 10^{-6}$, combined). Rs3812888, (adjusted $P = 0.0017$, GWA scan; $P = 4 \times 10^{-4}$, replication; $P = 10^{-5}$, combined) was the only SNP where replication results would remain significant following the stringent Bonferroni correction for multiple tests ($P = 0.048$). The OR of the rs3812888*C allele was 1.38 (95% CI: 1.15-1.66). The rs3812888*C allele was found at frequencies of 0.43 in cases and 0.35 in controls.

COG6 is a component of the conserved oligomeric golgi (COG) complex and is involved in intracellular transport and glycoprotein modification [149]. The glycosylation pathways in the golgi apparatus must be intact for protein secretion to continue unabated. In *C. elegans*, a COG complex is required to glycosylate an ADAM protease (a disintegrin and metalloprotease) [150]. In humans, variants within some ADAM genes lead to inflammatory diseases. For example, ADAM33 is an asthma susceptibility gene whose catalytic domain undergoes glycosylation [151]. Recent genetic studies suggest that ADAM33 is a psoriasis susceptibility gene as well [152]. Hence, COG6 could be involved in glycosylation of ADAM33 or other ADAM proteases. Staining of skin

sections with a COG6 antibody revealed cytoplasmic staining in the epidermis as well as strong T-cell staining (**Figure 2.5**). There was variable expression of the protein in non-lesional skin samples, but there was uniformly strong expression in all lesional sections. Very little is known about LHFP. It is a subset of the superfamily of tetraspan transmembrane protein encoding gene. Expression analysis from SOURCE (<http://smd-www.stanford.edu/>) indicates that highest levels are found in the ear and spinal cord.

When the PsA discovery cohort (n=91) was analyzed separately, four SNPs from a region on chromosome 15q21 between ubiquitin specific protease 8 (USP8) and tumor necrosis factor, alpha-induced protein 8-like 3 (TNFAIP8L3) were associated with $P < 5 \times 10^{-5}$. In the case of the most highly associated SNP (rs4775919), adjusted $P = 6.7 \times 10^{-6}$. Following replication genotyping, this and rs3803369 were associated with PS with P values consistent with replication (for rs3803369, adjusted $P = 2.5 \times 10^{-4}$, GWA scan; $P = 0.013$, replication; $P = 2.9 \times 10^{-5}$ combined; **Table 2.3, Figure 2.6**). The rs3803369*A allele was found at a frequencies of ~0.2 in cases and 0.15 in controls (OR 1.43, 95% CI: 1.21-1.69). Other genes in this region include the transient receptor potential melastatin 7 (TRPM7) [153], signal peptide peptidase like 2a (SPPL2A) [154] and AP4E1, a member of the heterotetrameric adaptor protein (AP) complexes (**Figure 2.6**). TNFAIP8L3 is a novel protein. It harbors a domain (DUF758) that is found in several proteins induced by tumor necrosis factor alpha (TNFA), but its function is unknown. However, the most plausible candidate is SPPL2A that catalyzes the intramembrane cleavage of TNFA, triggering the expression of IL12 by activated human dendritic cells [155]. Staining of skin sections with an SPPL2A antibody (**Figure 2.5**) revealed profound staining of the epidermis, and staining of some dermal cells in both lesional and non-lesional skin.

This region is also of interest however, because a processed pseudogene for one of the genes in this region (USP8) is found upstream from HLA-C [24]. As discussed earlier, this region is one that is most likely to harbor the PSORS1 variant. To ensure that our observations were not due to cross-hybridization of chromosome 15q21 SNPs with PSORS1 SNPs, we investigated alignment of genes from this region of chromosome 15q21 with the remainder of genome. We did not detect any significant identity with any other region, including the MHC. This, and the fact that chromosome 15 associated SNPs are in Hardy-Weinberg equilibrium indicate that our results are unlikely to be artifactual and due to amplification of PSORS1 sequences. The similarity between the PSORS1 and 15q21 variants and their biological consequences need to be investigated further since it may provide important insights into the nature of the PSORS1 variant. However, it is worth noting that our PsA cases which provided strongest evidence for association with 15q21 all had at least one first degree relative with PS, and association with this locus may be harder to detect in “sporadic” cases.

We also observed association of PS with a region of the Epidermal Differentiation Complex (EDC), which harbors a previously established psoriasis locus (PSORS4) [156]. In this instance, rs6701216 yielded a combined $P = 6.2 \times 10^{-5}$ (OR 1.45) (**Table 2.3**). This SNP lies within the late cornified envelope 1C gene (LCE1C) [157], and is one of a family of genes that are transcribed very late in epidermal differentiation.

There were two other regions selected for follow-up where P values were < 0.05 in the PS replication cohort, and where evidence for association increased in the combined cohort (**Table 2.3**). One was an intergenic region located between granulysin (GNLY) and atonal (ATOH) on chromosome 2p11. The most highly associated SNP was

rs2164807 (adjusted $P = 0.0015$, GWA scan; $P = 0.0039$, replication; $P = 1.6 \times 10^{-5}$ combined). For this SNP, the G allele was found at a frequency of ~ 0.47 in cases and ~ 0.39 in controls (OR: 1.35, 95% CI: 1.18-1.54). GNLY (Protein NKG5, Lymphokine LAG-2) is of considerable interest with respect to psoriasis. It is present in cytotoxic granules of cytotoxic T lymphocytes and natural killer cells, and is released upon antigen stimulation[158]. It has been shown to have antimicrobial activity against a broad range of microbes including Gram-positive and Gram-negative bacteria, fungi, and parasites including *M. tuberculosis* and other organisms. Priming of granulysin in CD4 is dysregulated in the CD4+ T cells of HIV-infected patients [159].

Other genes that should be evaluated in additional PS cohorts on the basis of replication P values < 0.05 , and increased significance in combined cohorts (**Table 2.3**) are calponin-like transmembrane domain protein (CLMN) [160], the gene for the catenin member, CTNNA2 [161], and a gene desert on chromosome 3q13.

2.3.5 Association with Chromosome 4q27

In the discovery cohort there were 3 SNPs from chromosome 4q27 with $P < 5 \times 10^{-5}$ (adjusted $P < 10^{-4}$). These were rs13151961, rs6822844 and rs6840978. The most significant of these was rs13151961, where adjusted $P = 4 \times 10^{-5}$ (**Table 2.4**). Association of this region with PsA was confirmed in an independent cohort of patients from the UK (for rs13151961, $P = 0.003$, **Table 2.4**), where the frequency of the associated T allele was ~ 0.25 in cases and ~ 0.31 in controls (OR: 0.64; 95% CI: 0.49-0.84). Association with additional SNPs in high LD with rs13151961 (rs6840978 and rs6822844), was also replicated in this cohort. Although association could not be replicated in our U.S. cohort

with a case/control approach, the trend in allele frequencies in cases versus controls was similar to that seen in the U.K. cohort (for rs13151961 the frequency of the T allele was ~0.26 in U.S. cases and ~0.29 in our U.S. controls, **Table 2.4**).

A recent study from the Wellcome Trust Case Control Consortium [132] identified this 4q27 region in a search for risk factors for type 1 diabetes (T1D). In a follow-up study [144], some support for this association with T1D was provided. In the latter study, association of this region with Grave's disease (GD) was also tested, and some evidence for association with the complementary allele of rs17388568 to that seen in T1D was obtained. The same region was also reported to be associated with Celiac disease (CeD) [162]. Recent evidence is also provided for a role for this region in rheumatoid arthritis and T1D from the Netherlands [163]. In that study the rs6822844*A allele was reported to be a perfect proxy for a haplotype that is highly associated with autoimmune diseases [163] with frequencies of 0.13 in cases versus 0.19 in N. European controls.

The risk alleles of rs13151961, rs13119723, rs6822844 and rs6840978 associated with PS and PsA in the current study are similar to those reported for CeD [162]. Overall, the risk alleles/haplotype of GD, CeD, PS and PsA appear to be the same, and of similar frequency. For example the frequency of the rs6822844*A allele is 0.14 in U.S. PS patients, ~0.16 in U.K. PsA patients, 0.14 in RA patients [163], ~0.13 in T1D patients from the Netherlands [163], 0.12 in Celiac Disease patients from the Netherlands [162], and ~0.14 in Irish CD patients [162]. This contrasts with frequencies of 0.19-0.20 in European control populations [162,163]. Although the frequency of the rs6822844*A allele was 0.14 in our PS cases and hence similar to frequencies seen in CeD, RA and

T1D cases, its frequency in our combined cohort of U.S. controls was 0.16 (**Table 2.4**). This is lower than that reported for European controls. However, it has been previously reported that geographic variability exists at this locus across Europe [162]. Our U.S. “European” controls are likely to be more disparate in origin, and are likely to account for our inability to obtain significant evidence for association with PS and 4q27.

To explore association of this region in our cohort of U.S. PS cases, without the possible confounding influence of subtle geographic variability at this locus in Europeans, we performed family based association tests in our 242 psoriasis nuclear families which are described elsewhere [126,164]. This approach provided evidence for replication of association of PS with rs6822844 and rs6840978 (PDT P = 0.029 and 0.007 respectively). For these SNPs, the over-transmitted rs6822844*G and rs6840978*C alleles were also the risk allele from case/control studies. Haplotype studies in families also revealed association with the rs6822844*G/rs6840978*C haplotype (multipoint TDT P = 0.006). Hence, our findings support chromosome 4q27 as harboring a variant/haplotype for PsA and PS.

As reported elsewhere the 4q27 locus that contains these associated SNPs corresponds to two closely correlated ~439kb and ~40kb haplotype blocks [162]. This extensive LD makes it very difficult to determine the predisposing variant. Chromosome 4q is also the location of PSORS3, which is generally placed slightly more distally [165]. However, the locus identified here may contribute in part to the previous observations of linkage. The long region of LD at chromosome 4q27 contains several genes [162]: Testis nuclear RNA-binding protein (TENR), a gene encoding a protein of unknown function (KIAA1109), and genes encoding the interleukin-2 (IL2) and interleukin-21 (IL21)

cytokines. TENR is expressed primarily in testis and KIAA1109 transcripts are ubiquitous, hence their roles in autoimmunity are not particularly compelling. However, IL2 and IL21 are of particular interest with respect to PS. IL2 is considered to be a pathogenic cytokine for PS[166], and blockade of the IL2 receptor with therapeutic antibodies has induced disease resolution in some cases [167]. IL2 is a survival factor for T cells and promotes the differentiation of cytotoxic T-lymphocytes and NK cells. Both of these cell types are present in psoriasis lesions. Moreover, many IL2 receptor (IL2R) positive T-cells that fit the phenotypic definition of regulatory T cells (T_{reg} s) are also present in psoriasis lesions. IL-2 may influence how a common precursor T-cell differentiates into either a T_{reg} or a T_h17 T-cell, since addition of IL-2 has been shown to suppress the differentiation of T_h17 T-cells in mice [168]. IL-2 antibodies stain normal epidermis and psoriatic epidermis, with generally lower staining in the dermis (**Figure 2.5**). The pattern of staining appears to be to dendritic cells (DCs) which are likely to be epidermal Langerhans cells. This pattern of IL-2 staining is probably due to DC activation and upregulation of IL2R. Cells with IL2 receptors include T-cells, B-cells, NK-cells, and dendritic cells.

The epidermal staining for IL-21 is much lower than for IL-2 (**Figure 2.5**) and appears to be mainly on dendritic cells in the superficial dermis. IL-21 is a product of activated T cells (under conditions of T_h17 polarization). It then acts in an autocrine or paracrine fashion on T-cells to up-regulate expression of the IL23 receptor which has already been implicated in psoriasis pathogenesis. IL23R sensitizes cells to IL-23 which stimulates IL17 synthesis and/or prolongs the survival of T_h17 cells [169]. Blocking IL21 reduces the progression of lupus erythematosus [170] which is one of the

autoimmune diseases that is now being considered as a “T_h17” mediated disease. Therefore IL21 may play a role in T_h17 formation in this and other autoimmune diseases where these cells are pathogenic.

Extensive resequencing of IL2 coding and flanking regions has already been performed in T1D samples and no coding or obvious regulatory/splice variants were identified [144]. As stated previously, this region needs to be resequenced thoroughly followed by comprehensive genotyping in larger numbers of samples to identify the autoimmune associated variants [144].

2.4 Conclusions

The observed associations in the current study are of interest for several reasons. It is noteworthy that the strongest association is with the MHC. Even in PsA, where associations are reportedly less than with PS (without PsA), associations with the class I region appear to be more significant than with any other region. We were also able to replicate previously reported associations with IL12B and IL23R and detected a potentially novel association upstream from IL12RB2. Novel associations within COG6 and the region on chromosome 15q21 harboring USP8 and SPPL2A are of interest. These and other regions reported here are worthy of follow up in other cohorts. Moreover, the association with chromosome 4q27 provides further evidence that this region is a common locus for multiple forms of autoimmune disease.

A recent study reported the IL13/IL4 region from chromosome 5q31 as being associated with PS [171]. Overall, the risk contributed by the MHC, the IL13/IL4 region and the IL23R and IL12B variants was estimated to be 3.83. With the COG6 and chromosome 15q21 loci described here, the risk would be increased. However, PS is a complex disease, and overlapping subsets of risk factors may be sufficient for susceptibility, so that risk effects cannot be computed in an additive manner.

The ability to identify risk variants of modest effect for common diseases such as PS and PsA will be limited by the cohort size, and larger numbers of cases and controls will be necessary to identify the majority of genetic factors for these diseases. Moreover, some of the SNPs with borderline discovery P values are also likely to be genetic risk factors for disease. It is worth noting that our discovery P value for the associated IL23R R381Q SNP did not reach significance ($P = 0.057$, adjusted $P = 0.081$) although allele

frequencies in the discovery dataset revealed a M.A.F. of 0.041 in cases and 0.065 in controls, which is similar to what has been reported in other studies [172,173]. Additional genome-wide association scans and replication studies are required to identify additional variants and to confirm some of those found in the current study. Such studies include a genome-wide scan for psoriasis variants from the Genome Wide Association Network (GAIN) consortium [133]. Genes such as these are important for determining the pathogenesis of PS and PsA and in identifying novel drug targets for these inflammatory diseases of the skin and joints.

Table 2.1 Summary of association with previously reported PS susceptibility loci (MHC, IL23R and IL12B) in U.S. PS cohort (810 cases, 1256 controls).

Cyto. locn	SNP	Location (hg18)	Gene/region	Disc. P	Disc. P (adj.)	Rep. P	Combined P	Minor allele	Freq. U.S. PS cases	Freq. U.S. Controls	OR US PS combined (95% C.I.)
6p21	rs10484554	31382534	MHC	8.7×10^{-12}	7.8×10^{-11}	1.82×10^{-30}	1.81×10^{-39}	T	0.32	0.151	2.8 (2.4, 3.3)
6p21	rs2395029	31539759	MHC	1.4×10^{-7}	5.3×10^{-7}	2.51×10^{-19}	2.13×10^{-26}	C	0.12	0.033	4.1 (3.1, 5.3)
1p31	rs11465804	67475114	IL23R	0.4	0.42	ND	0.0072	G	0.044	0.065	0.67 (0.50,0.9)
1p31	rs11209026	67478546	IL23R	0.067	0.081	0.0039	0.00014	T	0.039	0.066	0.56 (0.41,0.76)
1p31	rs12131065	67541594	IL23R	0.0025	0.0039	0.074	0.001	A	0.197	0.243	0.78 (0.66,0.91)
5q33	rs3212217	158687708	IL12B	N.D.	N.D.	0.012	N.D.	G	0.158	0.199	N.D.
5q33	rs6887695	158755223	IL12B	N.D.	N.D.	0.00005	N.D.	C	0.221	0.294	N.D.

Cyto. Locn: Cytogenetic location of SNP; Disc P: Trend P values for the GWA scan; Disc. P (adj.): Genomic control adjusted trend P values for the GWA scan; Rep. P: Trend P values obtained with the U.S. PS replication cohort; Combined P: Trend P values for U.S. discovery and replication cohorts combined ; Freq. U.S. PS cases: Frequency of the minor allele in U.S. psoriasis cases; Freq. U.S. Controls: Frequency of minor allele in the control population. OR: Odds Ratio; C.I.: confidence interval; N.D.: Not Done.

Table 2.2 Summary of association with previously reported PS susceptibility loci in U.K. PsA cohort (576 cases, 480 controls).

SNP	Gene/Region	Trend P value	Minor allele	Freq. UK PsA cases	Freq. UK PsA controls	OR UK PsA (95% C.I.)
rs10484554	MHC	6.86×10^{-11}	T	0.19	0.07	2.4 (1.8, 3.1)
rs2395029	MHC	1.86×10^{-10}	C	0.12	0.04	3.2 (2.2, 4.6)
rs11209026	IL23R	0.00083	T	0.043	0.079	0.52 (0.35, 0.77)
rs12131065	IL23R	0.31	A	0.21	0.23	0.89 (0.72, 1.11)
rs3212217	IL12B	N.D.	G	N.D.	N.D.	N.D.
rs6887695	IL12B	0.0013	C	0.213	0.28	0.69 (0.56, 0.85)

UK PsA: U.K. Psoriasis cases with arthritis; OR: Odds Ratio; C.I.: confidence interval; N.D.: Not Done.

Table 2.3 Potential novel loci from GWA scan.

Cytogenet. location	SNP	Location (hg18)	Gene/region	Disc. P	Disc. P (adj.)	Rep. P	Combined P	Minor allele	Freq cases	Freq. Controls	OR (95% C.I.)
1q21	rs6701216	151045150	EDC	0.0034	0.0053	0.0069	0.00005	T	0.174	0.127	1.45 (1.21,1.75)
2p11	rs2164807	85816062	GNLY- ATOH8	0.00084	0.0015	0.0039	0.000016	G	0.467	0.394	1.35 (1.18,1.54)
2p12	rs11126740	79754603	CTNNA2	0.0018	0.0018	0.019	0.00014	T	0.311	0.37	0.77 (0.67,0.88)
3q13	rs6804331	105237077	Gene desert	0.00073	0.0013	0.0455	0.0003	C	0.445	0.381	1.30 (1.13,1.50)
13q13	rs3812888	39128294	COG6	0.001	0.0017	4×10^{-4}	0.00001	C	0.43	0.35	1.38 (1.15,1.66)
13q13	rs7993214	39248912	COG6	4.7×10^{-5}	0.0001	0.0033	2×10^{-6}	T	0.279	0.351	0.71 (0.62,0.82)
14q32	rs2282276	94730882	CLMN	0.0096	0.014	0.047	0.0031	G	0.099	0.073	1.40 (1.12,1.76)
15q21	rs4775912	49068271	USP8- TNFAIP8L3	0.00034	0.00065	0.0136	5.6×10^{-5}	G	0.194	0.146	1.41 (1.19, 1.67)
15q21	rs3803369	49163121	USP8- TNFAIP8L3	0.00012	0.00025	0.0138	2.9×10^{-5}	A	0.195	0.145	1.43 (1.21,1.69)

Trend P values from discovery, replication and combined analyses for U.S. PS samples (810 cases, 1256 controls) are shown. The minor allele, its frequency in cases and controls and its odds ratios and 95% C.I.s from the combined data-set are also shown. Abbreviations are described in footnotes to **Table 2.1**.

Table 2.4 Association results at chromosome 4q27 in PsA and PS cohorts (U.S.: 810 PS cases, 1256 U.S. controls; U.K.: 576 PsA cases, 480 controls).

SNP	Location (hg18)	Disc. P	Disc. P (adj.)	UK PsA P value	Minor allele	Freq UK PsA cases	Freq UK controls	OR UK (95% CI)	Rep. US PS P value	Freq US PS. cases	Freq US controls	OR US PS cases.(95% CI)
rs13151961	123334952	1.6×10^{-5}	3.98×10^{-5}	0.0016	G	0.121	0.175	0.64(0.49,0.84)	0.17	0.138	0.158	0.86 (0.71,1.04)
rs7684187	123560609	1.6×10^{-4}	3.3×10^{-4}	0.001	T	0.247	0.313	0.72 (0.59,0.87)	0.52	0.259	0.298	0.82 (0.71,0.95)
rs6822844	123728871	4.3×10^{-5}	9.6×10^{-5}	0.008	A	0.155	0.203	0.72 (0.56,0.92)	0.29	0.143	0.164	0.85 (0.71,1.03)
rs6840978	123774157	2.9×10^{-5}	6.6×10^{-5}	0.013	A	0.192	0.236	0.77 (0.62,0.95)	0.87	0.172	0.204	0.81 (0.68,0.96)

The minor allele, its frequency in cases and controls and its odds ratios and 95% C.I.s from the combined dataset are also shown. Abbreviations are described in footnotes to **Table 2.1**.

Figure 2.1. Summary of genome-wide association scan results for all cases and the PsA subgroup. Negative $\text{LOG}_{10} P$ values for the Cochran-Armitage test of trend for genome-wide association across the genome and by chromosome are shown. Trend P values were adjusted with the genomic control (GC) method. The spacing between SNPs on the plot is uniform and does not reflect actual physical distances. Adjacent chromosomes are shown in red and then in blue. The horizontal dashed lines display a cutoff of $P = 5 \times 10^{-5}$. A: Results obtained with all cases. B: Results obtained with the subgroup of 91 psoriatic arthritis cases (PsA).

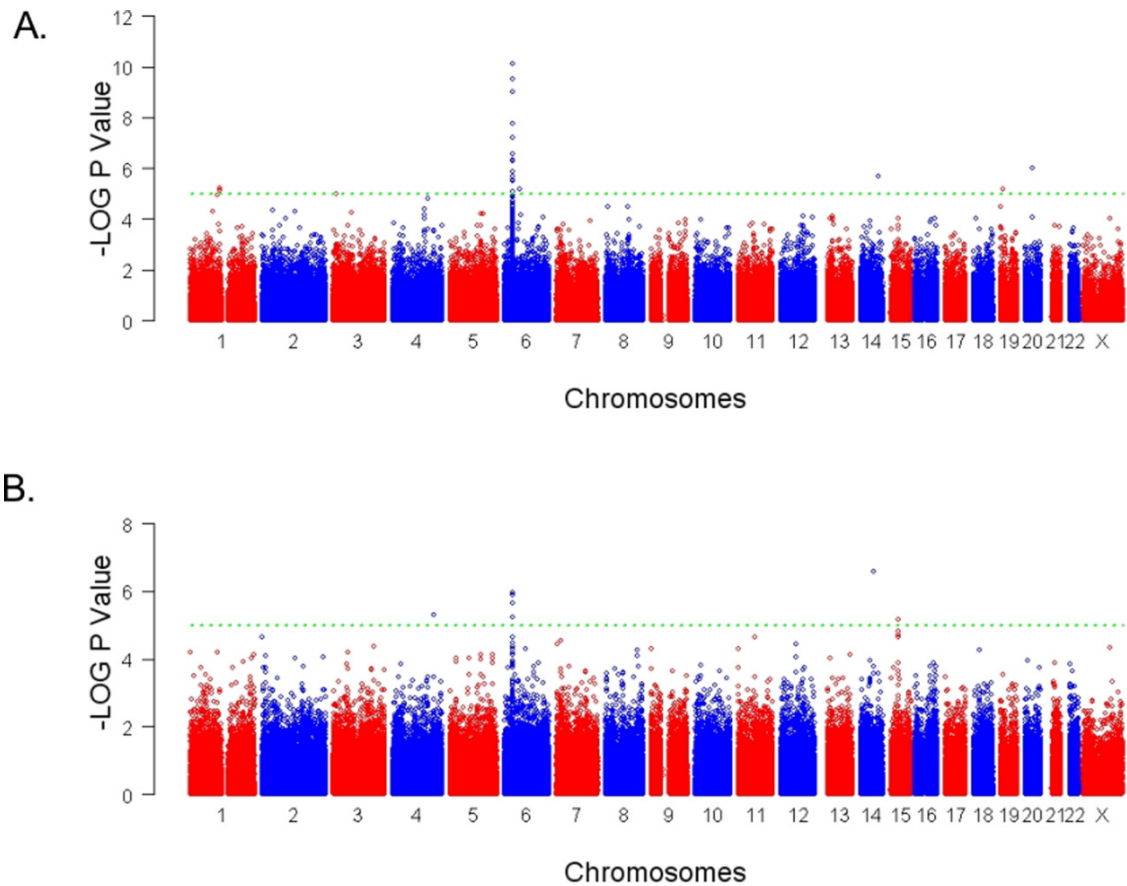


Figure 2.2. Association localization plots for the MHC following discovery and replication phases. Results for SNPs used in the discovery phase (adjusted for GC) are presented as diamonds. Negative LOG P values are provided on the Y axis. The X axis corresponds to the locations of SNPs. The P value for all samples (original GWA scan plus replication samples) are shown as circles. The P value obtained with the most highly associated SNP (from the original GWA scan plus the replication samples) is shown as a red circle. The SNPs shown as orange diamonds are in $r^2 > 0.8$ (European HapMap CEPH (CEU) samples) with the most significant SNP identified in our study. The recombination rate based on the CEU HapMap is shown in light blue along the x axis (scale on the right). The LD relationship of Illumina discovery SNPs derived from CEU HapMap genotypes are shown below the graph. The most highly associated SNPs are indicated with an asterisk. The green arrows indicate the locations of select genes.

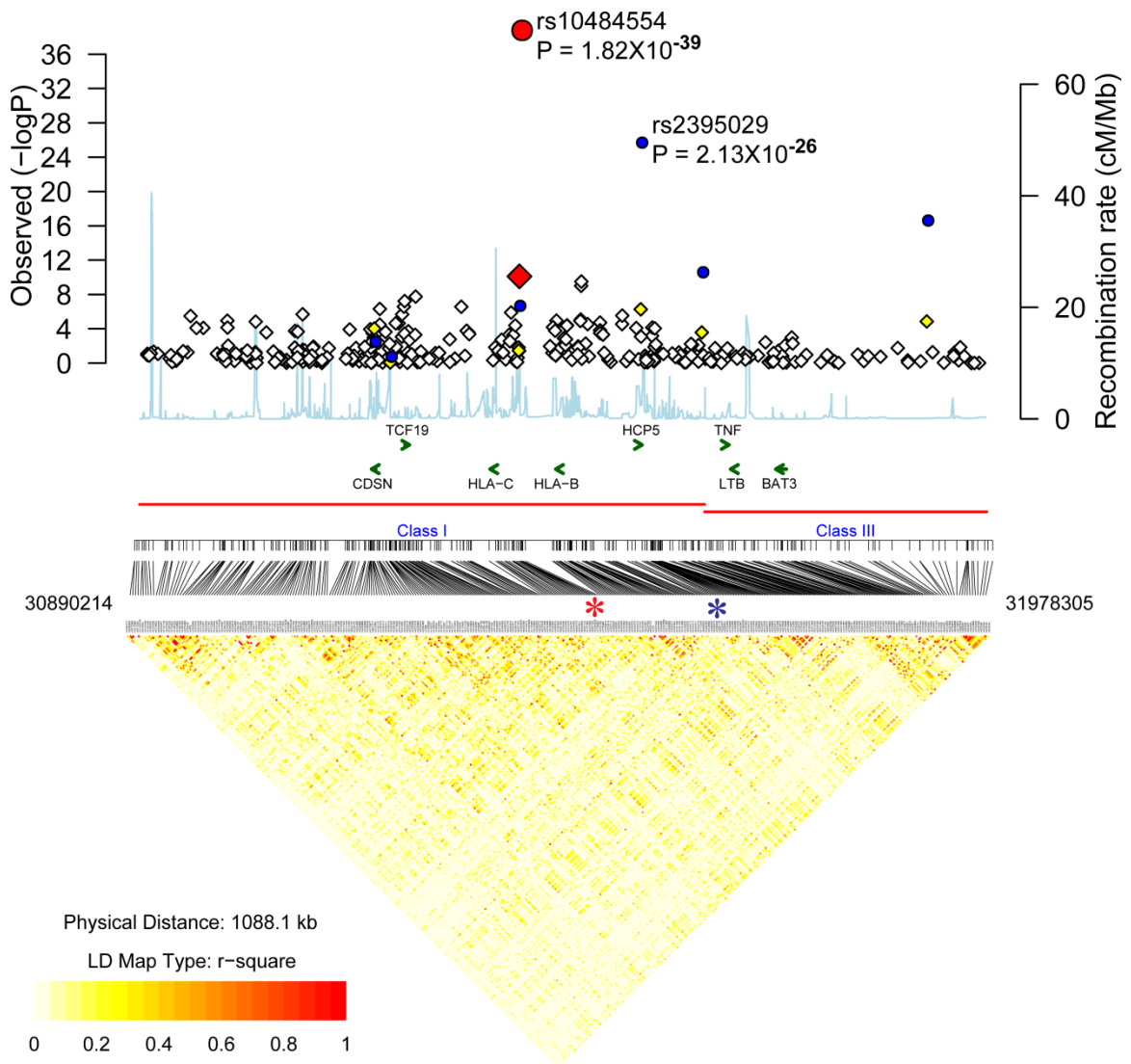


Figure 2.3. Association localization plots for the ILI23R region on chromosome 1. Symbols are the same as those used in **Figure 2.2**. SNPs indicated with an asterisk are rs11465804, rs11209026 (R381Q) and rs12131065.

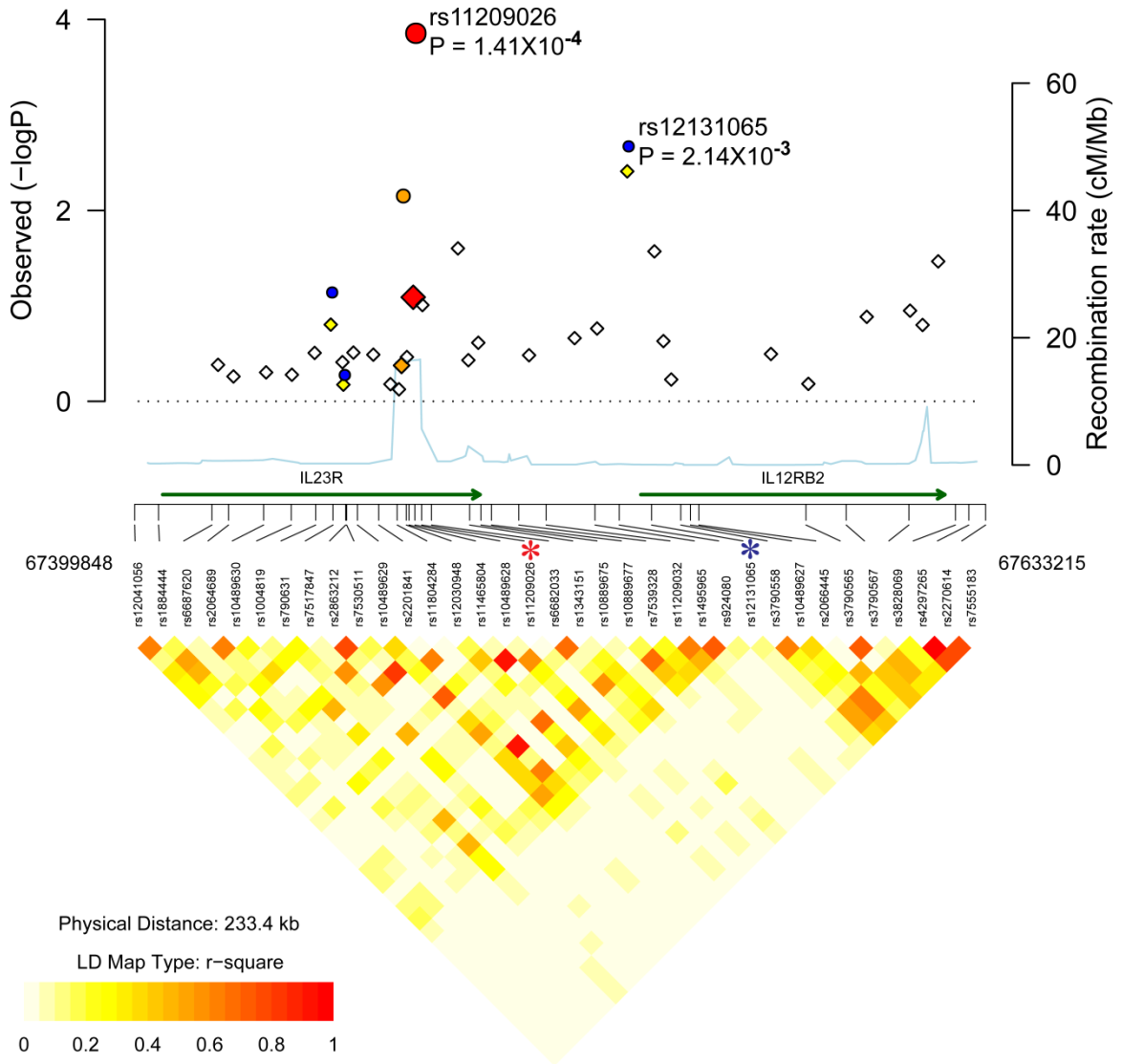


Figure 2.4. Association localization plots for novel replicated region on chromosome 13. Symbols are the same as those used in Figure 2.2. SNPs indicated with an asterisk are rs3812888 and rs7993214.

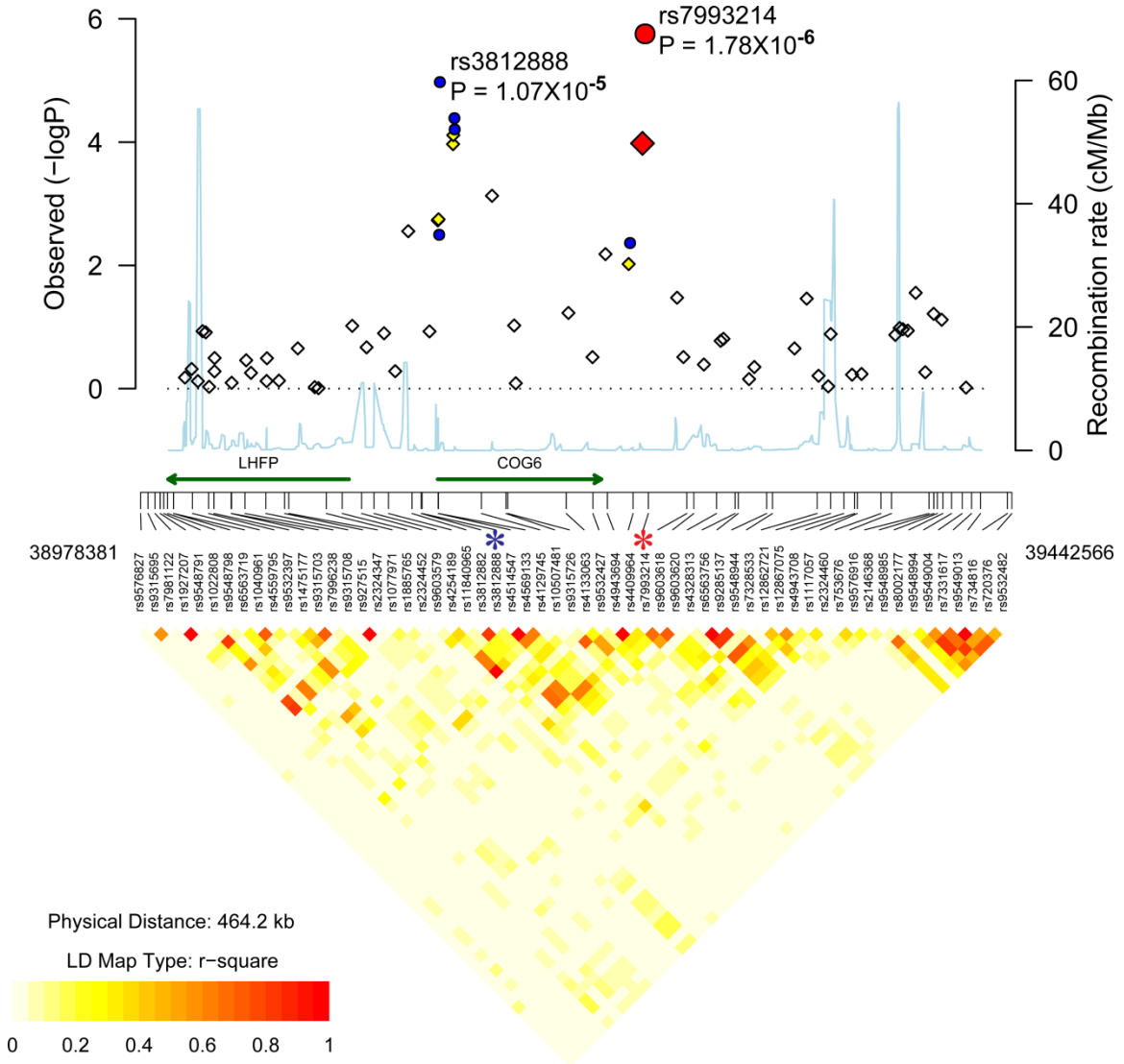


Figure 2.5. Immunostaining of normal, non-lesional and lesional skin for IL2, IL21, COG6 and SPPL2A proteins.

