expert reviews
in molecular medicine

Gene–environment interaction in autoimmune disease

Justine A Ellis[1,2,*], Andrew S Kemp[3] and Anne-Louise Ponsonby[2,3]

Autoimmune disease manifests in numerous forms, but as a disease group is relatively common in the population. It is complex in aetiology, with genetic and environmental determinants. The involvement of gene variants in autoimmune disease is well established, and evidence for significant involvement of the environment in various disease forms is growing. These factors may act independently, or they may interact, with the effect of one factor influenced by the presence of another. Identifying combinations of genetic and environmental factors that interact in autoimmune disease has the capacity to more fully explain disease risk profile, and to uncover underlying molecular mechanisms contributing to disease pathogenesis. In turn, such knowledge is likely to contribute significantly to the development of personalised medicine, and targeted preventative approaches. In this review, we consider the current evidence for gene–environment (G–E) interaction in autoimmune disease. Large-scale G–E interaction research efforts, while well-justified, face significant practical and methodological challenges. However, it is clear from the evidence that has already been generated that knowledge on how genes and environment interact at a biological level will be crucial in fully understanding the processes that manifest as autoimmunity.

Autoimmune disease, characterised by a host immune response against self-antigens, affects more than 8% of the world's population (Ref. 1). As for complex diseases in general, autoimmune disease is said to be caused by a combination of genes and environmental factors. The involvement of genes in autoimmune disease is unequivocal, with hundreds of risk loci identified across the more than 80 different diseases, many shared between diseases (Refs 2, 3, 4). The involvement of environmental factors in autoimmune disease is less well established (Ref. 5), but is becoming increasingly recognised. Most research effort to date has focused on

[1]Genes, Environment and Complex Disease, Murdoch Childrens Research Institute, Parkville, Victoria, Australia

[2]Department of Paediatrics, University of Melbourne, Victoria, Australia

[3]Environmental and Genetic Epidemiology Research, Murdoch Childrens Research Institute, Parkville, Victoria, Australia

*Corresponding author: Dr Justine Ellis, Genes, Environment and Complex Disease, Murdoch Childrens Research Institute, Royal Children's Hospital, 50 Flemington Road, Parkville, Victoria 3052, Australia. E-mail: justine.ellis@mcri.edu.au

Accession information: doi:10.1017/erm.2014.5; Vol. 16; e4; March 2014
© Cambridge University Press 2014

https://doi.org/10.1017/erm.2014.5 Published online by Cambridge University Press

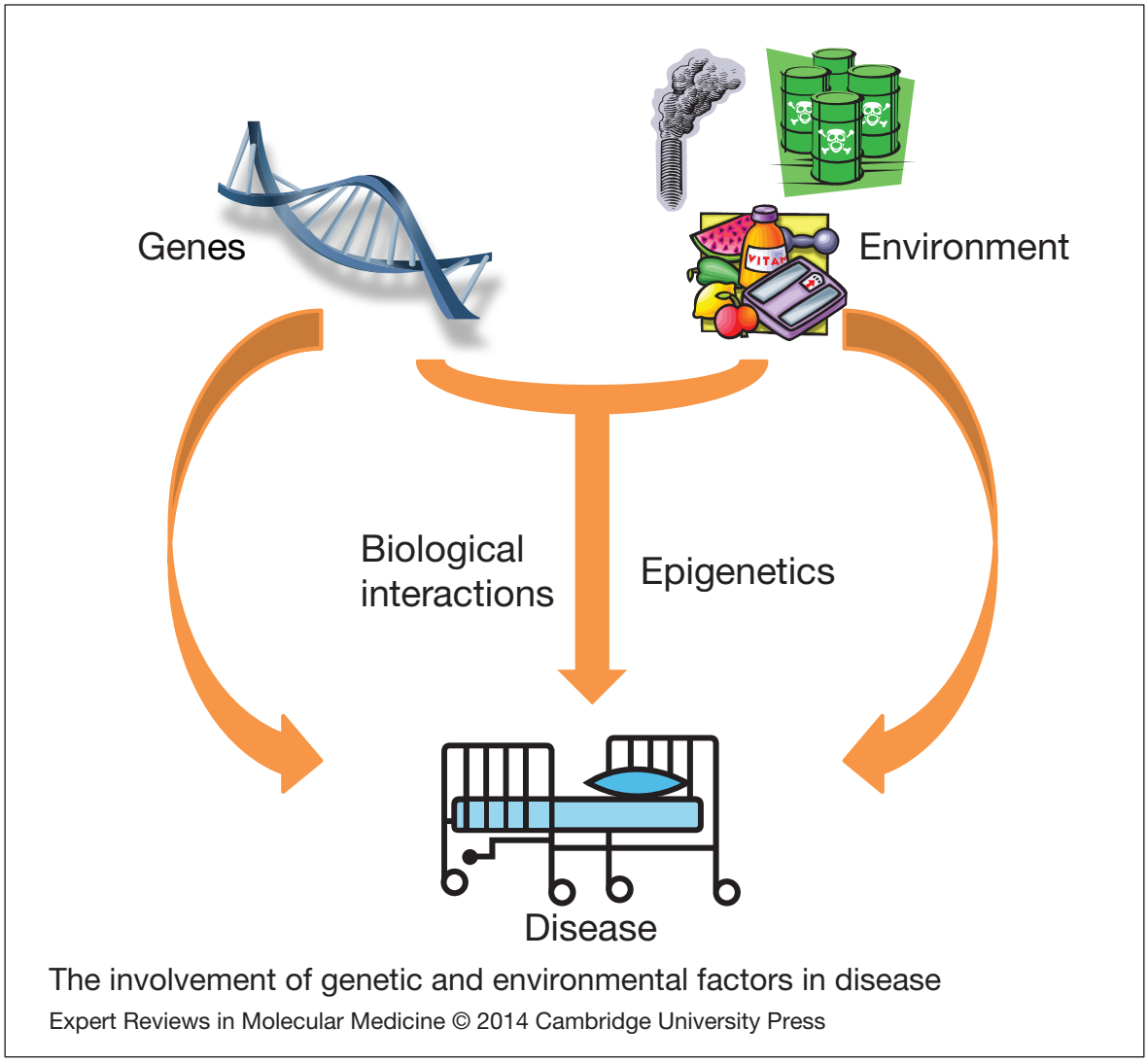


Figure 1. The involvement of genetic and environmental factors in disease. Genes and the environment may influence disease risk independently, or they may interact to determine risk. There are many models for such an interaction, for example, the environment might influence the risk conferred by the genes, the genes might influence the risk conferred by the environment, or both factors might be necessary for any impact on disease risk. The environment may impact the genome via epigenetic mechanisms such as alterations to DNA methylation.

identifying these factors in isolation. However, it is likely that at least a component of overall risk (or conversely, protection) will be determined by the *interaction* of the genes with the environment; that is, the degree of disease risk contributed by one factor will be influenced by the presence or absence of another (Fig. 1). Evidence for such interactions in autoimmune disease is now beginning to emerge, and is providing important mechanistic insight into disease causation.

In this review, following a brief summary of current knowledge as to the genetic and environmental risk factors for autoimmune disease, we will focus on the data generated to date that support the concept of gene–environment (G–E) interaction as an important component to overall autoimmune disease risk. While such interactions are often initially identified using statistical approaches, importantly, various molecular mechanisms of interaction between these intrinsic and extrinsic

Accession information: doi:10.1017/erm.2014.5; Vol. 16; e4; March 2014
© Cambridge University Press 2014

https://doi.org/10.1017/erm.2014.5 Published online by Cambridge University Press

factors are also being studied to support the statistical findings. Thus, our understanding of the causes of autoimmune disease is growing significantly through multifaceted approaches to research, and an ongoing focus on G–E interaction, although challenging, is highly justified.

What is autoimmune disease?

'Autoimmunity' has been defined as immune activation in the absence of known external drivers such as infection (Ref. 6). 'Autoimmune disease' refers to the progression of autoimmunity to a pathogenic state. Both the adaptive and innate immune systems are thought to play a role in autoimmunity. In contrast to autoinflammatory conditions, where innate immune dysregulation is predominant, in autoimmune disease, adaptive dysregulation of the T cell and/or B cell populations also occurs (Refs 7, 8). In relation to the adaptive immune system, self-tolerance is a highly regulated immune process in that self-reactive T and B cells are normally produced by the immune system but are selected against via various central and peripheral tolerance mechanisms. For example, centrally in the thymus, thymocytes are exposed to a vast array of self-peptides. Approximately 98% of thymocytes are found to be self-reactive and are destroyed. In the periphery, production of T cells from the remaining functionally competent (and non-self-reactive) thymocytes is tightly controlled by genomic factors, such as *FOXP3*. Comparable processes occur for the regulation of B cell production from the bone marrow (Ref. 9). Most autoimmune diseases are believed to be caused by a breakdown of these immune tolerance processes via mechanisms that remain largely unclear. This results in inflammation and tissue destruction in a systemic- or organ-specific fashion (Ref. 3). Examples of systemic autoimmune diseases include systemic lupus erythematosus (SLE), which may involve damage to skin, kidney, lungs, brain and heart (Ref. 10); and rheumatoid arthritis (RA) that primarily affects the joints but can also involve inflammation of organs, such as the lungs and heart (Ref. 11). Examples of organ-specific autoimmune diseases include multiple sclerosis (MS) in which demyelination of the central nervous system occurs (Ref. 12); and type 1 diabetes (T1D) that involves destruction of pancreatic beta cells (Ref. 13). The classification of a disorder as autoimmune can be controversial, and Crohn's disease (CD) represents one disease where such controversy exists (Ref. 14).

Genetic risk factors for autoimmune disease

A key feature of most autoimmune diseases is familial clustering, indicating the presence of shared genes and/or shared environmental risk factors. A 2006 systematic review demonstrated familial clustering of RA, T1D, MS and autoimmune thyroid disease (Ref. 15). It is assumed that this multi-disease clustering indicates an overlap in causal pathways. It is perhaps therefore not surprising that many identified genetic risk variants are common to more than one autoimmune disease, whereas others are specific to a single disease (Refs 2, 4).

The association of genetic variation within the human leukocyte antigen (HLA) region on chromosome 6 with multiple autoimmune diseases has been well established for some time, and is often the single greatest contributory factor to genetic predisposition (Ref. 9). Mechanistically, this likely reflects that HLA genes encode cell-surface antigen-presenting proteins. The HLA region is exceptionally variable at a sequence level, providing variability that is crucial in providing the capacity to present a variety of foreign, but also self-antigens, to the immune system. This may impact on individual immune tolerance in two ways – through the efficiency of presentation of self-antigens to thymocytes in order to eliminate those that show self-reactivity, and in the determination of the repertoire of antigens able to be recognised by T cells in the periphery (Ref. 16). Associations of particular HLA alleles are often specific to a particular autoimmune disease (Ref. 9), reflecting that the relevant self-antigen is likely to be different for each form of disease.

