Genome-wide Analysis of Atopic Dermatitis

Annika M. Sääf, Patrick O. Brown
Stanford University

Introduction

Atopic dermatitis (AD) or eczema is a common skin disorder that usually begins in infancy or early childhood with a significant proportion of children having continued problems into adult life. Patients with AD suffer from itchy, dry and inflamed skin, sometimes in combination with evolving asthma. Twin studies indicate a strong genetic contribution in the development of atopic dermatitis (Larsen 1986, Schultz-Larsen 1993) and up to date four genome screens have been performed that have identified several chromosomal regions linked to AD (Lee 2000, Cookson 2001, Bradley 2002, Haagerup 2004). However, little is known about the role of specific genes in the pathogenesis of AD and the underlying mechanism for developing this pathological condition still remains mysterious.

Over the past decades, much research has been focused on advancing the knowledge about the role and action of immune cells and inflammatory molecules in AD pathogenesis. Infiltrates of activated T-helper cells, CLA+ lymphocytes, esinophils, macrophages and Langerhans cells are characteristics of AD skin (Leung 2001, Leung 2003), and it is well established that there is an imbalance between Th1 and Th2 cells toward an increase in production of a Th2 cytokine profile at early phases of the disease. A network of cytokines and their receptors have been identified that is characteristically expressed in patients with AD (Nomura 2003 [A], Pivarcsi 2005). A current model is that high immunoglobulin E (IgE) levels driven by Th2 cell responses account for the pathogenesis of AD. However, it is still not clear whether the inflammatory response found in AD skin is the basic cause of AD, or if it is a secondary effect caused by other disease triggering factors.

A new genome-wide approach using the so-called DNA-chip technology has recently been undertaken in the search for new candidate genes involved in the pathogenesis of atopic dermatitis. Nomura et al (2003 [B]) used the Affymetrix gene chip to search for distinct gene-expression patterns between AD skin lesions and psoriasis. Nel-like 2, PPP1R5 and members of the CCL family (CCL18, CCL27, CCL13) are among the genes identified in this report to be significantly higher expressed in AD skin as compared to psoriasis. Among the genes identified over-expressed in psoriasis skin were genes encoding members of the cornified epithelial skin layer such as small proline rich protein 2C, S100A12, and members of the kallikrein protein family (KLK6, KLK13).

Sugiura et. al. (2005) compared gene-expression profiles in AD skin to healthy control skin and found that a set of so called “epidermal differentiation genes” involved in forming a protective cornified envelope structures (S100A7, S100A8, loricrin, filaggrin) are significantly differentially expressed in AD skin as compared to skin from healthy individuals. We here report the first genome-wide study where expression profiles of all human genes at once were analyzed in AD skin as compared to the global gene signature

in healthy control skin (Sääf et al. manuscript in preparation). We used human cDNA microarrays containing ~42,000 elements that represent approximately 24,500 unique genes (based on Unigene clusters) to identify a molecular picture of the programmed responses of the human genome to the pathological condition of AD. Gene-expression patterns in skin (involved and non-involved) from AD patients were compared to the gene-expression signature identified in healthy non-atopic skin. In parallel, we analyzed the transcriptional responses to *M. sympodialis* patch-test both in AD and healthy control skin. Complementary techniques (immunohisto chemistry and in situ hybridization) were used to identify *in vivo* expression in skin of potential AD susceptibility genes that were identified by the microarray technique. We also performed a genetic association study to further analyze one of the potential AD susceptibility genes (*SOCS3*) identified in our microarray study (Ekelund et al. manuscript in preparation).

Results

Global portrait of the gene-signature in atopic dermatitis skin

We used cDNA microarrays representing approximately 24,500 unique genes (based on Unigene clusters) to identify a detailed molecular picture of the programmed responses of the human genome to the pathological condition of atopic dermatitis. We also asked the question weather *Malassezia sympodiali* patch-tested skin is similar to lesional AD skin on a transcriptional level. The *M. sympodialis* patch-test gives an allergic reaction in the skin from AD patients similar to that observed in lesional skin. Members of the *Malassezia* fungi family are suggested to be a potential triggering factor in developing AD. However, nothing is known about the underlying transcriptional response to *M. sympodialis* in the skin.