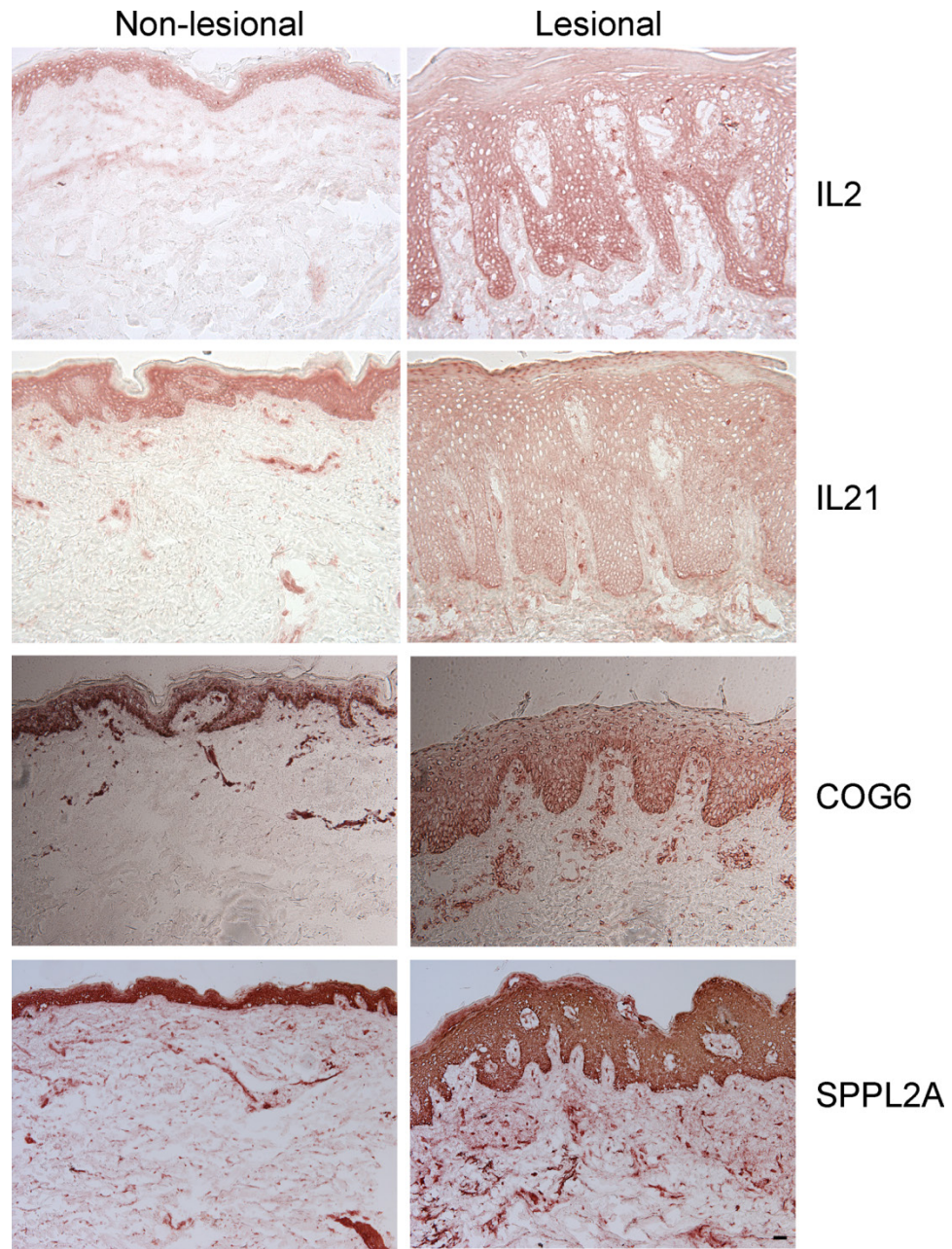
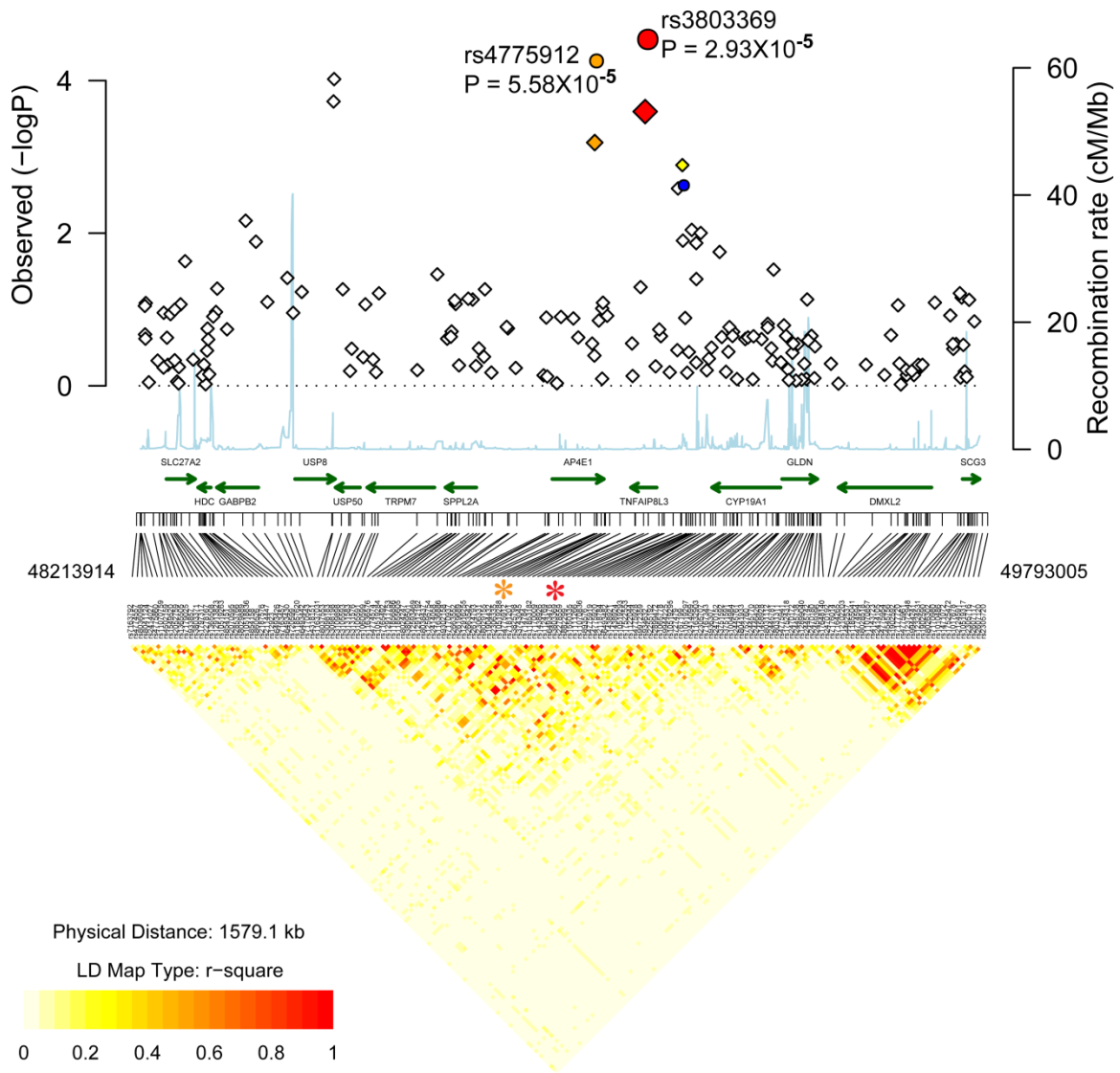


Figure 2.6. Association localization plots for novel replicated region on chromosome 15. Symbols are the same as those used in Figure 2.2. SNPs indicated with an asterisk are rs4775912 and rs3803369.



CHAPTER 3

IDENTIFICATION OF FUNCTIONAL GENETIC

VARIATIONS IN THE PSORS1 REGION

3.1 Introduction

3.1.1 Human genetic variations

An individual's physiologic response to environmental stimuli can be modulated by genetic variation. Millions of genetic alterations (or polymorphisms) occur at >1% in the human population and certain variants can have strong impact on disease susceptibility. Known genetic polymorphisms include tandem repeated segments, copy number variants, small segmental deletions/insertions/duplications, and single nucleotide polymorphisms (SNPs). SNPs, with an average of one SNP every 100–300 base pairs, are the most common genetic variations. The recent build of NCBI's SNP database (dbSNP 127) lists over 11 million identified SNPs in the human genome, with over 5 million validated by multiple investigators. Although most of them have little or no effect on gene regulation and protein activity, there are many circumstances where base changes can have deleterious effects, and therefore affect the way an individual responds to the environment and modify disease risk. SNPs located within the coding region of genes, especially the non-synonymous SNPs, which cause amino acid codon alterations, have been extensively studied because of their obvious impact on protein activities. However, coding SNPs only cover approximately 1.5% of the human genome and in fact the vast majority of the SNPs found to be associated with diseases via GWAS analyses reside in non-coding and/or possibly uncharted regulatory regions of the genome, termed as regulatory SNPs (rSNPs). These variations have become more prevalent in recent literatures due to their potential effect on gene regulatory sequences such as promoters, enhancers, and silencers [174]. Indeed, there are numerous examples of rSNPs associated with disease susceptibility, including hypercholesterolemia,

hyperbilirubinemia, myocardial infarction, acute lung injury, and asthma [175,176,177,178,179,180].

3.1.2 Genetic Variants in the PSORS1 Region

PSORS1 (psoriasis susceptibility gene 1) is a well-confirmed major susceptibility locus for psoriasis on chromosome 6p21.3. It is so far the strongest susceptibility locus [181], accounting for about 30% to 50% of the genetic contribution to the disease. In 1980, Tiilikainen *et al.* first reported a 20-fold increased risk of developing psoriasis in HLA-Cw6 carriers [13], providing the evidence that psoriasis is a T cell-mediated autoimmune disorder. By using HLA haplotype analysis, Schmitt-Egenolf *et al* found an association between psoriasis and the extended haplotype (EH) 57.1, with a 26-fold increased risk of developing early-onset psoriasis compared to EH-57.1-negative individuals [182]. And within EH-57.1, HLA class I antigens (Cw6-B57) were associated to a much greater extent with early-onset psoriasis than the HLA class II alleles. Since then, linkage disequilibrium (LD) mapping of the PSORS1 interval by using microsatellite markers were carried out by three independent research groups [91,183,184]. The objective was to identify the shortest ancestral haplotype segment that retains the PSORS1 disease allele. Although all of them reported extended LD pattern in PSORS1, the exact location and boundaries of the minimal critical PSORS1 region remain poorly defined. By examining a sample of 78 north European families by typing 14 markers spanning the entire MHC, Balendran *et al* suggested a 285-kb critical region between the makers tn821 and HLA-C [183]. In a case/control study of Japanese population, Oka *et al* obtained peaked signal within a 111 kb interval telomeric to HLA-

C [184]. In 2000, Nair *et al* typed 62 markers in 478 psoriasis families and narrowed the candidate interval for PSORS1 to a 60-kb interval telomeric to HLA-C [91].

To further break down the extensive LD and fine mapping of the genetic variation in the PSORS1 region, a more dense and high-resolution single-nucleotide polymorphism-based map is desired. In 2002, Veal *et al* re-sequenced a 150-kb region around HLA-C and identified a total of 59 high-frequency SNPs that were genotyped in 171 independently parent-affected offspring trios [24]. Family-based association study led to identification of strongest association at two SNPs (n. 7 and n. 9), 4 and 7 kb centromeric to HLA-C, respectively. Our lab also performed a comprehensive case/control and family-based association study by using microsatellite and SNP markers that spans a 772-kb segment of the HLA class I region. It pinpointed the location of the PSORS1 to a haplotype block harboring HLA-C that was distinct from CDSN and HCR [126]. Taken together, the extensive LD within this region has made it difficult to identify the true causative variant of PSORS1. A ~300 kb critical PSORS1 interval that contains HLA-C and at least 10 other biologically plausible candidate genes has been proposed to harbor PSORS1. Although recent studies suggested that HLA-Cw6 is the susceptibility allele [86,185], it alone cannot explain the full linkage evidence at the PSORS1 locus, implying that other risk alleles exist within the region [185].

Among all the candidate PSORS1 genes, the HLA-C gene has been intensively investigated as it is well known to be involved in many autoimmune disorders [186]. The leading hypothesis involves its antigen-presenting capacity. Evidence showed that skin-homing, CD8 T cells of HLA-Cw6-positive psoriatics respond more strongly to peptides

common to both keratin 17 and streptococcus than do CD8 T cells of HLA-Cw6-negative psoriatics [187]. A recent study also indicated a spatially distinct HLA-C expression pattern in psoriasis and eczema, implying a functional role of HLA-C in psoriasis-related immune response rather than a general feature of inflammation [188]. In addition to HLA-C, several other genes within the PSORS1 locus, including CDSN [189], HCR [190] and PSORS1C3 [191], also showed strong associations with psoriasis. Some of them are expressed in skin cells which made them very plausible candidates for the PSORS1 gene.

As discussed earlier, the strong LD characteristic of this region has complicated the search for the true functional PSORS1 variant. To overcome this difficulty, functional studies of the known genetic variants within this region may shed light on this problem. The Aim 2 of this thesis was to perform follow-up functional analyses to identify the PSORS1 variant as described below. We hypothesized that the PSORS1 risk alleles in a regulatory sequence can alter transcription of a nearby gene.

3.1.3 Experimental assessment of regulatory SNPs

Successful mapping of functional rSNPs requires the steps of prioritizing putative regulatory sequences or motifs, the co-location of SNPs in these sequences, as well as predicting the allele-specific impacts of rSNPs on transcription factor (TF) binding. Computational methods for the identification of cis-regulatory sequences have been applied in yeast, worm and mammals [192]. While some methods were plagued by high false positive rates in mammals primarily because of the large quantity of intergenic sequence present, many recent new bioinformatics algorithms have improved prediction

[193,194,195].

To validate the predictions of candidate rSNPs, we need an array of experimental approaches to quantitatively examine the molecular properties, i.e. polymorphic effect of rSNPs on gene transcription, binding affinity of a nuclear protein to TF binding sites and the impact of rSNPs in that region. Standard techniques such as luciferase reporter constructs [196] and Electrophoretic gel-Mobility Shift Assay (EMSA) [197] are widely used to assess effects of rSNPs on binding and gene transcription [198]. In this study, functional rSNPs within the PSORS1 region on chromosome 6 were identified and characterized by with these techniques in an attempt to identify the true PSORS1 variant.

3.2 Materials and Methods

3.2.1. Plasmid Construction (Luciferase Assay)

Oligos containing three tandem repeats of 31-bp genomic DNA fragments with the two different alleles (G or A alleles) of SNP9 (rs10456057) and (T or C alleles) of SNP7 (rs12208888) were cloned into the *Kpn1/BgIII* sites of pTA-Luc vector (Clontech). To verify the DNA sequence and orientation, the new constructs were sequenced from both forward and reverse directions (forward sequencing primer: 5'-GTCCCCAGTGCAAGTG-3'; reverse sequencing primer: 5'-GGCGTCTTCCATGGTG-3').

3.2.2 Transient Transfection and Measurement of Luciferase activity

0.5×10^6 293 cells per well were seeded in 12-well plates and transfected with pTA-luc vector alone (control) or pTA-luc vector containing either the psoriasis-risk G and T alleles for SNP9-PS and SNP7-PS, respectively or the normal reference A and C alleles for SNP9-RF and SNP7-RF, respectively, by using TransIT-LT-1 reagent (Mirus). All transfections were carried out in triplicate and repeated in an independent assay. After 24 hour of incubation, the transfected cells were collected, lysed and analyzed for luciferase activity by using the manufacturer's recommended protocol (Luciferase Assay System, Promega).

3.2.3 Electrophoretic mobility assay

[³²P]-labeled (Amersham) double-stranded oligonucleotide were incubated with Jurkat cell nuclear extract (Oncogene Research Products) at 25°C for 20 min using the Gel Shift

Assay System from Promega. We separated the reaction mixture by 4% PAGE and visualized the products by autoradiography. We added unlabeled oligonucleotide at 50-fold molar excess to the reaction for competition. The sequences of the oligonucleotide probes are as follows: SNP9-PS: 5'-TTAGGAAATGCTTGTTATA-3'; SNP9-RF: 5'-TTAGGAAATACTTGTTATA-3'; SNP7-PS: 5'-TCCAGAAATATTATTTGACCC-3' and SNP7-RF : 5'-TCCAGAAATACTATTTGACCC-3'.

3.2.4 Luciferase Assay in Mouse Keratinocyte Cells

In collaboration with Dr. Cristina Strong from NIH, normal and risk allele PSORS1 SNPs (**Table 3.1**) including respective 50-base-pair flanking sequences were cloned into the KpnI/Bgl II sites upstream of a keratinocyte-minimal *Sprr1a* promoter in pGL3 firefly luciferase plasmid (Promega). Dual luciferase assays were performed as described previously [199]. Briefly, mouse keratinocyte cells (SP-1) transfected (Lipofectamine, Invitrogen) with either normal or risk allele SNP pGL3 plasmid were assayed for luciferase activity (Promega) under proliferating (0.05mM Ca²⁺) and differentiating (1.2mM Ca²⁺) conditions at 48 and 72 hours, respectively. Luciferase activity was normalized to *Renilla* luciferase activity and to promoter only vector control.

3.2.5 Bioinformatic Evaluation of PSORS1 variants

To predict possible functional relevance of the detected PSORS1 variants, we used different publicly available bioinformatic tools for predicting transcription factor binding sites in DNA sequences (<http://www.gene-regulation.com>). The programs use different approaches to utilize the library of mononucleotide weight matrices in the TRANSFAC®

[200] (<http://www.gene-regulation.com/cgi-bin/pub/programs/match/bin/match.cgi>) and
Jaspar databases [201] (<http://jaspar.genereg.net>).

3.3 Results

3.3.1 Allele-specific repressor activity of SNP9

Associations with alleles from the HLA class I region on chromosome 6p21 (known as PSORS1, psoriasis susceptibility 1), particularly HLA-Cw*0602, were described over 20 years ago. A variety of genes and regions from a 238-kb interval extending from HLA-B to corneodesmosin (CDSN) have been proposed to harbor PSORS1. However, the exact location of PSORS1 gene remains controversial owing to extensive linkage disequilibrium across this region. In order to identify the minimum block of LD in the MHC class I region associated with psoriasis, our group previously conducted a comprehensive case/control and family-based association study on 242 Northern European psoriasis families and two separate European control populations [126]. Results showed that association was the strongest with single markers and haplotypes from a block of LD harboring HLA-C and SNP9 (rs10456057). SNP9 is an intergenic SNP between HLA-C and HLA-B, lying 4 kb proximal to HLA-C, and SNP9-G allele had been previously reported to be the most highly associated with UK psoriasis families [24]. These studies also highlighted another SNP, SNP7 (rs12208888), a non-coding SNP lying 7 kb proximal to HLA-C. Indeed, SNP7 and SNP9 are in complete LD ($|D'|=1$) and the T allele of SNP7 was always inherited with the G allele of SNP9 on an overtransmitted haplotype (HLA-Cw*0602, SNP9-G and SNP7-T) [24,126].

In order to investigate the potential regulatory role of different SNP9 or SNP7 alleles in nearby gene expression, I subcloned three tandem repeats of nucleotide segment around these SNPs into the pTA-Luc luciferase reporter vector, representing one of the two alleles, psoriasis-risk alleles (G for SNP9-PS and T for SNP7-PS) or normal

reference alleles (A for SNP9-RF and C for SNP7-RF) (**Figure 3.1A**). The constructs were then transiently transfected into human 293 cells. **Figure 3.1B** summarizes the outcomes of report gene activities with respect to different allele constructs. Compared with empty vector controls, the allele-specific construct containing the SNP9-RF allele had a significantly decrease in reporter gene activity by approximately 80%, while the construct containing psoriasis-risk allele (SNP9-PS) completely reversed this effect (**Figure 3.1B**). Constructs for SNP7 alleles (PS or RF) has no effect on the reporter gene activity. The assay was repeated several times and same trend was observed. This result suggests potential allele-specific repressor activity of SNP9, while an A→G mutation may abolish this activity.

3.3.2 Oligos containing SNP9 binds to nuclear extra proteins

To further assess the differential transcriptional activity of SNP9 alleles, electrophoretic mobility shift assay (EMSA) was performed according to standard protocols [202]. **Figure 3.2B** showed an allelic-specific binding of nuclear extracts from a human T cell line (Jurkat) to the oligos containing SNP9-RF (**Figure 3.2B, lane 8**), while others failed to show binding activity to the nuclear extracts from Jurket cells (**Figure 3.2B**). Binding activity for the oligonucleotide containing SNP9-RF resulted in two distinct complexes (arrow heads, **Figure 3.2B, lane 8**). The specificity of binding was further confirmed by competing with 50-fold excess of unlabeled SNP9-RF oligonucleotides (**Figure 3.2B, lane 9**). Taken together, these results clearly suggested a potential allele-specific regulatory role of SNP9 (rs10456057) in nearby gene expression via binding to nuclear transcription factors.

3.3.3 Systematic evaluation of enhancer/repressor activities of potential rSNPs in PSORS1

To systematically assess the enhancer activity of all candidate PSORS1 genetic variants, I expanded the list of SNPs of interest to include all nearby highly-correlated SNPs in PSORS1 region. . **Figure 3.3A** shows the schematic map of genomic locations of these SNPs as well as the surrounding known genes. The LD structure of these SNPs were obtained by querying the SNAP (SNP Annotation and Proxy Search) database, where the pairwise LD was pre-calculated based on phased genotype data from HapMap project (<http://www.broadinstitute.org/mpg/snap/ldsearchpw.php>). **Figure 3.3B** shows the pairwise LD either in Lewontin's |D'| or allelic correlation r^2 . As expected, the plots showed strong LD among these SNPs.

In collaboration with Dr. Cristina Strong from NIH, we further queried these risk SNPs for potential regulatory activity. Each of these SNPs (risk or non-risk alleles) along with the 19-bp flanking sequences on either side was cloned into pGL3i1 luciferase vector (**Table 3.1**). Each construct was then assayed for its ability to enhance/repress luciferase expression driven from mouse minimal Sprr1a keratinocyte promoter [199]. Interestingly, this study identified enhancer activity (> 2-fold luciferase activity) in the risk allele (T) of SNP rs13191343 compared to normal (C) (**Figure 3.4**). This activity was observed under differentiating conditions (**Figure 3.4B**) and not during the cell proliferative (**Figure 3.4A**) state. Notably, this SNP rs13191343, located 1.5 kb upstream of HLA-C (**Figure 3.3A**), was repeatedly reported to be highly associated with psoriasis and psoriatic arthritis in GWAS studies ($p=2.32\times 10^{-72}$, OR = 2.37 [2.16,2.61]) [203,204], and it has strong LD ($r^2 = 0.95$) with SNP rs10484554, the most significantly

associated variant described in our GWAS study described in Chapter 2 (**Table 2.1**). Together, this data suggests a gain in regulatory function in the risk allele of SNP rs13191343 that could likely play a role in the disease.

3.3.4 Genotypes of SNP rs13191343 are associated with CDSN expression

It is highly desirable to find out what gene(s) is affected by this polymorphism. Notably, CDSN, one of the candidate genes for psoriasis susceptibility, encodes the only PSORS1 transcript to be specifically expressed in terminally differentiated keratinocytes [22] and its dysregulation can lead to severe skin barrier defect [205]. It has been reported that its product, corneodesmosin, was up-regulated in the skin lesions of psoriatic patients [189]. Data from GAIN study [204] also confirmed that CDSN is differentially expressed when involved, uninvolved and normal skin are compared (**Figure 3.5A**) with the Kruskal-Wallis test (p -value = 2.82×10^{-5} , significant after multiple comparison correction). We further performed the association analyses for the genotypes (numbers of risk allele) of the PSORS1 SNP rs13191343 and 12 genes in the PSORS1 interval (C6orf15, CCHCR1, CDSN, HCG27, HLA-B, HLA-C, MICA, MICB, PSORS1C1, PSORS1C2, PSORS1C3 and TCF19). Interestingly, CDSN was the only gene whose expression level is significantly increased with PSORS1 risk alleles (**Figure 3.5B**, $p < 0.0001$) after the Bonferroni correction for multiple test. The SNP rs13191343 lies ~1.5 kb to the transcription starting site of HLA-C gene, and ~150kb apart from the CDSN gene. Based on the above evidence, a reasonable hypothesis is formulated: the PSORS1 SNP rs13191343 may alter the expression level of CDSN gene in skin lesions of psoriasis patients via a long range effect. This might affect skin barrier formation; a process which

is disrupted in psoriasis. Interestingly, the expression level of another candidate gene in POSRS1 interval, MICA, was negatively associated with PSORS1 risk allele (**Figure 3.5B**, $p=0.009$). The MICA gene is located about 117 kb centromeric to HLA-C and it encodes the highly polymorphic MHC class I chain-related A, a 43 kDa cell-surface protein expressed in response to stresses such as heat, viral infection, inflammation and DNA damage. Recent studies further suggested that certain MICA alleles are associated with psoriasis and/or PsA in Asian and European populations [206,207,208]. Thus, we can't exclude the possibility that PSORS1 risk alleles may also regulate the MICA gene expression.

3.3.5 Bioinformatic Evaluation of PSORS1 Variants

In order to examine whether any other known transcription factor(s) could bind preferentially to either of the two alleles (risk or non-risk) of rs10456057 (SNP9) and rs13191343, we queried the alleles of these two SNPs along with 5 bp or 9 bp flanking sequences on either side in the MATCH™ 1.0 database [209]. MATCH™, closely interconnected with the TRANSFAC database, is a weight matrix-based tool for searching putative transcription factor binding sites in DNA sequences. By choosing the options of vertebrate matrix and of minimizing both false positive (minFP) and negative rates (minFN), we did not obtain any known binding for all of the queried sequences. We further examined these sequences in JASPAR CORE databases [201]. Again, no known binding site was obtained for all queried sequences with a cutoff value of relative profile score threshold= 80%.

3.4 Discussion

Many DNA variants have been identified on more than 300 diseases and traits using Genome-Wide Association Studies (GWASs). Some have been validated using deep sequencing, but many fewer have been validated functionally. Functional studies primarily focused on those residing in coding regions, specifically non-synonymous coding SNPs (nsSNPs) that result in amino acid changes in proteins, as they are generally easy to interpret in terms of their obvious impact on protein activity. However, numerous SNPs that affect biological function are located outside the coding regions, i.e. in regulatory gene regions: promoters, enhancers, silencers, and introns. A recent study also suggested that synonymous coding SNPs and other non-coding SNPs shared similar likelihood and effect size for disease association with nsSNPs [210]. One appealing functional role for non-coding disease associated regulatory SNPs (rSNPs) is that they can alter the binding affinity of a transcription factor (TF) to the DNA, which in turn change the expression of certain genes, consequently contributing to the disease phenotype. Depending on localization, an SNP in a regulatory region may cause either complete elimination of the natural TF site [211], formation of a novel spurious TF site [212] or quantitative alteration in TF-binding efficiency [213].

Finding these rSNPs is difficult, since they can be obscured by the potentially large number of SNPs present in a linkage-disequilibrium block. This chapter described our effort to interpret previous findings in association studies, i.e. to investigate the potential regulatory role of genetic variations in PSORS1 region. Interestingly, both the luciferase assay on human 293 cells and electrophoretic mobility shift assay indicated an allele-specific regulatory role of SNP9 (rs10456057) in nearby gene expression via

binding to nuclear transcription factors. Thus, it is highly desirable to find out what gene(s) could be affected by SNP9 or other potential regulatory SNPs. As discussed previously, the CDSN gene, over-expressed in psoriatic lesions, could be one of the candidates. It appears that, for many genes, the two alleles possessed by an individual may produce different amounts of transcript [214]. When such allelic differences in transcription are observed for some individuals but not others, a plausible explanation is genetic variation in the *cis*-acting elements that regulate the gene in question on the same chromosome [215]. Thus, it is possible that *cis*-regulatory effect of SNP9 operating on CDSN at long range from a position ~150 kb upstream of the target. A number of novel methodologies have been proposed over past years to test such effect [216]. Recently, Forton *et al.* developed a new analytical approach, called the allelic transcript ratio (ATR), to allow them to calculate the ratio of the abundance of transcripts carrying different alleles within heterozygous individuals [217]. This method was considered relatively robust against environmental confounders, as allelic comparisons were only made within individuals. When applying it to the 5q31 chromosomal region of HapMap families, the authors successfully located a distal (~250kb upstream) highly significant *cis*-regulatory element for IL13 ($P = 2 \times 10^{-6}$) [217]. Actually, complex mechanisms of gene regulation and three-dimensional chromatin configuration had been proposed previously [218]. Of course, we could not rule out the possibility that certain genetic variants play a role on the stabilization of RNA transcripts and thus altered the rate of mRNA decay. Capon *et al.* demonstrated that a single synonymous SNP (CDSN*971T) in CDSN gene accounts for an observed increase in RNA stability. Compared with those transcribed from a neutral haplotype, mRNAs transcribed from risk haplotype bearing CDSN*971T

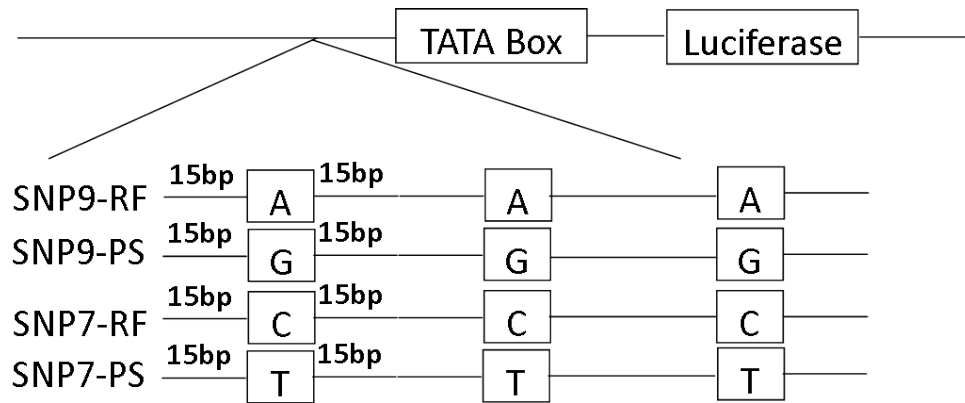
presented a 2-fold increase in stability [219]. Taken together, inspired by these findings, we proposed that variants in potential rSNPs, such as SNP9 or rs13191343, can result in dis-regulation of psoriasis-risk gene transcription. Future focus should be to apply validation techniques to assess the impact these polymorphisms on binding of regulatory machinery as well as functional consequences on transcription.

Table 3.1 Sequences of the plasmid constructs for luciferase assay in mouse keratinocyte cells.

Plasmid_id	SNP	Chrom	Position	Allele	Primer_seq(sense)
PGL3i1_SNP1	rs10456057	6	31353513	Non-risk	GGGGTACCCacttcgtaacttaggaatActgttatatgaactcaatGAAGATCTTC
PGL3i1_SNP2	rs10456057	6	31353513	Risk	GGGGTACCCacttcgtaacttaggaatGctgttatatgaactcaatGAAGATCTTC
PGL3i1_SNP3	rs12191877	6	31360904	Non-risk	GGGGTACCCtaagccaatcatctactcaCctatgatccagcaatagcaGAAGATCTTC
PGL3i1_SNP4	rs12191877	6	31360904	Risk	GGGGTACCCtaagccaatcatctactcaTctatgatccagcaatagcaGAAGATCTTC
PGL3i1_SNP5	rs13191343	6	31349088	Risk	GGGGTACCCctctctgccctctcatccCtcacacctcttccccttGAAGATCTTC
PGL3i1_SNP6	rs13191343	6	31349088	Non-risk	GGGGTACCCctctctgccctctcatccTtcacacctcttccccttGAAGATCTTC
PGL3i1_SNP7	rs13191519	6	31373731	Non-risk	GGGGTACCCtaacagctcctaaatagCctgtgtcctaatcctggaGAAGATCTTC
PGL3i1_SNP8	rs13191519	6	31373731	Risk	GGGGTACCCtaacagctcctaaatagTctgtgtcctaatcctggaGAAGATCTTC
PGL3i1_SNP9	rs13200571	6	31379490	Risk	GGGGTACCCggtgaagtgtgaagaatgaActgagagatgctgactgaGAAGATCTTC
PGL3i1_SNP10	rs13200571	6	31379490	Non-risk	GGGGTACCCggtgaagtgtgaagaatgaCctgagagatgctgactgaGAAGATCTTC
PGL3i1_SNP11	rs13203895	6	31352061	Risk	GGGGTACCCaggctgtaaaaactggttcCtttgcattcccagctatacGAAGATCTTC
PGL3i1_SNP12	rs13203895	6	31352061	Non-risk	GGGGTACCCaggctgtaaaaactggttcTtttgcattcccagctatacGAAGATCTTC
PGL3i1_SNP13	rs13208617	6	31356798	Risk	GGGGTACCCagctgccccgcaagaccaCcaaggcctgtgcaaggtggGAAGATCTTC
PGL3i1_SNP14	rs13208617	6	31356798	Non-risk	GGGGTACCCagctgccccgcaagaccaTcaaggcctgtgcaaggtggGAAGATCTTC
PGL3i1_SNP15	rs2524074	6	31352000	Non-risk	GGGGTACCCgccattgcacaggacagtCtacttcccacttgaaaaGAAGATCTTC
PGL3i1_SNP16	rs2524074	6	31352000	Risk	GGGGTACCCgccattgcacaggacagtTtacttcccacttgaaaaGAAGATCTTC
PGL3i1_SNP17	rs7751729	6	31379915	Risk	GGGGTACCCatgccaatccccgacttccCaggacccccagtaattttgGAAGATCTTC
PGL3i1_SNP18	rs7751729	6	31379915	Non-risk	GGGGTACCCatgccaatccccgacttccTaggacccccagtaattttgGAAGATCTTC
PGL3i1_SNP19	rs9468933	6	31373036	Non-risk	GGGGTACCCaaaagtatatattatgcataAttgattgtttcatttctGAAGATCTTC
PGL3i1_SNP20	rs9468933	6	31373036	Risk	GGGGTACCCaaaagtatatattatgcataTttgattgtttcatttctGAAGATCTTC
PGL3i1_SNP21	rs12204500	6	31184376	Non-risk	GGGGTACCCcggcctactcttaatttttAttttaagtgtctcaagctaaGAAGATCTTC
PGL3i1_SNP22	rs12204500	6	31184376	Risk	GGGGTACCCcggcctactcttaatttttTtttaagtgtctcaagctaaGAAGATCTTC
PGL3i1_SNP23	rs12207756	6	31194661	Risk	GGGGTACCCaggctaaagtcaataacaCggtctcagctcactgtaacGAAGATCTTC
PGL3i1_SNP24	rs12207756	6	31194661	Non-risk	GGGGTACCCaggctaaagtcaataacaTggtctcagctcactgtaacGAAGATCTTC
PGL3i1_SNP25	rs3094214	6	31193361	Non-risk	GGGGTACCCagaaaaccgagaggccgatGactgagataaggcagaagGAAGATCTTC
PGL3i1_SNP26	rs3094214	6	31193361	Risk	GGGGTACCCagaaaaccgagaggccgatTactgagataaggcagaagGAAGATCTTC
PGL3i1_SNP27	rs3132555	6	31190889	Risk	GGGGTACCCtctctctacttccagctCtcccacagaggaggaagaGAAGATCTTC
PGL3i1_SNP28	rs3132555	6	31190889	Non-risk	GGGGTACCCtctctctacttccagctGtcccacagaggaggaagaGAAGATCTTC
PGL3i1_SNP29	rs3778640	6	31202099	Non-risk	GGGGTACCCctgatgccagtggtgggcaCgaccgtgctgtatactttaGAAGATCTTC
PGL3i1_SNP30	rs3778640	6	31202099	Risk	GGGGTACCCctgatgccagtggtgggcaGgaccgtgctgtatactttaGAAGATCTTC
PGL3i1_SNP31	rs6929464	6	31205897	Non-risk	GGGGTACCCtttttctctcctaactaaTccacgttattgcttggGAAGATCTTC
PGL3i1_SNP32	rs6929464	6	31205897	Risk	GGGGTACCCtttttctctcctaactaaTtccacgttattgcttggGAAGATCTTC
PGL3i1_SNP33	rs12059256	1	150782055	Risk	GGGGTACCCggaaaattcctgagcagatAagagtgggaaatgaaagtGAAGATCTTC
PGL3i1_SNP34	rs12059256	1	150782055	Non-risk	GGGGTACCCggaaaattcctgagcagatGagagtgggaaatgaaagtGAAGATCTTC
PGL3i1_SNP35	rs6702463	1	150865335	Non-risk	GGGGTACCCttgaaattaaaatcacattAtgtagacaatatataggaGAAGATCTTC
PGL3i1_SNP36	rs6702463	1	150865335	Risk	GGGGTACCCttgaaattaaaatcacattTgttagacaatatataggaGAAGATCTTC

Figure 3.1 Repressor activity of SNP9 (rs10456057). 296 cells were transfected with luciferase reporter constructs containing either psoriasis-associated (SNP9-PS and SNP7-PS) or untransmitted reference alleles (SNP9-RF and SNP7-RF) of SNP9. Luciferase activity (mean \pm s.d., $n = 6$) is presented in relative luciferase units. Student T test yielded significant differing effects for the presence of reference allele A vs. psoriasis-associated G allele of SNP9 (p value = 5.37×10^{-10}).

A.



B.

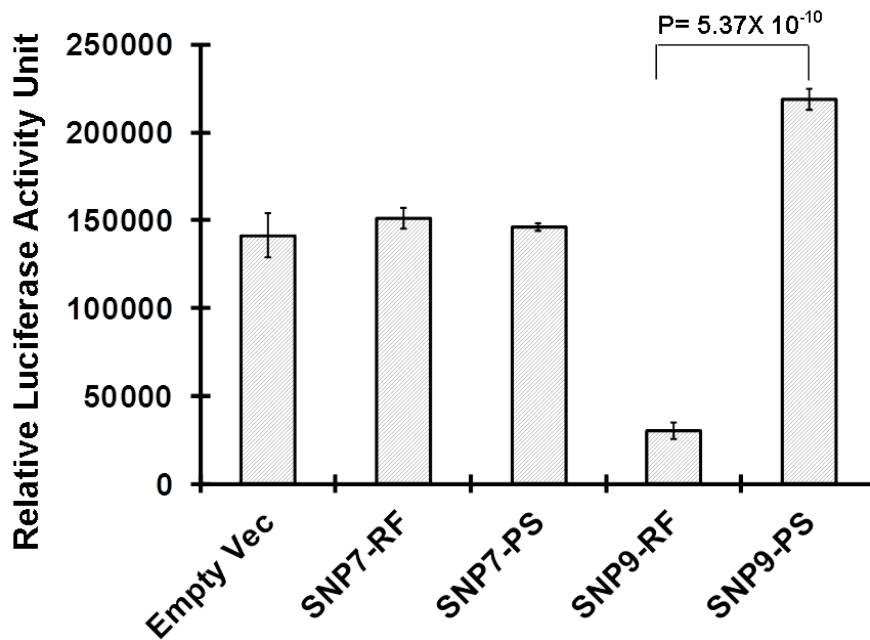


Figure 3.2. Electrophoretic mobility shift and competition assays with Jurkat nuclear cell extract for allelic variants of SNP9 and SNP7. A. The 19-bp oligonucleotide containing either A (SNP9-RF, normal reference allele) or G (SNP9-PS, psoriasis-risk allele) and 21-bp oligonucleotide containing either C (SNP7-RF) or T (SNP7-PS) were assayed with Jurkat cell extracts. The alleles were indicated in bold italic and underlined. B. Binding was present for the oligonucleotide containing SNP9-RF (lane 8), which resulted in two distinct complexes (arrow heads) that were competed by 50-fold excess of unlabeled SNP9-RF oligonucleotide (lane 9).

