Clustering and commonalities among autoimmune diseases

Ian R. Mackay*

Department of Biochemistry & Molecular Biology, Monash University, Clayton Victoria 3800, Australia

Abstract

The concept that autoimmune diseases are characterized by shared (common) threads is well illustrated by their propensity to co-associate in a patient or direct relatives, as coexistences or overlaps. Recognized are two major autoimmune clusters, “thyrogastric” (mostly organ-specific) and “lupus-associated” (mostly multisystem). Additionally, some autoimmune diseases distribute within either cluster and a few appear not to associate. Also, within each cluster there are overlaps constituting virtually a distinct syndrome. These patterns of coexistence/overlaps depend predominantly on genetic determinants as judged by data accruing from numerous highly powered genome-wide association studies. Gene polymorphisms thus revealed include those that may determine tissue targeting particularly HLA alleles (and others), the (numerous) genes that influence orderly progression (or tolerogenesis) among immune responses from innate immunity to effector processes, genes that influence pathways of apoptosis, and genes that influence vulnerability of target organs to immune-mediated damage. A telling illustration is provided by immune-mediated colonic diseases wherein there is now a demonstrated capability to allocate the multiple risk genes to either Crohn’s disease, ulcerative colitis, or both as inflammatory bowel disease. Precisely accurate clinical data-bases and serological reactivities, combined with deepening molecular genetic analyses, have much promise for a better understanding of autoimmunity.

Keywords:
Disease clusters
Autoimmunity
Immunogenetics
Overlap syndromes
Genome-wide associations

1. Introduction

The substantial oeuvre of Noel R Rose extending from the 1950s to the present spans many aspects of autoimmunity and contains telling precepts for students of autoimmune diseases. Earliest were the famous postulates developed in 1957 with his mentor Witebsky for specifying any given disease as autoimmune [1]. More recently Rose proposed that all of the autoimmune disorders are unified by ‘shared (common) threads’ [2]. He therefore advocated a comprehensive approach to the study and treatment of autoimmune diseases from those agencies less concerned hitherto with autoimmunity [3].

Yet, despite the currently intense genetic and population research on autoimmunity, few studies have addressed the actuality of, or explanation for patterns of autoimmune clustering across the spectrum of the 80 or more implicated diseases. So, prompted by Rose’s ideas, this review will examine exemplary associations and patterns of clustering and seek interpretations. Of historical interest, an early (1966) perception on associations of autoimmune disease is shown in Fig. 1 [5].

* Corresponding author. Tel.: +61 03 9905 1437.
E-mail address: ian.mackay@med.monash.edu.au

© 2009 Elsevier Ltd. All rights reserved.
1.1. Issues relevant to definition of autoimmune disease and clusters

1.1.1. Nomenclature of diseases

First, there is the haphazard nomenclature of disease, calling for a latter-day Linnaeus for a logical taxonomy in the clinical arena, applicable particularly to the autoimmune diseases. Nomenclature includes the ‘traditional’ (lupus, diabetes, rheumatoid), ‘eponymous’ (Addison, Hashimoto, Sjogren), and ‘descriptive’ (multiple sclerosis, myasthenia gravis, polymyositis); only for a few such diseases is the putative pathogenetic nomenclature in place i.e. autoimmune hemolytic anemia, autoimmune gastritis, autoimmun hepatitis. Two schemes of definition have been recognized [6]: ‘essentialist’ that hankers for a critical pathogenetic element as a linch-pin for diagnosis sometimes provided by histopathology, or ‘nominalist’ that pragmatically settles for a combination of descriptive elements (criteria) assembled by an expert committee. Expectedly, many of the autoimmune diseases, rheumatoid arthritis (RA), systemic lupus erythematosus (SLE, lupus), Sjogren’s syndrome (SiS), autoimmune hepatitis (AIH) and others are currently defined according to nominalist criteria but these of their nature fail to capture formes frustes or even early disease, at which modern therapy should be optimally effective: take RA as an example.

