Genetics in Atopic Dermatitis: Historical Perspective and Future Prospects

Sara J. BROWN1, Martina S. ELIAS1 and Maria BRADLEY2
1Skin Research Group, Division of Molecular and Clinical Medicine, School of Medicine, University of Dundee, Dundee, Scotland, and 2Dermatology and Venereology Unit, Department of Medicine Solna, Karolinska Institute, Stockholm, Sweden

Atopic dermatitis (AD) is a common, complex trait, arising from the interplay of multiple genetic and environmental factors. This review provides an overview of developments in the field of AD genetics. AD shows high heritability; strategies to investigate genetic risk include linkage, candidate gene studies, genome-wide association and animal modelling. Loss-of-function mutations in FLG, encoding the skin barrier protein filaggrin, remain the strongest genetic risk factor identified for AD, but variants influencing skin and systemic immune function are also important. AD is at the forefront of genetic research, from large-scale population studies to in vitro models and detailed molecular analyses. An understanding of genetic risk factors has considerably improved knowledge of mechanisms leading to atopic skin inflammation. Together this work has identified avenues for therapeutic intervention, but further research is needed to fully realise the opportunities of personalised medicine for this complex disease, to optimise patient benefit.

Key words: atopic dermatitis; eczema; filaggrin; genetic; genome-wide; risk; phenotype; transcriptome.

Accepted May 7, 2020; Epub ahead of print May 15, 2020
Acta Derm Venereol 2020; 100: adv00163.

Cor: Sara J Brown, Professor of Molecular and Genetic Dermatology & Wellcome Trust Senior Research Fellow in Clinical Science, Skin Research Group, Division of Molecular and Clinical Medicine, School of Medicine, University of Dundee, Dundee DD1 9SY, Scotland, UK. E-mail: s.j.brown@dundee.ac.uk

What have we learned about atopic dermatitis genetics over the past 100 years?

Heritability of AD: family and twin studies

It can clearly be observed that atopic diseases show a familial aggregation, with clustering of affected individuals within families, demonstrating the importance of genetic heritability. The term ‘heritability’ refers to the proportion of variation within a clinical feature that is attributable to genetic factors (5). A family history of atopic diseases, in particular AD, is the strongest of all risk factors. The presence of any atopic disease in one parent is estimated to increase a child’s risk of developing AD 1.5-fold, whereas the risk is increased ~3-fold and ~5-fold, respectively, if one or both parents have AD (6, 7). Familial aggregation can be due to shared environment and/or shared genes and a way to address the genetic component is to study twins. These studies have shown a concordance rate of 72–86% in monozygotic twins and 21–23% in dizygotic twins (8, 9). These data demonstrate that the genetic contribution to the development of AD is substantial and this heri-
tability has been estimated at 70–80% (10, 11) – a high heritability for a complex trait (12). For comparison, psoriasis heritability is approximately 68% (13) whilst other inflammatory barrier diseases show heritability of 7–38% for periodontitis (14) and approximately 67% in ulcerative colitis (15).

**Strategies for the investigation of genetic risk**

Various different strategies have been used to study genetic components in complex diseases such as AD. In broad genomic analyses (genome-wide linkage, genome-wide association studies) a pre-existing knowledge of the function of genes is not required, nor the biology of the trait in question; it is a ‘hypothesis-free’ approach. In contrast, directed genetic analysis such as a candidate gene approach is a strategy in which certain loci or genes considered to be of interest for the phenotype are selected for study. The selection can be based on earlier studies, “educated guesses” or knowledge of the pathogenesis and function of previously identified genes or loci; this is a ‘hypothesis-driven’ approach. Each of these strategies has been used to provide insight into AD.

**Linkage studies**

Genetic linkage is a method for mapping genes. It exploits the fact that a marker (often a microsatellite marker such as repeated DNA sequences, mostly di-, tri-, and tetra-nucleotide repeats) show variation between individuals. Informative markers have many alleles and are distributed at known locations throughout the genome. The first genome-wide study in AD identified a major susceptibility locus on chromosome 3q21 (16). During the following years, additional genome-wide studies in AD were performed and several more loci were identified including 1q21.3p, 17q, 18q, 11.13q. However, these loci were often too wide and they required labour intensive fine-mapping. Genome-wide association studies (GWAS) have subsequently replaced genome-wide linkage (17). The technique of GWAS is described in more detail below.