A.

SNP9-RF(sense) 5'-TTAGGAAAT**A**CTTGTTATA-3'
 SNP9-PS(sense) 5'-TTAGGAAAT**G**CTTGTTATA-3'
 SNP7-RF(sense) 5'-TCCAGAAATA**C**TATTTGACCC-3'
 SNP7-PS(sense) 5'-TCCAGAAATA**T**TATTTGACCC-3'

B.

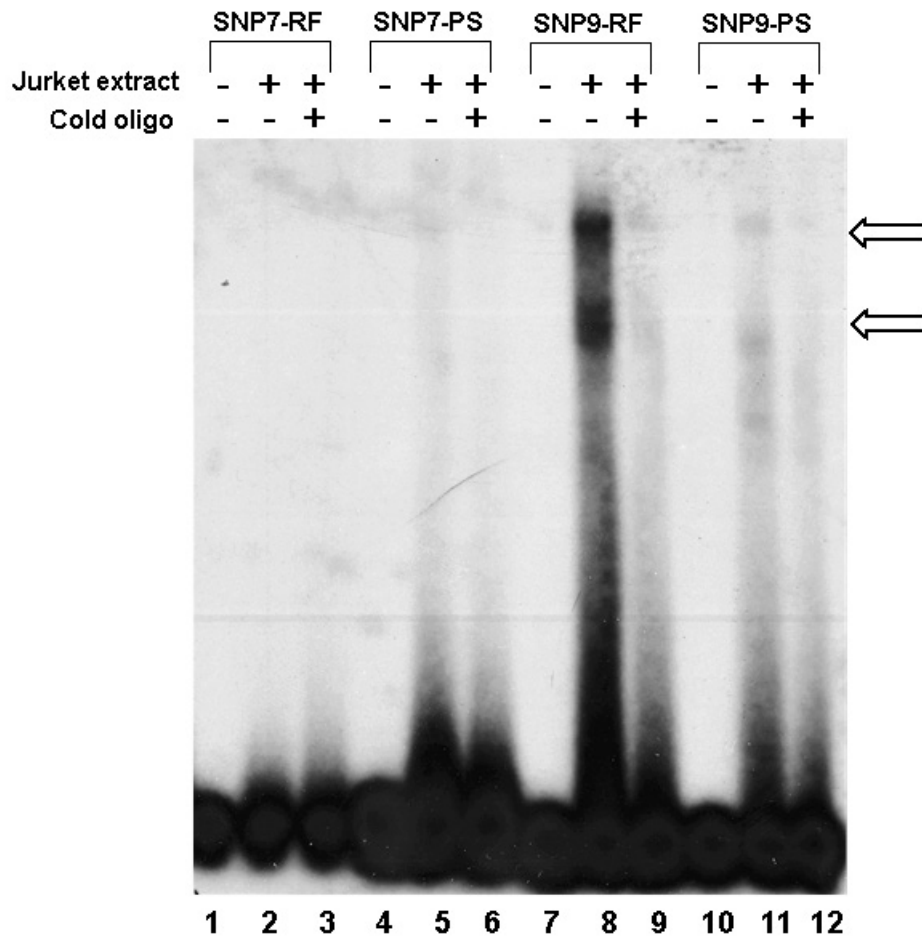
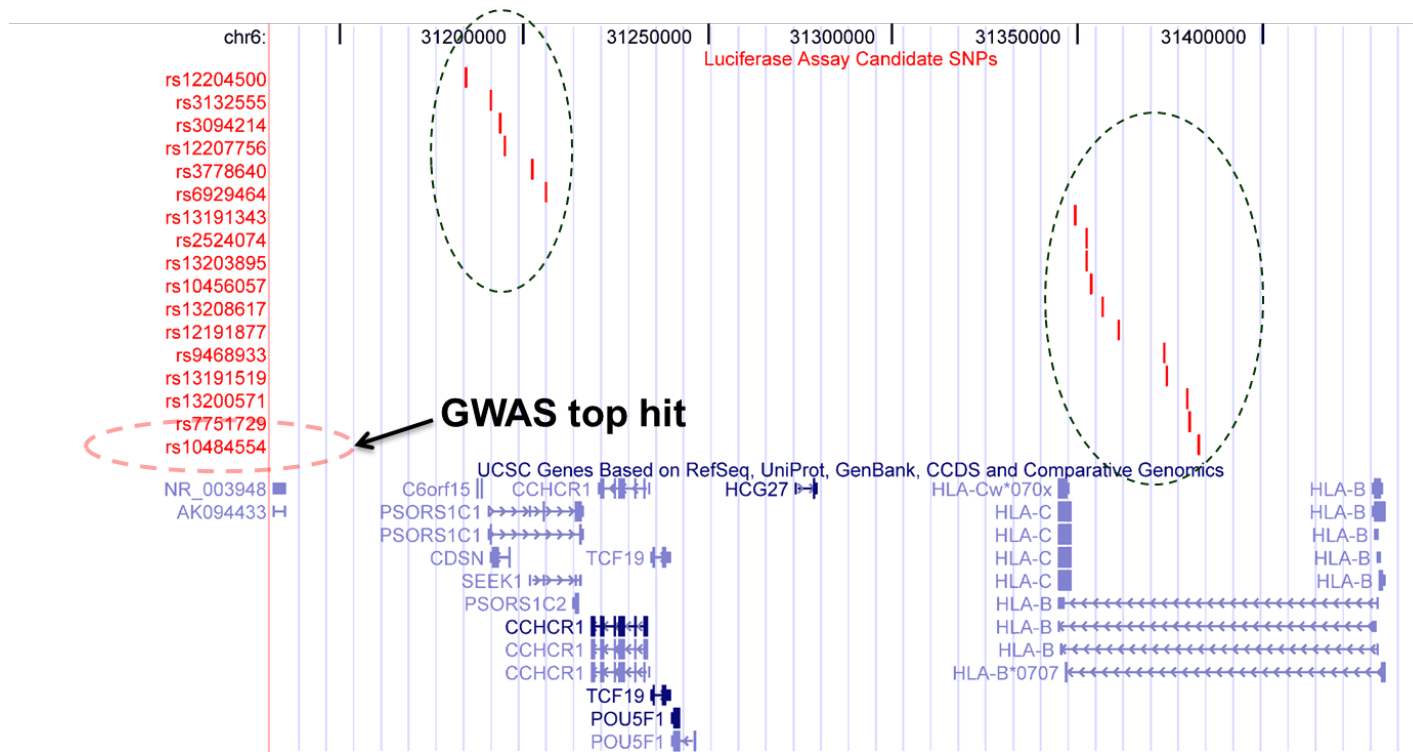


Figure 3.3 Scheme for candidate rSNPs and nearby genes in PSORS1 region. A. Highly correlated PSORS1 SNPs are shown with their genomic coordinates and candidate associated genes in this region. The top associated SNP identified in GWAS (Chapter 2) was indicated by red circle. B. LD patterns among PSORS1 SNPs in either $|D'|$ (left panel) or r^2 (right panel).

A.



B.

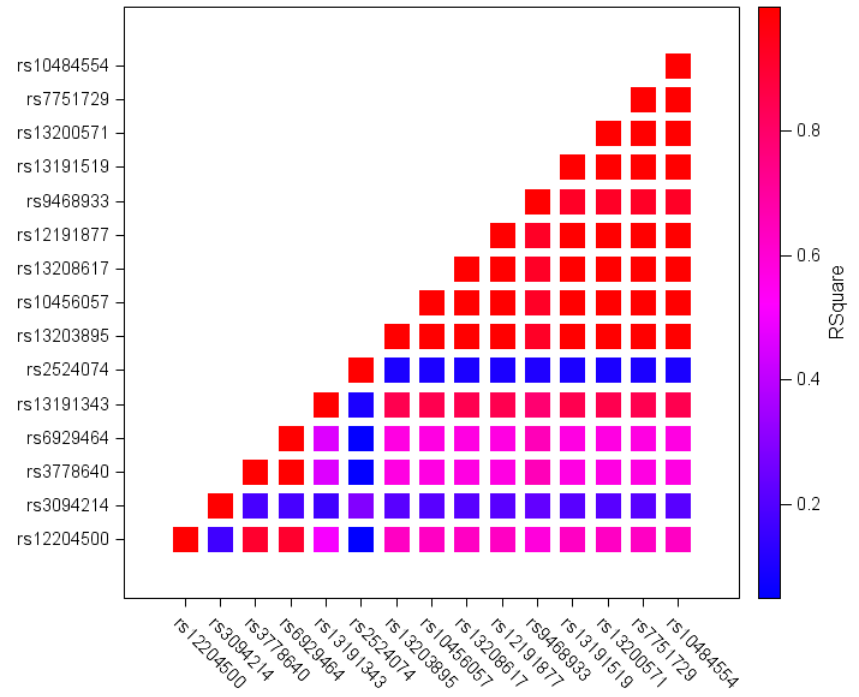
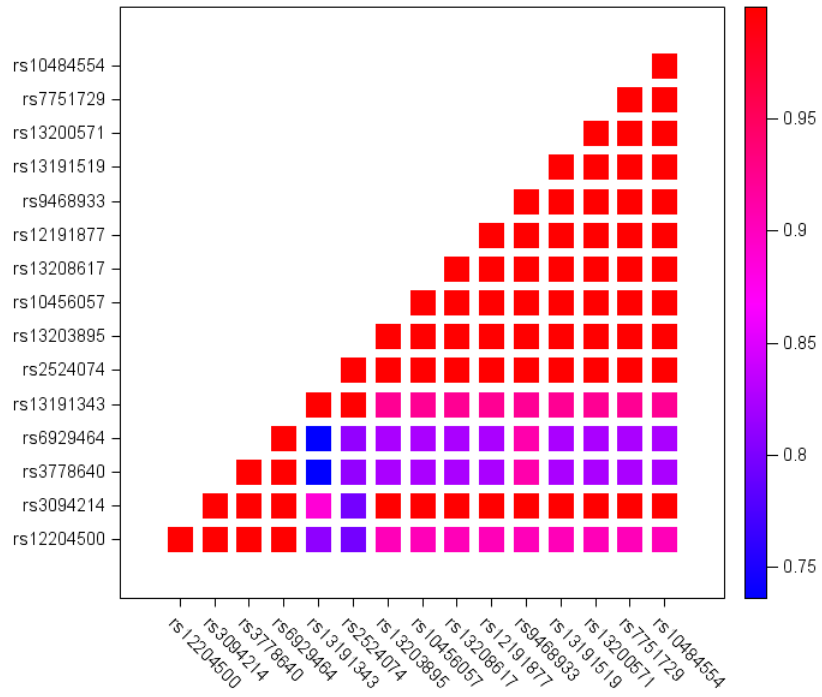
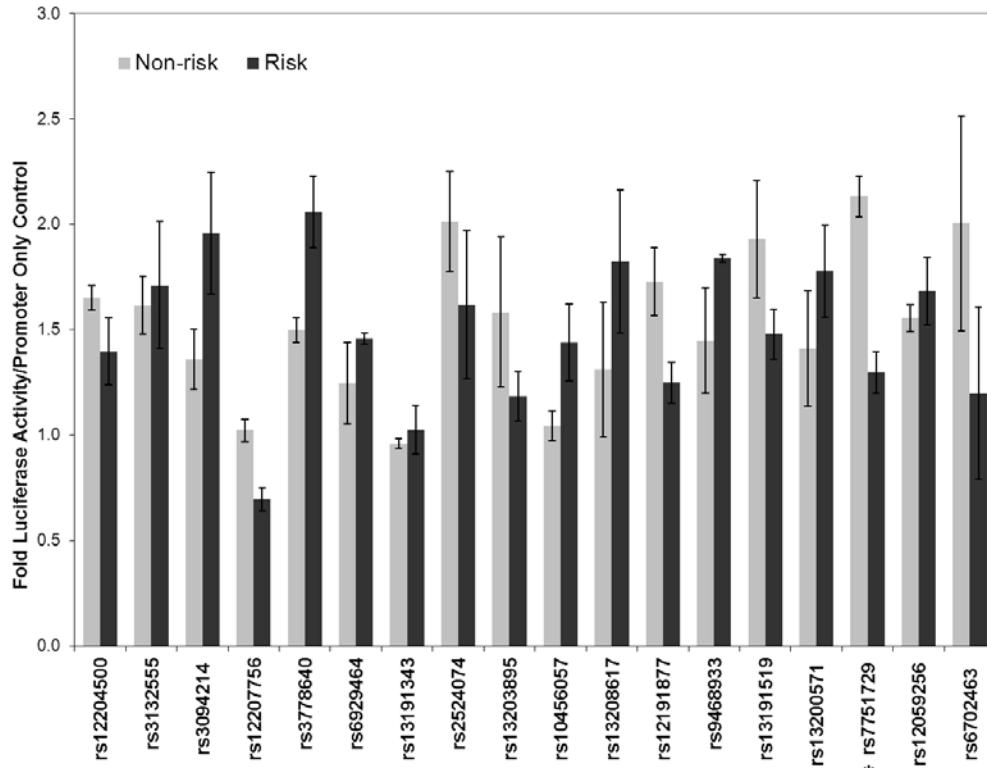


Figure 3.4 Reporter activity of PSORS1 alleles in keratinocytes. Luciferase readings represent an average of two independent, duplicate experiments that were performed under proliferating (A) or differentiating conditions (B). The SNPs whose differential reporter activity for the two alleles passed the Bonferroni correction for multiple test were indicated by asterisk.

A.



B.

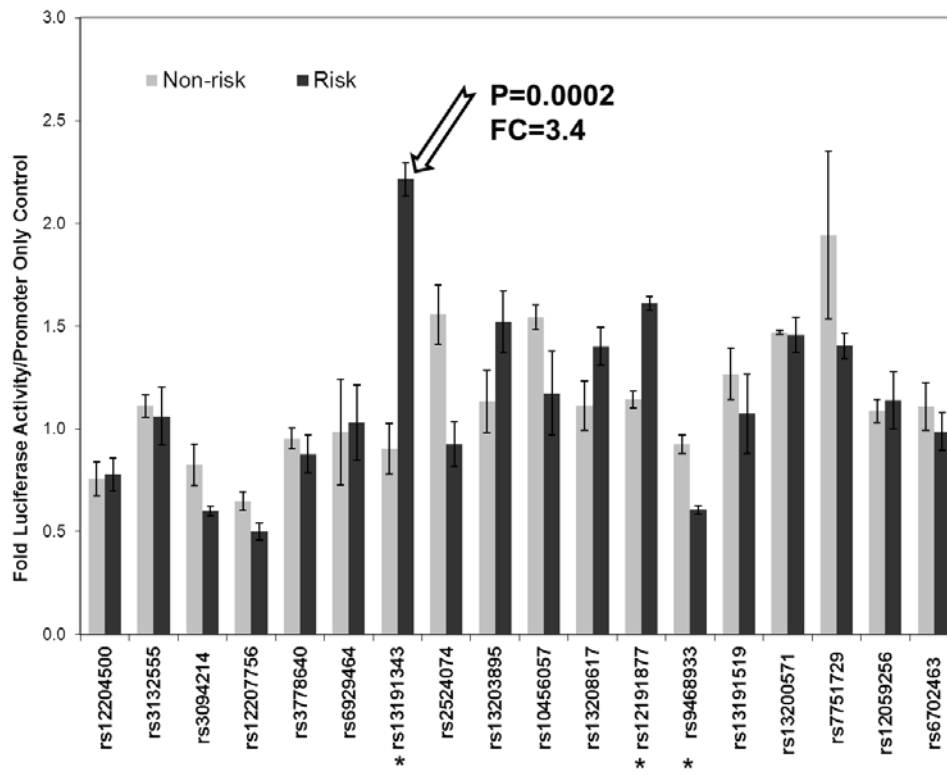
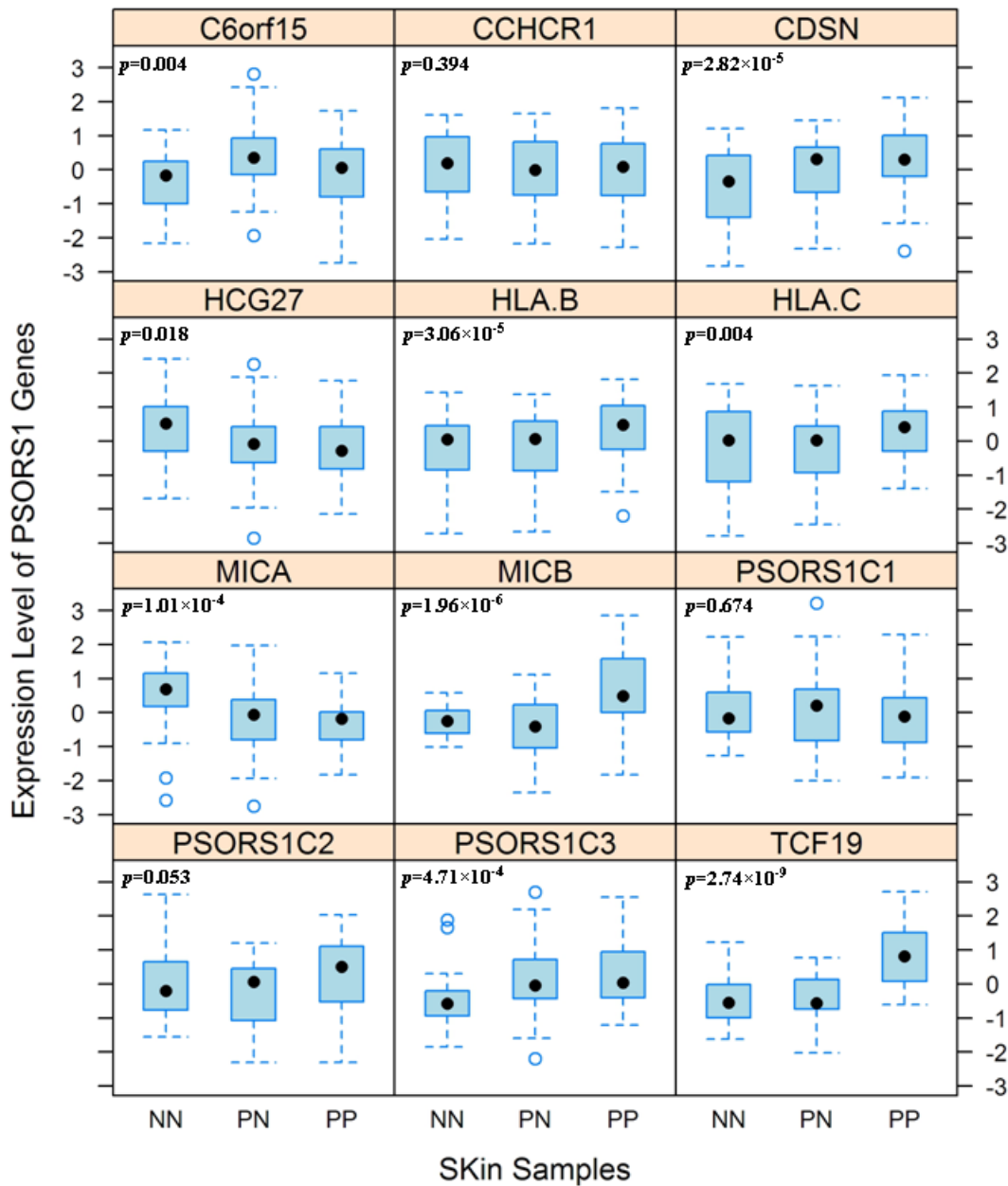
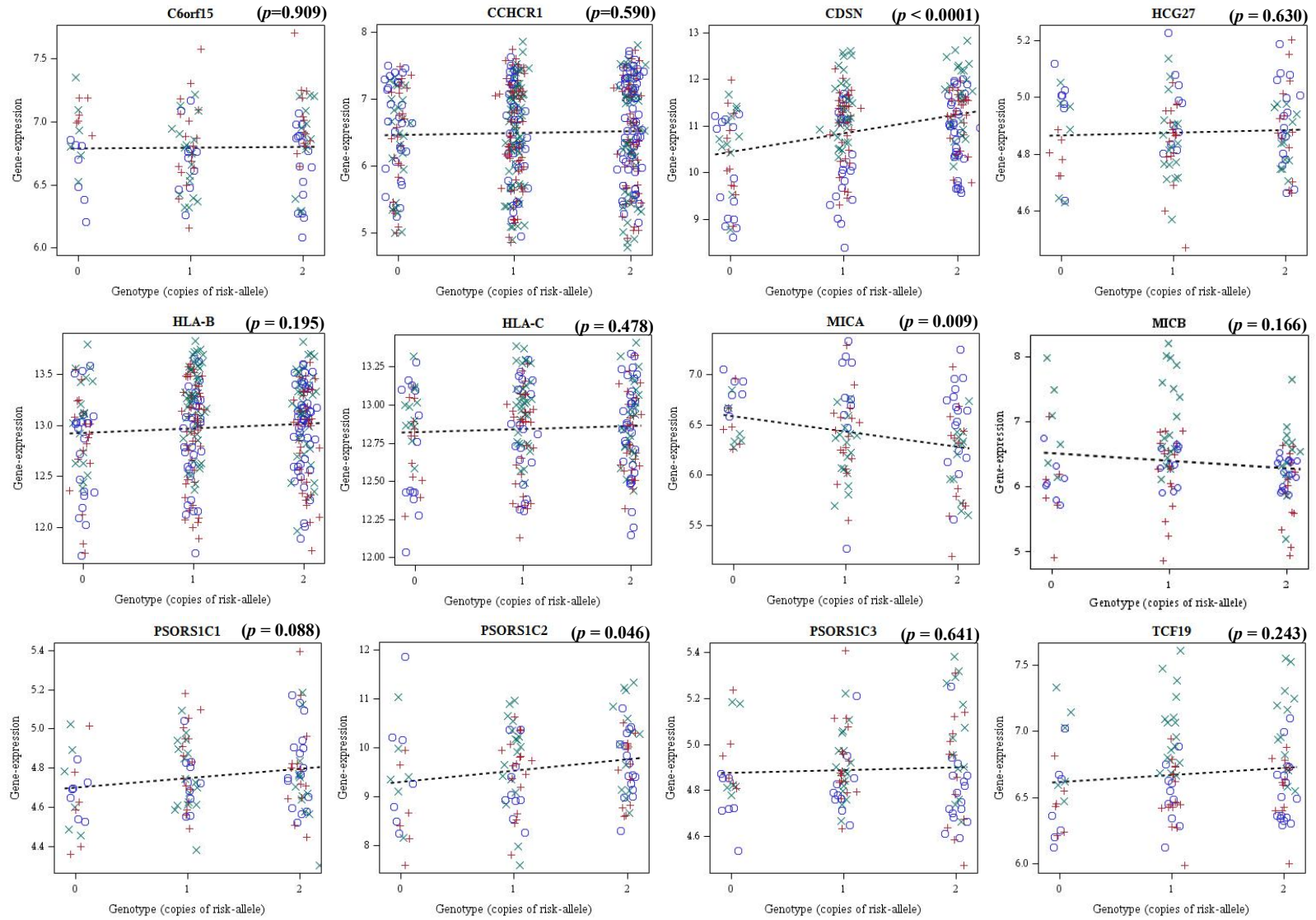


Figure 3.5 CDSN gene expression is associated with genotypes of PSORS1 SNPs. A. Boxplot panel shows the expression levels of 11 PSORS1 genes in psoriasis involved (PP), uninvolved (PN) and normal (NN) skin biopsies. The median values were indicated in black dots. Displayed p-values were derived from the non-parametric Kruskal-Wallis test for equality of medians between groups. B. Association of genotypes (number of risk alleles) of PSORS1 SNP with gene expression. P-values of testing the null hypothesis for slope of regression line equals zero were also shown.

A.



B.



CHAPTER 4

GLOBAL METHYLATION CHANGES IN PSORIATIC

SKIN

4.1 Introduction

Psoriasis is a chronic, relapsing inflammatory skin disease affecting 2-3% of the U.S. population [6,220,221]. Outbreaks or flares can occur at any time throughout an individual's lifetime, and typically follows a chronic cycle of remission and relapse. Flares can be exacerbated by stress, infection, or other environmental triggers [222]. In psoriasis, immune cell activation and altered epidermal differentiation are key pathogenic events [223,224] and these are correlated with major changes in the transcriptome [7,220,221,225,226,227,228].

Epigenetic alterations, such as DNA methylation and histone modification are correlated with gene expression changes [49,229,230,231]. Such alterations may be part of normal developmental or differentiation processes but can also be triggered by environmental factors [56,57,58,59,60]. In mammals, DNA methylation commonly occurs at CpG dinucleotides [50]. Approximately 70-80% of the CpG dinucleotides in the human genome are methylated, predominately in areas harboring repetitive sequences [51]. However, regions rich in CpGs, termed CpG islands (CGIs), are also found in promoters of more than 70% of annotated genes [52,53]. Approximately half of CGIs are associated with annotated gene transcription start sites [51], while others can have discrete set of CpG sites within their promoters. The methylation of these sites has direct effects on transcriptional levels, where methylation levels typically demonstrate an inverse correlation with expression level [232].

There have been only a few studies of epigenetic alterations in diseased tissue. Many of these have involved cancerous tissue where the methylation status of tumor genomes are compared to matched normal tissue [233,234,235,236]. However, studies of

methylation changes in the diseased tissues of patients with complex diseases, including those leading to autoimmunity, are limited since diseased tissue is often difficult to access. A study on epigenetic changes in the blood of systemic lupus erythematosus patients revealed altered methylation of several genes contributing to T-cell autoreactivity, B-cell overstimulation and macrophage killing [76]. Psoriasis is more tractable than many autoimmune diseases due to the accessibility of its target organ: the skin, and there have been a few reports of altered methylation within promoters of single genes in diseased skin. For example, the *SHP-1* (PTPN6) promoter is reported to be demethylated in psoriatic skin but not in Atopic Dermatitis (AD) or normal skin [77]. However, genome-wide studies of methylation changes in psoriasis have not been described.

This chapter of the dissertation describes a study on global changes of methylation in involved psoriatic skin versus uninvolved psoriatic and normal skin. This was performed by querying 27,578 CpG sites with Illumina bead arrays with DNA derived from skin of involved, uninvolved and normal skin. Hierarchical clustering of 50 of the top differentially methylated sites differentiated separated all psoriatic skin samples (involved and uninvolved) from normal skin. Methylation at 12 CpG sites was also significantly correlated with expression levels of a nearby gene. Research described in this chapter was accomplished in collaboration with Dr. Eli Roberson.

4.2 Materials and Methods

4.2.1 Skin biopsy samples

Three to six millimeter punch biopsies were obtained from the PP and PN skin of psoriasis patients and NN skin from healthy controls (**Appendix II, Table A2.1**). The transcriptomes of some of these samples were previously analyzed and are described elsewhere [7,228]. Skin biopsies were obtained from collaborating dermatologists at Washington University School of Medicine (Saint Louis, MO), Psoriasis clinic, Baylor Hospital (Dallas, TX) or from the University of California in San Francisco (CA). Informed consent was obtained from all individuals who donated skin biopsies. Protocols for obtaining patient biopsies were approved by Institutional Review Boards for the protection of human subjects.

4.2.2 DNA methylation profiling using bead array

Qiagen DNeasy Kits were used to isolate genomic DNA from skin biopsy samples according to the manufacturer's instructions. All samples were analyzed for DNA integrity, purity and concentration on a Nanodrop Spectrophotometer DN-100 (Nanodrop Technologies). The EZ DNA methylation kit (Zymo Research) was used for bisulfite conversion of all DNA samples (1 μ g) according to the manufacturer's recommendations [237]. Bisulfite-converted genomic DNA was then interrogated with the Illumina Infinium HumanMethylation27 Beadchip, with the recommended protocols provided by the manufacturer. After hybridization, the arrays were imaged with a BeadArray Reader scanner. Image processing, intensity data extraction and analyses were conducted with the BeadArray Reader.

4.2.3 Differential Methylation Analysis

All methylation data was loaded into a single GenomeStudio project. Non-normalized values were exported for statistical analysis. Exported data were analyzed using the R (v2.12.0) Bioconductor (Biobase v2.10.0) methylumi (v1.4.0) and lumi (v2.2.0) packages [238,239,240,241]. The Cy3 and Cy5 color channels within each array were balanced using the ‘lumiMethyC’ function to quantile normalize the intensities. The color balanced data was normalized using simple scaling normalization via the ‘lumiMethyN’ function. In some tables we report β -values along with M-values. β -values are more intuitive, corresponding to percent methylation. M-values are used for statistical tests since they better approximate a homoskedastic distribution. Let the intensity of the methylated and unmethylated alleles be I_{meth} and I_{unmeth} , respectively.

$$\beta\text{-value calculation (percent methylation): } \beta = \frac{I_{meth}}{I_{meth} + I_{unmeth}}$$

$$M\text{-value calculation (methylation ratio): } M = \log_2 \left(\frac{I_{meth}}{I_{unmeth}} \right)$$

Limma (v3.6.6) was used to fit linear models to each CpG region detected (detection p-value ≤ 0.01) in at least one sample [242]. After linear model fitting appropriate contrasts were defined for PP versus NN and PN versus NN. The log-odds of differential methylation was calculated from the linear fits for each CpG in each contrast using the ‘eBayes’ function.

Paired PP/PN samples were treated differently. In this case the paired samples were extracted from the rest of the data and filtered to probes detected in at least one sample. Linear models were fitted to the data. Separate factor vectors were created to represent the sample class and the sample identifier. A design matrix was created

combining the two, and 'eBayes' function was used to calculate the moderated paired t-test for the combination of the two factors. All p-values were adjusted for multiple tests using the false discovery rate method [243]. Power calculations were performed in R 2.12.0 [244].

4.2.4 Selection of the top 50 group discriminating differentially methylated CpG sites

Between-group analysis was used to determine a subset of CpG sites that most differentiate PP from NN skin. Normalized M-value data was reduced to CpG sites differentially methylated between PP and NN skin. Between-group analysis was performed using the reduced dataset and a training vector specifying the group membership of each sample with the 'bga' function of the MADE4 R package [245,246]. The discriminating method used was principal components analysis. The top sites were selected as the top 25 increased and top 25 decreased methylation sites on the first principal component axis. The top 50 sites were subsequently used to generate heatmaps showing the discriminatory power of these sites with Euclidean distance measures and complete hierarchical clustering on the axes. All heatmaps were generated using the Heatplus R package (v1.20.0) with a 50 color palette from the marray package maPalette function (v1.28.0). Pathway and network analyses for the top 50 genes/sites were carried out using GeneGo MetaCore[®] software (GeneGo, Inc., St. Joseph, MI, USA). The significance of biological pathways is estimated through a variation of Fisher's exact test as implemented in the MetaCore software package and adjusted for multiple testing using Benjamini-Hochberg FDR analysis.

4.2.5 Correlation with gene expression

Pearson product-moment correlations coefficients (r) and their 95% confidence intervals were calculated to evaluate the strength of linear dependence between methylation at specific CpG loci and the level of expression of a downstream target. The FDR adjusted p-values were calculated to test the null hypothesis of zero correlation. All analyses were performed using R statistical programming language (v2.10.1).

4.2.6 Pyrosequencing

CpG methylation at and around sites flanking the statistically significant Illumina CpG locus upstream from target genes were further validated by using Pyrosequencing approach. This allowed us to quantify methylation at multiple CpG sites individually [247]. Sample bisulfite treatment, PCR amplification, pyrosequencing, and extraction of percent methylation were performed at EpigenDx (Worcester, MA). Sequences analyzed were promoter regions of IFI27, LGALS3BP, SERPINB4 and C10orf99 genes (**Appendix II, Table A2.2**).

4.3 Results

4.3.1 Differential CpG site methylation in psoriatic skin

We used the high throughput genome-wide bead-array (Infinium HumanMethylation27 Beadchip, Illumina, Inc., USA) to obtain a global, quantitative measure of the methylation status of CpG sites in psoriatic involved (PP), psoriatic uninvolved (PN) and normal skin (NN). The array spanned 27,578 CpG loci selected from more than 14,000 genes, including more than 1,000 cancer-related genes and the promoter regions of 110 miRNAs. The vast majority of assayed CpG sites were located in the promoter regions of their cognate genes with an average distance of 365 bp (maximum ~1.5kb) from their transcription start sites.

PP skin samples were defined as skin biopsies collected from the site of an active psoriatic lesion. Conversely, PN skin samples were biopsies collected from skin that was not part of an active lesion. NN skin biopsies were defined as those biopsies collected from healthy volunteers with no clinically apparent skin lesions and no self-reported history of outbreaks. Our study included 12 PP, 8 PN and 10 NN skin samples. The PN samples were all derived from a donor who also contributed a PP sample, hence there were 8 “paired” PP/PN samples and 4 additional PP samples without a matched PN sample. For each CpG target on each array we calculated both β - and M-values. The M-value was calculated by taking the \log_2 ratio of detected signal intensities from methylated and unmethylated alleles after color channel balance and normalization (Methods). The percent methylation (β -value) was calculated by dividing the methylated allele intensity by the total intensity of alleles for each CpG, and therefore approximating percent methylation. Though more intuitive, the β values are quite heteroskedastic,

making them less appropriate for standard parametric statistical tests [248]. The M-values were used for statistical analysis since they are more homoskedastic (**Appendix II, Figure A2.2**).

We defined a CpG as differentially methylated if it had a FDR corrected p-value significance threshold of 0.05. CpG methylation in PP versus NN skin differed at 1,108 CpG sites, 88 of which demonstrated a greater than 2-fold change (**Figure 4.1**). The top differentially methylated sites for this comparison are shown in Table 1. A total of 27 CpG sites demonstrated differential methylation in PP skin compared to PN skin from the same individual and 2 of those sites had a greater than 2-fold change in M-value (**Figure 4.1**). Interestingly, PN skin compared to NN skin showed differential methylation at 15 CpGs, 8 of which were greater than 2-fold (**Figure 4.1**).

The largest number of methylation differences and the differences of the largest magnitude were seen in the comparison of PP skin compared to NN skin (**Table 4.1**). The observed alterations ranged from profound ($\geq 20\%$ difference) to more subtle ($\leq 10\%$ difference). A total of 96 genes had at least two CpG sites in their vicinity with a significant p-value. CCND1 and GATA4 had 4 significant sites each, while GPX3 and SFRP4 had 3 significant sites each. The most extreme change was found in cg16139316, which is lies ~ 400 bp upstream from S100A9 (p-value < 0.00001) and within the epidermal differentiation complex, a region key to epidermal development [199]. For this CpG site, the raw β -value in PP skin was 0.46 ± 0.08 , whereas the methylation level in NN skin was 0.84 ± 0.01 . It has been shown that S100A9 are absent or are expressed at minimal levels in normal epidermis, but strongly overexpressed in the basal and spinous layers in psoriasis-involved tissue [100]. Our methylation data indicated a decreased

methylation level in psoriatic skin, which is consistent with its enhanced expression. In total, twelve differentially methylated CpG sites close to genes upregulated in psoriasis (S100A3, S100A5, S100A7, S100A12, SPRR2A, SPRR2D, SPRR2E, LCE3A) mapped to the EDC (**Figure 4.2**).

Differences in PP versus PN skin were subtle compared to those observed in transcript levels which contrasts with what is seen with expression analysis, where PP and PN skin are very similar [7,249]. Out of all significant differences, methylation levels typically different by less than 10%. The largest fold increases in PP compared to PN skin was in MCF.2 cell line derived transforming sequence-like (MCF2L; FC = 2.40; $\beta_{PP} = 0.45 \pm 0.04$; $\beta_{PN} = 0.26 \pm 0.02$) and laminin alpha 4 (LAMA4; FC = 2.58 ; $\beta_{PP} = 0.37 \pm 0.05$; $\beta_{PN} = 0.19 \pm 0.02$). The largest fold decreases were in synaptopodin (SYNPO; FC = -1.91; $\beta_{PP} = 0.55 \pm 0.04$; $\beta_{PN} = 0.68 \pm 0.03$) and bone marrow stromal cell antigen 2 precursor (BST2; FC = -1.76; $\beta_{PP} = 0.20 \pm 0.03$; $\beta_{PN} = 0.30 \pm 0.03$).

Methylation differences in PN compared to NN skin were similarly few in number and even more subtle ($\leq 10\%$). The largest magnitude fold changes (≥ 2) were all increases in methylation in uninvolved skin. These included sites near GALR1, ZNF454, ZNF540, NEF3, RGS7, MLF1, FLJ42486 and NRIP2. The greatest decrease in methylation (-1.81 fold) was in a CpG site approximately 500bp upstream of the ZDHHC12 promoter.

4.3.2 Validation of differential methylation with pyrosequencing

We used a separate approach (pyrosequencing) to confirm methylation differences in CpG sites located in four genes (C10orf99, IFI27, LGALS3BP, SERPINB4) and to

investigate differential methylation at these regions further. By array analysis these sites had demonstrated decreased methylation of varying fold-change in PP skin compared to NN skin (C10orf99 = -1.35; LGALS3BP = -1.56; IFI27 = -2.74; SERPINB4 = -1.44). In all cases, the original CpG site determined to be differentially methylated with the Illumina bead assay was included in the pyrosequencing assay, along with nearby CpG sites.