Second, there is pathogenetic heterogeneity among broadly defined clinical entities that share the same designation. As a recent example, neuromylitis optica can now be separated off from the broad rubric of ‘multiple sclerosis’ [7]. Heterogeneity has been met by specifying subtypes of disease i.e., types 1 and 2 as for diabetes, or types A and B as for chronic gastritis, allowing for the appropriate subtype to be attributed to a relevant autoimmune category. Yet heterogeneity can exist even within subtypes, illustrated by autoimmune diabetes (type 1), diabetes, multiple sclerosis, rheumatoid arthritis, ulcerative colitis etc. Conversely, overlap specifies features of two different diseases affecting the same organ or system to such a degree that the phenotype is an inseparable combination of each of these entities, clinically and histopathologically, and often serologically as well. Examples include among the endocrine diseases the overlap of Hashimoto’s disease and thyrotoxicosis (‘hashitoxicosis’), among the multisystem rheumatic diseases the overlap of features of RA and SLE (‘rhupus’) and, among gastroenterological diseases, the overlap of AIH and primary biliary cirrhosis (PBC), and the overlap of ulcerative colitis (UC) and Crohn’s disease (CrD) as inflammatory bowel disease (IBD).

Fourth, the methods for, and reliability of data collection are critically important, in particular the reliability of data used to assess the true normal population prevalence of any given disease. Autoimmune diseases are not subject to any statutory notification although, for some, dedicated researchers have established regional or national disease registers for collection of data on prevalence. Even so, population prevalences, usually cited as cases per 10^5 individuals, do not necessarily inform on the representation of autoimmune disease within a relevant, and often relatively narrow, ‘susceptibility band’ of the population, say middle-aged to older women for PBC, or children/young adults for autoimmune diabetes. Furthermore, population prevalences of some diseases have been derived, for want of better, from hospital separations. However diagnoses thus ascertained are subject to various errors e.g. the Berkson bias [9], namely that diagnoses ascertained from patients with relatively weakly symptomatic diseases e.g. thyroiditis would have more associated diseases than diagnoses from patients with more highly symptomatic diseases. Finally results of autoimmune serological assays have been used as a surrogate index for a population prevalence of an autoimmune disease, but few have sufficient sensitivity/specificity for this purpose.

Thus, in summary, there are various issues that impact on the decision whether an apparent clustering is truly valid, or simply due to chance alone. Studies on humans face the difficulty of
generating sufficient case numbers for a statistically convincing association of two or more relatively rarely occurring diseases. A proportion of the published information in humans on autoimmune disease associations has come from limited case or family studies, or inadequately powered case collections from hospital record surveys. This is currently being overcome by multi-center consortia, particularly for ascertainment of genetic risk. These efforts are well worthwhile, since valid observations are more highly informative for an outbred population (humans) than for experimental (mostly inbred) laboratory animals. Of course, informative clustering of diseases does also characterize autoimmunity in various rodent strains (NOD and NZB mice, BB/W rats), and informative data on genetic determinants have been ascertained by inter-breeding and derivation of congenic strains [10,11].