**Candidate genes**

*Filaggrin (FLG).* Using a candidate gene approach, and the link between ichthyosis vulgaris and AD, the *FLG* gene was identified as a susceptibility gene for AD in 2006 (18). This was a major breakthrough and also established the impaired skin barrier function as having a key role in the development of AD. Filaggrin is involved in the development of keratinocytes to maintain epidermal integrity and it is an important marker of keratinocyte differentiation. During keratinocyte differentiation, profilaggrin is dephosphorylated and degraded into monomers, which condense in the cytoskeleton of keratin to form an intensive protein-lipid matrix. Consequently, these filaggrin monomers are degraded into amino acids, which contribute to the natural moisturising factors, maintaining skin hydration, a low pH and other aspect of the barrier function of the stratum corneum (Fig. 1).

Loss-of-function mutations in *FLG* are present in up to 10% in the Northern European population. They cause the common monogenetic dry skin disorder ichthyosis vulgaris. The most common loss-of-function mutations in Europe are R501X, 2282del4, R2447X and S3247X. Together these 4 null mutations account for >90% of null mutations in the population (21). Among European patients with moderate to severe AD up to 40% of the patients carry a *FLG* null mutation. In meta-analysis the risk of getting AD in a mutation carrier is increased 3-fold.

![Fig. 1. Filaggrin expression and processing in the epidermis.](image)
Additionally, filaggrin deficiency may confer an evolutionary advantage in higher latitudes (i.e. Northern Europe) through its role in increasing vitamin D biosynthesis. Vitamin D3 levels are 10% higher in German and Danish individuals with FLG null mutations, which may be due to a reduction in filaggrin’s role as an endogenous UVB filter in the skin (37).

Besides mutations there are intragenic repetitive gene sequences or ‘copy number variations’ in FLG that determine the amount of filaggrin monomer expressed in the skin. Having more repeats (12 compared to 10 on each allele) is associated with reduced risk of AD (24) by a dose-dependent effect within this repetitive gene sequence. The effects of cytokines, such as IL-4, IL-13, IL-17A, IL-22, IL-25, IL-31, and TNF-α have also been shown to suppress filaggrin expression in the skin, resulting in additional barrier impairment (39, 40).

Even though FLG mutations and the filaggrin protein are extremely important in AD pathogenesis, there must be yet unknown, additional factors genes or functions of gene involved in AD development that still to be found.

Some other candidate genes in atopic dermatitis. Other genes that has been detected through a candidate gene approach, supported by knowledge of AD biology, and replicated by GWAS are genes involved in the Th2 immune response, for instance IL-4, IL-13, IL-17A, IL-22, IL-25, IL-31, and TNF-α have also been shown to suppress filaggrin expression in the skin, resulting in additional barrier impairment (39, 40).

Moreover candidate genes have been detected through the study of monogenic diseases that have features that resemble AD. Netherton syndrome (OMIM #256500) is a rare monogenic disease with AD-like lesions in the skin and increased IgE levels. The gene mutation underlying Netherton is in the Serine Protease Inhibitor Kazal-Type 5 gene (SPINK5) located on chromosome 5q32. SPINK5 encodes a 15-domain protease inhibitor Lymphoepithelial Kazal-Type-Related Inhibitor (LEKTI) which is expressed in epithelial and mucosal surfaces and in the thymus. In several studies, there has been an association between SPINK5 variants and AD, also in different populations (42–45). Other candidate genes will be studied as a result of new approaches to assessing monogenic disorders and extreme phenotypes, as discussed below.

Animal models

Animal models have the advantages that one can more easily control the environment and create genetic homogeneity. Apart from humans, dogs have spontaneous AD that has been studied and documented (46).