Outside of the MHC, the advent of the genome-wide association study (GWAS) approach has seen the number of identified loci associated with various autoimmune diseases reach the hundreds in recent years (Ref. 9). Interestingly, these genes generally fall into distinct functional network groups, most prominently lymphocyte activation and intracellular signalling (e.g. *PTPN22*, *CTLA4*, *PTPN2*), cytokines and receptors (e.g. *IL23R*, *IL2RA*), innate immunity and microbial recognition (e.g. *NOD2*, *IRF5*) and

Accession information: doi:10.1017/erm.2014.5; Vol. 16; e4; March 2014
© Cambridge University Press 2014

https://doi.org/10.1017/erm.2014.5 Published online by Cambridge University Press

transcription factors (e.g. *STAT4*). Recently several analyses have shown that the degree of sharing of genetic effects between autoimmune diseases is highly significant (Refs 2, 4). For example, Cotsapas et al. examined 107 autoimmune disease loci in relation to RA, T1D, SLE, MS, celiac disease (CeD), CD and psoriasis. They presented evidence that single nucleotide polymorphisms (SNPs) at 47 loci are associated with more than one immune-mediated disease, where only one was expected by chance. One SNP at *SH2B3* (involved in T cell receptor activation signalling), showed evidence of association to all seven diseases. Others associated to various disease subgroups, for example a SNP at *PTPN22* associated with RA, SLE, CD and T1D; at *CTLA4* with RA, CeD and T1D; and at *IL23R* with psoriasis, MS and CD (Ref. 2).

Despite recent progress, for each distinct autoimmune disease, much of the heritability remains unaccounted for (Ref. 9). A 2009 review of the 'missing heritability' of various complex diseases indicated that only 15 and 20% of the heritability of SLE and CD respectively could be explained by identified risk loci (Ref. 17). Currently, it is still estimated that over half of the genetic predisposition to autoimmune disease is still unable to be explained (Ref. 9). It may be found within multiple yet-to-be-identified common variants contributing very modestly to overall risk, in rare variants that are not adequately captured by GWAS, or in gene–gene (G–G) interaction (Refs 17, 18). Furthermore, and the subject of this review, the missing heritability is likely to reflect that genes do not operate alone but also in the context of the environment, leading to interactions between risk loci and environmental risk factors (Refs 17, 18).

Environmental risk factors for autoimmune disease

The environmental risk factors associated with autoimmune disorders are varied, as is the strength of evidence for each. A comprehensive review of this field was recently published by an expert panel assembled by the US National Institute of Environmental Health Sciences (Ref. 19). The various exposures and the strength of the evidence in various diseases and further research requirements were outlined. Here, we briefly focus on a selection of these environmental factors as a necessary introduction to the consideration of G–E interaction in autoimmune disease.

Tobacco smoke

Tobacco smoke contains more than 6000 components, many of which are cytotoxic, carcinogenic, mutagenic and/or antigenic (Ref. 20). Studies of the role of tobacco smoke exposure in autoimmune disease are many and varied in terms of study quality, sample size and approach to exposure measurement. The most convincing and well-replicated association is with risk of developing RA. A recent meta-analysis of over 13 000 cases from 16 studies in which a primary purpose was to examine the relationship between smoking status and RA (Ref. 21), confirmed the relationship. The meta-analysis demonstrated that smoking is a risk factor for RA for both men (around twice the risk) and women (about 1.3 times the risk, when compared with non-smokers). It mattered little whether smoking status was classified as ever, current or past. For MS, evidence for association with smoking is also quite strong, as demonstrated by a recent meta-analysis of 14 studies (Ref. 22). For SLE, the evidence is somewhat weaker. A meta-analysis of nine studies demonstrated a significant increase in risk in those who were current smokers but there was no increase in risk among those who were former smokers (Ref. 23). In relation to autoimmune thyroid diseases, smoking has also been associated with an increased risk of Grave's hyperthyroidism and particularly with Grave's ophthalmopathy, in a dose-dependent fashion. It also appears associated with a decreased risk of the hypothyroid disease Hashimoto's thyroiditis (Ref. 24). The biological mechanisms for these associations remain unknown. Other autoimmune diseases examined in relation to smoking include CD and ulcerative colitis, however; studies have been more limited and the outcomes more variable [for review, see (Ref. 19)].

Sun exposure and vitamin D status

Sun exposure provides ultraviolet radiation (UVR) and thus assists the generation of UV-derived vitamin D. There is a growing body of evidence of association of low vitamin D and autoimmune disease risk but the specific roles of UVR and vitamin D are difficult to tease apart (Ref. 25). For MS, a striking ecological feature is the association between higher latitude, thus

Accession information: doi:10.1017/erm.2014.5; Vol. 16; e4; March 2014
© Cambridge University Press 2014

https://doi.org/10.1017/erm.2014.5 Published online by Cambridge University Press

reduced ambient UV levels, and increased prevalence [e.g. (Refs 26, 27, 28, 29)]. Markers of individual low past sun exposure (Ref. 30), season of birth (Refs 31, 32), and low maternal exposure to sun (Ref. 32), have all provided further support for a role of low sun exposure in increasing MS risk. Low vitamin D has also been associated with higher MS risk (Ref. 33). In a prospective US Military cohort, serum 25-hydroxyvitamin D was measured in samples collected prior to disease onset, and risk of MS significantly decreased with increasing vitamin D (Ref. 34). In a cohort of US nurses, vitamin D supplementation was also associated with reduced MS onset (Ref. 35).

For some other autoimmune diseases, similar but often less well-replicated associations with both sun exposure and vitamin D, have been identified. In relation to latitude, there is published evidence of an inverse association with T1D (Refs 36, 37), CD (Ref. 38) and RA (Refs 39, 40). A study by Disanto et al. recently examined the relationship between month of birth and a number of autoimmune diseases including RA, CD, UC and SLE. All diseases, with the exception of CD, showed seasonality (Ref. 41). T1D is also known to be related to season of birth (Refs 42, 43). Lower serum vitamin D levels, and some evidence of a positive effect of supplementation, have been reported in diseases including RA and T1D (Refs 44, 45, 46, 47, 48, 49, 50).

Microbial exposure

There are two hypotheses related to microbial exposure that are not mutually exclusive. The first pertains to microbial infection as a trigger for the onset of disease. The second deals with what is commonly referred to as the 'hygiene hypothesis', whereby microbial exposure during early life serves to 'prime' the developing immune system, providing increased protection against immune disorders later in life.

There are multiple lines of evidence that support an association between Epstein-Barr virus (EBV) infection and higher risk of subsequent MS (Ref. 51). A meta-analysis of more than 19 000 cases and 16 000 controls demonstrated an increased risk of MS in those with a past history of infectious mononucleosis (Ref. 52). Meta-analysis of 13 studies has reported an odds ratio (OR) of 13.5 for MS in those who are EBV seropositive (Ref. 53). Evidence of primary EBV infection coinciding with MS presentation is lacking (Ref. 54), suggesting that it is prior EBV infection that is of greatest importance to disease initiation (Ref. 51). There is some evidence that EBV may also increase risk of developing other autoimmune diseases, including SLE (Ref. 55) and RA (Ref. 56), but the data are generally less convincing (Ref. 19).

A number of other microbes have also been considered in relation to autoimmunity. For parvovirus, anti-B19 antibodies have been reported as more prevalent in the plasma and synovial fluid of RA cases (Ref. 57), however others failed to confirm these findings (Refs 58, 59). The observed co-occurrence of periodontitis and RA (Ref. 60) has led to the hypothesis that *Porphyromonas gingivalis* may be a contributory factor to RA, and some data exist to support this (Refs 61, 62). There have been a number of studies considering the role of the gram-negative bacterium *Helicobacter pylori* in autoimmune disease risk, including in RA and SLE, but relationships remain unconfirmed [for review, see (Ref. 63)]. Enterovirus infection appears to be associated with T1D onset, as shown by meta-analysis drawing together data from 26 case–control studies (Ref. 64).

The hygiene hypothesis, as proposed by Strachan in 1989 (Ref. 65), suggests that the decreasing incidence of early life infection through improvements to western hygiene is contributing to the increasing incidence of immune disorders such as allergic disease. The concept has also been postulated for autoimmune diseases (Ref. 66). Higher exposure to siblings, a proxy for a higher early life microbial load, has been associated with a reduced risk of MS (Ref. 67) and T1D (Ref. 68).

Diet

There is a well-established link between dietary elements, including vitamins, minerals and trace elements, and immune competence (Ref. 69). This suggests that the particular nutrients may associate with risk of autoimmune diseases. It has recently been shown that dietary micronutrients that affect DNA methylation can exacerbate or suppress murine lupus through genetic–epigenetic interactions (Ref. 70). In addition to vitamin D as discussed above, beneficial effects of supplementation of other nutrients have been demonstrated, such as vitamin A for psoriasis, and omega-3 fatty acids

Accession information: doi:10.1017/erm.2014.5; Vol. 16; e4; March 2014
© Cambridge University Press 2014

https://doi.org/10.1017/erm.2014.5 Published online by Cambridge University Press

for RA and possibly SLE (Ref. 69). Diet can also modify the intestinal microbiota, which have been postulated to influence autoimmune disease (Ref. 71). Recently, a high salt diet has been shown to promote the development of Th17 helper T cells, cells that are mediators of autoimmune inflammation (Ref. 72). Increased NaCl concentrations at physiological levels markedly boosted the induction of murine and human Th17 cells, driving autoimmune disease (Ref. 73). The consumption of cow's milk-based infant formulae has been implicated in the development of childhood T1D (Ref. 74). A more obvious role for diet in autoimmune disease is apparent for CeD, where ingestion of gluten triggers disease onset (Ref. 75).

The interaction of genes and environment – how is 'interaction' defined?

We now turn to current understanding of the ways in which genes are impacted by environment to determine risk of disease. Before considering the literature on G–E interactions in autoimmune disease, it is necessary to consider the meaning of 'interaction' in both a statistical and biological sense.

The term 'interaction' means, in general, a reciprocal action, effect or influence. In relation to genetic and environmental risk factors for disease therefore, it can be defined as the influence of one such risk factor on another, or a lack of independence between the two risk factors in determining disease risk. If two factors are independent, the presence or absence of the second risk factor has no influence over the degree of risk that is attributed by the first and vice versa. Put together, the attributable risk is the sum of the risk attributed by the two individual factors (Ref. 76). Rothman's sufficient cause 'pie-model' states that two risk factors can be said to be independent where there is no pathway to disease that requires the involvement of both factors. Conversely, if one pathway to disease requires the presence of both factors, those factors are said to interact (Refs 76, 77).

As reviewed by Wang et al., the term 'interaction' will have different interpretations depending on the field of reference. In the biological fields the term may imply physical interaction, such as G–G interaction that involves the physical contact of far-reaching genomic loci (Ref. 78). Proteins may interact, for example the physical contact of a ligand and its receptor. In the field of statistics, however, interaction is often defined as a statistically significant deviation from a specified model, although there are a number of other interpretations (Ref. 79). The combined effect of two factors may be synergistic or antagonistic, in that the combined effect might be greater or less than the *sum* (identified using an additive statistical model) or *product* (identified using a multiplicative statistical model) of the individual effects (Ref. 79). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement advises reporting the separate effect of each exposure as well as the joint effect compared with the unexposed group in order to provide sufficient information to interpret interaction on both additive (combined effect deviates significantly from the sum of the individual effects) and multiplicative (combined effect deviates significantly from the product of the individual effects) scales (Refs 80, 81). Unfortunately, such data have been presented in only a very small proportion of past articles reporting interactions (Ref. 82).