IFI27, which had the largest fold-change by methylation array, showed statistically significant decreased methylation at of PP compared to both PN and NN at 6 of 9 sites with pyrosequencing (**Figure 4.3**). An additional site, CpG7, was significant only for PP compared to NN. Two sites, CpG6 and CpG9, did not show evidence of differential methylation. All three C10orf99 CpG sites demonstrated decreased methylation in involved skin, versus normal skin by pyrosequencing (**Figure 4.4**). In the case of LGALS3BP only two of the seven CpG sites that were assayed (CpG3 and CpG4), showed any evidence for differential methylation of PP compared to PN (**Figure 4.5**). The single CpG site tested in SERPINB4 demonstrated significantly decreased methylation in PP compared to both PN and NN skin (**Figure 4.6**).

4.3.3 Methylation levels correctly classify involved, uninvolved, and normal skin samples

We hypothesized that methylation levels of differentially methylated CpG sites could be used to discriminate the different skin groups. Using between group analysis with principal component metrics we identified a subset of 50 sites (25 with increased methylation, 25 with decreased methylation) that differentiated PP from NN skin (**Appendix II, Table A2.3**). Data on an additional seven PP samples was obtained for

cross-validation of clustering validity. A heatmap of normalized M-values at the top 50 differentiating sites was generated with all PP, PN and NN samples (**Figure 4.7**). The confusion matrix showed in **Figure 4.7B** demonstrated excellent classifying power of hierarchical clustering by using these sites. Classification of psoriatic (PP or PN) versus NN were 100% accurate and 100% specific. PP clustered separately from both PN and NN skin, and performed better, with 100% sensitivity and 90% specificity. PN was classified with 75% sensitivity and 100% specificity. The lower sensitivity for PN samples was due to two PN samples (PN4, PN5) being classified as PP. Based on this dataset the classifying power of the global methylation data performed very well, especially at the classification of psoriatic versus normal, and may be as good a predictor of psoriasis as gene expression values.

To examine whether gene expression data from skin sample would equally well predict and classify the skin samples, I reanalyzed a previously published gene expression dataset [7] including paired biopsies of involved and uninvolved skin from 16 individuals. The expression values for each sample were normalized using simple scaling normalization via the 'scale' function in R package. Only the probe-sets with at least 2-fold changes in expression levels when comparing PP, PN and NN samples were included for further analysis. Between-group analysis was used to determine a subset of genes that most differentiate PP from NN skin (same procedures as I did for methylation study). A heatmap of top 50 differentiating genes was shown in **Appendix II, Figure A2.1**. These differential genes demonstrated excellent classifying power of classification of psoriatic PP versus (PN, NN) samples. All PP clustered separately from both PN and NN skin except only one PN being classified as PP. This result indicates that similar to

the top 50 differential methylation sites, expression data of top 50 differential genes can equally well separate and predict psoriatic skin from normal ones.

To identify functional relationships among the differentially methylated CpG sites and their cognated genes, I carried out pathway analysis using GeneGo MetaCore software. **Table 4.3** lists the top 4 functional pathways (FDR adj. $p < 0.10$) identified using this software as well as the p-value associated with the number of differentially methylated genes in each pathway. As shown in **Table 4.3**, the most significantly affected pathway was the cell adhesion and extracellular matrix (ECM) remodeling pathway ($p < 3.9 \times 10^{-4}$). Of the 52 genes/proteins in this pathway, 3 genes exhibited increased methylation in PP skin. Interestingly, the ECM pathway is involved in physiological processes, such as cell motility and adhesion, wound healing, angiogenesis, as well as in disease processes, such as arthritis and metastasis. I also performed a network analysis by using all the top 50 differential methylated genes in the MetaCore database, which also incorporates canonical pathways in its algorithm. This resulted in a list of 15 signaling networks (not shown) ordered according to their significance. The most significant network ($p < 1.4 \times 10^{-13}$) is shown in **Figure 4.8**. This network again highlights the interplay of two key processes: immune response (IL-13 signaling via JAK-STAT) and cell adhesion/ECM remodeling, implying that altered methylation levels in the genes of these processes may play an important role in the pathophysiology of psoriasis.

Finally, we plotted these top 50 sites, separated by methylation change observed in involved skin and by sample group (**Figure 4.9**). The medians of the three groups for sites with increased and sites with decreased methylation were significantly different by

the Kruskal-Wallis rank sum test. The trend is apparent for both the raw β -values and the normalized M-values. We also observed that in the case of these top 50 sites, PN skin had a methylation level intermediate to that of the NN and PP skin (**Figure 4.9**). These intermediate methylation levels contrast with the expression levels of mRNA transcripts in uninvolved skin which are usually very similar to that of normal skin [228]. These differences suggest that there are intrinsic epigenetic differences in uninvolved versus normal skin that may be reflective of a predisposition to psoriasis.

4.3.4 Correlation of methylation with nearby gene expression

Nine PP, five PN, and six NN samples used for methylation analysis had also been used for global transcriptome analysis with the Affymetrix U95 arrays [7]. We were therefore able to perform a direct correlation between methylation at specific CpG loci and the level of expression of a downstream target for these samples.

Correlations between methylation score values and nearby gene expression levels were performed with R and p-values were reported based on an FDR corrected p-value cutoff of 0.05. There were 12 CpG sites that significantly correlated with gene expression levels (adj. p-value \leq 0.05; **Table 4.2**). Among those, 9 demonstrated negative correlations and 3 showed positive correlations. These CpG sites were near the genes C10ORF99, OAS2 (3 sites), LGALS3BP, KYNU, GDPD3, IL1B, TRIM22, CCND1, TRIM14 and PHYHIP. Many of these genes (C10ORF99, OAS2, LGALS3BP, KYNU, IL1B, TRIM22) are highly up-regulated in psoriasis, and all of these exhibited negative correlation between expression and methylation [7], providing evidence of underlying methylation changes in the highly up-regulated genes in involved skin.

A consensus list of 890 down-regulated and 732 up-regulated genes in psoriatic skin determined across expression studies was recently published [220]. 128 differentially methylated CpG sites in PP compared to NN were near 113 genes in that consensus list. For example, the genes CCL27, DDAH2, TNS1 and TRIM2 all showed consistent down-regulation in psoriatic skin and we found consistently increased methylation in and near these genes. By contrast, IFI27, KYNU, OAS2, S100A9, SERPINB3 and TNIP3 all showed significantly increased expression in psoriasis, and we found significantly decreased methylation for sites near them. One gene in the consensus set, FCGBP, is significantly down-regulated in psoriasis lesions, but we found significantly decreased CpG methylation approximately 430bp upstream of this gene at cg19103704.

When correlations between gene expression and CpG methylation at the top 50 discriminating genes were made, only two genes were significant with adj. P value. These were c10orf99 and KYNU, which were both negatively correlated (**Figure 4.10**). Failure to correlate some of the other genes might be due to lack of power or paucity of coverage by the Illumina bead arrays. However, some differentially CpG sites lie near genes not reported to be differentially expressed in psoriasis (e.g. MTSS1, MCF2L). These sites are of interest for future studies.

Interestingly KYNU is one of a set of genes whose high transcriptional activation differentiates transcripts psoriasis from a second inflammatory disease such as atopic dermatitis [250]. This gene is upregulated ~100 fold compared to normal skin, and ~58 fold compared to atopic dermatitis lesional skin [250]. Our observations indicate that methylation status at loci such as this would be as valuable a predictor of psoriasis as

expression alterations.

4.4 Discussion

There have been many advances in our understanding of psoriasis histology, pathophysiology, genetics and molecular biology. Still, there are many questions yet to be answered, such as what triggers an initial outbreak and why do patients have a chronic, relapsing course? Relapses require either retreatment or persistent treatment in psoriasis patients. Those whose disease does not respond well to standard treatments require biologic treatment, such as with adalimumab or etanercept, that can cost tens of thousands of dollars per year [251]. Improving the treatment of psoriasis patients and developing long-term, cost-effective treatments requires an understanding of the molecular basis of disease initiation and progression. Genetic variants that contribute to psoriasis risk and the transcriptome of PP skin have been relatively well studied. Global methylation in psoriasis, however, has not. Methylation status can directly influence transcription levels, so knowledge of global methylation in involved skin is an important step forward.

Based on our analysis, there are extensive differences in global methylation in PP skin compared to NN. Furthermore, methylation differences, though fewer in number, are apparent between PP compared to PN skin as well as between PN compared to NN skin. The differentially methylated CpG sites that map to the epidermal differentiation complex (EDC) are particularly interesting, due to the critical role this region in the development and remodeling of the epidermis. The genes near the EDC differentially methylated sites, as well as genes near other differentially methylated CpG sites such as IFI27, OAS2 and KYNU have been shown to be highly upregulated in involved psoriatic skin [220,221,228,252,253]. Interestingly KYNU is one of a set of genes whose high

expression level differentiates psoriasis from other inflammatory skin diseases such as atopic dermatitis [250]. This gene is upregulated ~100 fold compared to normal skin, and ~58 fold compared to atopic dermatitis lesional skin [250].

The differential methylation of BST2 in PP compared to PN is particularly interesting, especially given its role in response to type-1 interferon signaling and location in a previously identified psoriasis susceptibility locus (19p13.11) [254]. Its product is an antiviral protein tetherin, also known as CD317, which is expressed on plasmacytoid dendritic cells. Tetherin inhibits virus budding [255] and also mediates the adhesion of monocytes to vascular endothelium [256]. Up-regulation of BST2 could therefore provide a mechanism to attract monocytes to psoriatic lesions that then could differentiate into macrophages and dendritic cells. In mouse, BST2 is mainly expressed in plasmacytoid dendritic cells initially, and in many other cells following stimulation with type-1 interferons [257]. The decreased methylation at this locus could reflect an increased expression of BST2 in response to the psoriasis T_H1 phenotype, as well as an increase in plasmacytoid dendritic cells in psoriatic lesions.

The large number of differentially methylated sites in PP versus NN contrasts with the relatively small number of sites identified by the PP to PN and PN to NN comparisons. One reason for this disparity may be lack of power to detect low percentage methylation changes (**Appendix II, Figure A2.3**). PP compared to NN typically had changes in methylation $\geq 10\%$ with standard deviations of less than 5%. A two-sample t-test (two-sided) with 8 samples in each group has 96% power to detect a difference of 10%. Since most of the observed changes had similar dispersion but greater magnitude difference, this suggests that the PP to NN comparison is well-powered to detect the

types of changes we observed. However, the PP to PN and PN to NN comparisons were testing methylation changes of much smaller percentage difference and reduced power. This in itself is interesting, however, since it reflects the intermediate methylation that PN skin exhibited compared to PP and NN skin. Since methylation status can be stably inherited, such differential methylation could persist even after resolution of symptoms, predisposing to recurrence of lesions at the same sites. Some of these intermediate methylation states were independently confirmed by pyrosequencing. Perturbations in methylation status may also account for the “pre-psoriatic” transcriptional signature which has been described in uninvolved psoriatic skin compared with normal skin.

Furthermore, when methylation levels of the most significant 50 CpG sites were used to cluster genes and samples there was efficient classification of PP, PN and NN samples. In fact, Classification of psoriatic (PP or PN) versus NN were 100% accurate and 100% specific. This is similar to what we have described elsewhere with respect to mRNA transcripts [228]. Therefore, methylation status may be useful in the classification of samples, and expression changes in key genes may be linked to their underlying methylation status.

In summary, the idealized future for psoriasis treatment would target the underlying molecular defect directly. There are many successful therapies from UV light treatment to biologic infusions. These treatments reduce symptoms but do not cure the disease. Based on the findings of this global methylation study, we may conclude that differential methylation could contribute to this chronic phenotype, and further study is warranted to elucidate the evolution and function of differential methylation in psoriasis pathogenesis.

Table 4.1 CPG sites exhibiting two-fold differences in methylation status in involved versus normal/uninvolved and involved/uninvolved versus normal skin.

CpG ID	Symbol	Chrom	Position hg18	PP vs. NN		PN vs. NN	
				Fold Change	FDR p-value	Fold Change	FDR p-value
cg14826683	SPRR2D	1	151,280,454	-2.49	≤ 0.001	-2.00	0.068
cg06509239	OSR1	2	19,422,050	2.31	0.025	2.17	0.162
cg03544320	CRMP1	4	5,945,592	3.54	0.019	2.94	0.179
cg15433631	IRX2	5	2,804,541	2.23	0.014	2.10	0.127
cg17820459	GPX3	5	150,380,724	2.14	0.032	2.54	0.077
cg03355526	ZNF454	5	178,301,021	2.54	0.001	2.60	0.014
cg18055007	DDAH2	6	31,806,205	2.05	0.008	2.20	0.057
cg26521404	HOXA9	7	27,171,506	2.94	0.023	2.58	0.179
cg23290344	NEF3	8	24,827,371	2.28	0.008	2.84	0.016
cg08441806	NKX6-2	10	134,449,139	2.12	0.027	2.22	0.118
cg05194726	NRIP2	12	2,814,741	2.72	≤ 0.001	2.11	0.045
cg22881914	NID2	14	51,605,897	2.64	0.022	2.59	0.137
cg03734874	FLJ42486	14	104,142,427	2.64	0.001	2.71	0.017
cg17861230	PDE4C	19	18,204,901	2.09	0.023	2.11	0.122
cg26267310	DHRS10	19	54,032,405	2.61	0.001	2.03	0.114
cg02440177	ZNF702	19	58,188,507	2.17	0.014	2.48	0.057
cg24713204	ZNF471	19	61,711,185	2.33	0.004	2.06	0.104
cg18074297	CLIC6	21	34,963,482	3.05	0.004	2.77	0.077

Table 4.2 CpG sites that correlated with expression levels. CpG sites with corresponding genes on the expression arrays were tested for correlation with gene expression. CpG ID: the Illumina CpG identifier. Chr & Position_hg18: the mapping information for the CpG site on build hg18. Gene: symbol of gene closest to the CpG site tested. TSS Distance (bp): the distance in base pairs of the CpG site to the transcriptional start site (TSS). Expression Probe: the Affymetrix identifier of the probe tested for correlation. Adj p-value: the p-value corrected by false-discovery rate. Correlation: the calculated Pearson correlation coefficient between methylation (normalized M-values) and expression presented as mean \pm 95% confidence intervals.

CpG ID	Chr	Position hg18	Gene	TSS Distance (bp)	Expression Probe	Adj p-value	Correlation
cg10045881	10	85,922,743	C10orf99	804	53747_at	0.0013	-0.749 \pm 0.152
cg04872689	12	111,900,901	OAS2	244	39263_at	0.0051	-0.772 \pm 0.138
cg23173910	17	74,487,605	LGALS3BP	51	37754_at	0.0066	-0.776 \pm 0.136
cg15836722	2	143,351,601	KYNU	64	40671_g_at	0.0113	-0.825 \pm 0.108
cg06131859	12	111,900,901	OAS2	244	39264_at	0.0135	-0.853 \pm 0.091
cg11134443	12	111,900,901	OAS2	244	90662_at	0.0187	-0.752 \pm 0.150
cg06806080	16	30,031,794	GDPD3	585	58504_at	0.0187	0.752 \pm 0.312
cg06131859	2	113,310,256	IL1B	571	1520_s_at	0.0187	-0.760 \pm 0.146
cg06131859	11	5,667,230	TRIM22	434	36825_at	0.0187	-0.743 \pm 0.155
cg18655915	11	69,170,515	CCND1	NA	2020_at	0.0274	0.755 \pm 0.311
cg00483154	9	99,922,197	TRIM14	888	33253_at	0.0274	0.778 \pm 0.290
cg05947740	8	22,145,723	PHYHIP	174	37191_at	0.0493	-0.792 \pm 0.127

Table 4.3 Top functional pathways identified by MetaCore pathway analysis.

Rank	GeneGo Pathways	P-value	Differentially methylated genes	Differentially methylated genes (n)	Genes/proteins in the pathway (n)
1	Cell adhesion_ECM remodeling	3.877E-04	KLK1, KLK2, LAMA4	3	52
2	G-protein signaling_RhoA regulation pathway	3.986E-03	MCF2L, EFNA3	2	34
3	Immune response_IL-13 signaling via JAK-STAT	6.903E-03	CCL17, SERPINB3	2	45
4	Development_PEDF signaling	8.145E-03	CFLAR(L), CFLAR(S)	2	49

Figure 4.1 Venn diagram of the CpG sites exhibiting differential methylation for each of three contrasts using a significance cutoff of 0.05 for the adjusted p-value. The contrasts are PP compared to PN (paired t-test), PP compared to NN and PN compared to NN. For each set the upper number is a count of the number of CpG sites with increased methylation, and the lower number is the count of CpG sites with decreased methylation. The total count of unique sites showing increased or decreased methylation in at least one comparison is shown at the bottom right.

Differentially Methylated CpG Sites

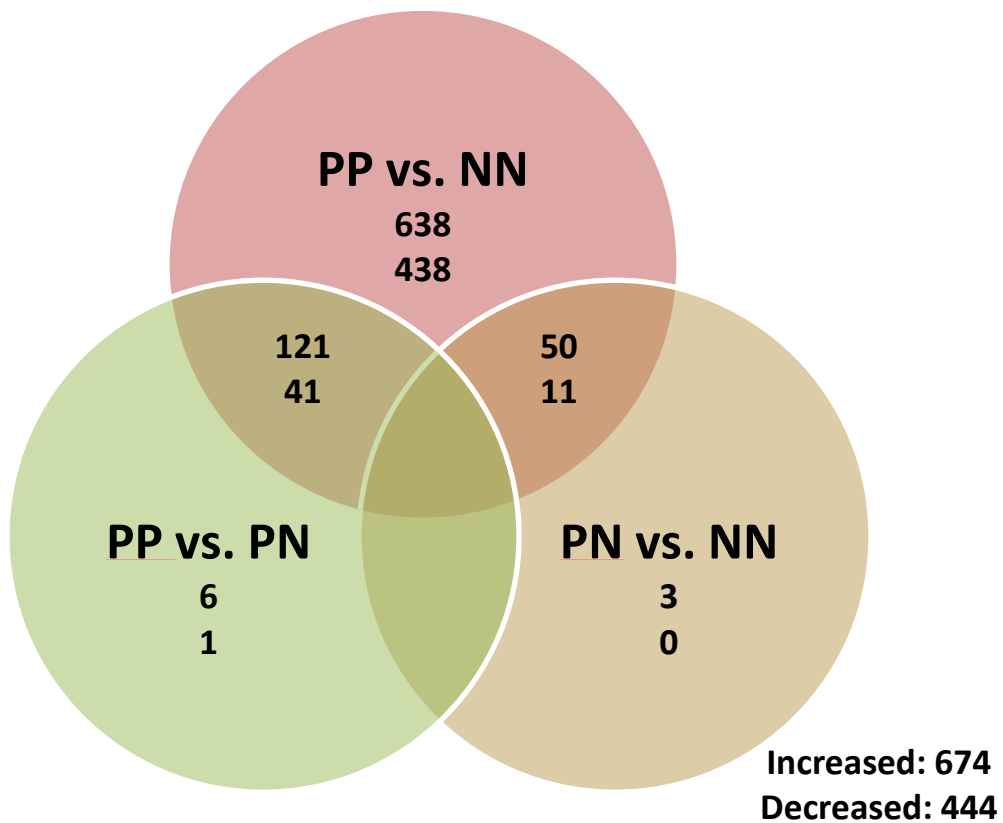


Figure 4.2 Differentially methylated CpGs that map to the epidermal differentiation complex (EDC). Genes of the EDC are critical to epidermal development. Twelve differentially methylated CpG sites in PP compared to NN map to this region of chromosome 1. The image was adapted from a postscript generated using the UCSC genome browser [258].

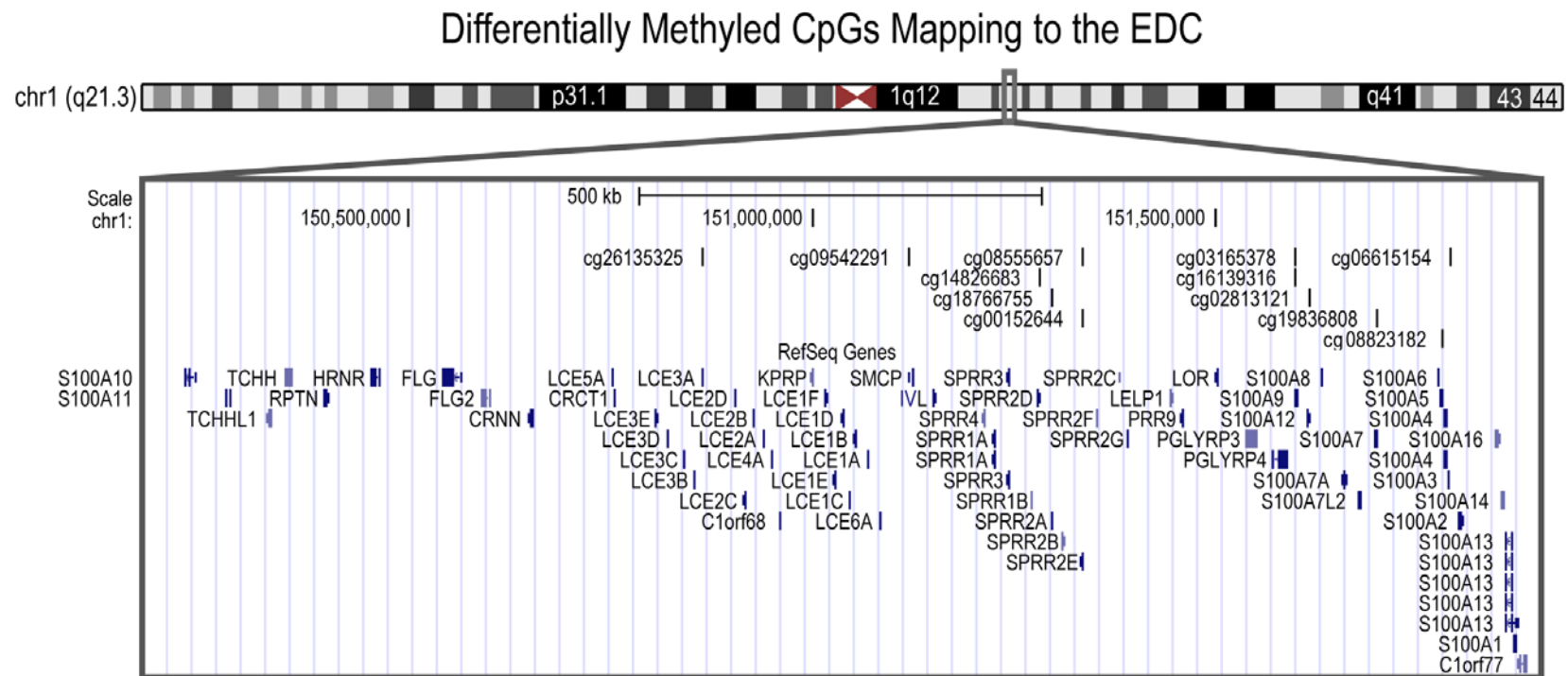


Figure 4.3 Pyrosequencing data in PP, PN and NN skin biopsies at 9 CpG sites in the gene IFI27. Methylation levels (%) with 95% confidence intervals are plotted for each CpG site by group. P-values calculated with a two-sample t-test. Methylation levels in PP samples were less than that of NN or PN samples. PN and NN skin had similar methylation levels. *, p-value < 0.05; **, p-value < 0.001.

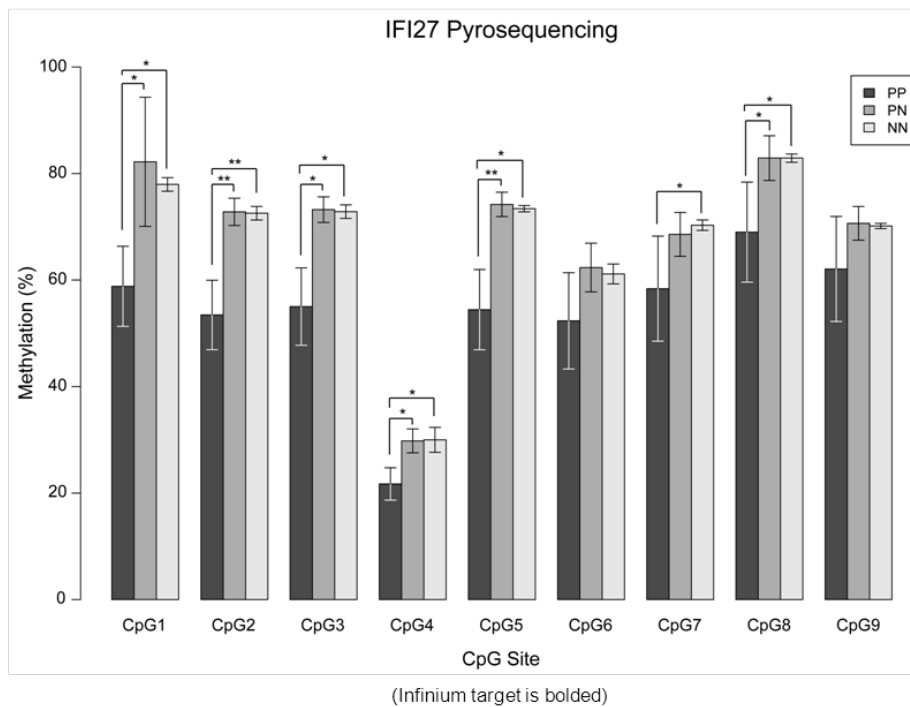
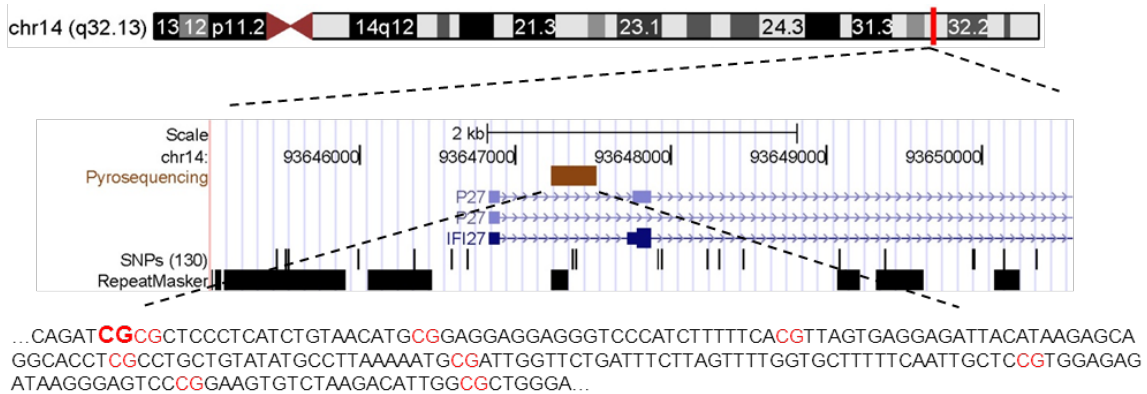


Figure 4.4 Pyrosequencing data for C10orf99 in PP, PN and NN skin biopsies at 3 CpG sites. Methylation levels (%) with 95% confidence intervals are plotted for each CpG site by group. P-values were calculated with a two-sample t-test. PN and NN biopsies were not separable, but both had increased methylation compared to PP. *, p-value < 0.05; **, p-value < 0.001.

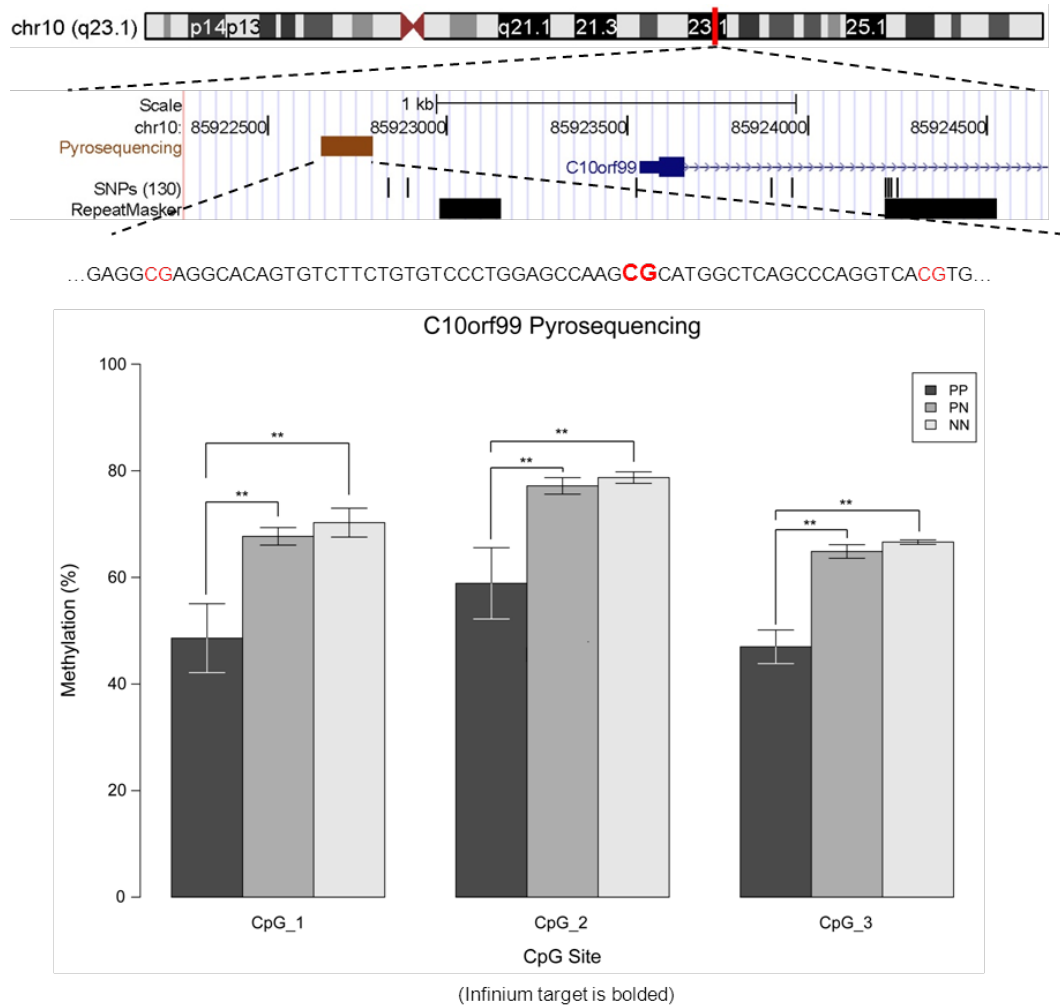
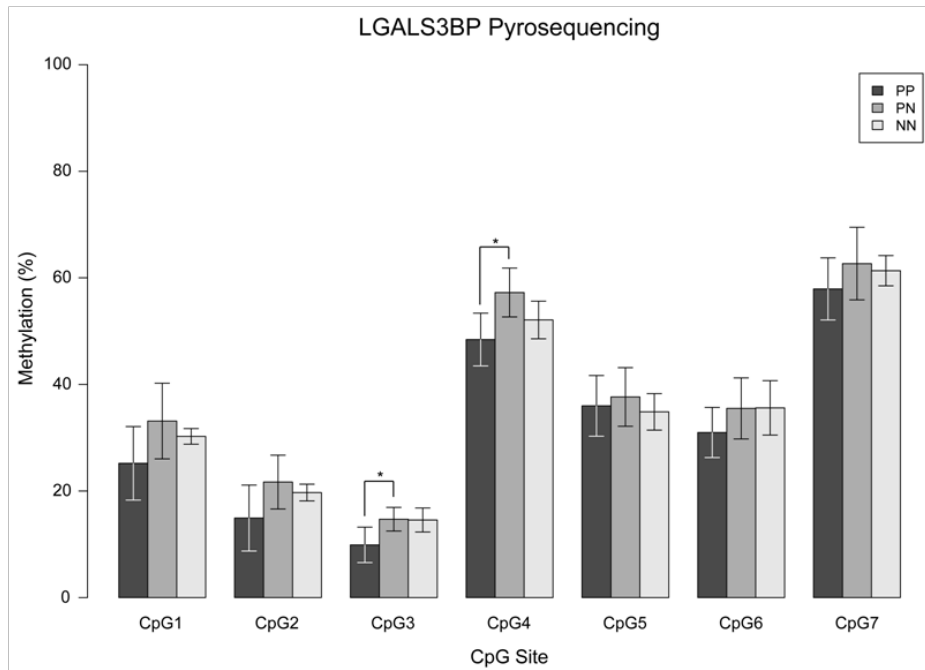
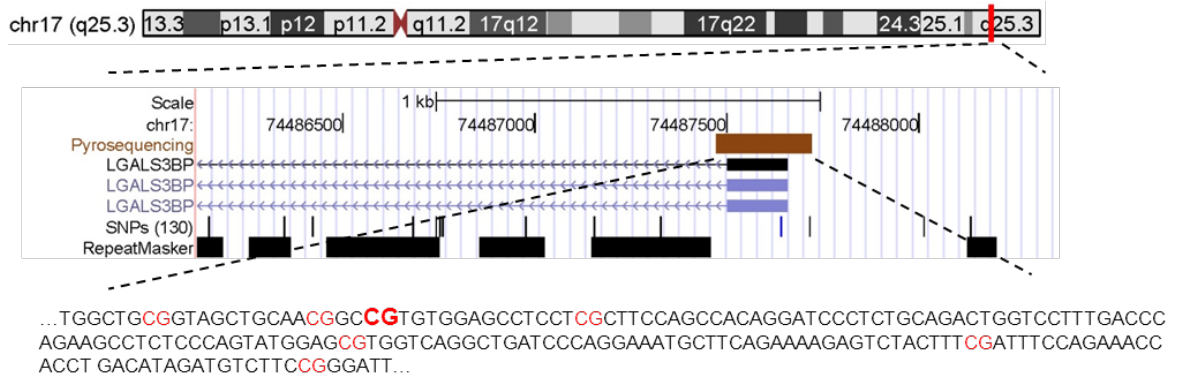


Figure 4.5 Pyrosequencing data for LGALS3BP in PP, PN and NN skin biopsies at 7 CpG sites. Methylation levels (%) with 95% confidence intervals are plotted for each CpG site by group. There was no obvious trend for lower methylation in these involved samples with only two CpG sites demonstrating sufficient evidence of differential methylation in PP vs. PN. *, p-value < 0.05; **, p-value < 0.001.



(Infinium target is bolded)

Figure 4.6 Pyrosequencing data for SERPINB4 in PP, PN and NN skin biopsies. Methylation levels (%) with 95% confidence intervals are plotted for each CpG site by group. There was no obvious trend for lower methylation in these involved samples with only two CpG sites demonstrating sufficient evidence of differential methylation in PP vs. PN. *, p-value < 0.05; **, p-value < 0.001.