There are also several AD mouse models that have been described and generated over the years, each focusing on one or more aspects of human AD. The mouse models can be divided into 3 main categories: (i) Inbred strains of mice that develop AD-like phenotypes. The most well-known of these are the flaky tail mouse and the NC/Nga mouse (47, 48). The Flaky tail (ft) recessive mouse mutation arose spontaneously on the background of a recessive matted (ma) trait (49). The ft mutation has been identified as a 1-bp deletion in the Flg gene resulting in a premature stop codon (50), analogous to the human FLG mutations. More recently the ma trait has been separated from the flaky tail mouse and identified as a nonsense mutation in the novel gene Matt encoding the protein mattrin which is also postulated to have a role in skin barrier biology (51). (ii) Genetically engineered models, in which genes can be silenced or be overexpressed, for example the claudin-1 (52) and Flg knockout mice (53). (iii) Models that can be induced by exogenous agents with for example the allergens ovalbumin and house dust mite (as recently reviewed (54)).

ATOPIC DERMATITIS IS AT THE FOREFRONT OF CURRENT GENETIC TECHNOLOGY AND ITS APPLICATION

The prevalence of AD and the accessibility of disease-relevant tissues – both skin and blood – has allowed AD research to be at the forefront of applying new technologies. This has been powerfully facilitated by the active collaboration of large consortia across Europe and throughout the world. The advance of genetic and genomic analysis techniques has occurred at a rapid pace over recent decades. Large-scale and more focused molecular analysis techniques provide complimentary information; an overview of these approaches is given in Fig. 2 and each is described below.
GWAS, PheWAS, WGS & WES: largescale population and DNA analysis

Fig. 2. Complimentary strategies for genetic analysis leading to therapy development. GWAS, genome-wide association study; PheWAS, phenome-wide association study; WGS, whole genome sequencing; WES, whole exome sequencing; AD: atopic dermatitis.

Genome-wide association studies

GWAS is a technique in which very large numbers of single nucleotide polymorphisms across the genome are compared between large numbers of cases and controls, to identify differences that are associated with disease status. GWAS have been conducted in several different populations worldwide, and a recent meta-analysis has synthesized these studies (55). Over 30 loci (regions of DNA) have been identified as showing association with AD risk. Some loci include well established genetic effects, such as the epidermal differentiation complex on chromosome 1q21.3 (which includes FLG) and the cytokine cluster on chromosome 5. Many of the other regions are between genes, meaning that their functions require detailed follow-up work to ascertain a functional mechanism. One example is the region on chromosome 11q13.5 which interacts with a gene, EMSY, >30 kilobases away; EMSY has recently been shown to have an effect on skin barrier formation and function of relevance to AD (56). Another gene, LRRC32, >60 kilobases away from the same locus on chromosome 11q13.5, may also play a role in AD pathogenesis (57), demonstrating the pleiotropic effects that arise from genetic variation.

Further, larger, meta-GWAS studies are on-going, because larger sample sizes allow the detection of additional risk loci, although their effect sizes are likely to be smaller.

Phenome-wide association studies

Phenome-wide association (PheWAS) is a technique in which large numbers of phenotypic traits are tested for association with single genetic variants. For example, a loss of function variant in FLG shows strong association with atopic phenotypes including AD, asthma, allergic rhinitis and food allergy in a PheWAS study, as expected (58). Unexpected or previously unknown associations with genotypes may be revealed using PheWAS and the technique may also be applied to drug repositioning (59).

Whole exome sequencing and whole genome sequencing

Whole exome sequencing (WES) is a technique that studies the genetic sequence of the DNA in exons that code for proteins, and also exonic regions in non-coding RNAs. WES focuses on exons because they are most likely to have a direct functional effect; however, each variant requires careful assessment to define which may lead to loss-of-function or other functional effect.

WES in 22 Ethiopian people with AD and ichthyosis vulgaris has revealed rare variants in FLG and several other genes within the epidermal differentiation complex, as well as nonsense and missense mutations in previously unreported candidate genes including GTF2H5, ADAM33, EVPL and NLRP1 (60). Some of these findings indicate population-specific variation rather than disease-associated variants. There was no evidence of recurrently-mutated causal genes in this population and AD appears to show considerable heterogeneity in genetic susceptibility (60).