Skin biopsy samples were collected from two groups of individuals; AD patients and non-atopic individuals (controls). Three different skin biopsies were taken from each individual; (1) lesional skin from AD patients (or healthy skin from control individuals), (2) *M. sympodialis* treated and (3) PBS patch-tested skin. Fluorescently labeled cDNA was hybridized to microarrays in a two-color comparative format; AD patient- or healthy control samples were labeled with Cy-5, and a reference pool of human mRNAs (Stratagene) was labeled with Cy-3. The abundance of each transcript measured in a skin biopsy sample relative to the common reference is represented by color; red for a relative increase in expression and green for a relative decrease in expression. We next used a Significance Analysis of Microarrays (SAM) analysis approach (Tusher 2001) to select a set of ~4,000 genes, consistently differentially expressed between AD skin and skin from healthy control individuals. A hierarchical clustering method was used to group these genes based on similarity in expression across the samples and to group experiments on the basis of similarities in gene-expression patterns. The clustering analysis clearly separates AD- and normal healthy skin samples into two distinct groups demonstrating that there is a complex variation in gene-expression patterns across the two groups. The most prominent feature of the AD gene-signature is a relative high level of expression of a group of genes involved in inflammation, and relatively low expression of a group of genes involved in synthesis and metabolism of fatty acids and sterols (like cholesterol). We also identified a set of novel genes involved in epithelial differentiation and cornified

envelope assembly to be significantly differently expressed in AD and non-atopic control skin. A set of genes was selected for further analysis by immunohistochemistry and *in situ* hybridization techniques to confirm and explore corresponding protein and RNA levels *in vivo*.

Induced expression of inflammatory genes in AD skin

The most prominent feature of up-regulated genes in AD skin as compared to healthy control skin is genes encoding immune and inflammatory genes. Among these genes are chemokine family members (CCL18, CCL21, CCL2, CXCL1, CXCL3) and genes encoding interleukin 32 (NK4) and members of the interleukin receptor family including IL2R(γ), IL4R, IL6R, IL7R, IL13R(α2). We identified higher relative amounts of the IgE receptor, gamma subunit (FCER1G) transcript in AD skin as compared with healthy skin, which supports the idea that high immunoglobulin E (IgE) levels play a role in AD pathogenesis. Genes encoding cell surface antigens, such as CD28, CD 37 and CD53, are also among the genes over-expressed in AD skin. CD 37 and CD53 are known to interact with integrins and may play an important role in T-cell-B-cell interactions. Notably is that members of the integrin family (ITGAX and ITGB2) show very similar expression pattern as the CD family members, across all samples. Familial deficiency of the *CD53* gene has previously been linked to an immunodeficiency associated with recurrent infectious diseases caused by bacteria, fungi and viruses. CD28 costimulation is essential for CD4-positive T-cell proliferation, survival, and interleukin-2 production, and T-helper type-2 (Th2) development. High relative amounts *CD28* in AD skin support the idea of AD being a Th2 type skin disease. Another sign of ongoing immune response in AD skin is the expression of *CD3* (zeta) gene, which plays an important role in coupling antigen recognition to several intracellular signal-transduction pathways.

A tight cluster of coordinately over-expressed genes encoding complement components (C1R, C1S and C1QB) is found in AD skin. We also identified genes encoding a complement regulator (SERPING1) and activator (CFHL1) to be up-regulated in AD skin compared to healthy skin. Complement is an essential component of the immune system and is of relevance for the destruction of invading microorganisms. However, excessive or uncontrolled complement activation contributes to undesired tissue damage and the role of complement in many inflammatory diseases has been reported (Szebeni 2004). Our genome-wide analysis approach of AD skin have confirmed expression patterns of several immune genes previously suggested in AD pathogenesis including chemokines like CCL18, and the interleukin 4- and 6 receptors. Among the immune genes presented in our microarray study are also several novel immune genes that may play a critical role in AD pathogenesis including the chemokine CCL21, CD family members (CD37, CD53) and SOCS3 (further discussed below). Notably is that immune response genes in general were found more strongly expressed in lesional and *M. sympodialis* patch-tested AD skin as compared to non-involved AD skin and normal healthy skin. However there are examples of infammatory genes, including CCL18 and CCL21 that are expressed in significant relative amounts also in non-involved AD skin.