The identification of G–E interaction is useful for several reasons. It is becoming increasingly recognised that the inclusion of G–E interactions can increase statistical power to detect gene–disease associations in some settings (Refs 83, 84, 85, 86, 87). Furthermore, for observational studies of adverse environmental factors, where randomised controlled trials are not feasible, demonstration of expected G–E interaction can provide further evidence to support a causal inference. Allowing for G–E interaction can also provide stratum-specific estimates of the environmental effect, which may be more relevant for clinical and public health applications than population-averaged effects (Ref. 85). It can also provide findings that lead to an understanding of the underlying biological interactions that are taking place at a molecular level.

Evidence for the interaction of genes and environment in determining autoimmune disease risk

Direct and indirect evidence of interaction

In reviewing G–E interaction in various autoimmune diseases it is clear that there are two broad forms of evidence. Direct evidence of interaction is apparent where studies have used various methods to statistically evaluate G–E interaction. That is, both the gene(s) and

Accession information: doi:10.1017/erm.2014.5; Vol. 16; e4; March 2014
© Cambridge University Press 2014

https://doi.org/10.1017/erm.2014.5 Published online by Cambridge University Press

environmental factor(s) were directly measured in the study, and the presence of interaction was specifically considered. The strength of the direct evidence varies widely, and in many cases, the power to detect interactions was likely inadequate due to small sample sizes. Few interactions could be categorised as 'convincing' according to recently proposed criteria for joint G–E effects on cancer (Ref. 88). Table 1 summarises the autoimmune disease data, concentrating specifically on those interactions that have been examined in more than one study.

In addition to the direct evidence, for many autoimmune diseases there is more indirect evidence of G–E interaction. This may be anecdotal rather than based on a formal test for interaction; derived from bioinformatics approaches; or assumed due to an apparent role of epigenetics in disease risk. Epigenetics concerns the chemical modification of DNA, for example DNA methylation, that has the capacity to alter gene transcription and thus cell function. Epigenetic marks are malleable, potentially altering across the life course, and there is considerable evidence that they are impacted by environmental exposures (Ref. 89). For example, smoking has been shown to alter DNA methylation at various loci across the genome (Refs 90, 91), and there is evidence that many other pollutants and chemicals can induce changes to the epigenome (for review, see (Ref. 92)). Therefore, epigenetics provides a conduit through which the environment can influence genes, and thus affords a biological mechanism for an interplay between environment and genes in autoimmune, and other, diseases (Ref. 3).

Disease-specific evidence

Rheumatoid Arthritis

Of all the autoimmune diseases, the direct evidence for G–E interaction in determining risk of disease is strongest for RA. The RA 'shared epitope' (SE) defines a group of HLA types that share a common amino acid sequence in the third hypervariable region of the HLA DRβ1 chain, encoded by the *DRB1* gene (Ref. 93). The SE incorporates a number of different alleles of *HLA–DRB1* that have been associated with RA in different populations. The statistical interaction of smoking with the *HLA–DR* SE was first reported in a Swedish population-based case–control study by Padyukov et al. in 2004 (Ref. 94). It was observed that smoking was only a modest risk factor for RA in those not carrying an SE allele (SE-negative), however, in those SE-positive, the contribution of smoking to RA risk was striking. A dose-dependent effect was also seen for SE, with risk increasing with a single, or double SE allele (current smokers with no SE allele relative risk (RR) = 1.1, 95% CI 0.7–1.9; current smokers with a single SE allele RR = 2.3, 95% CI 1.4–3.7; current smokers with double SE allele RR = 5.6, 95% CI 2.9–11.1; all compared to never smokers with no SE allele) (Ref. 94). This interaction has now been widely replicated (Refs 95, 96, 97, 98, 99, 100, 101, 102, 103).

The study by Padyukov et al. also reported another striking finding – that the association of the combination of smoking and SE with RA was only true for the subtype of RA characterised by rheumatoid factor (RF) seropositivity (Ref. 94). This finding was further extended by data that showed citrullination of certain peptides renders them more able to bind to the HLA-DR SE (Ref. 104), and knowledge that RF positivity is highly correlated with anti-cyclic citrullinated peptide (anti-CCP) antibody positivity in RA. When anti-CCP positivity was used to subclassify RA, the association of the combination of smoking and HLA-DR SE was even stronger, suggesting that the association with RF positivity was likely secondary to a primary association with anti-CCP positivity (Ref. 105). Again, this has been widely confirmed (Refs 97, 99, 100, 101, 102). It has been proposed that the likely biological mechanism behind this interaction involves the tobacco smoke-induced citrullination of self-peptides in the lungs. Citrullination involves the post-translational modification of protein that is catalysed by the enzyme peptidylarginine deiminase (PAD) (Ref. 106). Thus, smoking may preferentially induce an autoimmune response to citrullinated peptides in those who are HLA-DR SE carriers (Ref. 107) (Fig. 2). Thus, self-peptides may be chemically modified by tobacco smoke, rendering them more open to recognition by the immune system as non-self. The exact peptide(s) that is rendered arthritogenic by citrullination is the subject of considerable research effort. It has also been noted that the presence of anti-CCP can often be detected years before the onset of any clinical signs of disease, potentially providing the opportunity for preventive intervention (Ref. 76).

Accession information: doi:10.1017/erm.2014.5; Vol. 16; e4; March 2014
© Cambridge University Press 2014
https://doi.org/10.1017/erm.2014.5 Published online by Cambridge University Press

https://doi.org/10.1017/erm.2014.5 Published online by Cambridge University Press

Table 1. Direct evidence of G–E interaction in autoimmune disease – selected interactions examined by two or more studies

Disease	Interacting factors	Study design	Results	Comments	References
Rheumatoid arthritis, seropositive	Smoking, *HLA-DR* SE	Case–control	**RR/OR for smokers carrying at least 1 SE allele range from 7.5 to 53,** compared with never smokers with no SE. AP for interaction ranged from 0.4 to 0.7	Studies varied by definition of smoker and stratification by RF/anti-CCP status, but interaction appears confirmed	(94-100)
	Smoking, *PTPN22*	Case–control	**One study reported no interaction, while a second reported a heavy smoking interaction** (AP = 0.42, 95% CI 0.01–0.84)		(95, 113)
		Case only	**No evidence of interaction** in determining anti-CCP positive RA compared with anti-CCP negative RA	Smoking-*PTPN22* interaction remains unclear	(98)
	Smoking, *HLA-DR* SE, *PTPN22*	Case–control	**Study showed highest ORs were evident when all three factors were considered together**, but outcomes of interaction tests were not reported		(95)
		Case only	Combination of SE and smoking predictive of anti-CCP positivity irrespective of *PTPN22* status	Three-way interaction not apparent	(98)
Rheumatoid arthritis	Smoking, *GSTM1*	Case–control	**Significant ($P = 0.047$) interaction reported by one study, but not replicated** by two subsequent studies	Studies performed in various ethnic groups. Requires further replication efforts	(109, 110, 111)
	Smoking, *PADI4*	Case–control	**No interaction reported in Caucasian women. Significant interaction ($P = 0.34$) reported in Japanese men**	Studies focused on different genders and ethnicities. Inconclusive	(112, 113)

(continued on next page)

Accession information: doi:10.1017/erm.2014.5; Vol. 16; e4; March 2014

© Cambridge University Press 2014

https://doi.org/10.1017/erm.2014.5 Published online by Cambridge University Press

Table 1. Direct evidence of G–E interaction in autoimmune disease – selected interactions examined by two or more studies *(continued)*

Disease	Interacting factors	Study design	Results	Comments	References
Multiple sclerosis	Smoking, *HLA-DR15*	Case–control	One study reported the **relationship between *HLA-DR15* and MS is independent of smoking**. A second study identified a significant interaction, but **only in those lacking a protective *HLA-A**02 allele** (AP = 0.6, 95% CI 0.3–0.8)	Further confirmatory work taking into account *HLA-A* required to confirm relationship	(132, 133)
	EBV, *HLA-DR15*	Case–control	Two studies reported no interaction. One study reported **2.9-fold increased risk conferred by *HLA-DR15* in presence of IM** history. Another reported **interaction of *HLA-DR15* with both EBV antibodies** (AP = 0.60, 95% CI 0.45–0.76) **and IM history** (AP = 0.34, 95% CI 0.01–0.68)	Studies varied in their definition of EBV exposure. Interaction remains unconfirmed	(120, 121, 122, 123)
	Sibling exposure, *HLA-DR15*	Case–control	Study reported a **3.9-fold increase in risk conferred by *HLA-DR15* when infant sibling exposure was low**.		(122)
		Case only	**No difference** in frequency of *HLA-DR15* between cases grouped by low or high sibling exposure	Further confirmation required	(124)

(continued on next page)

Accession information: doi:10.1017/erm.2014.5; Vol. 16; e4; March 2014

© Cambridge University Press 2014

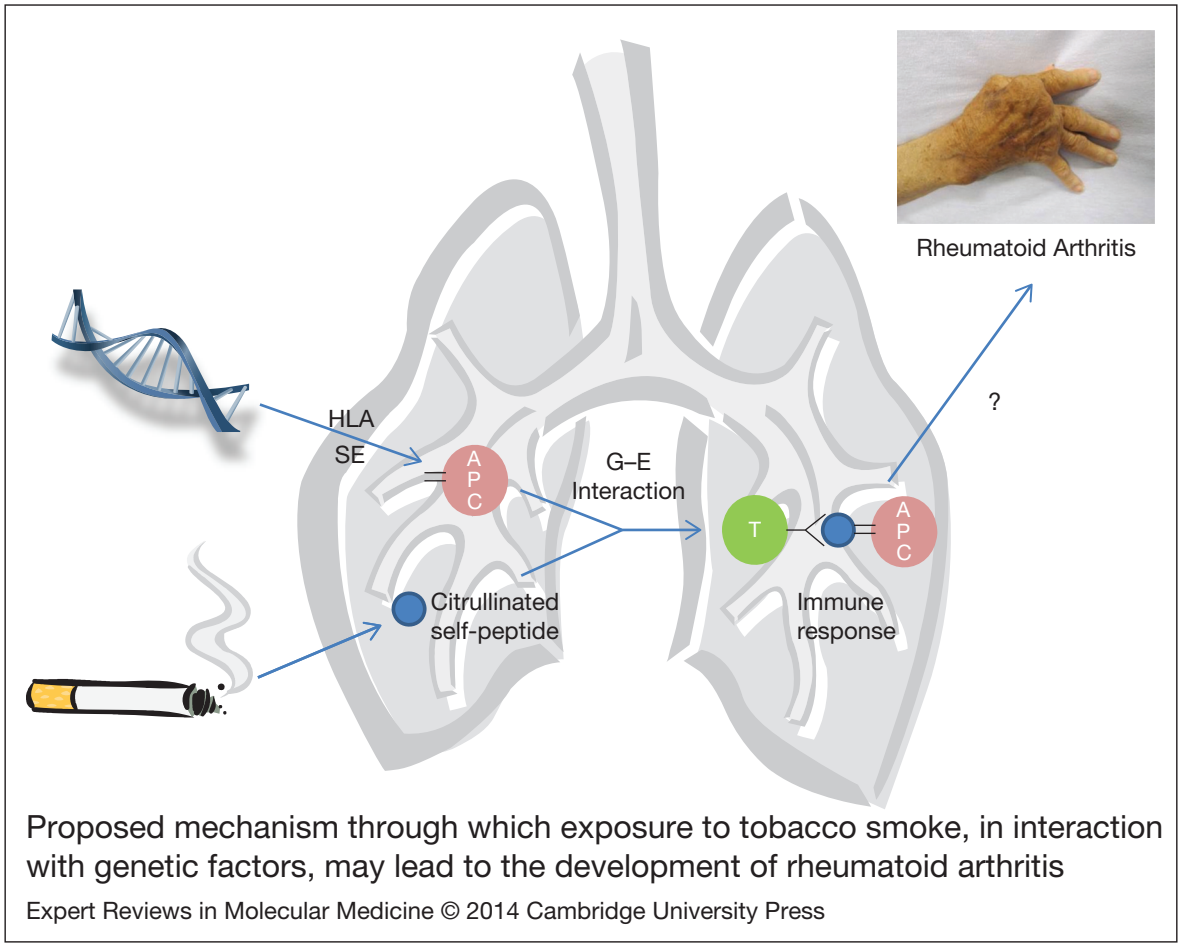
https://doi.org/10.1017/erm.2014.5 Published online by Cambridge University Press