Whole genome sequencing (WGS) sequences intergenic regions as well as exons, because many of the regulatory mechanisms are situated in intergenic DNA. WGS generates more data and is potentially more powerful than WES, but the interpretation of non-coding variants on a large scale remains very challenging as their functional effects are not well defined. The cost of WGS is also a limiting factor to sample size and to date no large WGS have been reported in AD.

Epigenetic studies

‘Epigenetic’ refers to heritable changes in gene expression that occur without alteration to the DNA sequence. In the context of AD, there are multiple environmental and pathophysiological effects which could impact on skin cells via epigenetic mechanisms, ranging from ma-
ternal factors \textit{in utero}, to early life exposures, irritant and allergic effects. Two important epigenetic mechanisms are histone modifications and DNA methylation. These regulate chromatin structure and DNA accessibility to transcription factors and polymerases (61). Specific histone modifications can be used to predict and delineate regulatory features such as promoters and enhancers in the genome. Epigenetic mechanisms are central to the precise control of skin development and homeostasis (reviewed (62)). A number of studies have linked abnormal epigenetic control of the immune system and skin barrier to AD pathogenesis (63). Key differences in DNA methylation are observed between lesional and non-lesional AD epidermis and these correlate with changes in the expression of skin barrier and innate immune genes (64). Non-coding RNA including micro RNAs (miRNAs) confer an additional level of epigenetic control by regulating mRNA translation or degradation. Differential expression of number of miRNAs has been reported in lesional AD skin (63). Considerable further work is needed to fully understand epigenetic control in AD.

Three-dimensional DNA analyses

DNA may be represented diagrammatically as if it were a straight linear molecule, but in vivo it is extensively folded and wrapped around protein structures in three-dimensional space. Due to this folding, genomic regions that are far from each other in the linear DNA are brought in close proximity in the 3D genome (65). This complex and dynamic process facilitates long range control of gene expression by bringing distant promoter and enhancer elements together (66). Recent technological advances including chromosomal conformation capture (5C) and Hi-C or Hi-Cap, have allowed these interacting regions to be delineated. The techniques crosslink DNA with formaldehyde prior to digestion and sequencing so that interacting regions are sequenced together (65, 67). HiCap uses probes to capture promoters across the genome and regions important in gene regulation such as enhancers. Then, selected promoter–enhancer interactions can be sequenced. This analysis is performed in different cell lines and at different timepoints to reveal the dynamic process and identify candidate genes (68). Importantly, since the 3D interactions are cell-type as well as cell-state-specific, Hi-C analysis has been applied to differentiating keratinocytes, to characterise spacial control of promoter-enhancer interactions likely to be of relevance to AD (56, 67).

Transcriptome analysis

Transcriptomic analysis describes the study of RNA molecules that are present in a cell or tissue, having recently been transcribed from DNA. These molecules include protein-coding messenger RNA (mRNA), ribosomal RNA, transfer RNA, long non-coding RNA (lncRNA), micro RNA (miRNA) and others; their half-lives range from seconds to minutes. The transcriptome is a highly dynamic system and it is cell-type and cell state-dependent; differentiated cells show different gene expression compared to undifferentiated cells. Transcriptomic analysis performed on skin itself is most relevant for dermatological conditions, but transcriptomics of serum or blood may also provide valuable insight into skin-related inflammatory conditions, including AD. Transcriptomic analyses are very sensitive; skin biopsy samples from so-called ‘non-lesional’ (clinically uninflamed) skin from an AD patient show profound abnormalities in the transcriptome, including barrier impairment, dysregulation of lipid metabolism and an activated stress response (69). The AD lesional skin transcriptome shows a disease signature (70) that improves after treatment (71).

Single cell analysis

Most of the molecular analyses on skin to date have been carried out using whole skin biopsies, or epidermal samples. However single cell analysis is now feasible, for DNA and RNA sequencing, as well as protein analysis (72). These techniques offer the prospect to study individual cells, define new cell types and gain insight into the functional and structural heterogeneity of skin as a complex organ. The Human Cell Atlas is an international collaboration to make single cell analytical data available to researchers (73) and the skin component of this atlas is eagerly awaited. Several research laboratories have already released published data and tools to allow the interrogation of skin transcriptome analysis, for example murine data from the Kasper lab (74).