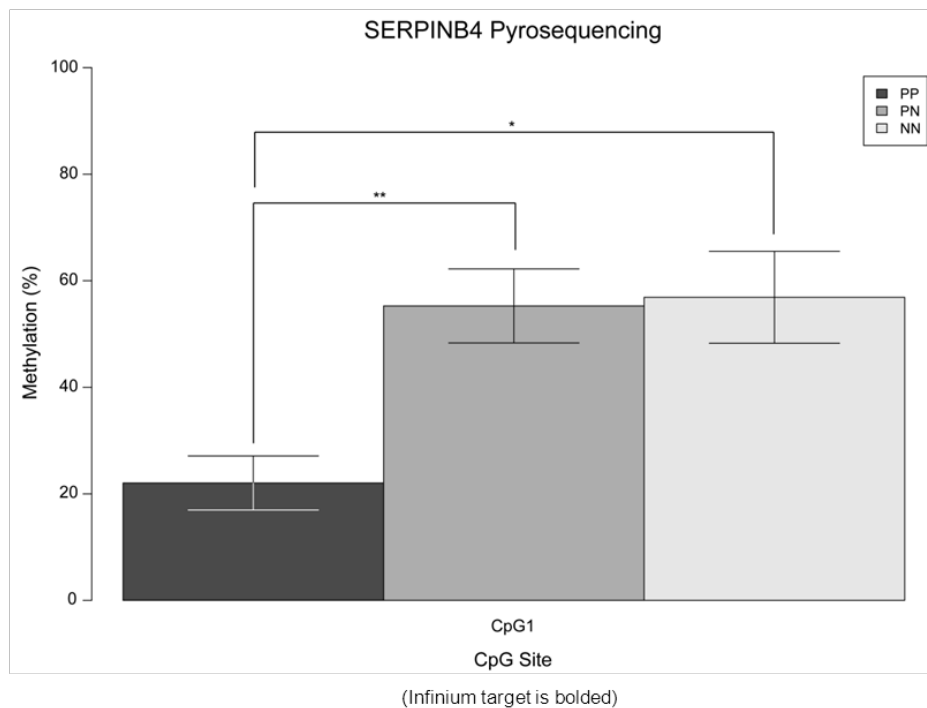
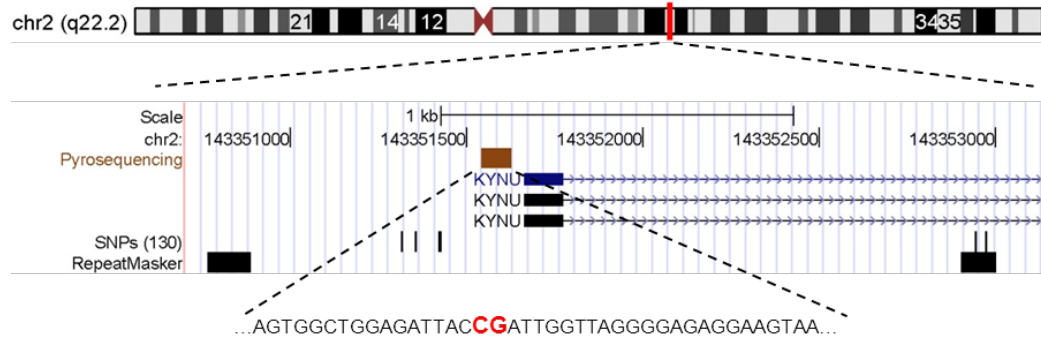
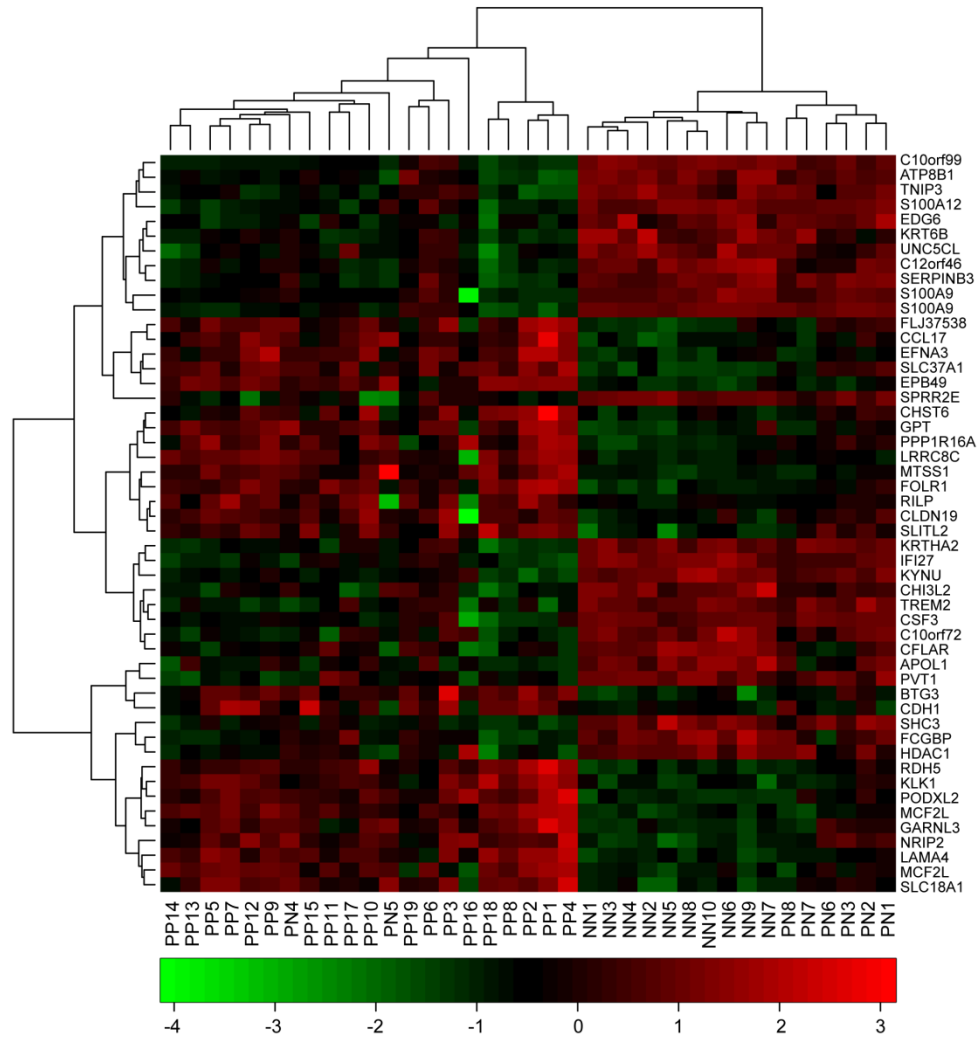


Figure 4.7 Heatmap showing PP, PN, and NN samples clustered using the top 50 CpG sites that differentiate PP from NN skin. A. Image was generated using normalized M-values. Using only those sites resulted in clustering that performs well at separating PP from PN and NN, and further separates NN and PN. Red values indicate relatively increased methylation while green indicates relatively decreased expression. B. Confusion matrix of sample classification of hierarchical clustering.

A.



B.

		Predicted			
		PP	PN	NN	
Actual	PP	19	0	0	19
	PN	2	6	0	8
	NN	0	0	10	10
		21	6	10	

Figure 4.8 The top scored network identified by using the top 50 differential methylation genes. Thick cyan lines indicate the fragments of canonical pathways; red circles indicate genes with decreased methylation and blue circles indicate genes with increased methylation.

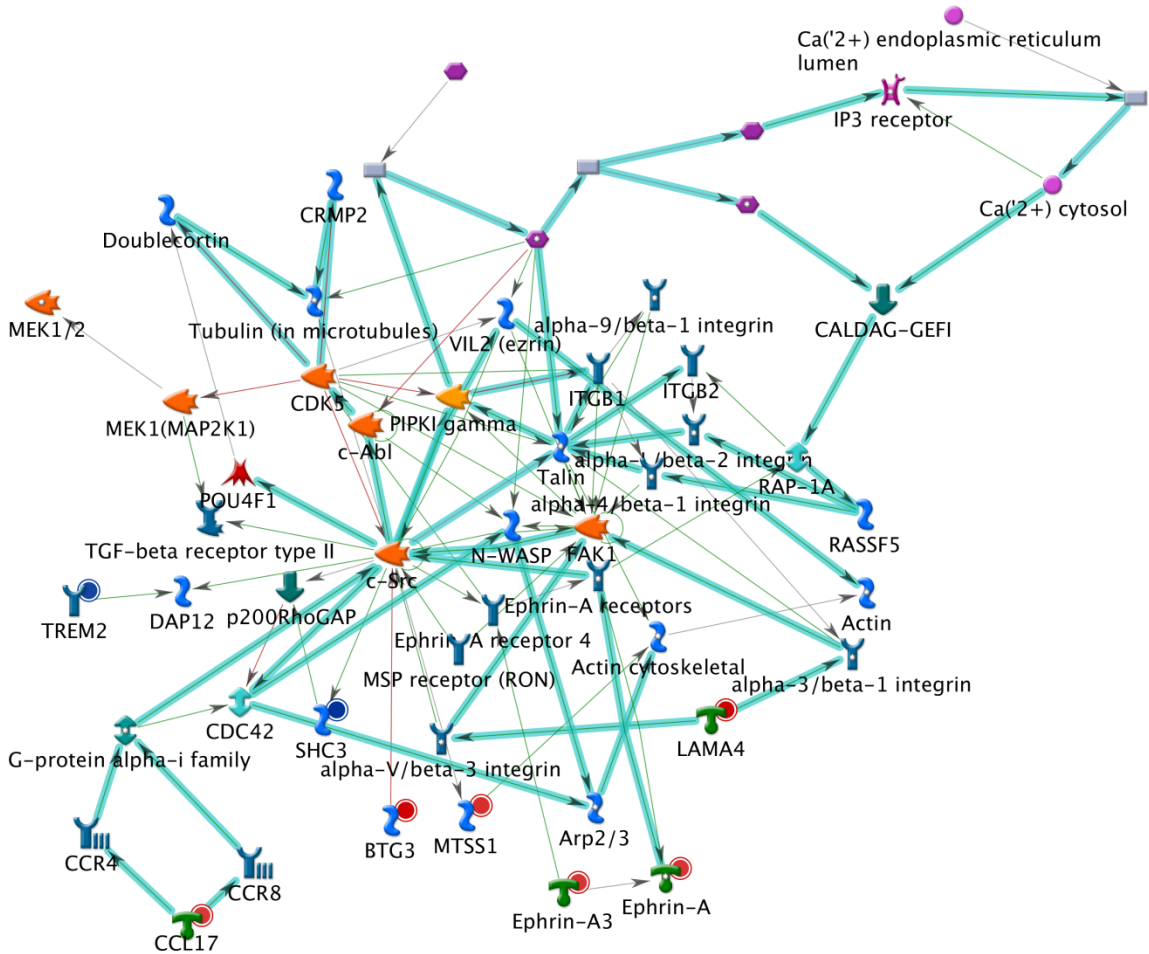


Figure 4.9 Boxplots of methylation levels in three sample groups. The top 50 sites that most differentiate involved from NN skin were determined by principal components analysis (see Methods). The upper panel shows the methylation levels for the top 25 CpG sites that show increased methylation, and the lower panel shows the top 25 CpG sites with decreased methylation. Displayed p-values were derived from the non-parametric Kruskal-Wallis test for equality of medians between groups. Dark lines represent the median of each group. The bottom and top borders of each box are defined by the first and third quartiles. Whiskers reach out to data points up to 1.5 times the interquartile range above or below the appropriate quartile. Data points outside of that range are considered outliers and are represented by circles.

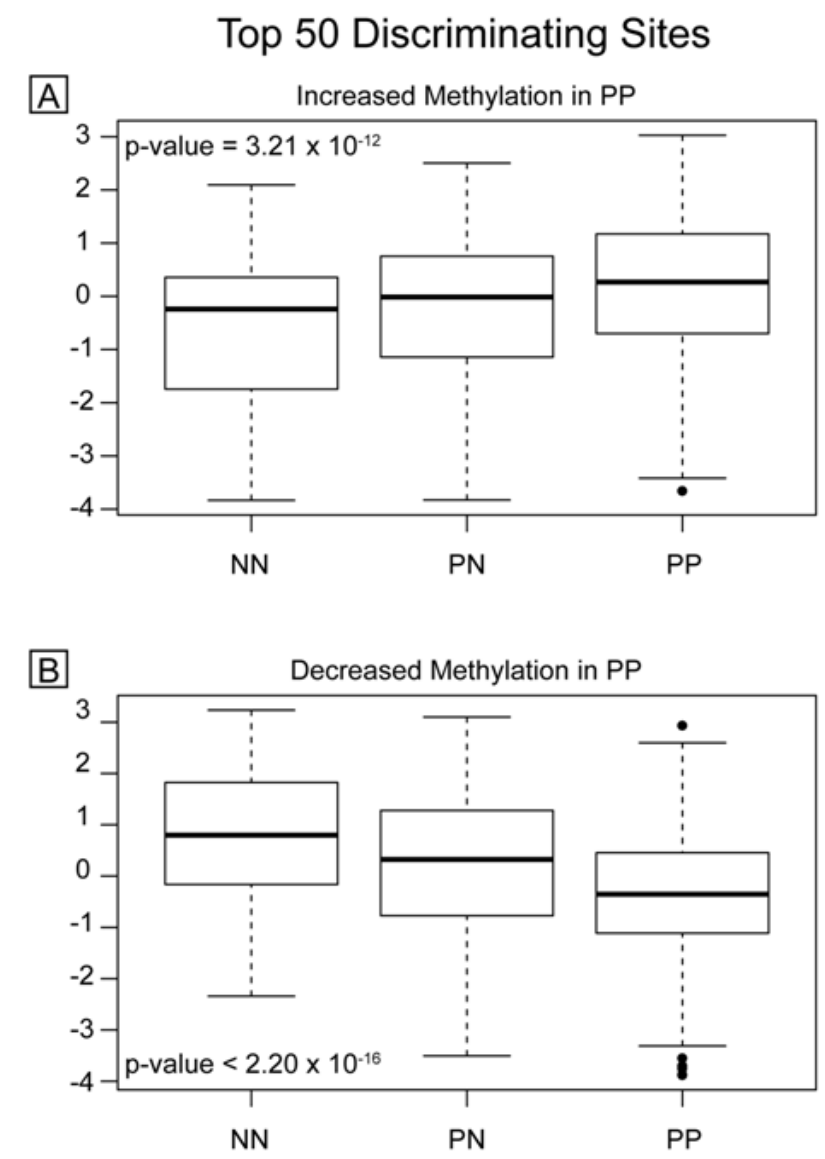
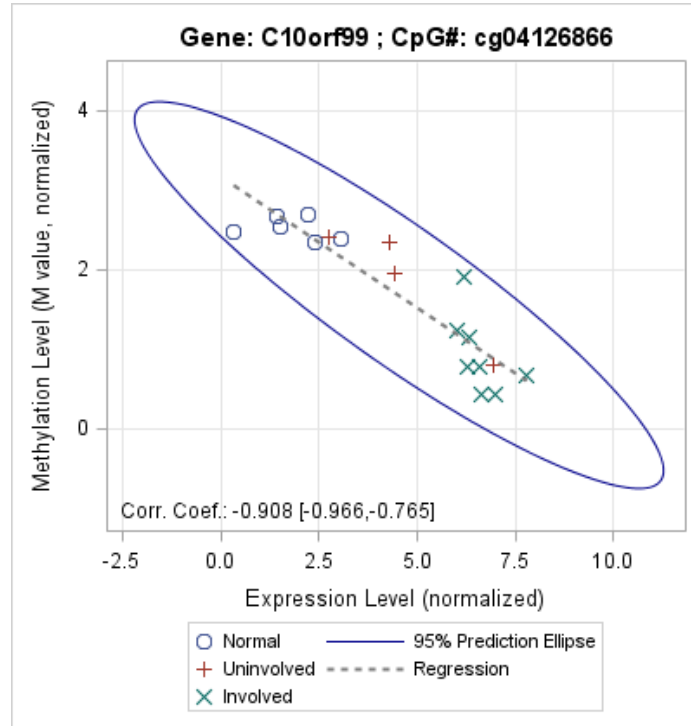
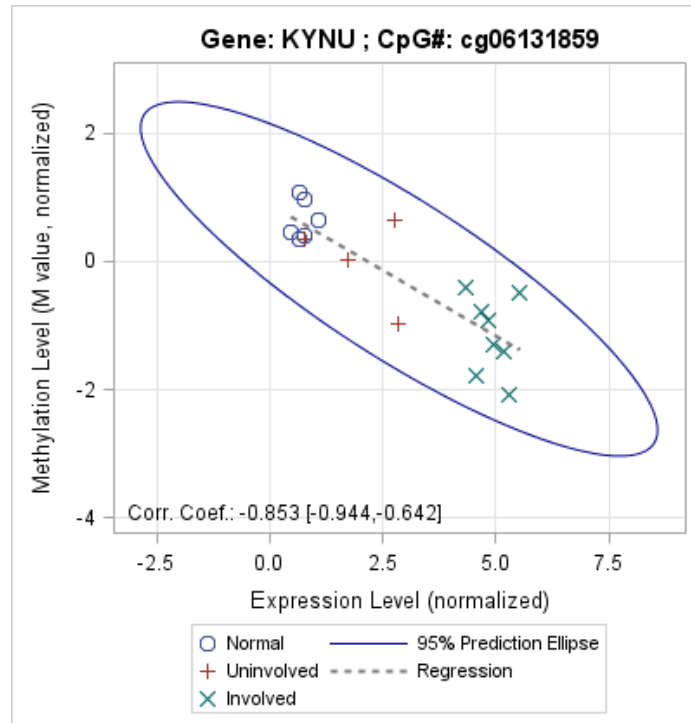


Figure 4.10 Correlation between methylation level of CpG sites and expression of C10orf99 (A) and KYNU (B) genes.

A.



B.



CHAPTER 5
SUMMARY AND FUTURE DIRECTIONS

5.1 Summary

Psoriasis is a complex disease of the skin. A complete understanding of how it develops will require studies of both genetic and environmental factors. In this thesis I attempted to explore both genetic underpinnings of this disease through genome wide association studies to identify genetic risk factors, and the possible effects of the environment through epigenetic modifications. What is more challenging is how to integrate findings from the two studies.

The GWAS study in Chapter 2 describes the first large scale genome-wide association scan for novel PS and PsA susceptibility genes. The study was conducted in two stages: the discovery stage was consisted of typing 233 cases (91 have PsA) and 519 controls with a 300k SNP platform. For the replication stage, a subset of candidate SNPs from 120 regions were further validated on an independent cohort of 577 PS cases and 737 controls, where 94 of these cases had also been diagnosed with PsA. The results of this study revealed that the MHC is truly the most important risk factor for PS and that it plays a very major role in PsA. In addition, it also confirmed recently identified associations with interleukin-23 receptor and interleukin 12B in both PS and PsA, and identified new associations. These include a region on chromosome 4q27 that contains genes for interleukin 2 and interleukin and seven additional regions that include chromosome 13q13 and 15q21.

In the last 3 years, with the advent of large-scale GWAS studies of psoriasis and psoriatic arthritis, several psoriasis risk factors that are common alleles in the general genetic variants, including SNPs [102,204,259,260,261,262] and CNVs (copy number variation) [263], have been identified and validated. The findings of these GWAS and

others are detailed below and summarized in **Table 5.1**. In general, these results provided important insights into the genetics of psoriasis and led to the discovery of new molecular pathways contributing to psoriasis and PsA.

First of all, most of these GWAS studies detected strong association in the vicinity of HLA-C, and additional signals in genes whose products are components of the IL-23 ligand-receptor complex (**Table 5.1**). Variants in or near genes that encode subunits of cytokines (IL12B, IL23A) or cytokine receptors (IL23R) are interesting given that the gene product of IL12B, p40, is the target of a recently approved monoclonal antibody therapy for psoriasis (ustekinumab) [264,265]. Expression of IL12B and IL23A shows highly significant differences in expression between involved and uninvolved skin ($p < 10^{-9}$) and IL23A is more highly expressed in the uninvolved skin of people with psoriasis than in the normal skin of controls ($p < 0.0003$) [204]. IL-23 signaling promotes cellular immune responses by promoting the survival and expansion of a recently identified subset of T cells expressing IL-17 that protects epithelia against microbial pathogens [266]. It is reasonable to speculate that aberrant IL-23 signaling renders certain individuals susceptible to inappropriate immune responses targeting epithelial cells, thus contributing to the chronic and relatively skin-specific inflammation seen in psoriasis (see Figure 3 in the reference [12]). Notably, inflammatory bowel disease has numerous clinical, immunologic, and genetic parallels to psoriasis. Psoriasis is more prevalent in patients with inflammatory bowel disease than in the general population and is thought to share common immunologic/inflammatory pathways, particular that of the Th17 pathway, in its pathogenesis. The same IL23R variant found in psoriasis has also been repeatedly associated with Crohn's disease [129]. In addition, the nuclear factor kappa B (NF- κ B)

pathway plays a central role in numerous processes in psoriasis [267]. Proteins modulating this pathway such as TNF- α -induced protein 3 (TNFAIP3) or A20 and TNFAIP3-interacting protein 1 (TNIP1) are also encoded by genes associated with psoriasis and are likely to enhance NF- κ B activation [204]. A20, a cytoplasmic zinc finger protein encoded by TNFAIP3, can be rapidly induced by NF- κ B activation, triggering A20's de-ubiquitinating and ubiquitinating activities of key signaling molecules, ultimately preventing the activation of NF- κ B. In GWAS, an intronic polymorphism of TNFAIP3 was found to associate with psoriasis and PsA [204]. Polymorphisms in TNIP1, another inhibitor of TNF-induced NF- κ B activation, were also identified in GWAS as being associated with both psoriasis and PsA [268]. A UK psoriasis population GWAS identified polymorphisms in yet another ubiquitin ligase, ZNF313, on chromosome 20q13 [269]. ZNF313, or zinc finger 313, is believed to regulate T-cell activation and expressed in skin, CD4⁺ T lymphocytes and dendritic cells.

The GWAS studies also provided the evidence that compromised skin barrier function plays a role in psoriasis susceptibility. The epidermal differentiation complex (EDC), located on chromosome 1q21 within PSORS4, has long been an area of investigation for both psoriasis and atopic dermatitis. Many genes of the EDC are upregulated in psoriatic lesions suggesting underlying alterations in coordinate regulation of genes of this complex [228]. A deletion polymorphism of ~30 kb that removes LCE3B and LCE3C genes within the late cornified envelope (LCE) complex of this region have been found to be associated with psoriasis in both Europeans and in Chinese [102,103]. It has been suggested that the absence of intact LCE3C and LCE3B genes could lead to an inappropriate repair response following barrier disruption [103]. Further investigation of

LCE expression in skin revealed an increased expression of the LCE3 following tape stripping, suggesting a link between this CNV and the abnormal differentiation and epidermal hyperproliferation seen in psoriasis.

Another set of genes with variations associated with psoriasis are the β -defensin cluster on chromosome 8p21. This region was recently found to harbor psoriasis-associated CNV (increased copy numbers) of the genes that encode DEFB4, DEFB103, and DEFB104 [270]. These genes have been shown to stimulate keratinocytes to release IL-8, IL-18, and IL-20, all of which are pro-inflammatory cytokines with a role in psoriasis [271].

Taken together, identifying the genetic basis of psoriasis is a challenging task and further complicated by the fact that psoriasis clearly involves an interaction between the immune system and the skin, leading to the question about where the primary defect resides. Recent GWAS has started to reveal genetic susceptibility factors of both immune and epidermis systems. This leads to the identification of a number of signaling pathways that are crucial to the formation of particular immune cells and skin barrier component in psoriasis lesions, and ultimately to the development of new forms of treatment.

The next stage, which would be a more challenging task, is to pinpoint the causative variants highlighted by GWAS. The third chapter of the thesis describes a follow-up study aimed to investigate the potential regulatory roles of certain genetic variations in PSORS1 region. The results implied a potential allele-specific regulatory role of one polymorphism (rs10456057) via binding to nuclear transcription factors. Further study with more SNPs in that region identified an “enhancer” activity in the risk

allele of another intergenic SNP rs13191343, which is highly associated with psoriasis and psoriatic arthritis. Interestingly, this activity was only observed under differentiating conditions of the transfected cells.

It is highly desirable to find out what transcription factor(s) bind to these rSNPs and what gene(s) is affected by this polymorphism. Notably, CDSN encodes the only PSORS1 transcript to be specifically expressed in terminally differentiated keratinocytes. And its product, corneodesmosin, has been reported to be up-regulated in the skin lesions of psoriatic patients. Genotypes (numbers of risk allele) of PSORS1 rSNPs are associated with CDSN gene expression (**Figure 3.5**). In addition, a mouse strain with a targeted epidermal deletion of CDSN has also been created [272]; this lead to detachment of the upper layers of the skin. When grafted onto immune-deficient mice, CDSN-deficient skin undergoes rapid hair loss together with epidermal abnormalities resembling psoriasis [272]. Based on the above evidence, a reasonable hypothesis is formulated: the PSORS1 rSNPs may alter the expression level of CDSN gene at skin lesions of psoriasis patients, which in turn disrupts the skin barrier and trigger keratinocyte hyper-proliferation.

It have been widely accepted that epigenetics offers an important window to understanding the role of the environment's interactions with the genome in causing disease. Despite increasing evidence for and interest in the role of epigenetics in human disease, virtually no epigenetic information is systematically measured at the genome level for psoriatic skin. The fourth chapter of the thesis described the first global study of altered CpG methylation in psoriatic (involved PP and uninvolved PN) and normal (NN) skin. We determined the methylation levels at 27,578 CpG sites in these samples. Involved skin differed from normal skin at 1,108 CpG sites, among which twelve mapped

to the epidermal differentiation complex close to genes that are highly up-regulated in psoriasis (S100A3/5/7/9/12, SPRR2A/D/E, LCE3A). Hierarchical clustering of 50 of the top differentially methylated sites 100% accurately separated all psoriatic skin samples (involved and uninvolved) from normal skin. Methylation at 12 CpG sites was significantly correlated with expression levels of a nearby gene. This included decreased methylation of sites at KYNU, OAS2, S100A12, and SERPINB3, whose strong transcriptional upregulation in psoriatic skin has previously been shown to be an important discriminator of psoriasis versus other inflammatory skin diseases such as atopic dermatitis. We also observed intrinsic epigenetic differences in uninvolved versus normal skin that may be reflective of a predisposition to psoriasis.

In conclusion, this thesis reports the first large-scale genome wide association study of psoriasis and of altered CpG methylation in psoriatic skin and has led to the identification of both genetic and epigenetic risk factors for psoriasis. The MHC (PSORS1) conferred the strongest effect upon disease risk, while most of the non-MHC risk factors identified from GWAS have modest effects. These included genes of the immune system and of the skin barrier. Follow-up functional studies identified a risk allele in the PSORS1 region with enhancer activity in differentiating keratinocytes. It was correlated with CDSN expression, a candidate gene that is ~150kb apart from the variant location. Upregulation of CDSN in the presence of PSORS1 would affect skin barrier formation and could predispose to psoriasis following environmental triggers such as infection. The global methylation study identified 1,108 differential methylated CpG sites and 12 are mapped to the Epidermal Differentiation Complex region on chromosome 1q21 which encodes genes involved in barrier formation. The top 50 differentially

methylation markers accurately clustered all psoriatic skin samples from normal skin and methylation at 12 CpGs was negatively correlated with nearby gene expression. Therefore, based on these results, we conclude that genetic and epigenetic risk factors of psoriasis lead to alterations in genes of the barrier and immune system. These act together to trigger the disease.

5.2 Future Perspectives on the Genetics of Psoriasis

In terms of psoriasis genetics, and more broadly chronic inflammatory disease genetics, there is much challenging work that lies ahead. First of all, as with most other complex diseases, studies of disease heritability strongly suggest a substantial genetic component of PS and PsA has not yet been identified. Although recent GWAS have provided insight into the genetic susceptibility component derived from common alleles having moderate effects, association signals can only account for a sibling recurrent risk of 1.35 (including 1.25 due to HLA) [204], where the overall recurrent risk has been estimated at between 3 to 6 fold [11]. So, what accounts for the missing heritability of psoriasis?

The recent work discovering and unraveling the rich structural variation/rearrangements in the human genome may form the basis of the next wave of genome-wide studies. In fact, a recent study reported that genomic copy number alterations at beta-defensin genes were susceptible to psoriasis [270]. Two other independent studies showed that deletion of the LCE3B and LCE3C genes, members of the late cornified envelope (LCE) gene cluster, was associated with increased risk of both psoriasis and rheumatoid arthritis [103,273]. Secondly, meta-analysis is a well-established method to synthesize results and draw conclusions from different studies for a set of related research hypotheses and can have greater impact than to other study designs [274]. When performed appropriately, meta-analysis may enhance the precision of the estimates of the effects of risk alleles, leading to reduced probability of false negative results. The increased availability of information can also lead to rejection of null hypotheses at lower levels of type I error, thus reducing the false discovery rate [275]. A successful example of such approach was reported in a recent study where the authors performed a meta-

analysis to investigate the contribution of the deletion of genes LCE3C and LCE3B to psoriasis susceptibility in multiple populations [276]. In this study, the authors confirmed the deletion of LCE3C and LCE3B as a common genetic factor for susceptibility to psoriasis in European, Chinese and Mongolian populations (OR(overall)=1.21, 1.27 and 1.34, respectively). The interaction analysis with HLA-Cw6 locus also highlighted significant differences in the epistatic interaction with the LCE3C and LCE3B deletion in some European populations, suggesting possible epistatic effects between these two major genetic contributors to psoriasis.

Finally, although emphasis has been placed on mapping common variants, recent studies have demonstrated that rare variants also play an important role in complex trait etiology and their identification should have a greater impact on risk assessment, disease prevention, and treatment due to their large genetic effects [277]. The development of second-generation sequencing (NexGen) technologies has made it possible to identify the rare causative variants/mutations responsible for disease [278,279]. These efforts will lead to identification of new psoriasis susceptibility genes and associated signal pathways, which eventually pave the way for new therapeutic targets of the disease.

5.3 Future Perspectives on Epigenetic Study of Psoriasis

Given that epigenetics is at the heart of phenotypic variation in health and disease, it seems likely that understanding and manipulating the epigenome holds enormous promise for preventing and treating common human illness. Epigenetics also offers an important window to understanding the role of the environment's interactions with the genome in causing disease, and in modulating those interactions to improve human health. Despite increasing evidence for and interest in the role of epigenetics in human disease, virtually no epigenetic information is systematically measured at the genome level. The current population-based approach to common disease relates common DNA sequence variants to disease status. Although this purely sequence-based approach is powerful, there is currently no conceptual framework to integrate epigenetic information. Bjornsson and colleagues recently proposed an hypothetical framework which integrated both the epigenetic and genetics factors in the etiology of common disease [280]. The common disease genetic and epigenetic (CDGE) hypothesis argues that in addition to genetic variation, epigenetics provides an added layer of variation that might mediate the relationship between genotype and internal and external environmental factors [280]. First of all, the epigenetic component in the model could potentially help to explain the marked increase in common diseases with age. A common characteristic of ageing is a time-dependent decline in responsiveness or adaptation to the environment, a form of loss of phenotypic plasticity. This loss of phenotypic plasticity could be mediated epigenetically if loss of the normal balance between gene-promoting and gene-silencing factors occurred across the genome. Secondly, epigenetic variation might also help to explain the quantitative nature of common disease phenotypes, which could show a

Gaussian distribution based on the quantitative nature of epigenetic variation (level of methylation at CpG loci). Finally, this hypothesis could also partially explain how the environmental risk factors interact with genetic predisposition to the disease. Thus, common diseases may involve phenotypic variants with both genetic variation and environmentally triggered epigenetic change that modulates the effects of DNA sequence variation. These epigenetic modifiers are, in turn, affected by variation in the genes that encode them, and environmental factors (hormones, growth factors, toxins and dietary methyl donors) influence both the genome and epigenome. This idea can be tested by incorporating an assessment of the epigenome into population epidemiological studies, rather than simply stratifying risk for environmental exposures as is done currently [281,282].

Finally, I would like to discuss the implications and prospects for epigenetic therapy. As epigenetic mechanisms for human disease are identified, epigenetic therapies become possible. One of the clinically relevant aspects of epigenetic alterations in diseases is the possibility of reversion by using various enzymatic inhibitors. Although this field is in its infancy, it carries great promise. Potential targets for drug development are histone modification and DNA methylating and demethylating enzymes [63]. In fact, some drugs that have an effect on the epigenome are already in widespread use, but their epigenetic effect has only recently been discovered. For example, valproic acid is used to treat various disorders, including seizures, bipolar disorder and cancer, and valproic acid was recently found to be a potent histone deacetylase inhibitor [283]. Some drugs have been tested in treatment of diseases specifically because of their known effects on the epigenome. Experimental evidence unequivocally shows that treatment with class I and

class II histone deacetylase (HDAC) inhibitors extends life span of several organisms [284]. Kypreou et al. [285] reported that treatment of peripheral blood phytohemagglutinin-stimulated lymphocytes from donors of different age groups (young, mid-aged, senior, and elderly) with trichostatin A (TSA), an HDAC inhibitor, resulted in hyperacetylation of histone H4 with increasing donor age. A recent study also indicated that in TSA-treated mice, the in vivo production and suppressive function of Foxp3+ CD4+ CD25+ regulatory T cells (Tregs) were increased [286]. Such knowledge has direct implications for the development of in vivo approaches to treat autoimmune and other inflammatory diseases.

Although the most advanced set of drugs in clinical development are histone deacetylase (HDAC) inhibitors, the prevalence of tissue-specific dysregulation of DNA methylation in various human diseases, especially in cancer and autoimmune diseases, suggests that it should shift the focus from HDAC inhibitors to DNA methylation/demethylation inhibitors [63]. In fact, most attention has been directed at developing DNA methylation inhibitors. The driving force behind this effort was the idea that tumor suppressor genes are silenced in cancer by DNA methylation of their promoters. Therefore, blocking DNMT during DNA synthesis would result in passive demethylation in dividing cancer cells and reactivation of these genes triggering suppression of tumor growth [287]. For example, two classes of epigenome-modifying agent are currently in clinical trials for cancer. DNA methyltransferase inhibitors such as decitabine, was used for the treatment of myelodysplasia [62]. The overall response rate with decitabine in a phase III study showed a small but statistically significant difference

for myelodysplasia and half of the clinically responsive patients showed a cytogenetic response [288].

The field of epigenetic pharmacology in autoimmune disease is still underdeveloped. Recent studies showed that DNA demethylation is critical for launching the changes in gene expression programming during differentiation of T cells into Th1 and Th2 cells [289] and proposed that methylated binding protein 2 (MBD2) is possibly involved in demethylation in T cells in lupus [290]. The global methylation profiling study in this thesis implied persistent epigenetic alterations in psoriatic skin, despite the evidence of symptoms improvement after TNF- α blockade treatment for a month. These results pave a new way to identify potent epigenetic signatures of psoriasis, and therefore, may make it possible to gauge whether certain therapeutics could truly restore the abnormally regulated epigenomes to a more normal state through epigenetic reprogramming.

5.4 Integrated Genetic and Epigenetic Approach

In summary, psoriasis is a complex disease with polygenic susceptibility. Recent GWAS studies have provided significant insight towards gene discovery and, to date, more than 20 common variant loci have been associated with psoriasis and/or PsA (**Table 5.1**). Despite these significant advances, known genetic variants explain only a small portion of the heritable component of disease, and little is known about their contribution to etiology. Ongoing studies are attempting to explain this ‘missing heritability’ in complex diseases via the detection of rare and structural variants (NexGen sequencing for rare functional variations), interactions between discovery SNPs and causal variants, and identification of associated stable epigenetic modifications. The epigenomic approach described in Chapter 4 of the thesis provides insights into the epigenetic factors underlying the etiology of psoriasis. However, the relationship between specific epigenetic modifications and genomic features still is still poorly understood. It is reasonable to hypothesize that genotype-epigenotype interactions may underlie the etiopathogenesis of psoriasis, and thus using a multi-dimensional integrative approach may reveal powerful insights into the molecular mechanisms of such interaction. The challenge, of course, is to combine these possibilities into a framework that is useful for considering these factors simultaneously to test particular hypotheses. By combining SNP genotype with DNA methylation data, for instance, potential novel genotype-epigenotype interactions within disease-associated loci could be uncovered to understand the sources of methylation variability. The flow diagram shown in **Figure 5.1** illustrated an integrated genetic and epigenetic approach for this purpose. In this thesis, I have discussed examples of interplay between sequence variation and reporter gene expression

(SNP9 in Chapter 3) and inverse correlations between promoter methylation and expression levels of target genes (Chapter 4). A successful example of evaluating the cis-impact of genotype on methylation status was reported in a recent study on Type 2 Diabetes [291], in which the authors assayed CpG methylation in 60 females, stratified according to disease susceptibility haplotype. Absolute methylation levels were then quantified across LD blocks. This led to identifying a 7.7 kb region of haplotype-specific methylation that had previously been validated as a long-range enhancer [291]. In fact, Bell et al. did a whole-genome survey on the CpG methylation levels from 77 HapMap Yoruba individuals, for which genome-wide gene expression and genotype data were also available [64]. They identified 180 CpG-sites in 173 genes that were associated with nearby SNPs that were within a distance of 5 kb. Interestingly, there was also a significant overlap of SNPs that were associated with both methylation and gene expression levels. These results implied a strong genetic component to inter-individual variation in DNA methylation profiles, and there was an enrichment of SNPs that affect both methylation and gene expression, providing evidence for shared mechanisms in a fraction of genes. Thus, inspired by these findings, it is believed that an integrated approach will pave the way for the functional interpretation of mechanisms underlying association of genetic variants with psoriasis.

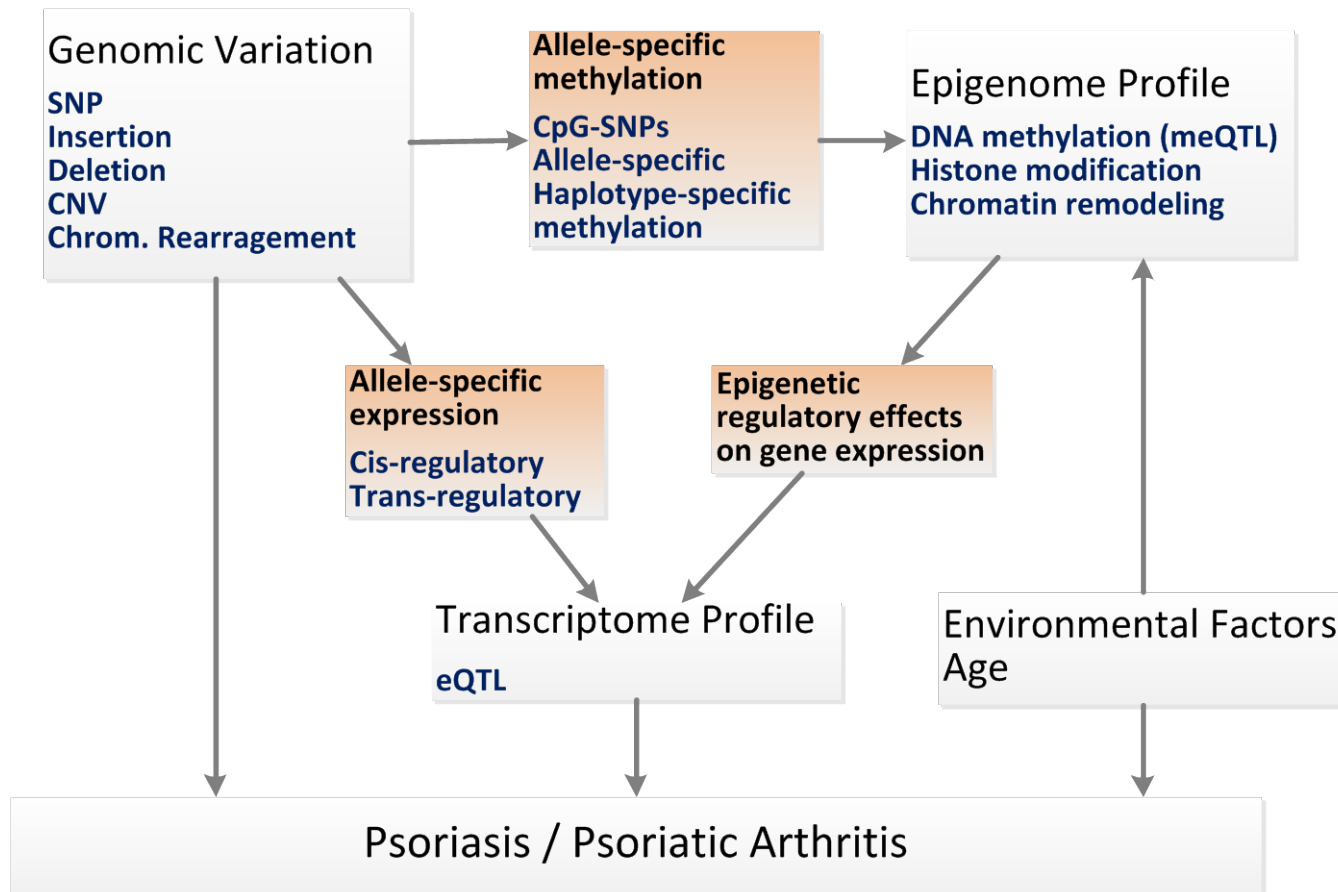
Table 5.1 Psoriasis-associated genes identified by GWAS studies.