Table 1. Direct evidence of G–E interaction in autoimmune disease – selected interactions examined by two or more studies *(continued)*

Disease	Interacting factors	Study design	Results	Comments	References
	Season of birth, *HLA-DR15*	Case–control	One study reported no case–control differences in month of birth, but observed an **excess of spring births/ deficit of autumn births in *DR15*-carrying MS cases** compared with *DR15*-non-carrying cases. Pattern not seen in controls. No formal test for interaction. A further study found that **spring births increased risk of MS regardless of *DR15* status**	Studies that formally test for interaction are required to confirm	(126, 129)
Crohn's disease	Smoking, *NOD2*	Case–control	**Association with *NOD2* differed by smoking status** of cases. No formal tests of interaction presented.		(148)
		Case only	***NOD2* risk allele more frequent in those cases who had ever smoked**, compared with never smoked. Interaction $P = 9 \times 10^{-4}$	Interactions at different *NOD2* SNPs between studies. Not yet confirmed	(147)

AP, Attributable proportion; EBV, Epstein-Barr virus; IM, infectious mononucleosis; OR, odds ratio; RR, relative risk; SE, Shared epitope.

Accession information: doi:10.1017/erm.2014.5; Vol. 16; e4; March 2014
© Cambridge University Press 2014



Proposed mechanism through which exposure to tobacco smoke, in interaction with genetic factors, may lead to the development of rheumatoid arthritis

Expert Reviews in Molecular Medicine © 2014 Cambridge University Press

Figure 2. Proposed mechanism through which exposure to tobacco smoke, in interaction with genetic factors, may lead to the development of rheumatoid arthritis. Smoking may cause citrullination of peptides present in the lung. Antigen-presenting cells (APCs) of genetically susceptible individuals who carry the HLA-DR shared epitope (SE) may be more prone to presenting these citrullinated self-peptides to T cells, resulting in an immune response against the peptides. The mechanism through which this process then initiates joint inflammation remains unclear. However, a 'second hit' such as a joint trauma may induce joint peptide citrullination, to which the immune system is already primed to respond (Ref. 76).

Further work, introducing a second genetic locus, *PTPN22*, into the interaction analysis has demonstrated that this gene too may play an important role in the interaction. *PTPN22* encodes a tyrosine phosphatase known to play a role in lymphocyte activation (for review, see (Ref. 108)), and contains a non-synonymous single nucleotide variant associated with autoimmune disease. Across three distinct RA case–control sample sets (Sweden, US, Netherlands), a G–G interaction was found between *HLA-DR* SE and *PTPN22* in determining risk of anti-CCP positive RA. No interaction between *PTPN22* and smoking was identified (Ref. 95).

A number of other smoking-gene interaction analyses in RA risk have been described. A 2006 report demonstrated interaction between smoking and the glutathione s-transferase M1 gene (*GSTM1*) (Ref. 109). This was not replicated by later studies (Refs 110, 111), although one such later study did detect interaction with other genes also encoding enzymes known to be involved in carcinogen detoxification and protection from oxidative stress (*GSTT1*, *HMOX1*) (Ref. 110). Interaction of smoking with *PADI-4*, encoding a PAD enzyme involved in post-translational citrullination, was reported in Japanese (Ref. 112), although this was in contrast to a previous study in Caucasian

Accession information: doi:10.1017/erm.2014.5; Vol. 16; e4; March 2014
© Cambridge University Press 2014

https://doi.org/10.1017/erm.2014.5 Published online by Cambridge University Press

women where no interaction was demonstrated (Ref. 113). Some evidence also exists to suggest interaction of smoking with complement-activating mannan-binding lectin (*MBL*, involved in innate immune defences) in Caucasians (Ref. 114) and with N-acetyltransferase 2 (*NAT2*, a carcinogen detoxifier) in African Americans (Ref. 111).

Indirect evidence for G–E interaction in RA comes from a growing number of studies reporting changes to the DNA methylation profile of various tissues from RA cases compared to unaffected controls. Earlier studies demonstrated a global hypomethylation of RA synovial fibroblast DNA (Ref. 115), and changes to the methylation state of various candidate loci such as *IL6* (Ref. 116) and *DR3* (Ref. 117) in disease-relevant tissues. More recent genome-scale studies have reported numerous loci across the genome that appear differentially methylated in cases compared to controls (e.g. (Refs 118, 119)). Collectively, this evidence directly implicates a role for epigenetics in the RA disease state, and indirectly suggests that the environment may be interacting with the genome at numerous loci in this disease.

Multiple sclerosis

Direct evidence of G–E interaction is also relatively strong for MS. Following from the knowledge that both *HLA-DRB1*15* and EBV exposure are associated with risk of MS, the interaction of these two factors was examined. A 2008 nested case–control study of 148 female MS cases and 296 matched controls from the US Nurses' Health Study demonstrated expected single factor associations of *DR15* and measures of EBV exposure (infectious mononucleosis, anti-EBV antibody titre) with disease (Ref. 120). However, statistical analysis of *DR15* and anti-EBV antibody titres revealed no evidence of interaction in determining MS disease risk (Ref. 120). The results of independent follow-up studies have been mixed. A 2009 study examining Danish and Swedish cohorts identified a significant interaction between IM and *HLA-DRB1*15*, with risk of MS associated with *HLA-DRB1*15* being 2.9 times higher in those with history of IM (95% CI 1.3–6.5, interaction $P = 0.007$) (Ref. 121). However, a lack of statistical interaction between anti-EBV antibodies and *DR15* was reported by a 2010 Australian study (Ref. 122). A 2012 Swedish case–control study (Ref. 123) looked for interaction between presence of risk allele *HLA-DRB1*15*, absence of protective allele *HLA-A*02* and anti-EBV antibody titres. They found significant interaction, as defined by departure from additivity, between EBV antibody titres and *DRB1*15* (AP = 0.6, 95% CI 0.45–0.76). This interaction persisted when data were restricted to those who were *A*02* negative or *A*02* positive. Similarly, interaction was found between EBV antibodies and absence of *A*02* (AP = 0.39, 95% CI 0.18–0.61), which persisted only in the *DRB1*15* negative restricted group. They argued that previous studies with a negative outcome had reported interaction only on the multiplicative scale, and that their detection of departure from additivity is more biologically relevant (Ref. 123). Thus detection of the interaction may be quite sensitive to study design and analytical approaches.

Studies have also examined interaction of HLA with infant sibling exposure (as a proxy for early life microbial exposure) in MS. In cases and controls recruited from Tasmania, Australia (Ref. 122), an interaction between low infant sibling exposure and *DR15* was detected, such that the combined effect of these two factors was 3.9 times greater on a multiplicative scale than that expected if no interaction were present (interaction $P = 0.019$). Therefore, infant sibling exposure may be more protective against MS in those who are *HLA-DR15* positive. No interaction of anti-EBV antibodies with *HLA-DR15* was detected, nor was there a statistically significant three-way interaction between the factors (Ref. 122). This was followed by a report of 197 Canadian MS probands with cumulative sibling exposure <1 year, and 1112 probands with higher exposure (Ref. 124). No difference in the frequency of *DRB1*15* alleles was observed between the exposure groups, suggesting a lack of interaction. The authors point out that unlike in the Australian study, they were unable to identify a single-factor association between MS and infant sibling exposure, and suggest that the reported phenomena may be population specific (Ref. 124).

A very recent study took a bioinformatics approach to indirectly considering G–E interaction in MS (Ref. 125). The authors defined candidate 'interactomes' – sets of genes for which some evidence exists for interactions with various viral factors in determining disease risk.

Accession information: doi:10.1017/erm.2014.5; Vol. 16; e4; March 2014
© Cambridge University Press 2014
https://doi.org/10.1017/erm.2014.5 Published online by Cambridge University Press

They then tested for statistical enrichment of genes within these interactomes amongst loci identified as associated with MS by GWAS. In particular, they found genes of the EBV interactome were significantly enriched within GWAS-identified loci, providing further support that G–E interaction is an important component of MS risk (Ref. 125).

The interaction of genes and sun exposure/vitamin D in determining risk of MS has also been considered. Season/month of birth has been used as a proxy measure of sun exposure/vitamin D in early life, although this measure does not capture variations in individual behaviours, skin pigmentation, or dietary intake of vitamin D (Ref. 32), and is likely associated with other seasonal factors. An analysis of month of birth in relation to *HLA-DRB1*15* status was carried out in a large sample of MS patients from Canada, Sweden and Norway, and in two groups of unaffected controls (proband siblings and unrelated individuals from Sweden) (Ref. 126). A higher proportion of *DRB1*15*-positive MS cases were born in April (spring) compared to *DRB1*15*-negative cases, and conversely, a lower proportion of *DRB1*15*-positive cases were born in November (autumn) compared to *DRB1*15*-negative cases. No such pattern was observed for controls. While no formal statistical test for interaction was performed, the opposing effect suggests that the combination of season of birth and *DRB1*15* is important in MS (Ref. 126). The reason for this may lie in the recent discovery of *DRB1* allele-specific functionality of an upstream vitamin D response element (VDRE). In *HLA-DRB1*15*-bearing haplotypes, the VDRE is functional, and vitamin D is capable of controlling the expression of the gene. On the background of other *DRB1* alleles not associated with MS, the VDRE is non-functional and vitamin D does not stimulate gene expression (Ref. 127). Therefore, exposure to vitamin D at a crucial early life time point that is captured by season of birth may be more relevant to MS risk in those carrying the *HLA*DRB1*15* allele. Furthermore, a recent study characterised the functionality of the VDRE associated with various *HLA-DRB1* alleles, then tested for association of MS with alleles carrying/not carrying a functional VDRE (Ref. 128). Increased risk of MS was associated with the functional VDRE-carrying *DRB1* allele. DRB1 alleles carrying non-responsive VDREs were associated with reduced MS risk (Ref. 128). One proposed mechanism is that thymic deletion of autoreactive T cells may be impaired when early life vitamin D is insufficient in those carrying *HLA-DRB1* alleles with responsive VDREs (Ref. 127). Not all studies though have reported consistent findings. In over 8000 Finnish MS cases and controls, a direction of effect opposite to that previously observed was reported, in that the increase in May (spring) births was more pronounced in *DR15* − MS cases than in *DR15*+ MS cases (Ref. 129). Similarly, no relationship between month of birth and *HLA*DRB15* was found within samples utilised for a large MS GWAS (Ref. 130).