CRISPR-cas9 gene editing

CRISPR (clustered regularly interspaced palindromic repeat) sequences are found in bacterial DNA and form part of their immune response to phage infection. Cas9 (CRISPR-associated protein 9) is an enzyme that cleaves DNA selectively at sequences containing the CRISPR motif. In 2012 it was reported that this mechanism can be exploited for genetic engineering; guide-RNAs are used to direct the cas9 enzyme to cleave DNA in precisely-targeted editing. Application of CRISPR-cas9 allows the effects of genetic variation to be tested directly and the technique has revolutionised molecular biology. This cost-effective and relatively easy-to-use technology has allowed researchers to precisely and efficiently target, edit, modify and mark genomic loci in a wide range of cells and organisms (75). Within dermatology, CRISPR-cas9 editing has been used to correct the genetic defects in several forms of epidermolysis bullosa and of relevance to AD, the technique can be used to investigate candidate genes \textit{in vitro} (see below).
Functional analyses in vivo

Clinical observation followed-up with genetic analysis has increased our understanding of severe phenotypes which include features of AD. Following on from Netherton syndrome, these ‘human knock-out’ models include CARD11 mutations (causing systemic atopic inflammation), DSG1 and DSP mutations (causing severe dermatitis, multiple allergies and metabolic wasting) and various immunodeficiency syndromes with AD-like skin inflammation (such as Wiskott-Aldrich, caused by mutations in WAS) (76).

Functional analysis of the skin of AD patients in vivo also offers opportunities to gain understanding of the pathophysiology. Transepidermal water loss (TEWL) (77) measures the ‘inside-to-outside’ barrier function and in vivo it is proportional to skin inflammation; capacitance or conductance of the stratum corneum give a quantitative measure of water content; and tape-stripping can be used as a relatively non-invasive methods for sampling the skin transcriptome, proteome and lipids of relevance to AD (78).

Organotypic models of atopic dermatitis

Three-dimensional organotypic models of human skin bridge the gap between cultured cells in monolayer and animal models. Multi-layered organotypic models recapitulate many features of human epidermis including: morphology, spatiotemporal expression of terminal differentiation/proliferative markers and an appropriate complement of epidermal lipids (79, 80). Several organotypic models of AD have been described which generally use one of two basic approaches: the first involves the treatment of organotypic models derived from normal healthy cells with AD-relevant cytokines and the second models FLG deficient AD through gene silencing or the use of FLG-mutant keratinocytes (81). Th2 cytokines (IL-4 and IL-13) stimulate a spongiotic epidermal morphology, similar to that observed in AD (82). Organotypic models deficient in filaggrin expression broadly recapitulate many of the structural, molecular and functional defects observed in AD skin. These include a lack of keratohyalin granules, increased paracellular permeability (83, 84) and protein expression signatures consistent with AD skin (85, 86). Filaggrin deficient organotypic skin, therefore, mirrors many changes observed in the AD skin and thus represents a useful model for the study of AD disease mechanisms and therapeutic options.

Organotypic models allow the investigation of tissue-specific genetic effects and the opportunity for testing other AD candidate genes, by knockdown, over-expression, or CRISPR-cas9 editing of genes of interest.

Functional analyses in vitro

Organotypic skin models grown at the air liquid interface develop a competent bidirectional epidermal barrier with similar biophysical properties to human skin. They offer the advantage over monolayer cell cultures, that they are tractable for physiologically relevant functional analysis (87). The outside-in barrier can be quantified in organotypic models using topically applied hydrophilic dye such as Lucifer yellow. This is naturally excluded from the epidermis by the lipid-rich stratum corneum but can permeate into the deeper epidermal and dermal layers if the skin barrier is immature or impaired (83). Analogous to the in vivo situation described above, the inside-outside barrier of organotypic cultures can also be determined by measuring the rate of TEWL (56). These techniques have been used successfully to investigate both the effect of previously uncharacterized genes and the FLG deficiency on skin barrier function (56, 83, 85).