Gene Candidate	Chrom Location	Function	Functional Category	Pleiotropy	References
<i>ADAM33</i>	20p13	A disintegrin and metalloprotease 33, membrane-anchored protein involving cell-cell and cell-matrix interactions	Inflammatory	Psoriasis Asthma COPD	[152,292,293] [294,295,296,297] [298]
<i>CDKAL1</i>	6p22	579-residue, 65-kD protein shares domain and amino acid homology with CDK5RAP1, an inhibitor of CDK5 (protein kinase), mRNA detected in human pancreatic islet and skeletal muscle		Psoriasis Diabetes mellitus type II Crohn's	[299,300] [299,301] [299,301]
<i>COG6</i>	13q14.11	Component of oligomeric golgi complex 6, required for normal Golgi function		Psoriasis Juvenile idiopathic arthritis	[181] [302]
<i>DEFB</i>	8p23.1	Human β -defensin, antimicrobial peptides	Inflammatory	Psoriasis Crohn's Atopic dermatitis Asthma	[270] [303] [304] [305]
<i>HLA-C</i>	6p21.3	Major histocompatibility complex, class I, C; involved in the presentation of foreign antigens to the immune system	Inflammatory	Psoriasis	[181,204]
<i>IL12B</i>	5q33.3	p40 subunit of IL-12 and IL-23	Inflammatory	Psoriasis Psoriatic arthritis Atopic dermatitis	[112,128,181,306,307] [181] [127]
<i>IL13</i>	5q31.1	A Th2 cytokine	Inflammatory	Psoriasis Psoriatic arthritis Asthma COPD	[204,308,309] [310] [311,312,313,314,315,316]]
<i>IL15</i>	4q31	Cytokine that affects T-cell activation and proliferation	Inflammatory	Psoriasis	[317] [318]
<i>IL2/IL21</i>	4q26-27	IL2 involved in regulation of T-cell clonal expansion; IL12, a cytokine promoting the transition between innate and adaptive immunity	Inflammatory	Psoriatic arthritis Rheumatoid arthritis Diabetes mellitus Ulcerative colitis Juvenile idiopathic arthritis	[181] [163] [163,319] [320,321] [322]
<i>IL23A</i>	12q13.2	p19 subunit of IL-23	Inflammatory	Psoriasis Psoriatic arthritis	[181,204] [181,204]

Table 5.1 (Continue).

Gene Candidate	Chrom Location	Function	Functional Category	Pleiotropy	References
<i>IL23R</i>	1p31.3	IL-23 receptor	Inflammatory	Psoriasis Crohn's Ulcerative colitis Ankylosing spondylitis	[112,128,181,204,306,307]] [129,323,324] [325] [326,327]
<i>KIR2DS1/ KIR2DL1/ KIR2DL5</i>	19q34	Killer immune-globulin-like receptors, ligand is HLA-C, regulate NK cell response	Inflammatory	Psoriasis Psoriatic arthritis	[328] [44]
<i>LCE</i>	1q21	Late cornified envelope. LCE3C and LCE3B highly expressed in psoriasis after tape stripping	Epidermal	Psoriasis	[102,103]
<i>PTPN22</i>	1p13.2	Intracellular protein tyrosine phosphatase; expressed primarily in lymphoid tissues; involved in T cell receptor signaling	Inflammatory	Psoriasis Rheumatoid arthritis SLE Addison's disease	[111,300,329] [330] [331] [332]
<i>SLC12A8</i>	3q21.2	Cation/chloride cotransporter that may play a role in the control of keratinocyte proliferation	Epidermal	Psoriasis	[36]
<i>SUMO4</i>	6q25.1	SMT3 suppressor of mif two 3 homolog 4; encodes small ubiquitin-related modifiers that are attached to proteins and control the target proteins' subcellular localization, stability, or activity. This protein specifically modifies IKBA, leading to negative regulation of NF-kappa-B-dependent transcription of the IL12B gene.	Inflammatory	Psoriasis Diabetes mellitus	[333] [334,335,336]
<i>TNFAIP3</i>	6q23.3	Ubiquitin-editing protein A20 that dampens TNF-induced NF-κB activation, inhibiting inflammation	Inflammatory	Psoriasis Psoriatic arthritis Crohn's Rheumatoid arthritis SLE Diabetes mellitus type I Celiac disease	[181,204] [181,204] [337] [338,339,340] [341,342,343] [344] [345]
<i>TNIP1</i>	5q33.1	ABIN-1, down-regulating TNF-induced NF-κB activation	Inflammatory	Psoriasis Psoriatic arthritis SLE	[181,204] [181,204] [346,347]
<i>ZNF313</i>	20q13	Ubiquitin ligase, expressed in human skin	Inflammatory	Psoriasis	[269]

Figure 5.1 An integrated genetic and epigenetic approach to psoriasis and PsA. A schematic summary of how genetic and epigenetic factors as well as the interactions might contribute to psoriasis and PsA. The sources of epigenetic variation (genetic, environmental and age) are also represented. SNP: single nucleotide polymorphism; CNV: copy number variation; eQTL: expression quantitative trait loci; meQTL: methylation quantitative trait loci.



References

1. Bhalerao J, Bowcock AM (1998) The genetics of psoriasis: a complex disorder of the skin and immune system. *Hum Mol Genet* 7: 1537-1545.
2. Boyman O, Hefti HP, Conrad C, Nickoloff BJ, Suter M, et al. (2004) Spontaneous development of psoriasis in a new animal model shows an essential role for resident T cells and tumor necrosis factor-alpha. *J Exp Med* 199: 731-736.
3. Nograles KE, Brasington RD, Bowcock AM (2009) New insights into the pathogenesis and genetics of psoriatic arthritis. *Nat Clin Pract Rheumatol* 5: 83-91.
4. Arnett FC, Reveille JD, Duvic M (1991) Psoriasis and psoriatic arthritis associated with human immunodeficiency virus infection. *Rheum Dis Clin North Am* 17: 59-78.
5. Rongioletti F, Fiorucci C, Parodi A (2009) Psoriasis induced or aggravated by drugs. *J Rheumatol Suppl* 83: 59-61.
6. Bowcock AM, Krueger JG (2005) Getting under the skin: the immunogenetics of psoriasis. *Nat Rev Immunol* 5: 699-711.
7. Zhou X, Krueger JG, Kao MC, Lee E, Du F, et al. (2003) Novel mechanisms of T-cell and dendritic cell activation revealed by profiling of psoriasis on the 63,100-element oligonucleotide array. *Physiol Genomics* 13: 69-78.
8. Liu Y, Krueger JG, Bowcock AM (2007) Psoriasis: genetic associations and immune system changes. *Genes Immun* 8: 1-12.
9. Lin HW, Wang KH, Lin HC (2011) Increased risk of acute myocardial infarction in patients with psoriasis: A 5-year population-based study in Taiwan. *J Am Acad Dermatol* 64: 495-501.
10. Gelfand JM, Dommasch ED, Shin DB, Azfar RS, Kurd SK, et al. (2009) The risk of stroke in patients with psoriasis. *J Invest Dermatol* 129: 2411-2418.

11. Elder JT, Nair RP, Guo SW, Henseler T, Christophers E, et al. (1994) The genetics of psoriasis. *Arch Dermatol* 130: 216-224.
12. Roberson ED, Bowcock AM (2010) Psoriasis genetics: breaking the barrier. *Trends Genet* 26: 415-423.
13. Tiilikainen A, Lassus A, Karvonen J, Vartiainen P, Julin M (1980) Psoriasis and HLA-Cw6. *Br J Dermatol* 102: 179-184.
14. Jenisch S, Henseler T, Nair RP, Guo SW, Westphal E, et al. (1998) Linkage analysis of human leukocyte antigen (HLA) markers in familial psoriasis: strong disequilibrium effects provide evidence for a major determinant in the HLA-B/-C region. *Am J Hum Genet* 63: 191-199.
15. Nair RP, Henseler T, Jenisch S, Stuart P, Bichakjian CK, et al. (1997) Evidence for two psoriasis susceptibility loci (HLA and 17q) and two novel candidate regions (16q and 20p) by genome-wide scan. *Hum Mol Genet* 6: 1349-1356.
16. Trembath RC, Clough RL, Rosbotham JL, Jones AB, Camp RD, et al. (1997) Identification of a major susceptibility locus on chromosome 6p and evidence for further disease loci revealed by a two stage genome-wide search in psoriasis. *Hum Mol Genet* 6: 813-820.
17. Asumalahti K, Veal C, Laitinen T, Suomela S, Allen M, et al. (2002) Coding haplotype analysis supports HCR as the putative susceptibility gene for psoriasis at the MHC PSORS1 locus. *Hum Mol Genet* 11: 589-597.
18. Allen MH, Veal C, Faassen A, Powis SH, Vaughan RW, et al. (1999) A non-HLA gene within the MHC in psoriasis. *Lancet* 353: 1589-1590.
19. Guerrin M, Vincent C, Simon M, Tazi Ahnini R, Fort M, et al. (2001) Identification of six novel polymorphisms in the human corneodesmosin gene. *Tissue Antigens* 57: 32-38.

20. Tazi Ahnini R, Camp NJ, Cork MJ, Mee JB, Keohane SG, et al. (1999) Novel genetic association between the corneodesmosin (MHC S) gene and susceptibility to psoriasis. *Hum Mol Genet* 8: 1135-1140.
21. Simon M, Jonca N, Guerrin M, Haftek M, Bernard D, et al. (2001) Refined characterization of corneodesmosin proteolysis during terminal differentiation of human epidermis and its relationship to desquamation. *J Biol Chem* 276: 20292-20299.
22. Jonca N, Guerrin M, Hadjiolova K, Caubet C, Gallinaro H, et al. (2002) Corneodesmosin, a component of epidermal corneocyte desmosomes, displays homophilic adhesive properties. *J Biol Chem* 277: 5024-5029.
23. Gallinaro H, Jonca N, Langbein L, Vincent C, Simon M, et al. (2004) A 4.2 kb upstream region of the human corneodesmosin gene directs site-specific expression in hair follicles and hyperkeratotic epidermis of transgenic mice. *J Invest Dermatol* 122: 730-738.
24. Veal CD, Capon F, Allen MH, Heath EK, Evans JC, et al. (2002) Family-based analysis using a dense single-nucleotide polymorphism-based map defines genetic variation at PSORS1, the major psoriasis-susceptibility locus. *Am J Hum Genet* 71: 554-564.
25. Helms C, Saccone NL, Cao L, Daw JAW, Cao K, et al. (2005) Localization of PSORS1 to a haplotype block harboring HLA-C and distinct from corneodesmosin and HCR. *Human Genetics* 118: 466-476.
26. Bowcock AM, Cookson WO (2004) The genetics of psoriasis, psoriatic arthritis and atopic dermatitis. *Hum Mol Genet* 13 Spec No 1: R43-55.
27. Tomfohrde J, Silverman A, Barnes R, Fernandez-Vina MA, Young M, et al. (1994) Gene for familial psoriasis susceptibility mapped to the distal end of human chromosome 17q. *Science* 264: 1141-1145.

28. Tzschach A, Hoffmann K, Hoeltzenbein M, Bache I, Tommerup N, et al. (2006) Molecular characterization of a balanced chromosome translocation in psoriasis vulgaris. *Clin Genet* 69: 189-193.
29. Hwu WL, Yang CF, Fann CS, Chen CL, Tsai TF, et al. (2005) Mapping of psoriasis to 17q terminus. *J Med Genet* 42: 152-158.
30. Yang CF, Hwu WL, Yang LC, Chung WH, Chien YH, et al. (2008) A promoter sequence variant of ZNF750 is linked with familial psoriasis. *J Invest Dermatol* 128: 1662-1668.
31. Birnbaum RY, Zvulunov A, Hallel-Halevy D, Cagnano E, Finer G, et al. (2006) Seborrhea-like dermatitis with psoriasiform elements caused by a mutation in ZNF750, encoding a putative C2H2 zinc finger protein. *Nat Genet* 38: 749-751.
32. Helms C, Cao L, Krueger JG, Wijsman EM, Chamian F, et al. (2003) A putative RUNX1 binding site variant between SLC9A3R1 and NAT9 is associated with susceptibility to psoriasis. *Nat Genet* 35: 349-356.
33. Prokunina L, Castillejo-Lopez C, Oberg F, Gunnarsson I, Berg L, et al. (2002) A regulatory polymorphism in PDCD1 is associated with susceptibility to systemic lupus erythematosus in humans. *Nat Genet* 32: 666-669.
34. Tokuhira S, Yamada R, Chang X, Suzuki A, Kochi Y, et al. (2003) An intronic SNP in a RUNX1 binding site of SLC22A4, encoding an organic cation transporter, is associated with rheumatoid arthritis. *Nat Genet* 35: 341-348.
35. Hewett D, Samuelsson L, Polding J, Enlund F, Smart D, et al. (2002) Identification of a psoriasis susceptibility candidate gene by linkage disequilibrium mapping with a localized single nucleotide polymorphism map. *Genomics* 79: 305-314.
36. Huffmeier U, Lascorz J, Traupe H, Bohm B, Schurmeier-Horst F, et al. (2005) Systematic linkage disequilibrium analysis of SLC12A8 at PSORS5 confirms a role in susceptibility to psoriasis vulgaris. *J Invest Dermatol* 125: 906-912.

37. Hebert SC, Mount DB, Gamba G (2004) Molecular physiology of cation-coupled Cl⁻ cotransport: the SLC12 family. *Pflugers Arch* 447: 580-593.
38. Capon F, Novelli G, Semprini S, Clementi M, Nudo M, et al. (1999) Searching for psoriasis susceptibility genes in Italy: genome scan and evidence for a new locus on chromosome 1. *J Invest Dermatol* 112: 32-35.
39. Moll JM, Wright V (1973) Familial occurrence of psoriatic arthritis. *Ann Rheum Dis* 32: 181-201.
40. Myers A, Kay LJ, Lynch SA, Walker DJ (2005) Recurrence risk for psoriasis and psoriatic arthritis within sibships. *Rheumatology (Oxford)* 44: 773-776.
41. Gladman DD, Farewell VT (2003) HLA studies in psoriatic arthritis: current situation and future needs. *J Rheumatol* 30: 4-6.
42. Ravindran JS, Owen P, Lagan A, Lewis J, Korendowych E, et al. (2004) Interleukin 1alpha, interleukin 1beta and interleukin 1 receptor gene polymorphisms in psoriatic arthritis. *Rheumatology (Oxford)* 43: 22-26.
43. Rahman P, Sun S, Peddle L, Snelgrove T, Melay W, et al. (2006) Association between the interleukin-1 family gene cluster and psoriatic arthritis. *Arthritis Rheum* 54: 2321-2325.
44. Williams F, Meenagh A, Sleator C, Cook D, Fernandez-Vina M, et al. (2005) Activating killer cell immunoglobulin-like receptor gene KIR2DS1 is associated with psoriatic arthritis. *Hum Immunol* 66: 836-841.
45. Parham P (2005) MHC class I molecules and KIRs in human history, health and survival. *Nat Rev Immunol* 5: 201-214.
46. Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, et al. (2005) Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci U S A* 102: 10604-10609.

47. Humpherys D, Eggan K, Akutsu H, Hochedlinger K, Rideout WM, 3rd, et al. (2001) Epigenetic instability in ES cells and cloned mice. *Science* 293: 95-97.
48. Esteller M (2006) The necessity of a human epigenome project. *Carcinogenesis* 27: 1121-1125.
49. Reik W (2007) Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* 447: 425-432.
50. Bestor TH, Coxon A (1993) Cytosine methylation: the pros and cons of DNA methylation. *Curr Biol* 3: 384-386.
51. Bird A (2002) DNA methylation patterns and epigenetic memory. *Genes Dev* 16: 6-21.
52. Bird A, Taggart M, Frommer M, Miller OJ, Macleod D (1985) A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. *Cell* 40: 91-99.
53. Saxonov S, Berg P, Brutlag DL (2006) A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc Natl Acad Sci U S A* 103: 1412-1417.
54. Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, et al. (2007) Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet* 39: 457-466.
55. Herman JG, Baylin SB (2003) Gene silencing in cancer in association with promoter hypermethylation. *N Engl J Med* 349: 2042-2054.
56. Morgan HD, Santos F, Green K, Dean W, Reik W (2005) Epigenetic reprogramming in mammals. *Hum Mol Genet* 14 Spec No 1: R47-58.
57. Santos F, Dean W (2004) Epigenetic reprogramming during early development in mammals. *Reproduction* 127: 643-651.

58. Suter MA, Aagaard-Tillery KM (2009) Environmental influences on epigenetic profiles. *Semin Reprod Med* 27: 380-390.
59. Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, et al. (2006) DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat Genet* 38: 1378-1385.
60. Weber M, Davies JJ, Wittig D, Oakeley EJ, Haase M, et al. (2005) Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat Genet* 37: 853-862.
61. Jones PA, Martienssen R (2005) A blueprint for a Human Epigenome Project: the AACR Human Epigenome Workshop. *Cancer Res* 65: 11241-11246.
62. Mack GS (2006) Epigenetic cancer therapy makes headway. *J Natl Cancer Inst* 98: 1443-1444.
63. Szyf M (2010) Epigenetic therapeutics in autoimmune disease. *Clin Rev Allergy Immunol* 39: 62-77.
64. Bell JT, Pai AA, Pickrell JK, Gaffney DJ, Pique-Regi R, et al. (2011) DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. *Genome Biol* 12: R10.
65. Feinberg AP (2007) Phenotypic plasticity and the epigenetics of human disease. *Nature* 447: 433-440.
66. Taberlay PC, Jones PA (2011) DNA methylation and cancer. *Prog Drug Res* 67: 1-23.
67. Esteller M (2008) Epigenetics in cancer. *N Engl J Med* 358: 1148-1159.
68. Feinberg AP, Vogelstein B (1983) Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature* 301: 89-92.
69. Feinberg AP, Tycko B (2004) The history of cancer epigenetics. *Nat Rev Cancer* 4: 143-153.

70. Esteller M, Silva JM, Dominguez G, Bonilla F, Matias-Guiu X, et al. (2000) Promoter hypermethylation and BRCA1 inactivation in sporadic breast and ovarian tumors. *J Natl Cancer Inst* 92: 564-569.
71. Esteller M (2007) Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat Rev Genet* 8: 286-298.
72. Saito Y, Liang G, Egger G, Friedman JM, Chuang JC, et al. (2006) Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells. *Cancer Cell* 9: 435-443.
73. Lujambio A, Ropero S, Ballestar E, Fraga MF, Cerrato C, et al. (2007) Genetic unmasking of an epigenetically silenced microRNA in human cancer cells. *Cancer Res* 67: 1424-1429.
74. Krupanidhi S, Sedimbi SK, Sanjeevi CB (2008) Epigenetics and epigenetic mechanisms in disease with emphasis on autoimmune diseases. *J Assoc Physicians India* 56: 875-880.
75. Qaddus J, Johnson KJ, Gavalchin J, Amento EP, Chrisp CE, et al. (1993) Treating activated CD4+ T cells with either of two distinct DNA methyltransferase inhibitors, 5-azacytidine or procainamide, is sufficient to cause a lupus-like disease in syngeneic mice. *J Clin Invest* 92: 38-53.
76. Strickland FM, Richardson BC (2008) Epigenetics in human autoimmunity. Epigenetics in autoimmunity - DNA methylation in systemic lupus erythematosus and beyond. *Autoimmunity* 41: 278-286.
77. Ruchusatsawat K, Wongpiyabovorn J, Shuangshoti S, Hirankarn N, Mutirangura A (2006) SHP-1 promoter 2 methylation in normal epithelial tissues and demethylation in psoriasis. *J Mol Med* 84: 175-182.
78. Botstein D, Risch N (2003) Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nat Genet* 33 Suppl: 228-237.

79. Risch N, Merikangas K (1996) The future of genetic studies of complex human diseases. *Science* 273: 1516-1517.
80. (2005) A haplotype map of the human genome. *Nature* 437: 1299-1320.
81. Manolio TA (2010) Genomewide association studies and assessment of the risk of disease. *N Engl J Med* 363: 166-176.
82. Bowcock AM (2010) Genome-wide association studies and infectious disease. *Crit Rev Immunol* 30: 305-309.
83. Ameen M, Allen MH, Fisher SA, Lewis CM, Cuthbert A, et al. (2005) Corneodesmosin (CDSN) gene association with psoriasis vulgaris in Caucasian but not in Japanese populations. *Clin Exp Dermatol* 30: 414-418.
84. Chang YT, Shiao YM, Chin PJ, Liu YL, Chou FC, et al. (2004) Genetic polymorphisms of the HCR gene and a genomic segment in close proximity to HLA-C are associated with patients with psoriasis in Taiwan. *Br J Dermatol* 150: 1104-1111.
85. Foerster J, Nolte I, Junge J, Bruinenberg M, Schweiger S, et al. (2005) Haplotype sharing analysis identifies a retroviral dUTPase as candidate susceptibility gene for psoriasis. *J Invest Dermatol* 124: 99-102.
86. Nair RP, Stuart PE, Nistor I, Hiremagalore R, Chia NV, et al. (2006) Sequence and haplotype analysis supports HLA-C as the psoriasis susceptibility 1 gene. *Am J Hum Genet* 78: 827-851.
87. Chang YT, Hsu CY, Chou CT, Lin MW, Shiao YM, et al. (2007) The genetic polymorphisms of POU5F1 gene are associated with psoriasis vulgaris in Chinese. *J Dermatol Sci* 46: 153-156.
88. Kramer U, Illig T, Grune T, Krutmann J, Esser C (2007) Strong associations of psoriasis with antigen processing LMP and transport genes TAP differ by gender and phenotype. *Genes Immun* 8: 513-517.

89. Holm SJ, Carlen LM, Mallbris L, Stahle-Backdahl M, O'Brien KP (2003) Polymorphisms in the SEEK1 and SPR1 genes on 6p21.3 associate with psoriasis in the Swedish population. *Exp Dermatol* 12: 435-444.
90. Hohler T, Kruger A, Schneider PM, Schopf RE, Knop J, et al. (1997) A TNF-alpha promoter polymorphism is associated with juvenile onset psoriasis and psoriatic arthritis. *J Invest Dermatol* 109: 562-565.
91. Nair RP, Stuart P, Henseler T, Jenisch S, Chia NV, et al. (2000) Localization of psoriasis-susceptibility locus PSORS1 to a 60-kb interval telomeric to HLA-C. *Am J Hum Genet* 66: 1833-1844.
92. Capon F, Helms C, Veal CD, Tillman D, Burden AD, et al. (2004) Genetic analysis of PSORS2 markers in a UK dataset supports the association between RAPTOR SNPs and familial psoriasis. *J Med Genet* 41: 459-460.
93. Samuelsson L, Enlund F, Torinsson A, Yhr M, Inerot A, et al. (1999) A genome-wide search for genes predisposing to familial psoriasis by using a stratification approach. *Hum Genet* 105: 523-529.
94. Enlund F, Samuelsson L, Enerback C, Inerot A, Wahlstrom J, et al. (1999) Analysis of three suggested psoriasis susceptibility loci in a large Swedish set of families: confirmation of linkage to chromosome 6p (HLA region), and to 17q, but not to 4q. *Hum Hered* 49: 2-8.
95. Zheng J, Jin S, Shi R (2003) Confirmation of PSORS psoriasis susceptibility loci in a Chinese population. *Arch Dermatol Res* 295: 14-18.
96. Parkinson J, Charon C, Baker BS, Powles AV, Rogers S, et al. (2004) Variation at the IRF2 gene and susceptibility to psoriasis in chromosome 4q-linked families. *J Invest Dermatol* 122: 640-643.
97. Foerster J, Nolte I, Schweiger S, Ehlert C, Bruinenberg M, et al. (2004) Evaluation of the IRF-2 gene as a candidate for PSORS3. *J Invest Dermatol* 122: 61-64.

98. Giardina E, Capon F, De Rosa MC, Mango R, Zambruno G, et al. (2004) Characterization of the loricrin (LOR) gene as a positional candidate for the PSORS4 psoriasis susceptibility locus. *Ann Hum Genet* 68: 639-645.
99. Chang YC, Wu WM, Chen CH, Hu CF, Hsu LA (2008) Association between P478S polymorphism of the filaggrin gene and risk of psoriasis in a Chinese population in Taiwan. *Arch Dermatol Res* 300: 133-137.
100. Broome AM, Ryan D, Eckert RL (2003) S100 protein subcellular localization during epidermal differentiation and psoriasis. *J Histochem Cytochem* 51: 675-685.
101. Benoit S, Toksoy A, Ahlmann M, Schmidt M, Sunderkotter C, et al. (2006) Elevated serum levels of calcium-binding S100 proteins A8 and A9 reflect disease activity and abnormal differentiation of keratinocytes in psoriasis. *Br J Dermatol* 155: 62-66.
102. Zhang XJ, Huang W, Yang S, Sun LD, Zhang FY, et al. (2009) Psoriasis genome-wide association study identifies susceptibility variants within LCE gene cluster at 1q21. *Nat Genet* 41: 205-210.
103. de Cid R, Riveira-Munoz E, Zeeuwen PL, Robarge J, Liao W, et al. (2009) Deletion of the late cornified envelope LCE3B and LCE3C genes as a susceptibility factor for psoriasis. *Nat Genet* 41: 211-215.
104. Giardina E, Sinibaldi C, Chini L, Moschese V, Marulli G, et al. (2006) Co-localization of susceptibility loci for psoriasis (PSORS4) and atopic dermatitis (ATOD2) on human chromosome 1q21. *Hum Hered* 61: 229-236.
105. Enlund F, Samuelsson L, Enerback C, Inerot A, Wahlstrom J, et al. (1999) Psoriasis susceptibility locus in chromosome region 3q21 identified in patients from southwest Sweden. *Eur J Hum Genet* 7: 783-790.
106. Samuelsson L, Stiller C, Friberg C, Nilsson C, Inerot A, et al. (2004) Association analysis of cystatin A and zinc finger protein 148, two genes located at the psoriasis susceptibility locus PSORS5. *J Invest Dermatol* 122: 1399-1400.

107. Zenz R, Eferl R, Kenner L, Florin L, Hummerich L, et al. (2005) Psoriasis-like skin disease and arthritis caused by inducible epidermal deletion of Jun proteins. *Nature* 437: 369-375.
108. Lee YA, Ruschendorf F, Windemuth C, Schmitt-Egenolf M, Stadelmann A, et al. (2000) Genomewide scan in german families reveals evidence for a novel psoriasis-susceptibility locus on chromosome 19p13. *Am J Hum Genet* 67: 1020-1024.
109. Hensen P, Windemuth C, Huffmeier U, Ruschendorf F, Stadelmann A, et al. (2003) Association scan of the novel psoriasis susceptibility region on chromosome 19: evidence for both susceptible and protective loci. *Exp Dermatol* 12: 490-496.
110. Huffmeier U, Lascorz J, Becker T, Schurmeier-Horst F, Magener A, et al. (2009) Characterisation of psoriasis susceptibility locus 6 (PSORS6) in patients with early onset psoriasis and evidence for interaction with PSORS1. *J Med Genet* 46: 736-744.
111. Huffmeier U, Steffens M, Burkhardt H, Lascorz J, Schurmeier-Horst F, et al. (2006) Evidence for susceptibility determinant(s) to psoriasis vulgaris in or near PTPN22 in German patients. *J Med Genet* 43: 517-522.
112. Smith RL, Warren RB, Eyre S, Ho P, Ke X, et al. (2008) Polymorphisms in the IL-12beta and IL-23R genes are associated with psoriasis of early onset in a UK cohort. *J Invest Dermatol* 128: 1325-1327.
113. Plant D, Young HS, Watson RE, Worthington J, Griffiths CE (2006) The CX3CL1-CX3CR1 system and psoriasis. *Exp Dermatol* 15: 900-903.
114. Karason A, Gudjonsson JE, Upmanyu R, Antonsdottir AA, Hauksson VB, et al. (2003) A susceptibility gene for psoriatic arthritis maps to chromosome 16q: evidence for imprinting. *Am J Hum Genet* 72: 125-131.
115. Zhang XJ, Yan KL, Wang ZM, Yang S, Zhang GL, et al. (2007) Polymorphisms in interleukin-15 gene on chromosome 4q31.2 are associated with psoriasis vulgaris in Chinese population. *J Invest Dermatol* 127: 2544-2551.

116. Zhang XJ, He PP, Wang ZX, Zhang J, Li YB, et al. (2002) Evidence for a major psoriasis susceptibility locus at 6p21(PSORS1) and a novel candidate region at 4q31 by genome-wide scan in Chinese hans. *J Invest Dermatol* 119: 1361-1366.
117. Asumalahti K, Laitinen T, Lahermo P, Suomela S, Itkonen-Vatjus R, et al. (2003) Psoriasis susceptibility locus on 18p revealed by genome scan in Finnish families not associated with PSORS1. *J Invest Dermatol* 121: 735-740.
118. Veal CD, Clough RL, Barber RC, Mason S, Tillman D, et al. (2001) Identification of a novel psoriasis susceptibility locus at 1p and evidence of epistasis between PSORS1 and candidate loci. *J Med Genet* 38: 7-13.
119. Bowcock AM, Krueger JG (2005) Getting under the skin: the immunogenetics of psoriasis.[erratum appears in *Nat Rev Immunol.* 2005 Oct;5(10):826]. *Nature Reviews Immunology* 5: 699-711.
120. Fitzgerald O, Dougados M (2006) Psoriatic arthritis: one or more diseases? *Best Practice & Research in Clinical Rheumatology* 20: 435-450.
121. Ritchlin CT (2005) Pathogenesis of psoriatic arthritis. *Current Opinion in Rheumatology* 17: 406-412.
122. Gladman DD, Farewell VT, Pellett F, Schentag C, Rahman P (2003) HLA is a candidate region for psoriatic arthritis. evidence for excessive HLA sharing in sibling pairs. *Hum Immunol* 64: 887-889.
123. Rahman P, Elder JT (2005) Genetic epidemiology of psoriasis and psoriatic arthritis. *Annals of the Rheumatic Diseases* 64 Suppl 2: ii37-39; discussion ii40-31.
124. Tiilikainen A, Lassus A, Karvonen J, Vartiainen P, Julin M (1980) Psoriasis and HLA-Cw6. *British Journal of Dermatology* 102: 179-184.
125. Nair RP, Henseler T, Jenisch S, Stuart P, Bichakjian CK, et al. (1997) Evidence for two psoriasis susceptibility loci (HLA and 17q) and two novel candidate regions (16q and 20p) by genome-wide scan. *Hum Molec Genet* 6: 1349-1356.

126. Helms C, Saccone NL, Cao L, Daw JA, Cao K, et al. (2005) Localization of PSORS1 to a haplotype block harboring HLA-C and distinct from corneodesmosin and HCR. *Hum Genet* 118: 466-476.
127. Tsunemi Y, Saeki H, Nakamura K, Sekiya T, Hirai K, et al. (2002) Interleukin-12 p40 gene (IL12B) 3'-untranslated region polymorphism is associated with susceptibility to atopic dermatitis and psoriasis vulgaris. *J Dermatol Sci* 30: 161-166.
128. Cargill M, Schrodi SJ, Chang M, Garcia VE, Brandon R, et al. (2007) A large-scale genetic association study confirms IL12B and leads to the identification of IL23R as psoriasis-risk genes. *Am J Hum Genet* 80: 273-290.
129. Duerr RH, Taylor KD, Brant SR, Rioux JD, Silverberg MS, et al. (2006) A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science* 314: 1461-1463.
130. Speckman RA, Wright Daw JA, Helms C, Duan S, Cao L, et al. (2003) Novel immunoglobulin superfamily gene cluster, mapping to a region of human chromosome 17q25, linked to psoriasis susceptibility. *Hum Genet* 112: 34-41.
131. Mitchell MK, Gregersen PK, Johnson S, Parsons R, Vlahov D (2004) The New York Cancer Project: rationale, organization, design, and baseline characteristics. *J Urban Health* 81: 301-310.
132. Consortium tWTCC (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447: 661-683.
133. Manolio TA, Rodriguez LL, Brooks L, Abecasis G, Ballinger D, et al. (2007) New models of collaboration in genome-wide association studies: the Genetic Association Information Network. *Nat Genet* 39: 1045-1051.
134. Rioux JD, Xavier RJ, Taylor KD, Silverberg MS, Goyette P, et al. (2007) Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. *Nat Genet* 39: 596-604.

135. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.
136. Plenge RM, Padyukov L, Remmers EF, Purcell S, Lee AT, et al. (2005) Replication of putative candidate-gene associations with rheumatoid arthritis in >4,000 samples from North America and Sweden: association of susceptibility with PTPN22, CTLA4, and PADI4. *Am J Hum Genet* 77: 1044-1060.
137. Devlin B, Roeder K, Wasserman L (2001) Genomic control, a new approach to genetic-based association studies. *Theor Popul Biol* 60: 155-166.
138. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, et al. (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38: 904-909.
139. Seldin MF, Shigeta R, Villoslada P, Selmi C, Tuomilehto J, et al. (2006) European population substructure: clustering of northern and southern populations. *PLoS Genetics* 2: e143.
140. Martin ER, Monks SA, Warren LL, Kaplan NL (2000) A test for linkage and association in general pedigrees: the pedigree disequilibrium test. *Am J Hum Genet* 67: 146-154.
141. Speckman RA, Wright Daw JA, Helms C, Duan S, Cao L, et al. (2003) Novel immunoglobulin superfamily gene cluster, mapping to a region of human chromosome 17q25, linked to psoriasis susceptibility. *Human Genetics* 112: 34-41.
142. Seldin MF, Shigeta R, Villoslada P, Selmi C, Tuomilehto J, et al. (2006) European population substructure: clustering of northern and southern populations. *PLoS Genet* 2: e143.
143. Pritchard JK, Rosenberg NA (1999) Use of unlinked genetic markers to detect population stratification in association studies. *Am J Hum Genet* 65: 220-228.

144. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81: 559-575.
145. Fellay J, Shianna KV, Ge D, Colombo S, Ledergerber B, et al. (2007) A whole-genome association study of major determinants for host control of HIV-1. *Science* 317: 944-947.
146. Chang JT, Shevach EM, Segal BM (1999) Regulation of interleukin (IL)-12 receptor beta2 subunit expression by endogenous IL-12: a critical step in the differentiation of pathogenic autoreactive T cells. *J Exp Med* 189: 969-978.
147. Airoidi I, Di Carlo E, Cocco C, Sorrentino C, Fais F, et al. (2005) Lack of Il12rb2 signaling predisposes to spontaneous autoimmunity and malignancy. *Blood* 106: 3846-3853.
148. Petit MM, Schoenmakers EF, Huysmans C, Geurts JM, Mandahl N, et al. (1999) LHFP, a novel translocation partner gene of HMGIC in a lipoma, is a member of a new family of LHFP-like genes. *Genomics* 57: 438-441.
149. Loh E, Hong W (2004) The binary interacting network of the conserved oligomeric Golgi tethering complex. *J Biol Chem* 279: 24640-24648.
150. Kubota Y, Sano M, Goda S, Suzuki N, Nishiwaki K (2006) The conserved oligomeric Golgi complex acts in organ morphogenesis via glycosylation of an ADAM protease in *C. elegans*. *Development* 133: 263-273.
151. Garlisi CG, Zou J, Devito KE, Tian F, Zhu FX, et al. (2003) Human ADAM33: protein maturation and localization. *Biochem Biophys Res Commun* 301: 35-43.
152. Lesueur F, Oudot T, Heath S, Foglio M, Lathrop M, et al. (2007) ADAM33, a new candidate for psoriasis susceptibility. *PLoS ONE* 2: e906.

153. Wykes RCE, Lee M, Duffy SM, Yang W, Seward EP, et al. (2007) Functional transient receptor potential melastatin 7 channels are critical for human mast cell survival. *Journal of Immunology* 179: 4045-4052.
154. Friedmann E, Lemberg MK, Weihofen A, Dev KK, Dengler U, et al. (2004) Consensus analysis of signal peptide peptidase and homologous human aspartic proteases reveals opposite topology of catalytic domains compared with presenilins. *J Biol Chem* 279: 50790-50798.
155. Friedmann E, Hauben E, Maylandt K, Schleegeer S, Vreugde S, et al. (2006) SPPL2a and SPPL2b promote intramembrane proteolysis of TNFalpha in activated dendritic cells to trigger IL-12 production. *Nat Cell Biol* 8: 843-848.
156. Capon F, Semprini S, Chimenti S, Fabrizi G, Zambruno G, et al. (2001) Fine mapping of the PSORS4 psoriasis susceptibility region on chromosome 1q21. *J Invest Dermatol* 116: 728-730.
157. Jackson B, Tilli CM, Hardman MJ, Avilion AA, MacLeod MC, et al. (2005) Late cornified envelope family in differentiating epithelia--response to calcium and ultraviolet irradiation. *J Invest Dermatol* 124: 1062-1070.
158. Krensky AM, Clayberger C (2005) Granulysin: a novel host defense molecule. *American Journal of Transplantation* 5: 1789-1792.
159. Zheng CF, Ma LL, Jones GJ, Gill MJ, Krensky AM, et al. (2007) Cytotoxic CD4+ T cells use granulysin to kill *Cryptococcus neoformans*, and activation of this pathway is defective in HIV patients. *Blood* 109: 2049-2057.
160. Ishisaki Z, Takaishi M, Furuta I, Huh N (2001) Calmin, a protein with calponin homology and transmembrane domains expressed in maturing spermatogenic cells. *Genomics* 74: 172-179.
161. Smith A, Bourdeau I, Wang J, Bondy CA (2005) Expression of Catenin family members CTNNA1, CTNNA2, CTNNB1 and JUP in the primate prefrontal cortex and hippocampus. *Brain Research Molecular Brain Research*. 135: 225-231.