Furthermore, a study showed that a promoter variant of the vitamin D receptor gene (*VDR*) was not associated with overall risk of MS in Tasmanian cases compared to controls. However, the *VDR* variant was associated with an increased risk of MS in those with low winter sun exposure (≤2 h/day), but not in those with high exposure. The interaction was statistically significant ($P = 0.012$) (Ref. 131). Again, this illustrates the potential to identify new genetic associations through stratification by relevant environmental exposures.

In relation to smoking, one study found that association with the *HLA-DRB1*15* risk allele was independent of smoking (Ref. 132). A later study identified an interaction between presence of *HLA-DRB1*15* and absence of *HLA-A*02* (protective allele) in disease risk that was observed only in smokers (Ref. 133). The biological mechanism(s) through which any HLA-smoking interaction might act in MS are unknown. Further work is required to confirm the findings and to increase mechanistic understanding.

Systemic lupus erythematosus

NAT2 has also been examined for interaction with smoking in SLE in Japanese. The combination of smoking and the genotype associated with slower detoxification significantly increased risk compared to the rapid genotype with no smoking history (AP = 0.05, 95% CI 0.12–0.88) (Ref. 134). Similarly, in Japanese women, an interaction between smoking and *TNFRSF1B* (a member of the tumour necrosis factor receptor family) has been reported (Ref. 135). Additionally, *CYP1A1* and *GSTM1* have been tested for interaction with smoking in Japanese

Accession information: doi:10.1017/erm.2014.5; Vol. 16; e4; March 2014
© Cambridge University Press 2014
https://doi.org/10.1017/erm.2014.5 Published online by Cambridge University Press

SLE. Both these genes encode enzymes that metabolically activate carcinogenic or toxic substances to produce reactive oxygen species (ROS) (Ref. 136). ROS have been implicated in SLE pathogenesis (Ref. 137). The combined 'at risk' *CYP1A1*/*GSTM1* genotype was associated with a 17.5-fold increase in disease risk in ever smokers (95% CI 3.2–95.9), but only a 2.8-fold increase in risk in non-smokers (95% CI 1.7–4.6) (Ref. 136). All these findings await further independent confirmation.

Fraser et al. looked for interaction between occupational sun exposure and *GSTM1* in SLE. They found a 3-fold increase in risk in those with high sun exposure who were homozygous for the *GSTM1* null allele, but there was no association with sun exposure in those who carried a functional *GSTM1* (Ref. 138).

There is also substantial evidence, gathered over many years, for a role for epigenetics in SLE (for recent review, refer to (Ref. 139)). In brief, early studies demonstrated decreased global DNA methylation levels in T cells of SLE patients (Ref. 140). Candidate gene studies have demonstrated altered SLE methylation states at various loci, including *ITGAL* (Ref. 141) and *PRF1* (Ref. 142). Recent genome-scale efforts comparing methylation between discordant twin pairs (Ref. 143) and unrelated cases and controls (e.g. (Ref. 144)) have also demonstrated differences in methylation patterns across the genome that are associated with SLE. Furthermore, recent work has demonstrated that, while exposure of peripheral blood mononuclear cells to ultraviolet B can induce an inhibition of DNA methylation, PBMCs from patients with active SLE are more sensitive to this process (Ref. 145). Photosensitivity is known to be associated with SLE (Ref. 146); therefore this finding provides a potential mechanism through which UVB, as an environmental exposure, may interact with the genome in disease.

An interesting example in which differential DNA methylation could be linked to both environment and altered gene expression was published very recently in a mouse model of SLE (Ref. 70). Genetically lupus-resistant and lupus-susceptible mouse strains were fed diets either high or low in dietary micronutrients important for the process of DNA methylation (methyl donors), such as methionine, choline, folate and other B vitamins. The authors showed that a diet higher in methyl donors decreased autoantibody levels and kidney disease severity in lupus-susceptible mice. They also showed that in mice on a methyl donor restricted diet, DNA methylation was reduced at the gene encoding CD40L, a protein that is elevated on T cells from human and mouse females with lupus. They concluded therefore that a diet high in methyl donors may be beneficial to lupus patients via an increase in methylation at lupus-susceptibility genes, reducing their expression (Ref. 70).

Crohn's disease

A small number of studies of CD have examined interactions of genes with smoking in disease risk. *NOD2* plays an important role in regulation of host response to intestinal bacteria, and has been robustly associated with CD. A large German case-only study recently identified interaction between smoking (current and ever) and a *NOD2* CD risk allele (Ref. 147). This was consistent with an earlier Dutch case–control study (Ref. 148). Other genes in the Dutch study apparently interacting with smoking included *DLG5*, *IL23R* and *PTPN2*. Association with *PTPN2* was only detectable if the smoking CD group alone was compared with controls, showing that stratification by relevant environmental phenotypes is likely to have the capacity to identify new risk genes (Ref. 148).

A moderate level of indirect evidence of G–E interaction in CD comes from a limited number of studies examining the CD epigenome. A genome-scale effort examining around 27 000 CpG sites identified more than 1000 differentially methylated loci in peripheral blood mononuclear cells (Ref. 149). The authors also reported significant enrichment of differential methylation near genes associated with CD by GWAS studies, including *NOD2*. This suggests that CD risk genes altered by sequence variation may be further impacted by epigenetic variability, possibly brought about by differing environmental exposures such as smoking. Smoking is known to impact DNA methylation (Ref. 150).

Thyroid autoimmunity

Age is likely a proxy measure for cumulative exposure to numerous non-genetic factors. For the autoimmune thyroid disease Grave's disease, there are data to suggest that some genetic associations with this condition differ by age.

Accession information: doi:10.1017/erm.2014.5; Vol. 16; e4; March 2014
© Cambridge University Press 2014

https://doi.org/10.1017/erm.2014.5 Published online by Cambridge University Press

For example, an age-dependent gene dosage effect of the risk allele of *PTPN22* has been reported (Ref. 151), as has association of variants of *HLA-DRB1* only in those diagnosed at a younger age (Ref. 152). Assuming that the use of age indirectly captures one or more relevant environmental factors, these data may indirectly indicate a G–E interaction in determining disease risk. Further indirect evidence comes from work showing that the thyroglobulin gene (*TG*) a gene associated with thyroid autoimmune disease, is up-regulated by interferon alpha (IFNα), which in turn is stimulated by environmental factors such as viral infection (Ref. 153).

Type 1 diabetes

There is some evidence of G–E interaction for T1D onset. However, this issue has been relatively understudied in T1D. In a case-only analysis, Badenhoop et al. examined the pattern of month of birth among those carrying various HLA-DQ genotypes, and observed differences in birth month distributions (Ref. 154). This may reflect an impact of environmental factors related to season, including vitamin D levels and rates of viral infections, on the degree of disease risk attributed by HLA alleles. No such variability in month of birth was noted between cases carrying non-HLA risk alleles, suggesting that the impact may be specific to HLA. Moreover, in 15 pairs of monozygotic twins discordant for T1D, over 100 loci across the genome were found to be differentially methylated in monocytes (Ref. 155), providing indirect epigenetic evidence for G–E interaction in this disease.

Celiac disease

An example of indirect evidence for G–E interaction, rather than having been formally tested for using statistical methods, is that of HLA and gluten in CeD. Virtually 100% of all individuals diagnosed with CeD carry a *HLA-DQ2* or *DQ8* haplotype. Additionally, following deamidation by the enzyme transglutamidase 2 (Ref. 156), gluten peptides are known to be the pivotal trigger of disease initiation, demonstrating that at least one causal pathway to disease must include both of these factors (Ref. 75). Furthermore, a gene-dosage effect has been shown in which those homozygous for the *HLA-DQ2* risk allele have a greater magnitude and breadth of gluten-specific T cell responses compared with heterozygotes (Ref. 157).

Other diseases for which indirect evidence exists

A number of other autoimmune diseases carry indirect epigenetic evidence of G–E interaction. Aberrant DNA methylation has been associated with psoriasis, systemic sclerosis, primary biliary cirrhosis and primary Sjögren's syndrome, as reviewed elsewhere (e.g. (Ref. 158)).

Clinical implications/applications

Ultimately, a better understanding of disease risk genes and environmental exposures, and the ways in which they interact, will enable the development of more effective approaches to disease treatment and prevention. In complex diseases, including the autoimmune diseases discussed in this review, various causal pathways are likely to exist in the population, and thus the number of contributing risk factors and their combinations may vary from person to person. An ability to determine an individual's potentially unique profile of risk factors and their interactions will likely shed light on the underlying biological mechanisms responsible for disease onset, and will therefore be a powerful tool in determining the most effective treatment strategies. Understanding the influence of environment on genetic profile is likely to enhance the field of pharmacogenomics, in which the aim is to predict the efficacy or toxicity of a drug based on genetic makeup (Ref. 159), moving closer to more refined personalised medicine approaches. Additionally, disease-related changes to the epigenome brought about by environmental exposure may be able to be pharmaceutically corrected (Ref. 160). In terms of prevention, understanding how genes influence risk in different environmental strata, and vice versa, will help better define groups of individuals in the population for which preventive interventions are more likely to be effective (Ref. 161).

Research progress and outstanding research questions

One of the biggest challenges facing G–E interaction research at the present time is that international, multi-study, harmonised measurement of environmental exposures is very difficult to achieve (Ref. 162). The explosion of genetic associations via GWAS approaches has been successful for the simple reason that

Accession information: doi:10.1017/erm.2014.5; Vol. 16; e4; March 2014
© Cambridge University Press 2014

https://doi.org/10.1017/erm.2014.5 Published online by Cambridge University Press

genome-wide genotyping is an established technique that can be carried out anywhere in the world with high precision and at low cost. DNA is easy to collect, and genotypes do not change over time so it does not matter when (pre- or post-disease onset) DNA was collected from study participants (Ref. 84). Thus, international disease consortia that bring together thousands of case DNA samples have become commonplace. Often samples are genotyped using various methods but there are well-established methodologies to harmonise genotype data between case groups (Ref. 163). Thus, it is relatively easy to collect GWAS sample sets of sufficient size to provide power to detect loci with modest risk effects. It is also straightforward to access independent sample sets within which to perform replication analyses, the gold standard for confirmation of associations (Ref. 163).

In contrast, extensive environmental data are often not collected alongside cohort or case–control DNA collections (Ref. 162). The resources required to collect such data are often significant, and thus not prioritised in studies aimed primarily at the largest possible DNA sample size. Even if some environmental data are captured, the way in which it is captured may be (and often is) different or imprecise between studies, making the task of data harmonisation to enable the assembly of large, well-powered, discovery and replication datasets a difficult one (Ref. 164). It has been concluded that the trade-off between bigger studies versus more precise and informative measurement should balance more towards the latter in G–E interaction studies of continuous traits (Ref. 84). Further work to develop more sophisticated methods of assessing G–E that account for gene allele dose responses and genotype risk scores within gene networks, rather than single gene variants, is likely to be worthwhile (Ref. 165).