OUTSTANDING QUESTIONS AND FUTURE WORK

The rapid progress made in recent years still leaves a large amount of work to fully capitalize on novel understanding for the benefit of patients.

More detailed genetic studies

The majority of heritability in AD remains unexplained. Improvements in technology have allowed more and more detailed interrogation of the coding and non-coding regions of the genome which are likely to hold important mechanistic information. Outstanding questions involve tissue-specific effects in skin; the relative accessibility of this tissue allows dermatological studies to take advantage of direct sampling for epigenetic studies and more detailed transcriptome analyses. Copy number variation within FLG has a dose-dependent effect on AD (38) and more detailed analyses are required to assess CNV in other risk loci. On a genome-wide level, even larger numbers of cases and controls will be required to achieve the statistical power to detect gene-gene interactions and gene-environment interactions of relevance to AD. These studies remain challenging in their financial cost and computational requirements.

More inclusive genetic research

As described above, the majority of genetic research to date in AD has been conducted in people of white European ancestry. However, the clinical phenotype of AD is different in different ethnicities and studies of genetic risk in African (35) and Asian (88) populations have provided valuable complimentary insight (89). There has been a call in the field to prioritise diversity in human genomics research because this will increase the accuracy, utility and acceptability of using genomic information for clinical care (90). The International Symposium on Atopic Dermatitis (ISAD) has recently...
published a position statement calling for more research on AD in Africa (91).

**Integration of -omics for personalised medicine**

Genetic studies have given important information for understanding AD mechanisms, particularly the initial or ‘root cause’ of atopic skin inflammation. However, the combination and integration of information provided by the full complement of techniques described above will be required to increase our understanding of AD pathophysiology sufficiently to allow translation for clinical impact. Furthermore, given the complexity and diversity of this trait, further developments in machine learning and more powerful in silico analyses (76) are likely to be required to gain full benefit from the wealth of molecular data.

**Application of genetic discoveries to drug development**

The quest for understanding genetic mechanisms in AD is not merely an academic exercise. Genetic studies can provide a causative link between a sequence variant and a phenotype and drugs developed to target a pathway informed by human genetic studies have above-average chances of clinical success (92). Filaggrin deficiency remains a challenging therapeutic target, even though the genetic discovery was made more than a decade ago, but genetic studies continue to identify causal pathways for AD in increasingly precise and personalised detail. The era of “personalised medicine” is expected to bring a new relationship between genomics and drug development, testing the physiological and molecular bases for disease, but success in this endeavour would ultimately transform drug development and clinical use (93).

**CONCLUSION: THE FUTURE LOOKS BRIGHT**

In 1952, Rosalind Franklin was the first to crystallise DNA fibres to study their structure using X-ray diffraction; in 1953 James Watson and Francis Crick reported the double helix structure of DNA; in 1990 the Human Genome Project began and in 2003 the Human Genome Project was completed, providing a sequence of the entire human genome – approximately 3 billion base pairs in length.

Since this time, we have progressed a long way in understanding more of the detail of how DNA sequence variation contributes to human health and disease. There has been a particularly rapid explosion of knowledge in the last 20 years, driven about by increased technical capacity for sequencing DNA and RNA. Whilst it is unlikely that another single gene exists with the impact of FLG upon AD risk, the future appears bright for AD patients: New techniques will refine understanding of genetic risk, with a multi-ethnic perspective, providing powerful insight to drive the development of new pharmacological interventions. These will increasingly be targeted to specific disease mechanisms for each individual patient with AD. The next 100 years is likely to see a step-change in the management of this challenging disease.

**ACKNOWLEDGEMENT**

MSE and SJB are funded by a Wellcome Trust Senior Research Fellowship (ref 106865/Z/15/Z) awarded to SJB. MB is funded by the Swedish Skin Foundation (Hudfonden).

The authors have no conflicts of interest to declare.

**REFERENCES**


study of 21,000 cases and 95,000 controls identifies new risk loci for atopic dermatitis. Nat Genet 2015; 47: 1449–1456.