Altered Lipogenic gene-expression program in AD skin

The most prominent feature of the AD down-regulated gene cluster is a set of genes with a well-defined role in lipid homeostasis including genes that are dedicated to the

synthesis and uptake of cholesterol and fatty acids such as ATP-Citrate lyase, Acetyl-CoA synthase, HMG-CoA synthetase and HMG-CoA reductase. Presented in the same cluster are genes encoding key enzymes in the pathway for poly unsaturated fatty acid (PUFA) metabolism, e.g. FADS1, FADS and ELOVL5. Biological active PUFA derivative lipids such as leukotrienes and prostaglandins are known mediators involved in inflammation. Recently, the PUFA lipids have also been designed an important role in regulating gene-transcription (reviewed in Sampath 2005).
We found lower transcriptional levels of the *Lipin-1* gene in AD skin as compared to healthy control skin. Lipin-1 encodes a transcription factor with a suggested role during normal adipose tissue development. It was previously demonstrated in mouse that lipin deficiency impairs adipocyte differentiation and causes lipodystrophy. We also found hydroxy steroid dehydrogenase family members and members of the cytochrome P450 family that are essential for steroid hormone biosynthesis (Payne 2004) to be down-regulated in AD skin.
Lipid molecules serve various biological roles in the cell, from being a source of energy to have a structural role in the cell membrane, and they can act as signaling molecules in various biological processes. The coordinately down-regulation of a large set of lipogenic genes found in AD skin could therefore have sever effects on numbers of different biological processes and be of critical importance in AD pathogenesis.

Altered expression of genes encoding structural components and enzymes involved in epithelial cell differentiation and cornified envelope assembly in atopic dermatitis skin

We found several genes up-regulated in AD skin that are involved in keratinocyte differentiation and cornified envelope assembly, including members of the S100 (S100A4, S100A9, S100A10 and S100A12) and small proline-rich protein (SPRR1B, SPRR2A, SPRR2G and SPRR4) family. Corneodesmosin (CDSN), a late differentiation epidermal glycoprotein putatively involved in keratinocyte adhesion, is one among the most significantly differentially expressed genes in lesional AD skin compared to healthy normal skin. Coordinately over-expressed in lesional AD skin is Calmodulin-like 5 (CALML5), which encodes a calcium-binding epidermis protein that directly related to keratinocyte differentiation. The cornified envelope is a covalently cross-linked structure that forms beneath the plasma membrane in differentiating keratinocytes. We identified two key enzymes (TGM1 and TGM3), involved in forming the advanced network of cross-linked proteins in the cornified envelope, to be significantly highly expressed in lesional AD skin compared to healthy normal skin. Transglutaminase 1 is localized to 14q11, a chromosomal region previously inked to AD, which makes this gene an interesting AD susceptibility gene. Several of the genes involved in keratinocyte differentiation, including S100- and SPRR family members, are localized to the epidermal differentiating complex on 1q21, a chromosomal region also previously linked to AD. Notably is that many of the cornified envelope transcripts we identified over-expressed in lesion AD skin is also significantly over-expressed in non-involved AD skin, but less expressed in the *M.sympodialis* treated skin where the inflammation signature is more dominant. Our data suggest that AD patients may have a predisposed genetic program that could result in altered regulation and construction of the protective

epithelial layer of the skin. These results may help explaining why AD patients are hyper sensitive to allergens that are innocuous to healthy individuals. The precise role that keratinocytes and epithelial barrier dysfunctions may have in AD pathogenesis remains to be further investigated.

Genetic Association Links the SOCS3 Gene to Atopic Dermatitis

SOCS3, located in a candidate AD region (17q25), was among the genes we found significantly more highly expressed in lesional and *M.sympodialis* treated skin from AD patients as compared to healthy individuals. SOCS3 protein levels were also found elevated in AD skin compared to healthy control individuals. Higher expression SOCS3 protein was identified in dendritic cells by using the IHC technique, and an increased number of SOCS3-positive cells were found in epidermis from lesional AD skin compared to non-lesional AD skin and normal skin from healthy controls