Efforts towards larger-scale G–E studies are building, with several international data collections and data harmonisation efforts now underway (Ref. 166). To provide further clarity in this field, it was recommended in 2012 that the causal inference for future G–E interactions could be reported using guidelines which incorporate: (i) a score for strength of evidence for the (a) environmental and (b) genetic factors; (ii) establish a prior score and decide on the pattern of interaction to be expected; (iii) score interaction strength; and (iv) evaluate the overall plausibility for interaction (Ref. 88). Progress in this field is likely to uncover new and more complete information about the breadth of autoimmune disease risk factors; help to further understand the proportion of risk that has not yet been explained by genetic research, and provide new insights into biological mechanisms behind disease development.

In summary, the number of established genetic and environmental risk factors associated with autoimmune disease is growing. Some of these factors appear shared between a number of autoimmune diseases and some appear specific for single diseases. Similarly, some particular combinations of G–E interactions appear to be shared between diseases, but the biological mechanisms behind the interactions may be different. However, a lack of well-established standardised analytical approaches, coupled with the inherent difficulties of large-scale collection of validated and harmonised environmental exposure data provide significant challenges in moving the field further forward. But the importance of identifying G–E interaction in autoimmune, and other complex diseases, is becoming increasingly acknowledged.

Financial support

This work was supported by the Victorian Government's Operational Infrastructure Support Programme. J. A. E. was supported by an Australian Research Council Future Fellowship, Project ID FT120100253. A. L. P. was supported by a National Health and Medical Research Council Senior Research Fellowship, ID 1008396.

Conflicts of Interest

None.

References

1 Cooper, G.S., Bynum, M.L. and Somers, E.C. (2009) Recent insights in the epidemiology of autoimmune diseases: improved prevalence estimates and understanding of clustering of diseases. Journal of Autoimmunity 33, 197-207
2 Cotsapas, C. et al. (2011) Pervasive sharing of genetic effects in autoimmune disease. PLoS Genetics 7, e1002254
3 Javierre, B.M., Hernando, H. and Ballestar, E. (2011) Environmental triggers and epigenetic deregulation in autoimmune disease. Discovery Medicine 12, 535-545

Accession information: doi:10.1017/erm.2014.5; Vol. 16; e4; March 2014
© Cambridge University Press 2014

https://doi.org/10.1017/erm.2014.5 Published online by Cambridge University Press

4 Richard-Miceli, C. and Criswell, L.A. (2012) Emerging patterns of genetic overlap across autoimmune disorders. Genome Medicine 4, 6
5 Ellis, J.A., Munro, J.E. and Ponsonby, A.L. (2010) Possible environmental determinants of juvenile idiopathic arthritis. Rheumatology (Oxford) 49, 411-425
6 Davidson, A. and Diamond, B. (2001) Autoimmune diseases. New England Journal of Medicine 345, 340-350
7 Doria, A. et al. (2012) Autoinflammation and autoimmunity: bridging the divide. Autoimmunity Reviews 12, 22-30
8 Hayter, S.M. and Cook, M.C. (2012) Updated assessment of the prevalence, spectrum and case definition of autoimmune disease. Autoimmunity Reviews 11, 754-765
9 Cho, J.H. and Gregersen, P.K. (2011) Genomics and the multifactorial nature of human autoimmune disease. New England Journal of Medicine 365, 1612-1623
10 Tsokos, G.C. (2011) Systemic lupus erythematosus. New England Journal of Medicine 365, 2110-2121
11 McInnes, I.B. and Schett, G. (2011) The pathogenesis of rheumatoid arthritis. New England Journal of Medicine 365, 2205-2219
12 Ramagopalan, S.V. et al. (2010) Multiple sclerosis: risk factors, prodromes, and potential causal pathways. Lancet Neurology 9, 727-739
13 Bluestone, J.A., Herold, K. and Eisenbarth, G. (2010) Genetics, pathogenesis and clinical interventions in type 1 diabetes. Nature 464, 1293-1300
14 Behr, M.A., Divangahi, M. and Lalande, J.D. (2010) What's in a name? The (mis)labelling of Crohn's as an autoimmune disease. Lancet 376, 202-203
15 Somers, E.C. et al. (2006) Autoimmune diseases co-occurring within individuals and within families: a systematic review. Epidemiology 17, 202-217
16 Thorsby, E. and Lie, B.A. (2005) HLA associated genetic predisposition to autoimmune diseases: genes involved and possible mechanisms. Transplant Immunology 14, 175-182
17 Manolio, T.A. et al. (2009) Finding the missing heritability of complex diseases. Nature 461, 747-753
18 Eichler, E.E. et al. (2010) Missing heritability and strategies for finding the underlying causes of complex disease. Nature Reviews Genetics 11, 446-450
19 Miller, F.W. et al. (2012) Epidemiology of environmental exposures and human autoimmune diseases: findings from a National Institute of Environmental Health Sciences Expert Panel Workshop. Journal of Autoimmunity 39, 259-271
20 Arnson, Y., Shoenfeld, Y. and Amital, H. (2010) Effects of tobacco smoke on immunity, inflammation and autoimmunity. Journal of Autoimmunity 34, J258-J265
21 Sugiyama, D. et al. (2010) Impact of smoking as a risk factor for developing rheumatoid arthritis: a meta-analysis of observational studies. Annals of the Rheumatic Diseases 69, 70-81
22 Antonovsky, A. et al. (1965) Epidemiologic study of multiple sclerosis in Israel. I. An overall review of methods and findings. Archives of Neurology 13, 183-193
23 Costenbader, K.H. et al. (2004) Cigarette smoking and the risk of systemic lupus erythematosus: a meta-analysis. Arthritis Rheumatology 50, 849-857
24 Wiersinga, W.M. (2013) Smoking and thyroid. Clinical Endocrinology (Oxf) 79, 145-151
25 Hart, P.H., Gorman, S. and Finlay-Jones, J.J. (2011) Modulation of the immune system by UV radiation: more than just the effects of vitamin D? Nature Reviews Immunology 11, 584-596
26 Hammond, S.R. et al. (1988) The epidemiology of multiple sclerosis in three Australian cities: Perth, Newcastle and Hobart. Brain 111(Pt 1), 1-25
27 Kurtzke, J.F., Beebe, G.W. and Norman, J.E., Jr. (1979) Epidemiology of multiple sclerosis in U.S. veterans: 1. Race, sex, and geographic distribution. Neurology 29(9 Pt 1), 1228-1235
28 Taylor, B.V. et al. (2010) MS prevalence in New Zealand, an ethnically and latitudinally diverse country. Multiple Sclerosis 16, 1422-1431
29 Vukusic, S. et al. (2007) Regional variations in the prevalence of multiple sclerosis in French farmers. Journal of Neurology, Neurosurgery and Psychiatry 78, 707-709
30 van der Mei, I.A. et al. (2003) Past exposure to sun, skin phenotype, and risk of multiple sclerosis: case–control study. BMJ 327, 316
31 Dobson, R., Giovannoni, G. and Ramagopalan, S. (2012) The month of birth effect in multiple sclerosis: systematic review, meta-analysis and effect of latitude. Journal of Neurology, Neurosurgery and Psychiatry 84, 427-32
32 Staples, J., Ponsonby, A.L. and Lim, L. (2010) Low maternal exposure to ultraviolet radiation in pregnancy, month of birth, and risk of multiple sclerosis in offspring: longitudinal analysis. BMJ 340, c1640
33 van der Mei, I.A. et al. (2007) Vitamin D levels in people with multiple sclerosis and community

Accession information: doi:10.1017/erm.2014.5; Vol. 16; e4; March 2014
© Cambridge University Press 2014

https://doi.org/10.1017/erm.2014.5 Published online by Cambridge University Press

controls in Tasmania, Australia. Journal of Neurology 254, 581-590
34 Munger, K.L. et al. (2006) Serum 25-hydroxyvitamin D levels and risk of multiple sclerosis. JAMA 296, 2832-2838
35 Munger, K.L. et al. (2004) Vitamin D intake and incidence of multiple sclerosis. Neurology 62, 60-65
36 Diabetes Epidemiology Research International Group (1988) Geographic patterns of childhood insulin-dependent diabetes mellitus. Diabetes Epidemiology Research International Group. Diabetes 37, 1113-1119
37 EURODIAB ACE Study Group (2000) Variation and trends in incidence of childhood diabetes in Europe. EURODIAB ACE Study Group. Lancet 355, 873-876
38 Armitage, E.L. et al. (2004) Incidence of juvenile-onset Crohn's disease in Scotland: association with northern latitude and affluence. Gastroenterology 127, 1051-1057
39 Ramos-Remus, C. et al. (2007) Latitude gradient influences the age of onset in rheumatoid arthritis patients. Clinical Rheumatology 26, 1725-1728
40 Vieira, V.M. et al. (2010) Association between residences in U.S. northern latitudes and rheumatoid arthritis: a spatial analysis of the Nurses' Health Study. Environmental Health Perspectives 118, 957-961
41 Disanto, G. et al. (2012) Month of birth, vitamin D and risk of immune mediated disease: a case control study. BMC Medicine 10, 69
42 Kahn, H.S. et al. (2009) Association of type 1 diabetes with month of birth among U.S. youth: the SEARCH for Diabetes in Youth Study. Diabetes Care 32, 2010-2015
43 Rothwell, P.M. et al. (1996) Seasonality of birth of patients with childhood diabetes in Britain. BMJ 312, 1456-1457
44 Bener, A. et al. (2009) High prevalence of vitamin D deficiency in type 1 diabetes mellitus and healthy children. Acta Diabetologia 46, 183-189
45 EURODIAB Substudy 2 Study Group (1999) Vitamin D supplement in early childhood and risk for Type I (insulin-dependent) diabetes mellitus. The EURODIAB Substudy 2 Study Group. Diabetologia 42, 51-54
46 Hypponen, E. et al. (2001) Intake of vitamin D and risk of type 1 diabetes: a birth-cohort study. Lancet 358, 1500-1503
47 Kroger, H., Penttila, I.M. and Alhava, E.M. (1993) Low serum vitamin D metabolites in women with rheumatoid arthritis. Scandinavian Journal of Rheumatology 22, 172-177
48 Merlino, L.A. et al. (2004) Vitamin D intake is inversely associated with rheumatoid arthritis: results from the Iowa Women's Health Study. Arthritis Rheumatology 50, 72-77
49 Oelzner, P. et al. (1999) Relationship between soluble markers of immune activation and bone turnover in post-menopausal women with rheumatoid arthritis. Rheumatology (Oxford) 38, 841-847
50 Stene, L.C. et al. (2000) Use of cod liver oil during pregnancy associated with lower risk of Type I diabetes in the offspring. Diabetologia 43, 1093-1098
51 Lucas, R.M. et al. (2011) Epstein-Barr virus and multiple sclerosis. Journal of Neurology, Neurosurgery and Psychiatry 82, 1142-1148
52 Handel, A.E. et al. (2010) An updated meta-analysis of risk of multiple sclerosis following infectious mononucleosis. PLoS One 5, e12496
53 Ascherio, A. and Munch, M. (2000) Epstein-Barr virus and multiple sclerosis. Epidemiology 11, 220-224
54 Wagner, H.J. et al. (2000) Altered prevalence and reactivity of anti-Epstein-Barr virus antibodies in patients with multiple sclerosis. Viral Immunology 13, 497-502
55 Draborg, A.H., Duus, K. and Houen, G. (2012) Epstein-Barr virus and systemic lupus erythematosus. Clinical and Developmental Immunology 2012, 370516
56 Meron, M.K. et al. (2010) Infectious aspects and the etiopathogenesis of rheumatoid arthritis. Clinical Reviews in Allergy and Immunology 38, 287-291
57 Chen, Y.S. et al. (2006) Parvovirus B19 infection in patients with rheumatoid arthritis in Taiwan. Journal of Rheumatology 33, 887-891
58 Kozireva, S.V. et al. (2008) Incidence and clinical significance of parvovirus B19 infection in patients with rheumatoid arthritis. Journal of Rheumatology 35, 1265-1270
59 Peterlana, D. et al. (2003) The presence of parvovirus B19 VP and NS1 genes in the synovium is not correlated with rheumatoid arthritis. Journal of Rheumatology 30, 1907-1910
60 Chen, H.H. et al. (2013) Association between a history of periodontitis and the risk of rheumatoid arthritis: a nationwide, population-based, case–control study. Annals of the Rheumatic Diseases 72, 1206-1211
61 Mikuls, T.R. et al. (2009) Antibody responses to *Porphyromonas gingivalis* (*P. gingivalis*) in subjects with rheumatoid arthritis and periodontitis. International Immunopharmacology 9, 38-42