162. van Heel DA, Franke L, Hunt KA, Gwilliam R, Zhernakova A, et al. (2007) A genome-wide association study for celiac disease identifies risk variants in the region harboring IL2 and IL21. *Nature Genet* 39: 827-829.
163. Zhernakova A, Alizadeh BZ, Bevova M, van Leeuwen MA, Coenen MJ, et al. (2007) Novel association in chromosome 4q27 region with rheumatoid arthritis and confirmation of type 1 diabetes point to a general risk locus for autoimmune diseases. *Am J Hum Genet* 81: 1284-1288.
164. Helms C, Cao L, Krueger JG, Wijsman EM, Chamian F, et al. (2003) A putative RUNX1 binding site variant between SLC9A3R1 and NAT9 is associated with susceptibility to psoriasis.[see comment]. *Nature Genetics* 35: 349-356.
165. Bowcock AM (2004) Psoriasis genetics: the way forward. *J Invest Dermatol* 122: xv-xvii.
166. Gottlieb AB (1988) Immunologic mechanisms in psoriasis. *J Am Acad Dermatol* 18: 1376-1380.
167. Krueger JG, Walters IB, Miyazawa M, Gilleaudeau P, Hakimi J, et al. (2000) Successful in vivo blockade of CD25 (high-affinity interleukin 2 receptor) on T cells by administration of humanized anti-Tac antibody to patients with psoriasis. *J Am Acad Dermatol* 43: 448-458.
168. Laurence A, Tato CM, Davidson TS, Kanno Y, Chen Z, et al. (2007) Interleukin-2 signaling via STAT5 constrains T helper 17 cell generation. *Immunity* 26: 371-381.
169. Zhou L, Ivanov, II, Spolski R, Min R, Shenderov K, et al. (2007) IL-6 programs T(H)-17 cell differentiation by promoting sequential engagement of the IL-21 and IL-23 pathways. *Nat Immunol* 8: 967-974.
170. Herber D, Brown TP, Liang S, Young DA, Collins M, et al. (2007) IL-21 has a pathogenic role in a lupus-prone mouse model and its blockade with IL-21R.Fc reduces disease progression. *J Immunol* 178: 3822-3830.

171. Chang M, Li Y, Yan C, Callis-Duffin KP, Matsunami N, et al. (2007) Variants in the 5q31 cytokine gene cluster are associated with psoriasis. *Genes Immun.*
172. Capon F, Di Meglio P, Szaub J, Prescott NJ, Dunster C, et al. (2007) Sequence variants in the genes for the interleukin-23 receptor (IL23R) and its ligand (IL12B) confer protection against psoriasis. *Hum Genet.*
173. Cargill M, Schrodi SJ, Chang M, Garcia VE, Brandon R, et al. (2007) A large-scale genetic association study confirms IL12B and leads to the identification of IL23R as psoriasis-risk genes. *American Journal of Human Genetics* 80: 273-290.
174. Ponomarenko JV, Orlova GV, Merkulova TI, Gorshkova EV, Fokin ON, et al. (2002) rSNP_Guide: an integrated database-tools system for studying SNPs and site-directed mutations in transcription factor binding sites. *Hum Mutat* 20: 239-248.
175. Bosma PJ, Chowdhury JR, Bakker C, Gantla S, de Boer A, et al. (1995) The genetic basis of the reduced expression of bilirubin UDP-glucuronosyltransferase 1 in Gilbert's syndrome. *N Engl J Med* 333: 1171-1175.
176. Jinnai N, Sakagami T, Sekigawa T, Kakihara M, Nakajima T, et al. (2004) Polymorphisms in the prostaglandin E2 receptor subtype 2 gene confer susceptibility to aspirin-intolerant asthma: a candidate gene approach. *Hum Mol Genet* 13: 3203-3217.
177. Marzec JM, Christie JD, Reddy SP, Jedlicka AE, Vuong H, et al. (2007) Functional polymorphisms in the transcription factor NRF2 in humans increase the risk of acute lung injury. *FASEB J* 21: 2237-2246.
178. Nakamura S, Kugiyama K, Sugiyama S, Miyamoto S, Koide S, et al. (2002) Polymorphism in the 5'-flanking region of human glutamate-cysteine ligase modifier subunit gene is associated with myocardial infarction. *Circulation* 105: 2968-2973.

179. Ono S, Ezura Y, Emi M, Fujita Y, Takada D, et al. (2003) A promoter SNP (-1323T>C) in G-substrate gene (GSBS) correlates with hypercholesterolemia. *J Hum Genet* 48: 447-450.
180. Sugatani J, Yamakawa K, Yoshinari K, Machida T, Takagi H, et al. (2002) Identification of a defect in the UGT1A1 gene promoter and its association with hyperbilirubinemia. *Biochem Biophys Res Commun* 292: 492-497.
181. Liu Y, Helms C, Liao W, Zaba LC, Duan S, et al. (2008) A genome-wide association study of psoriasis and psoriatic arthritis identifies new disease loci. *PLoS Genet* 4: e1000041.
182. Schmitt-Egenolf M, Eiermann TH, Boehncke WH, Stander M, Sterry W (1996) Familial juvenile onset psoriasis is associated with the human leukocyte antigen (HLA) class I side of the extended haplotype Cw6-B57-DRB1*0701-DQA1*0201-DQB1*0303: a population- and family-based study. *J Invest Dermatol* 106: 711-714.
183. Balendran N, Clough RL, Arguello JR, Barber R, Veal C, et al. (1999) Characterization of the major susceptibility region for psoriasis at chromosome 6p21.3. *J Invest Dermatol* 113: 322-328.
184. Oka A, Tamiya G, Tomizawa M, Ota M, Katsuyama Y, et al. (1999) Association analysis using refined microsatellite markers localizes a susceptibility locus for psoriasis vulgaris within a 111 kb segment telomeric to the HLA-C gene. *Hum Mol Genet* 8: 2165-2170.
185. Fan X, Yang S, Huang W, Wang ZM, Sun LD, et al. (2008) Fine mapping of the psoriasis susceptibility locus PSORS1 supports HLA-C as the susceptibility gene in the Han Chinese population. *PLoS Genet* 4: e1000038.
186. Blais ME, Dong T, Rowland-Jones S (2011) HLA-C as a mediator of natural killer and T-cell activation: spectator or key player? *Immunology* 133: 1-7.

187. Johnston A, Gudjonsson JE, Sigmundsdottir H, Love TJ, Valdimarsson H (2004) Peripheral blood T cell responses to keratin peptides that share sequences with streptococcal M proteins are largely restricted to skin-homing CD8(+) T cells. *Clin Exp Immunol* 138: 83-93.
188. Carlen L, Sakuraba K, Stahle M, Sanchez F (2007) HLA-C expression pattern is spatially different between psoriasis and eczema skin lesions. *J Invest Dermatol* 127: 342-348.
189. Allen M, Ishida-Yamamoto A, McGrath J, Davison S, Iizuka H, et al. (2001) Corneodesmosin expression in psoriasis vulgaris differs from normal skin and other inflammatory skin disorders. *Lab Invest* 81: 969-976.
190. Suomela S, Elomaa O, Asumalahti K, Kariniemi AL, Karvonen SL, et al. (2003) HCR, a candidate gene for psoriasis, is expressed differently in psoriasis and other hyperproliferative skin disorders and is downregulated by interferon-gamma in keratinocytes. *J Invest Dermatol* 121: 1360-1364.
191. Chang YT, Chou CT, Shiao YM, Lin MW, Yu CW, et al. (2006) Psoriasis vulgaris in Chinese individuals is associated with PSORS1C3 and CDSN genes. *Br J Dermatol* 155: 663-669.
192. Chang LW, Nagarajan R, Magee JA, Milbrandt J, Stormo GD (2006) A systematic model to predict transcriptional regulatory mechanisms based on overrepresentation of transcription factor binding profiles. *Genome Res* 16: 405-413.
193. Warner JB, Philippakis AA, Jaeger SA, He FS, Lin J, et al. (2008) Systematic identification of mammalian regulatory motifs' target genes and functions. *Nat Methods* 5: 347-353.
194. Torkamani A, Schork NJ (2008) Predicting functional regulatory polymorphisms. *Bioinformatics* 24: 1787-1792.
195. Macintyre G, Bailey J, Haviv I, Kowalczyk A (2010) is-rSNP: a novel technique for in silico regulatory SNP detection. *Bioinformatics* 26: i524-530.

196. Alam J, Cook JL (1990) Reporter genes: application to the study of mammalian gene transcription. *Anal Biochem* 188: 245-254.
197. Garner MM, Revzin A (1981) A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the *Escherichia coli* lactose operon regulatory system. *Nucleic Acids Res* 9: 3047-3060.
198. Chorley BN, Wang X, Campbell MR, Pittman GS, Noureddine MA, et al. (2008) Discovery and verification of functional single nucleotide polymorphisms in regulatory genomic regions: current and developing technologies. *Mutat Res* 659: 147-157.
199. de Guzman Strong C, Conlan S, Deming CB, Cheng J, Sears KE, et al. (2010) A milieu of regulatory elements in the epidermal differentiation complex syntenic block: implications for atopic dermatitis and psoriasis. *Hum Mol Genet* 19: 1453-1460.
200. Wingender E (2008) The TRANSFAC project as an example of framework technology that supports the analysis of genomic regulation. *Brief Bioinform* 9: 326-332.
201. Portales-Casamar E, Thongjuea S, Kwon AT, Arenillas D, Zhao X, et al. (2010) JASPAR 2010: the greatly expanded open-access database of transcription factor binding profiles. *Nucleic Acids Res* 38: D105-110.
202. Hellman LM, Fried MG (2007) Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nat Protoc* 2: 1849-1861.
203. Huffmeier U, Uebe S, Ekici AB, Bowes J, Giardina E, et al. (2010) Common variants at TRAF3IP2 are associated with susceptibility to psoriatic arthritis and psoriasis. *Nat Genet* 42: 996-999.
204. Nair RP, Duffin KC, Helms C, Ding J, Stuart PE, et al. (2009) Genome-wide scan reveals association of psoriasis with IL-23 and NF-kappaB pathways. *Nat Genet* 41: 199-204.

205. Oji V, Eckl KM, Aufenvenne K, Natebus M, Tarinski T, et al. (2010) Loss of corneodesmosin leads to severe skin barrier defect, pruritus, and atopy: unraveling the peeling skin disease. *Am J Hum Genet* 87: 274-281.
206. Cheng L, Zhang SZ, Xiao CY, Hou YP, Li L, et al. (2000) The A5.1 allele of the major histocompatibility complex class I chain-related gene A is associated with psoriasis vulgaris in Chinese. *Br J Dermatol* 143: 324-329.
207. Choi HB, Han H, Youn JI, Kim TY, Kim TG (2000) MICA 5.1 allele is a susceptibility marker for psoriasis in the Korean population. *Tissue Antigens* 56: 548-550.
208. Pollock R, Chandran V, Barrett J, Eder L, Pellett F, et al. (2011) Differential major histocompatibility complex class I chain-related A allele associations with skin and joint manifestations of psoriatic disease. *Tissue Antigens*.
209. Kel AE, Gossling E, Reuter I, Cheremushkin E, Kel-Margoulis OV, et al. (2003) MATCH: A tool for searching transcription factor binding sites in DNA sequences. *Nucleic Acids Res* 31: 3576-3579.
210. Chen R, Davydov EV, Sirota M, Butte AJ (2010) Non-synonymous and synonymous coding SNPs show similar likelihood and effect size of human disease association. *PLoS One* 5: e13574.
211. Vasiliev GV, Merkulov VM, Kobzev VF, Merkulova TI, Ponomarenko MP, et al. (1999) Point mutations within 663-666 bp of intron 6 of the human TDO2 gene, associated with a number of psychiatric disorders, damage the YY-1 transcription factor binding site. *FEBS Lett* 462: 85-88.
212. Knight JC, Udalova I, Hill AV, Greenwood BM, Peshu N, et al. (1999) A polymorphism that affects OCT-1 binding to the TNF promoter region is associated with severe malaria. *Nat Genet* 22: 145-150.
213. Langdon SD, Kaufman RE (1998) Gamma-globin gene promoter elements required for interaction with globin enhancers. *Blood* 91: 309-318.

214. Gimelbrant A, Hutchinson JN, Thompson BR, Chess A (2007) Widespread monoallelic expression on human autosomes. *Science* 318: 1136-1140.
215. Pastinen T, Hudson TJ (2004) Cis-acting regulatory variation in the human genome. *Science* 306: 647-650.
216. Won KJ, Agarwal S, Shen L, Shoemaker R, Ren B, et al. (2009) An integrated approach to identifying cis-regulatory modules in the human genome. *PLoS One* 4: e5501.
217. Forton JT, Udalova IA, Campino S, Rockett KA, Hull J, et al. (2007) Localization of a long-range cis-regulatory element of IL13 by allelic transcript ratio mapping. *Genome Res* 17: 82-87.
218. Spilianakis CG, Flavell RA (2004) Long-range intrachromosomal interactions in the T helper type 2 cytokine locus. *Nat Immunol* 5: 1017-1027.
219. Capon F, Allen MH, Ameen M, Burden AD, Tillman D, et al. (2004) A synonymous SNP of the corneodesmosin gene leads to increased mRNA stability and demonstrates association with psoriasis across diverse ethnic groups. *Hum Mol Genet* 13: 2361-2368.
220. Suarez-Farinas M, Lowes MA, Zaba LC, Krueger JG (2010) Evaluation of the Psoriasis Transcriptome across Different Studies by Gene Set Enrichment Analysis (GSEA). *PLoS ONE* 5: e10247.
221. Gudjonsson JE, Ding J, Johnston A, Tejasvi T, Guzman AM, et al. (2010) Assessment of the Psoriatic Transcriptome in a Large Sample: Additional Regulated Genes and Comparisons with In Vitro Models. *J Invest Dermatol*.
222. Langley RG, Krueger GG, Griffiths CE (2005) Psoriasis: epidemiology, clinical features, and quality of life. *Ann Rheum Dis* 64 Suppl 2: ii18-23; discussion ii24-15.

223. Lew W, Bowcock AM, Krueger JG (2004) Psoriasis vulgaris: cutaneous lymphoid tissue supports T-cell activation and "Type 1" inflammatory gene expression. *Trends Immunol* 25: 295-305.
224. Zaba LC, Fuentes-Duculan J, Eungdamrong NJ, Abello MV, Novitskaya I, et al. (2009) Psoriasis is characterized by accumulation of immunostimulatory and Th1/Th17 cell-polarizing myeloid dendritic cells. *J Invest Dermatol* 129: 79-88.
225. Mee JB, Johnson CM, Morar N, Burslem F, Groves RW (2007) The psoriatic transcriptome closely resembles that induced by interleukin-1 in cultured keratinocytes: dominance of innate immune responses in psoriasis. *Am J Pathol* 171: 32-42.
226. Quekenborn-Trinquet V, Fogel P, Aldana-Jamayrac O, Ancian P, Demarchez M, et al. (2005) Gene expression profiles in psoriasis: analysis of impact of body site location and clinical severity. *Br J Dermatol* 152: 489-504.
227. Nomura I, Gao B, Boguniewicz M, Darst MA, Travers JB, et al. (2003) Distinct patterns of gene expression in the skin lesions of atopic dermatitis and psoriasis: a gene microarray analysis. *J Allergy Clin Immunol* 112: 1195-1202.
228. Bowcock AM, Shannon W, Du F, Duncan J, Cao K, et al. (2001) Insights into psoriasis and other inflammatory diseases from large-scale gene expression studies. *Hum Mol Genet* 10: 1793-1805.
229. Wilson CB, Rowell E, Sekimata M (2009) Epigenetic control of T-helper-cell differentiation. *Nat Rev Immunol* 9: 91-105.
230. Champagne FA, Curley JP (2009) Epigenetic mechanisms mediating the long-term effects of maternal care on development. *Neurosci Biobehav Rev* 33: 593-600.
231. Shi L, Wu J (2009) Epigenetic regulation in mammalian preimplantation embryo development. *Reprod Biol Endocrinol* 7: 59.

232. Bell J, Pai A, Pickrell J, Gaffney D, Pique-Regi R, et al. (2011) DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. *Genome Biology* 12: R10.
233. Koga Y, Pelizzola M, Cheng E, Krauthammer M, Sznol M, et al. (2009) Genome-wide screen of promoter methylation identifies novel markers in melanoma. *Genome Res* 19: 1462-1470.
234. Ordway JM, Bedell JA, Citek RW, Nunberg A, Garrido A, et al. (2006) Comprehensive DNA methylation profiling in a human cancer genome identifies novel epigenetic targets. *Carcinogenesis* 27: 2409-2423.
235. Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, et al. (2009) The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat Genet* 41: 178-186.
236. Hu M, Yao J, Cai L, Bachman KE, van den Brule F, et al. (2005) Distinct epigenetic changes in the stromal cells of breast cancers. *Nat Genet* 37: 899-905.
237. Bibikova M, Lin Z, Zhou L, Chudin E, Garcia EW, et al. (2006) High-throughput DNA methylation profiling using universal bead arrays. *Genome Res* 16: 383-393.
238. Davis S, Du P, Bilke S (2010) methylumi: Handle Illumina methylation data. v1.4.0 ed.
239. Du P, Kibbe WA, Lin SM (2008) lumi: a pipeline for processing Illumina microarray. *Bioinformatics* 24: 1547-1548.
240. Gentleman R, Carey V, Bates D, Bolstad B, Dettling M, et al. (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology* 5: R80.
241. R Development Core Team (2010) R: A Language and Environment for Statistical Computing. v2.12.0 ed. Vienna, Austria: R Foundation For Statistical Computing.

242. Smyth G (2005) Limma: linear models for microarray data. In: Gentleman R, Carey V, Dudoit S, Irizarry R, Huber W, editors. *Bioinformatics and Computational Biology Solutions using R and Bioconductor*. New York: Springer. pp. 397-420.
243. Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B (Methodological)* 57: 289-300.
244. Champely S (2009) pwr: Basic functions for power analysis. v1.1.1 ed.
245. Culhane AC, Perriere G, Considine EC, Cotter TG, Higgins DG (2002) Between-group analysis of microarray data. *Bioinformatics* 18: 1600-1608.
246. Culhane AC, Thioulouse J, Perriere G, Higgins DG (2005) MADE4: an R package for multivariate analysis of gene expression data. *Bioinformatics* 21: 2789-2790.
247. Colella S, Shen L, Baggerly KA, Issa JP, Krahe R (2003) Sensitive and quantitative universal Pyrosequencing methylation analysis of CpG sites. *Biotechniques* 35: 146-150.
248. Du P, Zhang X, Huang CC, Jafari N, Kibbe WA, et al. (2010) Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics* 11: 587.
249. Gudjonsson JE (2007) Analysis of global gene expression and genetic variation in psoriasis. *J Am Acad Dermatol* 57: 365.
250. Guttman-Yassky E, Suarez-Farinas M, Chiricozzi A, Nograles KE, Shemer A, et al. (2009) Broad defects in epidermal cornification in atopic dermatitis identified through genomic analysis. *J Allergy Clin Immunol* 124: 1235-1244 e1258.
251. Beyer V, Wolverton SE (2010) Recent Trends in Systemic Psoriasis Treatment Costs. *Arch Dermatol* 146: 46-54.

252. Oestreicher JL, Walters IB, Kikuchi T, Gilleaudeau P, Surette J, et al. (2001) Molecular classification of psoriasis disease-associated genes through pharmacogenomic expression profiling. *Pharmacogenomics J* 1: 272-287.
253. Suomela S, Cao L, Bowcock A, Saarialho-Kere U (2004) Interferon alpha-inducible protein 27 (IFI27) is upregulated in psoriatic skin and certain epithelial cancers. *J Invest Dermatol* 122: 717-721.
254. Lee Y-A, Rüschenhoff F, Windemuth C, Schmitt-Egenolf M, Stadelmann A, et al. (2000) Genomewide Scan in German Families Reveals Evidence for a Novel Psoriasis-Susceptibility Locus on Chromosome 19p13. *The American Journal of Human Genetics* 67: 1020-1024.
255. Swiecki M, Scheaffer SM, Allaire M, Fremont DH, Colonna M, et al. (2011) Structural and Biophysical Analysis of BST-2/Tetherin Ectodomains Reveals an Evolutionary Conserved Design to Inhibit Virus Release. *Journal of Biological Chemistry* 286: 2987-2997.
256. Yoo H, Park S-H, Ye S-K, Kim M IFN- γ -induced BST2 mediates monocyte adhesion to human endothelial cells. *Cellular Immunology* 267: 23-29.
257. Blasius AL, Giurisato E, Cella M, Schreiber RD, Shaw AS, et al. (2006) Bone Marrow Stromal Cell Antigen 2 Is a Specific Marker of Type I IFN-Producing Cells in the Naive Mouse, but a Promiscuous Cell Surface Antigen following IFN Stimulation. *The Journal of Immunology* 177: 3260-3265.
258. Fujita PA, Rhead B, Zweig AS, Hinrichs AS, Karolchik D, et al. (2010) The UCSC Genome Browser database: update 2011. *Nucleic Acids Research* 39: D876-D882.
259. Strange A, Capon F, Spencer CC, Knight J, Weale ME, et al. (2010) A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. *Nat Genet* 42: 985-990.

260. Stuart PE, Nair RP, Ellinghaus E, Ding J, Tejasvi T, et al. (2010) Genome-wide association analysis identifies three psoriasis susceptibility loci. *Nat Genet* 42: 1000-1004.
261. Ellinghaus E, Ellinghaus D, Stuart PE, Nair RP, Debrus S, et al. (2010) Genome-wide association study identifies a psoriasis susceptibility locus at TRAF3IP2. *Nat Genet* 42: 991-995.
262. Duffin KC, Woodcock J, Krueger GG (2010) Genetic variations associated with psoriasis and psoriatic arthritis found by genome-wide association. *Dermatol Ther* 23: 101-113.
263. Huffmeier U, Bergboer JG, Becker T, Armour JA, Traupe H, et al. (2010) Replication of LCE3C-LCE3B CNV as a risk factor for psoriasis and analysis of interaction with other genetic risk factors. *J Invest Dermatol* 130: 979-984.
264. Yeilding N, Szapary P, Brodmerkel C, Benson J, Plotnick M, et al. (2011) Development of the IL-12/23 antagonist ustekinumab in psoriasis: past, present, and future perspectives. *Ann N Y Acad Sci* 1222: 30-39.
265. Krueger GG, Langley RG, Leonardi C, Yeilding N, Guzzo C, et al. (2007) A human interleukin-12/23 monoclonal antibody for the treatment of psoriasis. *N Engl J Med* 356: 580-592.
266. Bettelli E, Oukka M, Kuchroo VK (2007) T(H)-17 cells in the circle of immunity and autoimmunity. *Nat Immunol* 8: 345-350.
267. Lizzul PF, Aphale A, Malaviya R, Sun Y, Masud S, et al. (2005) Differential expression of phosphorylated NF-kappaB/RelA in normal and psoriatic epidermis and downregulation of NF-kappaB in response to treatment with etanercept. *J Invest Dermatol* 124: 1275-1283.
268. Mauro C, Pacifico F, Lavorgna A, Mellone S, Iannetti A, et al. (2006) ABIN-1 binds to NEMO/IKKgamma and co-operates with A20 in inhibiting NF-kappaB. *J Biol Chem* 281: 18482-18488.

269. Capon F, Bijlmakers MJ, Wolf N, Quaranta M, Huffmeier U, et al. (2008) Identification of ZNF313/RNF114 as a novel psoriasis susceptibility gene. *Hum Mol Genet* 17: 1938-1945.
270. Hollox EJ, Huffmeier U, Zeeuwen PL, Palla R, Lascorz J, et al. (2008) Psoriasis is associated with increased beta-defensin genomic copy number. *Nat Genet* 40: 23-25.
271. Niyonsaba F, Ushio H, Nagaoka I, Okumura K, Ogawa H (2005) The human beta-defensins (-1, -2, -3, -4) and cathelicidin LL-37 induce IL-18 secretion through p38 and ERK MAPK activation in primary human keratinocytes. *J Immunol* 175: 1776-1784.
272. Leclerc EA, Huchenoq A, Mattiuzzo NR, Metzger D, Chambon P, et al. (2009) Corneodesmosin gene ablation induces lethal skin-barrier disruption and hair-follicle degeneration related to desmosome dysfunction. *J Cell Sci* 122: 2699-2709.
273. Docampo E, Rabionet R, Riveira-Munoz E, Escaramis G, Julia A, et al. (2010) Deletion of the late cornified envelope genes, LCE3C and LCE3B, is associated with rheumatoid arthritis. *Arthritis Rheum* 62: 1246-1251.
274. Patsopoulos NA, Analatos AA, Ioannidis JP (2005) Relative citation impact of various study designs in the health sciences. *JAMA* 293: 2362-2366.
275. Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N (2004) Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. *J Natl Cancer Inst* 96: 434-442.
276. Riveira-Munoz E, He SM, Escaramis G, Stuart PE, Huffmeier U, et al. (2010) Meta-Analysis Confirms the LCE3C_LCE3B Deletion as a Risk Factor for Psoriasis in Several Ethnic Groups and Finds Interaction with HLA-Cw6. *J Invest Dermatol*.
277. Bansal V, Libiger O, Torkamani A, Schork NJ (2010) Statistical analysis strategies for association studies involving rare variants. *Nat Rev Genet* 11: 773-785.

278. Li B, Leal SM (2009) Discovery of rare variants via sequencing: implications for the design of complex trait association studies. *PLoS Genet* 5: e1000481.
279. Liu DJ, Leal SM (2010) Replication strategies for rare variant complex trait association studies via next-generation sequencing. *Am J Hum Genet* 87: 790-801.
280. Bjornsson HT, Fallin MD, Feinberg AP (2004) An integrated epigenetic and genetic approach to common human disease. *Trends Genet* 20: 350-358.
281. Appasani K (2007) Epigenomics and sequencing: an intertwined and emerging big science of the next decade. *Pharmacogenomics* 8: 1109-1113.
282. Callinan PA, Feinberg AP (2006) The emerging science of epigenomics. *Hum Mol Genet* 15 Spec No 1: R95-101.
283. Phiel CJ, Zhang F, Huang EY, Guenther MG, Lazar MA, et al. (2001) Histone deacetylase is a direct target of valproic acid, a potent anticonvulsant, mood stabilizer, and teratogen. *J Biol Chem* 276: 36734-36741.
284. Chang KT, Min KT (2002) Regulation of lifespan by histone deacetylase. *Ageing Res Rev* 1: 313-326.
285. Kypreou KP, Sourlingas TG, Sekeri-Pataryas KE (2004) Age-dependent response of lymphocytes in the induction of the linker histone variant, H1 degrees and histone H4 acetylation after treatment with the histone deacetylase inhibitor, trichostatin A. *Exp Gerontol* 39: 469-479.
286. Wang L, Tao R, Hancock WW (2009) Using histone deacetylase inhibitors to enhance Foxp3(+) regulatory T-cell function and induce allograft tolerance. *Immunol Cell Biol* 87: 195-202.
287. Szyf M (2009) Epigenetics, DNA methylation, and chromatin modifying drugs. *Annu Rev Pharmacol Toxicol* 49: 243-263.

288. Yee KW, Jabbour E, Kantarjian HM, Giles FJ (2005) Clinical experience with decitabine in North American patients with myelodysplastic syndrome. *Ann Hematol* 84 Suppl 1: 18-24.
289. Santangelo S, Cousins DJ, Winkelmann NE, Staynov DZ (2002) DNA methylation changes at human Th2 cytokine genes coincide with DNase I hypersensitive site formation during CD4(+) T cell differentiation. *J Immunol* 169: 1893-1903.
290. Balada E, Ordi-Ros J, Serrano-Acedo S, Martinez-Lostao L, Vilardell-Tarres M (2007) Transcript overexpression of the MBD2 and MBD4 genes in CD4+ T cells from systemic lupus erythematosus patients. *J Leukoc Biol* 81: 1609-1616.
291. Bell CG, Finer S, Lindgren CM, Wilson GA, Rakyan VK, et al. (2010) Integrated genetic and epigenetic analysis identifies haplotype-specific methylation in the FTO type 2 diabetes and obesity susceptibility locus. *PLoS One* 5: e14040.
292. Deng W, Liang WB, Gao LB, Wang YY, Lv ML, et al. (2010) Association of ADAM33 polymorphisms and susceptibility to psoriasis. *DNA Cell Biol* 29: 435-439.
293. Siroux V, Bouzigon E, Dizier MH, Pin I, Demenais F, et al. (2008) Replication of association between ADAM33 polymorphisms and psoriasis. *PLoS One* 3: e2448.
294. Awasthi S, Tripathi P, Ganesh S, Husain N (2010) Association of ADAM33 gene polymorphisms with asthma in Indian children. *J Hum Genet*.
295. Su D, Zhang X, Sui H, Lu F, Jin L, et al. (2008) Association of ADAM33 gene polymorphisms with adult allergic asthma and rhinitis in a Chinese Han population. *BMC Med Genet* 9: 82.
296. Hirota T, Hasegawa K, Obara K, Matsuda A, Akahoshi M, et al. (2006) Association between ADAM33 polymorphisms and adult asthma in the Japanese population. *Clin Exp Allergy* 36: 884-891.

297. Howard TD, Postma DS, Jongepier H, Moore WC, Koppelman GH, et al. (2003) Association of a disintegrin and metalloprotease 33 (ADAM33) gene with asthma in ethnically diverse populations. *J Allergy Clin Immunol* 112: 717-722.
298. Xiao J, Han J, Wang X, Hua D, Su D, et al. (2010) Association of ADAM33 gene with susceptibility to COPD in Tibetan population of China. *Mol Biol Rep*.
299. Quaranta M, Burden AD, Griffiths CE, Worthington J, Barker JN, et al. (2009) Differential contribution of CDKAL1 variants to psoriasis, Crohn's disease and type II diabetes. *Genes Immun* 10: 654-658.
300. Li Y, Liao W, Chang M, Schrodi SJ, Bui N, et al. (2009) Further genetic evidence for three psoriasis-risk genes: ADAM33, CDKAL1, and PTPN22. *J Invest Dermatol* 129: 629-634.
301. Wolf N, Quaranta M, Prescott NJ, Allen M, Smith R, et al. (2008) Psoriasis is associated with pleiotropic susceptibility loci identified in type II diabetes and Crohn disease. *J Med Genet* 45: 114-116.
302. Thompson SD, Sudman M, Ramos PS, Marion MC, Ryan M, et al. (2010) The susceptibility loci juvenile idiopathic arthritis shares with other autoimmune diseases extend to PTPN2, COG6, and ANGPT1. *Arthritis Rheum* 62: 3265-3276.
303. Kocsis AK, Lakatos PL, Somogyvari F, Fuszek P, Papp J, et al. (2008) Association of beta-defensin 1 single nucleotide polymorphisms with Crohn's disease. *Scand J Gastroenterol* 43: 299-307.
304. Prado-Montes de Oca E, Garcia-Vargas A, Lozano-Inocencio R, Gallegos-Arreola MP, Sandoval-Ramirez L, et al. (2007) Association of beta-defensin 1 single nucleotide polymorphisms with atopic dermatitis. *Int Arch Allergy Immunol* 142: 211-218.
305. Levy H, Raby BA, Lake S, Tantisira KG, Kwiatkowski D, et al. (2005) Association of defensin beta-1 gene polymorphisms with asthma. *J Allergy Clin Immunol* 115: 252-258.

306. Nair RP, Ruether A, Stuart PE, Jenisch S, Tejasvi T, et al. (2008) Polymorphisms of the IL12B and IL23R genes are associated with psoriasis. *J Invest Dermatol* 128: 1653-1661.
307. Capon F, Di Meglio P, Szaub J, Prescott NJ, Dunster C, et al. (2007) Sequence variants in the genes for the interleukin-23 receptor (IL23R) and its ligand (IL12B) confer protection against psoriasis. *Hum Genet* 122: 201-206.
308. Chang M, Li Y, Yan C, Callis-Duffin KP, Matsunami N, et al. (2008) Variants in the 5q31 cytokine gene cluster are associated with psoriasis. *Genes Immun* 9: 176-181.
309. Li Y, Chang M, Schrodi SJ, Callis-Duffin KP, Matsunami N, et al. (2008) The 5q31 variants associated with psoriasis and Crohn's disease are distinct. *Hum Mol Genet* 17: 2978-2985.
310. Duffin KC, Freeny IC, Schrodi SJ, Wong B, Feng BJ, et al. (2009) Association between IL13 polymorphisms and psoriatic arthritis is modified by smoking. *J Invest Dermatol* 129: 2777-2783.
311. Beghe B, Hall IP, Parker SG, Moffatt MF, Wardlaw A, et al. (2010) Polymorphisms in IL13 pathway genes in asthma and chronic obstructive pulmonary disease. *Allergy* 65: 474-481.
312. Arima K, Umeshita-Suyama R, Sakata Y, Akaiwa M, Mao XQ, et al. (2002) Upregulation of IL-13 concentration in vivo by the IL13 variant associated with bronchial asthma. *J Allergy Clin Immunol* 109: 980-987.
313. Black S, Teixeira AS, Loh AX, Vinall L, Holloway JW, et al. (2009) Contribution of functional variation in the IL13 gene to allergy, hay fever and asthma in the NSHD longitudinal 1946 birth cohort. *Allergy* 64: 1172-1178.
314. Howard TD, Koppelman GH, Xu J, Zheng SL, Postma DS, et al. (2002) Gene-gene interaction in asthma: IL4RA and IL13 in a Dutch population with asthma. *Am J Hum Genet* 70: 230-236.

315. Hunninghake GM, Soto-Quiros ME, Avila L, Su J, Murphy A, et al. (2007) Polymorphisms in IL13, total IgE, eosinophilia, and asthma exacerbations in childhood. *J Allergy Clin Immunol* 120: 84-90.
316. Sadeghnejad A, Karmaus W, Arshad SH, Kurukulaaratchy R, Huebner M, et al. (2008) IL13 gene polymorphisms modify the effect of exposure to tobacco smoke on persistent wheeze and asthma in childhood, a longitudinal study. *Respir Res* 9: 2.
317. Hegab AE, Sakamoto T, Saitoh W, Massoud HH, Massoud HM, et al. (2004) Polymorphisms of IL4, IL13, and ADRB2 genes in COPD. *Chest* 126: 1832-1839.
318. Weger W, Hofer A, Wolf P, El-Shabrawi Y, Renner W, et al. (2008) Role of the interleukin 15 96516A>T and IL15 96330C>A gene polymorphisms in Caucasian patients with chronic plaque psoriasis. *J Dermatol Sci* 51: 147-149.
319. Cooper JD, Walker NM, Healy BC, Smyth DJ, Downes K, et al. (2009) Analysis of 55 autoimmune disease and type II diabetes loci: further confirmation of chromosomes 4q27, 12q13.2 and 12q24.13 as type I diabetes loci, and support for a new locus, 12q13.3-q14.1. *Genes Immun* 10 Suppl 1: S95-120.
320. Glas J, Stallhofer J, Ripke S, Wetzke M, Pfennig S, et al. (2009) Novel genetic risk markers for ulcerative colitis in the IL2/IL21 region are in epistasis with IL23R and suggest a common genetic background for ulcerative colitis and celiac disease. *Am J Gastroenterol* 104: 1737-1744.
321. Festen EA, Goyette P, Scott R, Annese V, Zhernakova A, et al. (2009) Genetic variants in the region harbouring IL2/IL21 associated with ulcerative colitis. *Gut* 58: 799-804.
322. Albers HM, Kurreeman FA, Stoeken-Rijsbergen G, Brinkman DM, Kamphuis SS, et al. (2009) Association of the autoimmunity locus 4q27 with juvenile idiopathic arthritis. *Arthritis Rheum* 60: 901-904.

323. Tremelling M, Cummings F, Fisher SA, Mansfield J, Gwilliam R, et al. (2007) IL23R variation determines susceptibility but not disease phenotype in inflammatory bowel disease. *Gastroenterology* 132: 1657-1664.
324. Cummings JR, Ahmad T, Geremia A, Beckly J, Cooney R, et al. (2007) Contribution of the novel inflammatory bowel disease gene IL23R to disease susceptibility and phenotype. *Inflamm Bowel Dis* 13: 1063-1068.
325. Buning C, Schmidt HH, Molnar T, De Jong DJ, Fiedler T, et al. (2007) Heterozygosity for IL23R p.Arg381Gln confers a protective effect not only against Crohn's disease but also ulcerative colitis. *Aliment Pharmacol Ther* 26: 1025-1033.
326. Wang X, Huang J, Lin Z, Liao Z, Li C, et al. (2010) Single-nucleotide polymorphisms and expression of IL23R in Chinese ankylosing spondylitis patients. *Rheumatol Int* 30: 955-959.
327. Safrany E, Pazar B, Csongei V, Jaromi L, Polgar N, et al. (2009) Variants of the IL23R gene are associated with ankylosing spondylitis but not with Sjogren syndrome in Hungarian population samples. *Scand J Immunol* 70: 68-74.
328. Luszczyk W, Manczak M, Cislo M, Nockowski P, Wisniewski A, et al. (2004) Gene for the activating natural killer cell receptor, KIR2DS1, is associated with susceptibility to psoriasis vulgaris. *Hum Immunol* 65: 758-766.
329. Smith RL, Warren RB, Eyre S, Ke X, Young HS, et al. (2008) Polymorphisms in the PTPN22 region are associated with psoriasis of early onset. *Br J Dermatol* 158: 962-968.
330. Curtin K, Wong J, Allen-Brady K, Camp NJ (2007) Meta-genetic association of rheumatoid arthritis and PTPN22 using PedGenie 2.1. *BMC Proc* 1 Suppl 1: S12.
331. Lea WW, Lee YH (2011) The association between the PTPN22 C1858T polymorphism and systemic lupus erythematosus: a meta-analysis update. *Lupus* 20: 51-57.