SOCS3 is a negative regulator of cytokine signaling. It has recently been demonstrated that this gene is predominantly expressed in T-helper type 2 cells and has an important role in regulating the onset and maintenance of T_H2-mediated responses (Seki 2003). To further study the genetics of *SOCS3* in AD patients, a set of 409 Swedish families with at least one sibling with AD was analyzed (Ekelund et al. manuscript in preparation). Eight polymorphic SNPs were genotyped in the 3,5 kB *SOCS3* gene. Two of the SNPs were significantly associated with the AD phenotype (rs12952093 $p<0.01$and rs4969170 $p<0.007$). By using the HapBlock program we found that all the analyzed SNPs fell into one haplotype block. We then identified three haplotype tagging SNPs as representing this haplotype block. Two of them were in the promoter region of the gene (rs12952093 and rs4969170) and the third was in 3'UTR (rs4969168). To confirm the genetic association we analyzed the three haplotype tagging SNPs in two independent sets of patients and controls. The first group of patients was 555 children up to four years of age (328 affected with AD and 227 controls) and the second was a case-control study, comprising 187 adult AD patients and 230 age and sex matched controls from the UK. In our analysis of the Swedish case-control samples, we found a highly significant genetic association between *SOCS3* and the phenotype AD ($p<0.003$ for the positively associated haplotype and $p<0.001$ for the negatively). However, we could not confirm the association in the UK study. The lack of association in the latter UK study could be due to differences between the two populations in the relative contribution of *SOCS3* as a genetic risk factor. Another potential explanation for the discrepancy is that the importance of *SOCS3* as a risk factor might be age-dependent, since the most significant association of SOCS3 was seen in children with AD.

In conclusion, we have found that *SOCS3* is consistently more highly expressed in skin from AD patients than in normal control skin and that specific haplotypes of the *SOCS3* gene are significantly associated with the disease. *SOCS3* may play an important role in AD pathogenesis, at least for a Swedish population with AD and it may be an important target for future investigations of the molecular pathogenesis of AD and for therapeutic development.

Summary

We used the DNA microarray technique to provide a detailed molecular picture of the programmed responses of the human genome to the pathological condition of atopic dermatitis. The transcriptional program in AD skin was compared to that of healthy control individuals. We also compared the global transcriptional response identified in lesional AD skin to the gene-signature in M. sympodialis patch-test skin ("induced eczema") and found that M. sympodialis treated skin is amazingly similar to authentic AD skin lesions on a transcriptional level. Several potential susceptibility genes are identified in this study that may play a critical role in the pathological condition of atopic dermatitis. Many of these genes are localized on chromosomal regions previously associated with AD. The most prominent feature of the gene-signature identified in AD skin is a reciprocal expression of two large clusters of genes representing inflammatory genes and genes involved in lipid homeostasis. We also identified a set of novel genes involved in epithelial differentiation and cornified envelope assembly to be among the genes over-expressed in AD skin as compared to non-atopic skin. Interestingly, many of the genes that were found differentially expressed in lesional AD skin as compared to healthy control skin was also found to have corresponding elevated expression profiles in non-involved AD skin, which support the hypothesis that AD patients have a predisposed genetic program that potentially make them sensitive to allergens and other disease triggering factors in the environment. These results and all the original data will be made publicly available with free, unrestricted, open access.

SOCS3, which is among the over-expressed genes in AD skin and localized in a chromosomal AD region (17q25), was further analyzed in a genetic association study. It was recently demonstrated that *SOCS3* is predominantly expressed in T-helper type 2 cells and has an important role in regulating the onset and maintenance of T_H2-mediated responses (Seki 2003). We found that specific haplotypes of the *SOCS3* gene are significantly associated with atopic dermatitis in two independent groups of patients ($p<0.03$ and $p<0.003$ respectively) (Ekelund et al. manuscript in preparation).

Literature Cited

Bradley 2002, Hum Mol Genet 2002, 11:1539-1548
Cookson 2001, Nat Genet 27:372-373
Ekelund et al. 2006, manuscript in preparation
Haagerup 2004, Allergy 59:88-94
Larsen 1986, J Am Acad Dermatol 15:487-494
Lee 2000, Nat Genet 26:470-473
Leung 2001 J Am Acd Derm 44:S1-S12
Leung 2003 Lancet 361:151-160
Nomura 2003 (A), J Immunol 171:3262-3269
Nomura 2003(B), J Allergy Clin Immunol 112:1195-1202
Payne 2004, Nedocrine Rev 25 (6) 947-970
Pivarcsi 2005, Curr Allergy and Asthma Rep, 5:284-290
Sampath 2005, AnnuRev Nutr, 25:317-40
Schultz-Larsen 1993, J Am Acad Dermatol 28 :719-723
Seki 2003, Nature Medicine 9:1047-1054
Sugiura 2005, Brit J Dermatology 152:146-149
Szebeni 2004, Kluver Acad Publicer, Massachusetts
Sääf et al. 2006, manuscript in preparation
Tusher 2001, PNAS 98(9):5116-21