Accession information: doi:10.1017/erm.2014.5; Vol. 16; e4; March 2014
© Cambridge University Press 2014

https://doi.org/10.1017/erm.2014.5 Published online by Cambridge University Press

62 Mikuls, T.R. et al. (2012) *Porphyromonas gingivalis* and disease-related autoantibodies in individuals at increased risk of rheumatoid arthritis. Arthritis Rheumatology 64, 3522-3530
63 Hasni, S., Ippolito, A. and Illei, G.G. (2011) *Helicobacter pylori* and autoimmune diseases. Oral Diseases 17, 621-627
64 Yeung, W.C., Rawlinson, W.D. and Craig, M.E. (2011) Enterovirus infection and type 1 diabetes mellitus: systematic review and meta-analysis of observational molecular studies. BMJ 342, d35
65 Strachan, D.P. (1989) Hay fever, hygiene, and household size. BMJ 299, 1259-1260
66 Okada, H. et al. (2010) The 'hygiene hypothesis' for autoimmune and allergic diseases: an update. Clinical and Experimental Immunology 160, 1-9
67 Ponsonby, A.L. et al. (2005) Exposure to infant siblings during early life and risk of multiple sclerosis. JAMA 293, 463-469
68 Cardwell, C.R. et al. (2011) Birth order and childhood type 1 diabetes risk: a pooled analysis of 31 observational studies. International Journal of Epidemiology 40, 363-374
69 Selmi, C. and Tsuneyama, K. (2010) Nutrition, geoepidemiology, and autoimmunity. Autoimmunity Reviews 9, A267-A270
70 Strickland, F.M. et al. (2013) Diet influences expression of autoimmune associated genes and disease severity by epigenetic mechanisms in a transgenic lupus model. Arthritis Rheumatology 65, 1872-1881
71 Brown, K. et al. (2012) Diet-induced dysbiosis of the intestinal microbiota and the effects on immunity and disease. Nutrients 4, 1095-1119
72 Wu, C. et al. (2013) Induction of pathogenic TH17 cells by inducible salt-sensing kinase SGK1. Nature 496, 513-517
73 Kleinewietfeld, M. et al. (2013) Sodium chloride drives autoimmune disease by the induction of pathogenic TH17 cells. Nature 496, 518-522
74 Virtanen, S.M. et al. (2012) Food consumption and advanced beta cell autoimmunity in young children with HLA-conferred susceptibility to type 1 diabetes: a nested case–control design. American Journal of Clinical Nutrition 95, 471-478
75 Fasano, A. and Catassi, C. (2012) Clinical practice. Celiac disease. New England Journal of Medicine 367, 2419-2426
76 Klareskog, L. et al. (2011) Smoking, citrullination and genetic variability in the immunopathogenesis of rheumatoid arthritis. Seminars in Immunology 23, 92-98
77 Rothman, K.J. (1976) Causes. American Journal of Epidemiology 104, 587-592
78 Lanctot, C. et al. (2007) Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. Nature Reviews Genetics 8, 104-115
79 Wang, X., Elston, R.C. and Zhu, X. (2010) The meaning of interaction. Human Heredity 70, 269-277
80 de Mutsert, R. et al. (2009) The effect of joint exposures: examining the presence of interaction. Kidney International 75, 677-681
81 Vandenbroucke, J.P. et al. (2007) Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. PLoS Medicine 4, e297
82 Knol, M.J. et al. (2009) When one depends on the other: reporting of interaction in case–control and cohort studies. Epidemiology 20, 161-166
83 Dwyer, T. et al. (2004) Measuring environmental factors can enhance the search for disease causing genes? Journal of Epidemiology and Community Health 58, 613-615
84 Khoury, M.J. and Wacholder, S. (2009) Invited commentary: from genome-wide association studies to gene-environment-wide interaction studies – challenges and opportunities. American Journal of Epidemiology 169, 227-230; discussion 34–5
85 Lindstrom, S. et al. (2009) The impact of gene-environment dependence and misclassification in genetic association studies incorporating gene-environment interactions. Human Heredity 68, 171-181
86 Williamson, E. et al. (2010) Effect of including environmental data in investigations of gene-disease associations in the presence of qualitative interactions. Genetic Epidemiology 34, 552-560
87 Hancock, D.B. et al. (2012) Genome-wide joint meta-analysis of SNP and SNP-by-smoking interaction identifies novel loci for pulmonary function. PLoS Genetics 8, e1003098
88 Boffetta, P. et al. (2012) Recommendations and proposed guidelines for assessing the cumulative evidence on joint effects of genes and environments on cancer occurrence in humans. International Journal of Epidemiology 41, 686-704
89 Foley, D.L. et al. (2009) Prospects for epigenetic epidemiology. American Journal of Epidemiology 169, 389-400
90 Breitling, L.P. et al. (2011) Tobacco-smoking-related differential DNA methylation: 27K discovery and replication. American Journal of Human Genetics 88, 450-457
91 Wan, E.S. et al. (2012) Cigarette smoking behaviors and time since quitting are associated with

Accession information: doi:10.1017/erm.2014.5; Vol. 16; e4; March 2014
© Cambridge University Press 2014

https://doi.org/10.1017/erm.2014.5 Published online by Cambridge University Press

differential DNA methylation across the human genome. Human Molecular Genetics 21, 3073-3082
92 Feil, R. and Fraga, M.F. (2012) Epigenetics and the environment: emerging patterns and implications. Nature Reviews Genetics 13, 97-109
93 Gregersen, P.K., Silver, J. and Winchester, R.J. (1987) The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. Arthritis Rheumatology 30, 1205-1213
94 Padyukov, L. et al. (2004) A gene-environment interaction between smoking and shared epitope genes in HLA-DR provides a high risk of seropositive rheumatoid arthritis. Arthritis Rheumatology 50, 3085-3092
95 Kallberg, H. et al. (2007) Gene-gene and gene-environment interactions involving HLA-DRB1, PTPN22, and smoking in two subsets of rheumatoid arthritis. American Journal of Human Genetics 80, 867-875
96 Karlson, E.W. et al. (2010) Gene-environment interaction between HLA-DRB1 shared epitope and heavy cigarette smoking in predicting incident rheumatoid arthritis. Annals of the Rheumatic Diseases 69, 54-60
97 Mahdi, H. et al. (2009) Specific interaction between genotype, smoking and autoimmunity to citrullinated alpha-enolase in the etiology of rheumatoid arthritis. Nature Genetics 41, 1319-1324
98 Morgan, A.W. et al. (2009) Reevaluation of the interaction between HLA-DRB1 shared epitope alleles, PTPN22, and smoking in determining susceptibility to autoantibody-positive and autoantibody-negative rheumatoid arthritis in a large UK Caucasian population. Arthritis Rheumatology 60, 2565-2576
99 Pedersen, M. et al. (2007) Strong combined gene–environment effects in anti-cyclic citrullinated peptide-positive rheumatoid arthritis: a nationwide case–control study in Denmark. Arthritis Rheumatology 56, 1446-1453
100 Too, C.L. et al. (2012) Smoking interacts with HLA-DRB1 shared epitope in the development of anti-citrullinated protein antibody-positive rheumatoid arthritis: results from the Malaysian Epidemiological Investigation of Rheumatoid Arthritis (MyEIRA). Arthritis Research and Therapy 14, R89
101 van der Helm-van Mil, A.H. et al. (2007) The HLA-DRB1 shared epitope alleles differ in the interaction with smoking and predisposition to antibodies to cyclic citrullinated peptide. Arthritis Rheumatology 56, 425-432
102 van der Woude, D. et al. (2010) Gene-environment interaction influences the reactivity of autoantibodies to citrullinated antigens in rheumatoid arthritis. Nature Genetics 42, 814-816; author reply 16
103 Willemze, A. et al. (2011) The interaction between HLA shared epitope alleles and smoking and its contribution to autoimmunity against several citrullinated antigens. Arthritis Rheumatology 63, 1823-1832
104 Hill, J.A. et al. (2003) Cutting edge: the conversion of arginine to citrulline allows for a high-affinity peptide interaction with the rheumatoid arthritis-associated HLA-DRB1*0401 MHC class II molecule. Journal of Immunology 171, 538-541
105 Klareskog, L. et al. (2006) A new model for an etiology of rheumatoid arthritis: smoking may trigger HLA-DR (shared epitope)-restricted immune reactions to autoantigens modified by citrullination. Arthritis Rheumatology 54, 38-46
106 Wegner, N. et al. (2010) Autoimmunity to specific citrullinated proteins gives the first clues to the etiology of rheumatoid arthritis. Immunological Reviews 233, 34-54
107 Klareskog, L. et al. (2006) Mechanisms of disease: Genetic susceptibility and environmental triggers in the development of rheumatoid arthritis. Nature Clinical Practice Rheumatology 2, 425-433
108 Rhee, I. and Veillette, A. (2012) Protein tyrosine phosphatases in lymphocyte activation and autoimmunity. Nature Immunology 13, 439-447
109 Criswell, L.A. et al. (2006) Smoking interacts with genetic risk factors in the development of rheumatoid arthritis among older Caucasian women. Annals of the Rheumatic Diseases 65, 1163-1167
110 Keenan, B.T. et al. (2010) Effect of interactions of glutathione S-transferase T1, M1, and P1 and HMOX1 gene promoter polymorphisms with heavy smoking on the risk of rheumatoid arthritis. Arthritis Rheumatology 62, 3196-3210
111 Mikuls, T.R. et al. (2012) Impact of interactions of cigarette smoking with NAT2 polymorphisms on rheumatoid arthritis risk in African Americans. Arthritis Rheumatology 64, 655-664
112 Kochi, Y. et al. (2011) PADI4 polymorphism predisposes male smokers to rheumatoid arthritis. Annals of the Rheumatic Diseases 70, 512-515
113 Costenbader, K.H. et al. (2008) Genetic polymorphisms in PTPN22, PADI-4, and CTLA-4 and risk for rheumatoid arthritis in two longitudinal cohort studies: evidence of gene-environment interactions with heavy cigarette smoking. Arthritis Research and Therapy 10, R52