332. Skinningsrud B, Husebye ES, Gervin K, Lovas K, Blomhoff A, et al. (2008) Mutation screening of PTPN22: association of the 1858T-allele with Addison's disease. *Eur J Hum Genet* 16: 977-982.
333. Li H, Gao L, Shen Z, Li CY, Li K, et al. (2008) Association study of NFKB1 and SUMO4 polymorphisms in Chinese patients with psoriasis vulgaris. *Arch Dermatol Res* 300: 425-433.
334. Rudofsky G, Jr., Schlotterer A, Humpert PM, Tafel J, Morcos M, et al. (2008) A M55V polymorphism in the SUMO4 gene is associated with a reduced prevalence of diabetic retinopathy in patients with Type 1 diabetes. *Exp Clin Endocrinol Diabetes* 116: 14-17.
335. Lin HY, Wang CL, Hsiao PJ, Lu YC, Chen SY, et al. (2007) SUMO4 M55V variant is associated with diabetic nephropathy in type 2 diabetes. *Diabetes* 56: 1177-1180.
336. Guo D, Li M, Zhang Y, Yang P, Eckenrode S, et al. (2004) A functional variant of SUMO4, a new I kappa B alpha modifier, is associated with type 1 diabetes. *Nat Genet* 36: 837-841.
337. (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447: 661-678.
338. Bowes J, Lawrence R, Eyre S, Panoutsopoulou K, Orozco G, et al. (2010) Rare variation at the TNFAIP3 locus and susceptibility to rheumatoid arthritis. *Hum Genet* 128: 627-633.
339. Elsby LM, Orozco G, Denton J, Worthington J, Ray DW, et al. (2010) Functional evaluation of TNFAIP3 (A20) in rheumatoid arthritis. *Clin Exp Rheumatol* 28: 708-714.
340. Dieguez-Gonzalez R, Calaza M, Perez-Pampin E, Balsa A, Blanco FJ, et al. (2009) Analysis of TNFAIP3, a feedback inhibitor of nuclear factor-kappaB and the neighbor intergenic 6q23 region in rheumatoid arthritis susceptibility. *Arthritis Res Ther* 11: R42.

341. Adrianto I, Wen F, Templeton A, Wiley G, King JB, et al. (2011) Association of a functional variant downstream of TNFAIP3 with systemic lupus erythematosus. *Nat Genet* 43: 253-258.
342. Musone SL, Taylor KE, Lu TT, Nititham J, Ferreira RC, et al. (2008) Multiple polymorphisms in the TNFAIP3 region are independently associated with systemic lupus erythematosus. *Nat Genet* 40: 1062-1064.
343. Graham RR, Cotsapas C, Davies L, Hackett R, Lessard CJ, et al. (2008) Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus. *Nat Genet* 40: 1059-1061.
344. Fung EY, Smyth DJ, Howson JM, Cooper JD, Walker NM, et al. (2009) Analysis of 17 autoimmune disease-associated variants in type 1 diabetes identifies 6q23/TNFAIP3 as a susceptibility locus. *Genes Immun* 10: 188-191.
345. Trynka G, Zhernakova A, Romanos J, Franke L, Hunt KA, et al. (2009) Coeliac disease-associated risk variants in TNFAIP3 and REL implicate altered NF-kappaB signalling. *Gut* 58: 1078-1083.
346. Kawasaki A, Ito S, Furukawa H, Hayashi T, Goto D, et al. (2010) Association of TNFAIP3 interacting protein 1, TNIP1 with systemic lupus erythematosus in a Japanese population: a case-control association study. *Arthritis Res Ther* 12: R174.
347. He CF, Liu YS, Cheng YL, Gao JP, Pan TM, et al. (2010) TNIP1, SLC15A4, ETS1, RasGRP3 and IKZF1 are associated with clinical features of systemic lupus erythematosus in a Chinese Han population. *Lupus* 19: 1181-1186.

APPENDIX I

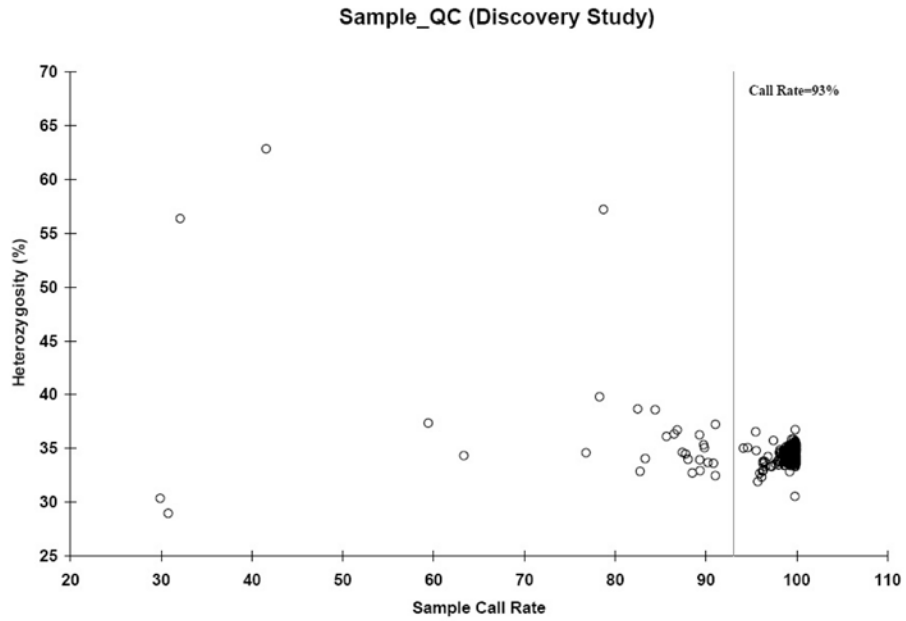
SUPPLEMENTARY TABLES/FIGURES FOR CHAPTER 2

Table A1.1 Summary of cases and controls used in discovery and replication stages.

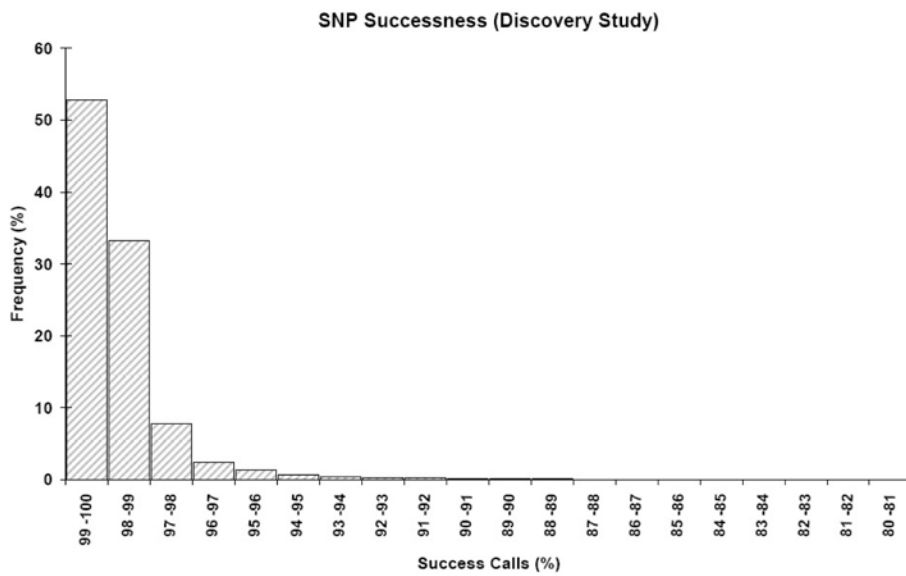
		Discovery Study (n = 742)		Replication Study (n = 2370)	
		Cases	Control	Cases	Control
n		223	519	1153	1217
Sex					
	Males	46.40%	44.51%	45.93%	48.64%
	Females	53.64%	55.49%	54.07%	51.36%
Age (years)					
	Mean \pm SD	49.8 \pm 14.6	46.0 \pm 11.9	44.2 \pm 16.4	56.5 \pm 15.7
	Range	12-92	21-81	2-94	20-94
Site					
	USA	223	519	577	737
	UK	NA	NA	576	480

Figure A1.1 Quality control measures for discovery study. A. Heterozygosity of sample versus genotyping call rate. B. Distribution of SNP success rate in the discovery study.

A.



B.



APPENDIX II

SUPPLEMENTARY TABLES/FIGURES FOR CHAPTER 4

Table A2.1 Sample information. A list of the information for each sample used in the study along with group membership and plot identifiers. Paired With: for paired samples (such as pre-/post-treatment and paired PP/PN) the identifier listed in this column is the paired-sample for that ID. Group: the group the sample belonged to (PP, PN, NN). Contrast: contrasts in which this sample was tested for differential expression. For example, a sample listing the contrasts “Paired PP vs. PN; PP vs. NN” was used to test for differential methylation with the paired PP/PN samples, and was also used to test for differential methylation in the test of all PP samples versus all NN samples.

ID	Paired With	Group	Contrasts
PP1	PN1	PP	Paired Involved vs. Uninvolved; Involved vs. Normal
PP2	PN2	PP	Paired Involved vs. Uninvolved; Involved vs. Normal
PP3	PN3	PP	Paired Involved vs. Uninvolved; Involved vs. Normal
PP4	PN4	PP	Paired Involved vs. Uninvolved; Involved vs. Normal
PP5	PN5	PP	Paired Involved vs. Uninvolved; Involved vs. Normal
PP6	PN6	PP	Paired Involved vs. Uninvolved; Involved vs. Normal
PP7	PN7	PP	Paired Involved vs. Uninvolved; Involved vs. Normal
PP8	PN8	PP	Paired Involved vs. Uninvolved; Involved vs. Normal
PP9	NA	PP	Involved vs. Normal
PP10	NA	PP	Involved vs. Normal
PP11	NA	PP	Involved vs. Normal
PP12	NA	PP	Involved vs. Normal
PP13	NA	PP	Cluster
PP14	NA	PP	Cluster
PP15	NA	PP	Cluster
PP16	NA	PP	Cluster
PP17	NA	PP	Cluster
PP18	NA	PP	Cluster
PP19	NA	PP	Cluster
NN1	NA	NN	Involved vs. Normal; Uninvolved vs. Normal
NN2	NA	NN	Involved vs. Normal; Uninvolved vs. Normal
NN3	NA	NN	Involved vs. Normal; Uninvolved vs. Normal
NN4	NA	NN	Involved vs. Normal; Uninvolved vs. Normal
NN5	NA	NN	Involved vs. Normal; Uninvolved vs. Normal
NN6	NA	NN	Involved vs. Normal; Uninvolved vs. Normal
NN7	NA	NN	Involved vs. Normal; Uninvolved vs. Normal
NN8	NA	NN	Involved vs. Normal; Uninvolved vs. Normal
NN9	NA	NN	Involved vs. Normal; Uninvolved vs. Normal
NN10	NA	NN	Involved vs. Normal; Uninvolved vs. Normal
PN1	PP1	PN	Paired Involved vs. Uninvolved; Uninvolved vs. Normal
PN2	PP2	PN	Paired Involved vs. Uninvolved; Uninvolved vs. Normal
PN3	PP3	PN	Paired Involved vs. Uninvolved; Uninvolved vs. Normal
PN4	PP4	PN	Paired Involved vs. Uninvolved; Uninvolved vs. Normal
PN5	PP5	PN	Paired Involved vs. Uninvolved; Uninvolved vs. Normal
PN6	PP6	PN	Paired Involved vs. Uninvolved; Uninvolved vs. Normal
PN7	PP7	PN	Paired Involved vs. Uninvolved; Uninvolved vs. Normal
PN8	PP8	PN	Paired Involved vs. Uninvolved; Uninvolved vs. Normal

Table A2.2 Pyrosequencing assay information. The information in this table relates to the assays used for confirmatory pyrosequencing. Illumina ID: if the CpG tested was part of the Illumina array the Illumina identifier is listed in this column. Gene: the symbol for the gene nearest the tested CpG. CpG Site: which CpG site is tested. Chr & Location: the chromosome and position mapping for the CpG site on the hg18 human genome build. TSS: distance from the CpG to the transcription start site (TSS). Strand: the strand the target is located on. Sequencing Primer Name: the primer name assigned by EpigenDX. Sequence for analysis: the sequence around the CpG of interest. For sequences with multiple CpGs the correct site is listed in bold red and is underlined.

Illumina ID	Gene	CpG Site	Chr	Location	Distance to TSS (bp)	Sequence for analysis
cg06131859	KYNU	CpG#1	2	143,351,601	64	ACAGTGGCTGGAGATTAC <u>CG</u> ATTGGTTAGGGGAGAGGAAGTAAAG
	C10orf99	CpG#1	10	85,922,708	839	ATGCTGTGAGACAGTCACTTGTCTGCACAGCTTCCACCAACAGGAGC TCCTTGAGG <u>CG</u> GAGGCACAGTGTCTTCTGTGTCCCTGGAGCCAAGCGCATG GCTCAGCCCAGGTCACGTGTCCAGTGAATGGGTGGCATCTGAGC
cg04126866	C10orf99	CpG#2	10	85,922,743	804	ATGCTGTGAGACAGTCACTTGTCTGCACAGCTTCCACCAACAGGAGC TCCTTGAGGCGAGGCACAGTGTCTTCTGTGTCCCTGGAGCCAAG <u>CG</u> CATG GCTCAGCCCAGGTCACGTGTCCAGTGAATGGGTGGCATCTGAGC
	C10orf99	CpG#3	10	85,922,764	783	ATGCTGTGAGACAGTCACTTGTCTGCACAGCTTCCACCAACAGGAGC TCCTTGAGGCGAGGCACAGTGTCTTCTGTGTCCCTGGAGCCAAGCGCATG GCTCAGCCCAGGTC <u>CG</u> TGTCCAGTGAATGGGTGGCATCTGAGC

Table A2.2 (Continue).

Illumina ID	Gene	CpG Site	Chr	Location	Distance to TSS (bp)	Sequence for analysis
cg20161089	IFI27	CpG#1	14	93,647,267	435	ACTGAGCCAGAT <u>CG</u> CGCTCCCTCATCTGTAACATGCGGAGGAGGAGGGT CCCATCTTTTTACGTTAGTGAGGAGATTACATAA
	IFI27	CpG#2	14	93,647,269	437	ACTGAGCCAGATCG <u>CG</u> CTCCCTCATCTGTAACATGCGGAGGAGGAGGGT CCCATCTTTTTACGTTAGTGAGGAGATTACATAA
	IFI27	CpG#3	14	93,647,290	458	ACTGAGCCAGATCGCGCTCCCTCATCTGTAACATG <u>CG</u> GAGGAGGAGGGT CCCATCTTTTTACGTTAGTGAGGAGATTACATAA
	IFI27	CpG#4	14	93,647,317	485	ACTGAGCCAGATCGCGCTCCCTCATCTGTAACATGCGGAGGAGGAGGGT CCCATCTTTTTCA <u>CG</u> TTAGTGAGGAGATTACATAA
	IFI27	CpG#5	14	93,647,351	519	CAGGCACCT <u>CG</u> CCTGCTGTATATGCCTTAAAAATGCGATTGGTTCTGATT TCTTAGTTTTG
	IFI27	CpG#6	14	93,647,377	545	CAGGCACCTCGCCTGCTGTATATGCCTTAAAAATG <u>CG</u> ATTGGTTCTGATT TCTTAGTTTTG
	IFI27	CpG#7	14	93,647,421	589	TGCTC <u>CG</u> TGGAGAGATAAGGGAGTCCCGGAAGTGTCTAAGACATTGGCG CTGGGACTTTCAGGAGAAAGAAAAG
	IFI27	CpG#8	14	93,647,442	610	TGCTCCGTGGAGAGATAAGGGAGTCC <u>CG</u> GGAAGTGTCTAAGACATTGGCG CTGGGACTTTCAGGAGAAAGAAAAG
	IFI27	CpG#9	14	93,647,463	631	TGCTCCGTGGAGAGATAAGGGAGTCCCGGAAGTGTCTAAGACATTGG <u>CG</u> CTGGGACTTTCAGGAGAAAGAAAAG

Table A2.3 Top 50 sites that best discriminate involved from normal skin. Sites listed were determined based on between-group analysis with principal components, and were taken from the first principal component axis. CpG ID: Illumina CpG identifier. Chr & Position_hg18: the chromosome and position mapping for the CpG on build hg18. Gene: symbol for the gene near or overlapping the CpG sites. Direction PP: the direction of methylation change in PP skin compared to NN. PP vs. NN p-value: the corrected p-value from the PP versus PN test of differential methylation. Unnormalized β -value: raw β -values for the CpG in the NN, PN and PP samples shown as mean \pm 95% confidence intervals; shown due to ease of interpretation. Normalized M-value: the normalized M-values for the CpG in the NN, PN and PP samples shown as mean \pm 95% confidence intervals.

CpG ID	Chr	Position_hg18	Gene	PP Direction	Adj p-value PP v. NN	Unnormalized β -value			Normalized M-value		
						NN	PN	PP	NN	PN	PP
cg24468890	1	32,529,881	HDAC1	DOWN	0.015	0.34 \pm 0.01	0.31 \pm 0.03	0.27 \pm 0.01	-0.97 \pm 0.08	-1.16 \pm 0.21	-1.47 \pm 0.09
cg10045881	1	111,571,814	CHI3L2	DOWN	\leq 0.001	0.54 \pm 0.03	0.46 \pm 0.02	0.38 \pm 0.03	0.23 \pm 0.19	-0.21 \pm 0.11	-0.68 \pm 0.18
cg00152644	1	151,334,860	SPRR2E	DOWN	\leq 0.001	0.89 \pm 0.00	0.86 \pm 0.04	0.81 \pm 0.03	3.03 \pm 0.07	2.66 \pm 0.36	2.15 \pm 0.24
cg03165378	1	151,596,506	S100A9	DOWN	0.004	0.79 \pm 0.03	0.73 \pm 0.05	0.57 \pm 0.05	1.91 \pm 0.24	1.45 \pm 0.34	0.44 \pm 0.32
cg16139316	1	151,597,382	S100A9	DOWN	\leq 0.001	0.84 \pm 0.01	0.75 \pm 0.10	0.46 \pm 0.08	2.35 \pm 0.11	1.75 \pm 0.72	-0.23 \pm 0.51
cg02813121	1	151,615,535	S100A12	DOWN	\leq 0.001	0.82 \pm 0.01	0.80 \pm 0.03	0.72 \pm 0.02	2.23 \pm 0.11	2.01 \pm 0.22	1.35 \pm 0.17
cg06131859	2	143,351,601	KYNU	DOWN	\leq 0.001	0.58 \pm 0.03	0.48 \pm 0.07	0.32 \pm 0.04	0.49 \pm 0.19	-0.13 \pm 0.42	-1.12 \pm 0.29
cg18119407	2	201,688,749	CFLAR	DOWN	0.004	0.48 \pm 0.02	0.35 \pm 0.05	0.36 \pm 0.03	-0.11 \pm 0.12	-0.89 \pm 0.30	-0.84 \pm 0.17
cg20950277	4	122,304,691	TNIP3	DOWN	\leq 0.001	0.77 \pm 0.03	0.71 \pm 0.05	0.58 \pm 0.05	1.80 \pm 0.21	1.35 \pm 0.38	0.47 \pm 0.29
cg22346765	6	41,115,500	UNC5CL	DOWN	0.001	0.68 \pm 0.02	0.60 \pm 0.04	0.54 \pm 0.02	1.08 \pm 0.15	0.57 \pm 0.22	0.21 \pm 0.14
cg01980222	6	41,238,895	TREM2	DOWN	0.005	0.50 \pm 0.01	0.48 \pm 0.04	0.41 \pm 0.02	0.02 \pm 0.06	-0.10 \pm 0.25	-0.53 \pm 0.14
cg13784855	8	128,876,120	PVT1	DOWN	0.002	0.21 \pm 0.02	0.16 \pm 0.04	0.11 \pm 0.02	-1.92 \pm 0.15	-2.46 \pm 0.47	-3.03 \pm 0.27
cg13351583	9	90,983,468	SHC3	DOWN	\leq 0.001	0.29 \pm 0.02	0.26 \pm 0.04	0.20 \pm 0.02	-1.28 \pm 0.13	-1.50 \pm 0.27	-2.02 \pm 0.14
cg08427977	10	49,993,838	C10orf72	DOWN	\leq 0.001	0.46 \pm 0.02	0.40 \pm 0.03	0.35 \pm 0.02	-0.24 \pm 0.11	-0.61 \pm 0.21	-0.91 \pm 0.14
cg04126866	10	85,922,743	C10orf99	DOWN	\leq 0.001	0.85 \pm 0.01	0.77 \pm 0.07	0.65 \pm 0.04	2.53 \pm 0.10	1.83 \pm 0.51	0.95 \pm 0.31
cg13435792	12	14,983,701	C12orf46	DOWN	\leq 0.001	0.75 \pm 0.03	0.62 \pm 0.07	0.54 \pm 0.04	1.59 \pm 0.23	0.74 \pm 0.43	0.21 \pm 0.24
cg12795208	12	51,132,205	KRT6B	DOWN	\leq 0.001	0.68 \pm 0.02	0.62 \pm 0.03	0.57 \pm 0.02	1.11 \pm 0.16	0.71 \pm 0.21	0.41 \pm 0.14
cg20161089	14	93,647,267	IFI27	DOWN	\leq 0.001	0.63 \pm 0.02	0.51 \pm 0.07	0.40 \pm 0.04	0.77 \pm 0.13	0.04 \pm 0.41	-0.58 \pm 0.27
cg21432842	17	35,425,223	CSF3	DOWN	0.002	0.48 \pm 0.01	0.46 \pm 0.03	0.39 \pm 0.02	-0.10 \pm 0.06	-0.22 \pm 0.19	-0.64 \pm 0.13
cg08611714	17	36,877,384	KRTHA2	DOWN	\leq 0.001	0.60 \pm 0.02	0.56 \pm 0.06	0.42 \pm 0.04	0.61 \pm 0.14	0.35 \pm 0.35	-0.46 \pm 0.26
cg25979644	18	53,550,001	ATP8B1	DOWN	\leq 0.001	0.83 \pm 0.02	0.73 \pm 0.09	0.59 \pm 0.05	2.36 \pm 0.24	1.48 \pm 0.56	0.56 \pm 0.28
cg10533434	18	59,480,206	SERPINB3	DOWN	\leq 0.001	0.71 \pm 0.03	0.67 \pm 0.07	0.51 \pm 0.05	1.30 \pm 0.18	1.04 \pm 0.46	0.09 \pm 0.29

Table A2.3 (Continue).

CpG ID	Chr	Position hg18	Gene	PP Direction	PP v. NN Adj p-value	Unnormalized β -value			Normalized M-value		
						NN	PN	PP	NN	PN	PP
cg20583073	19	3,129,759	EDG6	DOWN	0.001	0.66 ± 0.02	0.65 ± 0.02	0.59 ± 0.01	0.95 ± 0.11	0.90 ± 0.15	0.51 ± 0.08
cg19103704	19	45,132,803	FCGBP	DOWN	≤ 0.001	0.42 ± 0.03	0.33 ± 0.04	0.27 ± 0.02	-0.45 ± 0.17	-1.02 ± 0.30	-1.47 ± 0.13
cg07713361	22	34,979,090	APOL1	DOWN	0.006	0.20 ± 0.01	0.16 ± 0.02	0.14 ± 0.01	-2.02 ± 0.12	-2.45 ± 0.21	-2.68 ± 0.15
cg11075745	1	42,978,563	CLDN19	UP	0.025	0.48 ± 0.01	0.52 ± 0.01	0.54 ± 0.01	-0.09 ± 0.07	0.09 ± 0.05	0.21 ± 0.06
cg06641366	1	89,870,192	LRRRC8C	UP	0.008	0.53 ± 0.03	0.60 ± 0.06	0.73 ± 0.04	0.19 ± 0.18	0.62 ± 0.37	1.48 ± 0.27
cg17582777	1	153,316,724	EFNA3	UP	≤ 0.001	0.68 ± 0.02	0.69 ± 0.02	0.79 ± 0.02	1.06 ± 0.14	1.19 ± 0.16	1.96 ± 0.17
cg05781767	3	128,830,052	PODXL2	UP	≤ 0.001	0.30 ± 0.01	0.37 ± 0.04	0.48 ± 0.04	-1.26 ± 0.10	-0.77 ± 0.24	-0.14 ± 0.22
cg14289461	6	112,682,622	LAMA4	UP	≤ 0.001	0.19 ± 0.02	0.25 ± 0.03	0.37 ± 0.05	-2.08 ± 0.16	-1.62 ± 0.23	-0.78 ± 0.35
cg08626653	7	99,920,258	FLJ37538	UP	≤ 0.001	0.74 ± 0.02	0.80 ± 0.03	0.85 ± 0.02	1.50 ± 0.16	2.01 ± 0.24	2.51 ± 0.19
cg22441882	8	20,084,934	SLC18A1	UP	≤ 0.001	0.18 ± 0.02	0.26 ± 0.04	0.31 ± 0.03	-2.20 ± 0.18	-1.57 ± 0.30	-1.17 ± 0.23
cg14706739	8	21,972,301	EPB49	UP	≤ 0.001	0.60 ± 0.03	0.67 ± 0.06	0.78 ± 0.04	0.59 ± 0.19	1.06 ± 0.46	1.85 ± 0.30
cg18939260	8	125,810,527	MTSS1	UP	0.010	0.49 ± 0.01	0.56 ± 0.07	0.62 ± 0.03	-0.08 ± 0.07	0.38 ± 0.47	0.69 ± 0.18
cg12515371	8	145,692,804	PPP1R16A	UP	0.001	0.55 ± 0.01	0.64 ± 0.02	0.68 ± 0.03	0.29 ± 0.09	0.81 ± 0.14	1.12 ± 0.20
cg15983520	8	145,699,914	GPT	UP	≤ 0.001	0.54 ± 0.02	0.57 ± 0.03	0.63 ± 0.01	0.23 ± 0.11	0.43 ± 0.17	0.75 ± 0.09
cg00342530	9	129,066,841	GARNL3	UP	≤ 0.001	0.26 ± 0.01	0.34 ± 0.05	0.42 ± 0.05	-1.53 ± 0.11	-0.96 ± 0.32	-0.49 ± 0.29
cg03699566	11	71,578,300	FOLR1	UP	≤ 0.001	0.43 ± 0.03	0.51 ± 0.04	0.66 ± 0.04	-0.42 ± 0.16	0.07 ± 0.23	0.97 ± 0.25
cg05194726	12	2,814,741	NRIP2	UP	≤ 0.001	0.15 ± 0.01	0.28 ± 0.05	0.33 ± 0.05	-2.57 ± 0.18	-1.38 ± 0.38	-1.07 ± 0.31
cg02192520	12	54,400,422	RDH5	UP	≤ 0.001	0.26 ± 0.03	0.35 ± 0.04	0.52 ± 0.06	-1.49 ± 0.19	-0.93 ± 0.27	0.13 ± 0.38
cg03623878	13	112,703,561	MCF2L	UP	≤ 0.001	0.19 ± 0.02	0.25 ± 0.05	0.35 ± 0.04	-2.09 ± 0.17	-1.62 ± 0.36	-0.88 ± 0.23
cg11668844	13	112,703,623	MCF2L	UP	≤ 0.001	0.26 ± 0.02	0.31 ± 0.05	0.45 ± 0.04	-1.53 ± 0.14	-1.21 ± 0.32	-0.28 ± 0.21
cg09276451	16	4,361,604	SLITL2	UP	0.002	0.45 ± 0.02	0.51 ± 0.02	0.55 ± 0.02	-0.30 ± 0.13	0.08 ± 0.13	0.28 ± 0.13
cg07634706	16	55,996,025	CCL17	UP	≤ 0.001	0.75 ± 0.01	0.79 ± 0.02	0.82 ± 0.01	1.59 ± 0.11	1.90 ± 0.22	2.24 ± 0.15
cg20716119	16	67,329,264	CDH1	UP	0.005	0.08 ± 0.01	0.09 ± 0.02	0.13 ± 0.01	-3.43 ± 0.10	-3.38 ± 0.25	-2.72 ± 0.17
cg02298612	16	74,085,960	CHST6	UP	0.001	0.58 ± 0.02	0.62 ± 0.02	0.69 ± 0.03	0.49 ± 0.10	0.69 ± 0.14	1.15 ± 0.17
cg04049033	17	1,500,411	RILP	UP	0.035	0.50 ± 0.01	0.49 ± 0.03	0.56 ± 0.02	-0.02 ± 0.05	-0.05 ± 0.19	0.33 ± 0.09
cg26415633	19	56,019,168	KLK1	UP	0.020	0.34 ± 0.02	0.37 ± 0.02	0.46 ± 0.02	-0.97 ± 0.13	-0.78 ± 0.13	-0.25 ± 0.15
cg14380517	21	17,907,632	BTG3	UP	≤ 0.001	0.11 ± 0.01	0.15 ± 0.02	0.21 ± 0.03	-3.05 ± 0.23	-2.54 ± 0.23	-1.91 ± 0.26
cg13265003	21	42,792,092	SLC37A1	UP	≤ 0.001	0.60 ± 0.02	0.70 ± 0.03	0.72 ± 0.03	0.58 ± 0.09	1.20 ± 0.17	1.41 ± 0.19

Figure A2.1 Hierarchical clustering analysis of psoriatic skin samples. Heatmap showing PP, PN, and NN samples clustered using the top 50 genes that differentiate PP from NN skin. Image was generated using simple scaling normalized gene expression values. Red values indicate up-regulation while green indicates down-regulation of expression. The Affymetrix U95 microarray probeset ids and gene symbols were also shown at the right column.

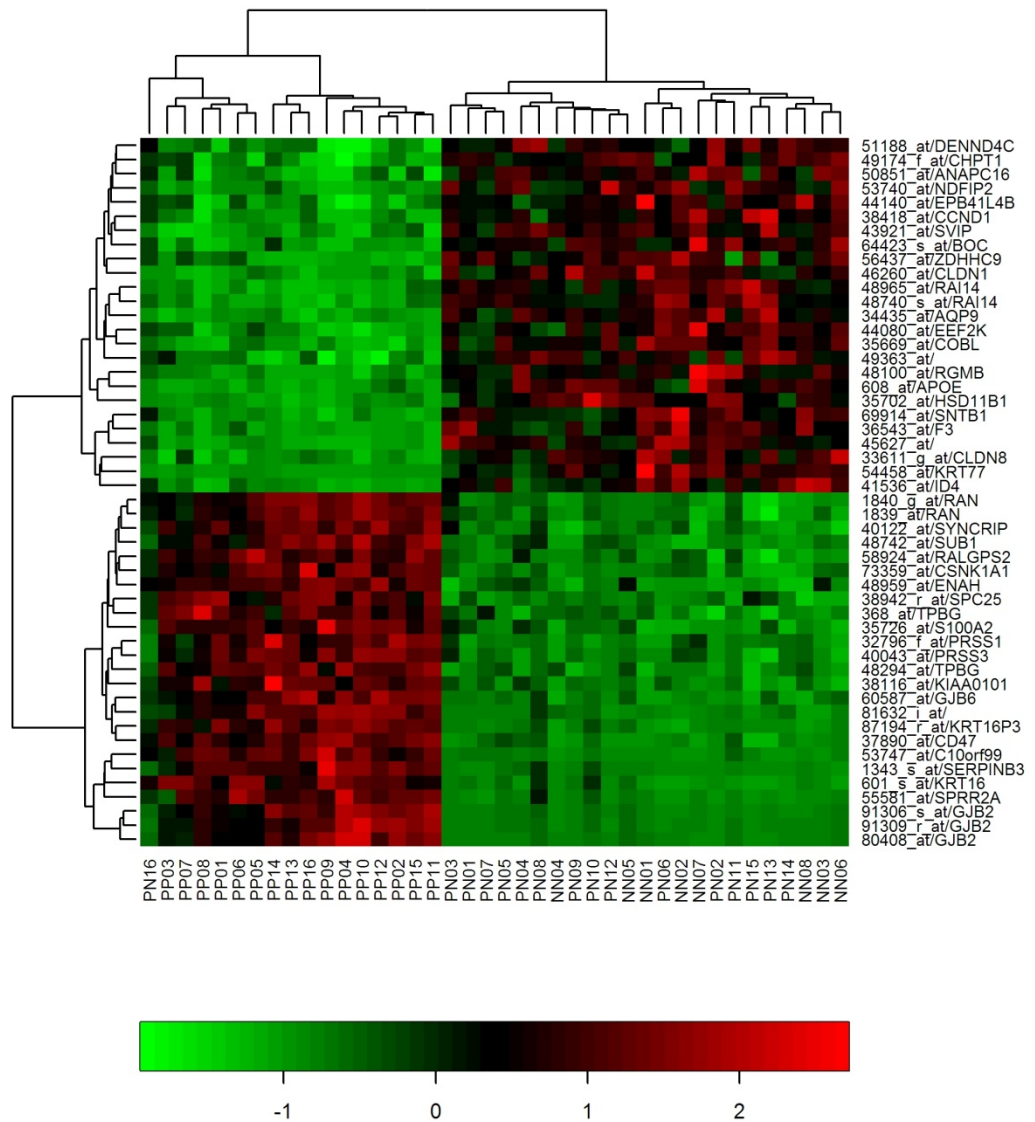


Figure A2.2 Comparison of raw and normalized distributions for β - and M-values. Each panel of the figure shows a plot of the mean rank versus standard deviation. Data distributions with no mean/variance relationship are flat on this type of plot. The transformed raw beta ($\beta_{\text{transformed}} = \text{sqrt}(\beta)$) demonstrates severe heteroskedascity, that is somewhat improved with variance stabilizing normalization, where $\beta_{\text{normalized}} = \text{asin}(\text{sqrt}(\beta))$. The raw M values demonstrate a lower degree of heteroskedascity as compared to the beta values. However, the normalized M-values (background subtracted and normalized by simple scaling normalization implemented in the methylumi package) are flatter and more evenly distributed. The greater homoskedascity for the normalized M-values makes them more appropriate for parametric statistical analysis.

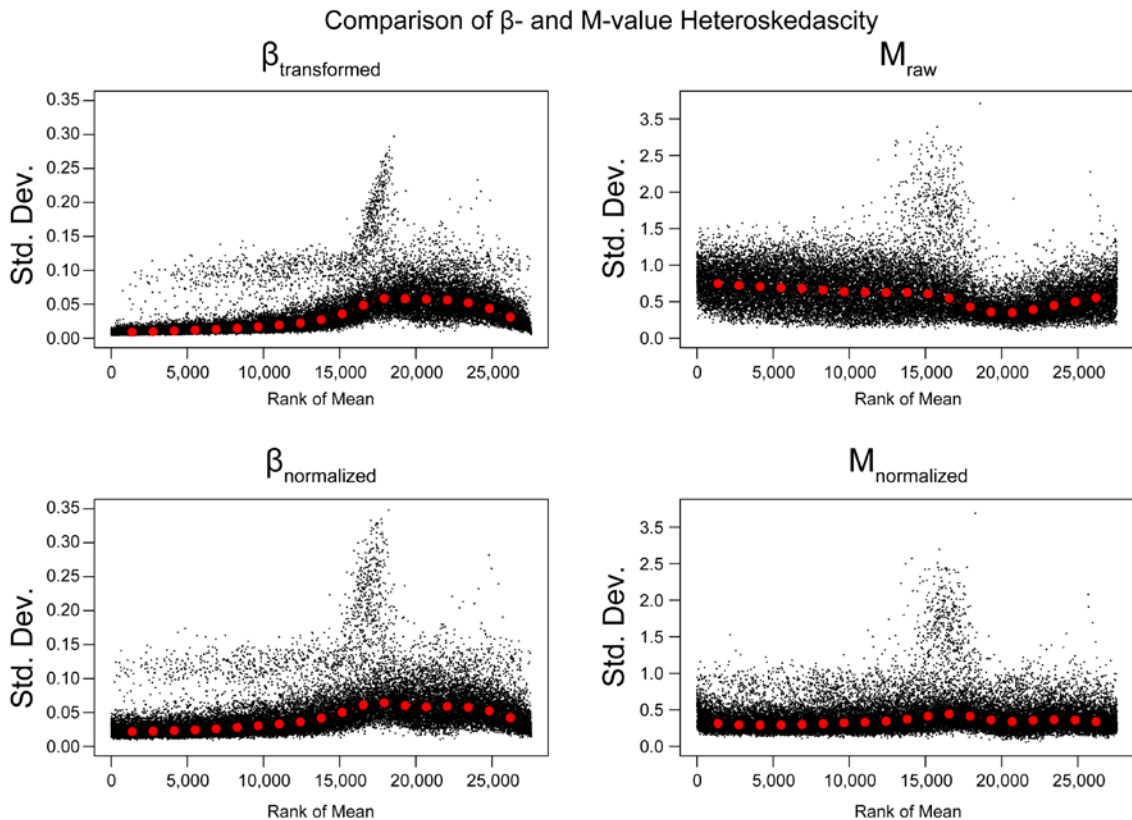


Figure A2.3 Theoretical power for detecting methylation changes. Two panels are shown. The upper panel is the power for a paired t-test, and the lower panel is the power for a two-sample t-test. Since we performed global unbiased analysis, all power calculations are for two sided t-tests. The x-axis is the true difference in methylation levels (β) between the sample groups. The y-axis is the power to detect the true difference. Three lines are shown in each plot, showing the power at 5% methylation standard deviation, 10% and 15%, represented as solid, dashed, and dotted lines. The paired t-test demonstrates great power compared to the two-sample t-test, as is expected. Large changes in methylation, on the order of 40%, are detectable even with large variance. The 5% and 10% standard deviation lines are representative of the majority of sites as few had standard deviations of raw β as larger than 10%.