1.2. Exemplary autoimmune disease clusters

1.2.1. Thyroiditis cluster and overlaps

Hashimoto-type thyroiditis was not the first of the autoimmune diseases to be so designated, but it soon became the paradigm following the classic Rose-Witebsky experiments in the 1950s showing thyroid autoimmunity in rabbits after immunization with autologous thyroid extract [1,12], concurrent with the discovery in patients with thyroiditis of highly raised levels of anti-thyroid antibodies [13]. Thyroiditis has the broadest propensity for autoimmune clustering, as shown in Fig. 1, beginning with studies on its association with pernicious anemia [14] as the terminal consequence of autoimmune chronic atrophic gastritis. Theretofollowing the widening use in the 1960s of indirect immunofluorescence [IFF] microscopy for autoimmune serological testing disclosed further associations with thyroiditis including Addisonian adrenal insufficiency and premature ovarian failure [15], and the coexistence of autoimmune thyroiditis/gastritis with juvenile–onset diabetes mellitus predicted autoimmunity in diabetes, well before the clinching detection of islet cell antibodies [14,16]. These associations generated new terminologies, including autoimmune polyendocrine syndrome (APS) [17] and thyroagastic (TG) cluster [18]. Still further broadening followed inclusion within the TG cluster of non-glandular diseases, notably skin diseases (alopecia, vitiligo), celiac disease (CeD) [19] and, surprisingly, the Lambert-Eaton syndrome of impaired presynaptic neuromuscular transmission [20]. Clinical observations, which came to prompt later fundamental discoveries on immune tolerance, segregated cases of APS into particular types [17]. APS-1 presented in childhood together with features of T-cell immune deficiency, with a recessive monogenic inheritance. APS-2 and the thyroiditis–diabetes variant (APS3v) presented in adult life, with polygenic inheritance. The monogenic error in APS-1 is homozygous inheritance of a deficiency mutant of the autoimmune regulator (AIRE), a transcriptional protein encoded by the AIRE gene and expressed particularly in the thymus; functionally, AIRE facilitates promiscuous tolerogenic expression of low-abundance tissue-specific autoantigens [21].

Apart from coexistence, there is also disease overlap within the TG cluster exemplified by hashitoxicosis (see above), and perhaps by the tightly associated polyendocrine combination of thyroiditis and diabetes, APS3v.

1.2.2. Lupus cluster and overlaps

The lupus cluster of non-organ-specific multisystem autoimmune diseases is mostly derivative from the now obsolete grouping of ‘connective tissue/collagen vascular’ diseases. Relatively less ‘compact’ than the thyroagastic cluster, this comprises SLE, RA, systemic sclerosis–scleroderma (SSc), polymyositis, varieties of vasculitis, Sjögren’s syndrome (SJS) and the later-defined entities, mixed connective tissue disease (MCTD) and anti-phospholipid syndrome (a-PLS). Each individually has unique pathognomonic features together with a relatively disease-specific autoimmune serological signature of predominant reactivity to one or another nuclear or cytoplasmic autoantigen. Among these diseases are discernible overlaps, often with their own particular serological signature (reviewed in [22]), and as well their individual immunogenetic signpost.

1.2.2.1. SLE and rheumatoid arthritis, ’rhupus’. The reality of rhupus is substantiated by overlap of both clinical and serological features of each of the partners [22–24]. Difficulty in judging the actual frequency is illustrated by a 25-year population study on 605 patients with well-established RA during which the cumulative incidence over time of cases of RA with > 4 features of SLE increased from 1.5% to 15.5% [25]. The sharing of immunogenetic features prompted the statement that ‘those patients who have both diseases do so because they have genetic determinants of both diseases’ [24].

1.2.2.2. Mixed connective tissues disease (MCTD). This controversial overlap designates combined features of SLE, SSc and polymyositis [26,27]. The basis for considering MCTD as a discrete disease entity includes a very high female: male sex ratio (16:1), dominant Raynauds syndrome due to the vascular–sclerotic narrowing seen in SSc, absence of anti-ds-DNA, low propensity for immune-complex formation, and infrequent renal disease relative to SLE. There is a (relatively) specific serological signature, autoantibody to the uridine (U)-rich-ribonucleoprotein (U1–RNP autoantigen [22]), itself associated with immunogenetic determinants, HLA DR2/DR4 [28].

1.2.2.3. SSc-polymyositis. This overlap attracted attention during the 1980s after recognition of a unique serological specificity, anti-PM-Scl, detectable in some 50% of cases [29]. An overall laboratory experience was that, among all instances of sero-positivity for PMScl, 59% had this overlap disease, and the remainder mostly had one or other of the two diseases singly [30]. The molecularly large PM-Scl autoantigen, referred to as the exosome, comprises a core complex of nine proteins and several associated proteins, located in many cellular compartments, mainly the nucleus, and has several autoantigenic reactants. The overlap disease, and also anti-PM-Scl, is immunogenetically linked to HLA alleles, DQA1*0501, DQB1*02 and DRB1*0301 [30].