Accession information: doi:10.1017/erm.2014.5; Vol. 16; e4; March 2014
© Cambridge University Press 2014

https://doi.org/10.1017/erm.2014.5 Published online by Cambridge University Press

114 Saevarsdottir, S. et al. (2011) Mannan Binding Lectin (MBL) genotypes coding for high MBL serum levels are associated with rheumatoid factor negative rheumatoid arthritis in never smokers. Arthritis Research and Therapy 13, R65
115 Karouzakis, E. et al. (2009) DNA hypomethylation in rheumatoid arthritis synovial fibroblasts. Arthritis Rheumatology 60, 3613-3622
116 Nile, C.J. et al. (2008) Methylation status of a single CpG site in the IL6 promoter is related to IL6 messenger RNA levels and rheumatoid arthritis. Arthritis Rheumatology 58, 2686-2693
117 Takami, N. et al. (2006) Hypermethylated promoter region of DR3, the death receptor 3 gene, in rheumatoid arthritis synovial cells. Arthritis Rheumatology 54, 779-787
118 Liu, Y. et al. (2013) Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. Nature Biotechnology 31, 142-147
119 Nakano, K. et al. (2013) DNA methylome signature in rheumatoid arthritis. Annals of the Rheumatic Diseases 72, 110-117
120 De Jager, P.L. et al. (2008) Integrating risk factors: HLA-DRB1*1501 and Epstein-Barr virus in multiple sclerosis. Neurology 70(13 Pt 2), 1113-1118
121 Nielsen, T.R. et al. (2009) Effects of infectious mononucleosis and HLA-DRB1*15 in multiple sclerosis. Multiple Sclerosis 15, 431-436
122 van der Mei, I.A. et al. (2010) Human leukocyte antigen-DR15, low infant sibling exposure and multiple sclerosis: gene-environment interaction. Annals of Neurology 67, 261-265
123 Sundqvist, E. et al. (2012) Epstein-Barr virus and multiple sclerosis: interaction with HLA. Genes and Immunity 13, 14-20
124 Ramagopalan, S.V. et al. (2010) HLA-DRB1*15, low infant sibling exposure, and multiple sclerosis gene-environment interaction. Annals of Neurology 67, 694-695
125 Mechelli, R. et al. (2013) A "candidate-interactome" aggregate analysis of genome-wide association data in multiple sclerosis. PLoS One 8, e63300
126 Ramagopalan, S.V. et al. (2009) HLA-DRB1 and month of birth in multiple sclerosis. Neurology 73, 2107-2111
127 Ramagopalan, S.V. et al. (2009) Expression of the multiple sclerosis-associated MHC class II Allele HLA-DRB1*1501 is regulated by vitamin D. PLoS Genetics 5, e1000369
128 Nolan, D. et al. (2012) Contributions of vitamin D response elements and HLA promoters to multiple sclerosis risk. Neurology 79, 538-546
129 Saastamoinen, K.P., Auvinen, M.K. and Tienari, P.J. (2012) Month of birth is associated with multiple sclerosis but not with HLA-DR15 in Finland. Multiple Sclerosis 18, 563-568
130 International Multiple Sclerosis Genetics Consortium et al. (2011) Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature 476, 214-219
131 Dickinson, J.L. et al. (2009) Past environmental sun exposure and risk of multiple sclerosis: a role for the Cdx-2 vitamin D receptor variant in this interaction. Multiple Sclerosis 15, 563-570
132 Simon, K.C. et al. (2010) Combined effects of smoking, anti-EBNA antibodies, and HLA-DRB1*1501 on multiple sclerosis risk. Neurology 74, 1365-1371
133 Hedstrom, A.K. et al. (2011) Smoking and two human leukocyte antigen genes interact to increase the risk for multiple sclerosis. Brain 134(Pt 3), 653-664
134 Kiyohara, C. et al. (2009) Cigarette smoking, N-acetyltransferase 2 polymorphisms and systemic lupus erythematosus in a Japanese population. Lupus 18, 630-638
135 Kiyohara, C. et al. (2009) Cigarette smoking, STAT4 and TNFRSF1B polymorphisms, and systemic lupus erythematosus in a Japanese population. Journal of Rheumatology 36, 2195-2203
136 Kiyohara, C. et al. (2012) Risk modification by CYP1A1 and GSTM1 polymorphisms in the association of cigarette smoking and systemic lupus erythematosus in a Japanese population. Scandinavian Journal of Rheumatology 41, 103-109
137 Griffiths, H.R. (2008) Is the generation of neo-antigenic determinants by free radicals central to the development of autoimmune rheumatoid disease? Autoimmunity Reviews 7, 544-549
138 Fraser, P.A. et al. (2003) Glutathione S-transferase M null homozygosity and risk of systemic lupus erythematosus associated with sun exposure: a possible gene-environment interaction for autoimmunity. Journal of Rheumatology 30, 276-282
139 Zhang, Y. et al. (2013) Impaired DNA methylation and its mechanisms in CD4(+)T cells of systemic lupus erythematosus. Journal of Autoimmunity 41, 92-99
140 Richardson, B. et al. (1990) Evidence for impaired T cell DNA methylation in systemic lupus erythematosus and rheumatoid arthritis. Arthritis Rheumatology 33, 1665-1673
141 Lu, Q. et al. (2002) Demethylation of ITGAL (CD11a) regulatory sequences in systemic lupus

Accession information: doi:10.1017/erm.2014.5; Vol. 16; e4; March 2014
© Cambridge University Press 2014

https://doi.org/10.1017/erm.2014.5 Published online by Cambridge University Press

erythematosus. Arthritis Rheumatology 46, 1282-1291
142 Kaplan, M.J. et al. (2004) Demethylation of promoter regulatory elements contributes to perforin overexpression in CD4+ lupus T cells. Journal of Immunology 172, 3652-3661
143 Javierre, B.M. et al. (2010) Changes in the pattern of DNA methylation associate with twin discordance in systemic lupus erythematosus. Genome Research 20, 170-179
144 Jeffries, M. et al. (2011) Genome-wide DNA methylation patterns in CD4+ T cells from patients with systemic lupus erythematosus. Epigenetics 6, 593-601
145 Wang, G.S. et al. (2009) Ultraviolet B exposure of peripheral blood mononuclear cells of patients with systemic lupus erythematosus inhibits DNA methylation. Lupus 18, 1037-1044
146 Wysenbeek, A.J., Block, D.A. and Fries, J.F. (1989) Prevalence and expression of photosensitivity in systemic lupus erythematosus. Annals of the Rheumatic Diseases 48, 461-463
147 Helbig, K.L. et al. (2012) A case-only study of gene-environment interaction between genetic susceptibility variants in NOD2 and cigarette smoking in Crohn's disease aetiology. BMC Medical Genetics 13, 14
148 van der Heide, F. et al. (2010) Differences in genetic background between active smokers, passive smokers, and non-smokers with Crohn's disease. American Journal of Gastroenterology 105, 1165-1172
149 Nimmo, E.R. et al. (2012) Genome-wide methylation profiling in Crohn's disease identifies altered epigenetic regulation of key host defense mechanisms including the Th17 pathway. Inflammatory Bowel Disease 18, 889-899
150 Zeilinger, S. et al. (2013) Tobacco smoking leads to extensive genome-wide changes in DNA methylation. PLoS One 8, e63812
151 Skorka, A. et al. (2005) Lymphoid tyrosine phosphatase (PTPN22/LYP) variant and Graves' disease in a Polish population: association and gene dose-dependent correlation with age of onset. Clinical Endocrinology (Oxf) 62, 679-682
152 Jurecka-Lubieniecka, B. et al. (2013) Association between age at diagnosis of Graves' disease and variants in genes involved in immune response. PLoS One 8, e59349
153 Hasham, A. and Tomer, Y. (2012) Genetic and epigenetic mechanisms in thyroid autoimmunity. Immunology Research 54, 204-213
154 Badenhoop, K. et al. (2009) MHC-environment interactions leading to type 1 diabetes: feasibility of an analysis of HLA DR-DQ alleles in relation to manifestation periods and dates of birth. Diabetes, Obesity and Metabolism 11(Suppl 1), 88-91
155 Rakyan, V.K. et al. (2011) Identification of type 1 diabetes-associated DNA methylation variable positions that precede disease diagnosis. PLoS Genetics 7, e1002300
156 Sollid, L.M. and Jabri, B. (2011) Celiac disease and transglutaminase 2: a model for posttranslational modification of antigens and HLA association in the pathogenesis of autoimmune disorders. Current Opinion in Immunology 23, 732-738
157 Vader, W. et al. (2003) The HLA-DQ2 gene dose effect in celiac disease is directly related to the magnitude and breadth of gluten-specific T cell responses. Proceedings of the National Academy of Sciences of the United States of America 100, 12390-12395
158 Lu, Q. (2013) The critical importance of epigenetics in autoimmunity. Journal of Autoimmunity 41, 1-5
159 Yu, Q. and Huang, J.F. (2013) The DEER database: a bridge connecting drugs, environmental effects, and regulations. Gene 520, 98-105
160 Kelly, T.K., De Carvalho, D.D. and Jones, P.A. (2010) Epigenetic modifications as therapeutic targets. Nature Biotechnology 28, 1069-1078
161 Song, M., Lee, K.M. and Kang, D. (2011) Breast cancer prevention based on gene-environment interaction. Molecular Carcinogenesis 50, 280-290
162 Bookman, E.B. et al. (2011) Gene-environment interplay in common complex diseases: forging an integrative model-recommendations from an NIH workshop. Genetic Epidemiology 35, 217-225
163 Ioannidis, J.P., Thomas, G. and Daly, M.J. (2009) Validating, augmenting and refining genome-wide association signals. Nature Reviews Genetics 10, 318-329
164 Cornelis, M.C. et al. (2010) The Gene, Environment Association Studies consortium (GENEVA): maximizing the knowledge obtained from GWAS by collaboration across studies of multiple conditions. Genetic Epidemiology 34, 364-372
165 Lin, R. et al. (2013) Novel modulating effects of PKC family genes on the relationship between serum vitamin D and relapse in multiple sclerosis. Journal of Neurology, Neurosurgery and Psychiatry. Published online first July 18, doi:10.1136/jnnp-2013-305245
166 Aschard, H. et al. (2012) Challenges and opportunities in genome-wide environmental interaction (GWEI) studies. Human Genetics 131, 1591-1613

Accession information: doi:10.1017/erm.2014.5; Vol. 16; e4; March 2014
© Cambridge University Press 2014

https://doi.org/10.1017/erm.2014.5 Published online by Cambridge University Press

Features associated with this article

Figures

Figure 1. The involvement of genetic and environmental factors in disease.
Figure 2. Proposed mechanism through which exposure to tobacco smoke, in interaction with genetic factors, may lead to the development of rheumatoid arthritis.

Table

Table 1. Direct evidence of G–E interaction in autoimmune disease – selected interactions examined by two or more studies.

Citation details for this article

Justine A Ellis, Andrew S Kemp and Anne-Louise Ponsonby (2014) Gene–environment interaction in autoimmune disease. Expert Rev. Mol. Med. Vol. 16, e4, March 2014, doi:10.1017/erm.2014.5

Accession information: doi:10.1017/erm.2014.5; Vol. 16; e4; March 2014
© Cambridge University Press 2014

https://doi.org/10.1017/erm.2014.5 Published online by Cambridge University Press