1.2.2.4. Secondary Sjögren’s syndrome. Primary SJS itself could well claim inclusion in both of the major disease clusters, since the tissue primarily affected is epithelium of exocrine salivary and lacrimal glands, and there exist multiple extra-salivary features as well. Familial occurrence is prominent [31]. Some 5–10% of all cases of SJS are ‘secondary’, meaning coexistence with either SLE or RA, but SJS can coexist also with various other lupus-associated diseases and with thyroiditis, PBC and AIH. ’Secondary’ is actually a misnomer since SJS is not necessarily consequential to the ‘primary’ partner, SLE or RA. Indeed the reverse could pertain since SJS is the predecessor in SLE-SJS, although appears as the successor ‘primary’ partner, SLE or RA. The reality of rhupus could pertain since SJS is the predecessor in SLE-SJS, although appears as the successor in RA-SJS. The signature autoantibodies are aligned with those of the lupus–associated cluster, the molecularly characterized ribonucleoproteins (RNP)s Ro (SS-A) and La (SS-B) [32], which are (to a degree) disease-specific but not tissue (salivary) specific, and the autoantibody profile reflects that of the primary partner, meaning that SLE-SJS expresses the shared specificity of anti-Ro although lacks high affinity autoantibodies to dsDNA [32], whereas RA-SJS expresses both antibodies to RNP and to citrullinated peptides (anti-CP). Immunogenetic studies so far have not discriminated primary and secondary SJS.
1.2.2.5. Secondary anti-phospholipid syndrome (a-PLS). As pertains for SjS, a-PLS is designated ‘secondary’ when coexisting with SLE, as pertains for an estimated 10–20% of cases of a-PLS, varying according to diagnostic criteria and serological test systems. The autoantibody most associated with a-PLS, and often also demonstrable but without features of a-PLS in about one third of SLE patients, is directed against beta2-glycoprotein-I, of which the activity in vivo is implicated in the thrombotic features of a-PLS [33].

1.2.2.6. SLE and autoimmune hepatitis (formerly chronic active hepatitis). This overlap, described in the 1950s as ‘lupoid hepatitis’, gained currency following publication of numerous albeit small case series [34]. However among larger case surveys of either disease alone, the frequency of true overlap is usually <5%, although some features of SLE may accumulate over time during the course of AIH, much as pertains in rhusus. Serological overlap includes production of autoantibody to nucleosome (chromatin), but rarely to dsDNA. On the other hand, in ‘classical’ SLE, the liver histopathology merely shows ‘non-specific reactive hepatitis’, and anti-F-actin which characterizes AIH is lacking [34]. There are no immunogenetic markers that specify the overlaps cited in 1.2.2.5 and 1.2.2.6.

1.2.3. Autoimmune overlaps affecting a single organ exemplified by liver, colon

Considering the liver, the special feature of overlap is involvement of two structurally different cell systems, hepatocytes and biliary ductular epithelial cells. The overlaps involved are PBC-AIH (usually), or primary sclerosing cholangitis (PSC)-AIH. Overlap between PBC and PSC is scarcely recognized, for inexplicable reasons. Since the respective population prevalences for PBC and AIH per 10^5 (according to region) are only 25–40 and ~17 respectively, co-occurrence by chance alone is excluded. The PBC-AIH overlap, repeatedly reviewed since the 1970s [35,36], expresses mixed clinical, biochemical, histological and serological features of both diseases, and successful therapy requires a combination of the drug regimens optimal for either. The frequency of ‘overlap’ cases among case series reported as PBC or AIH alone is up to 10% for either diagnosis, perhaps rather higher for case series of PBC. Also, in overlaps, expressions of PBC usually precede those of AIH, suggesting that PBC is the ‘dominant’ partner, with features of AIH conferred by carriage of the susceptibility HLA haplotype for AIH, particularly joints and skin (psoriasis). The coexistence of UC with AIH (~10% of cases of AIH) is well recognized [34], and PSC coexists with UC in as many as 70–80% of cases [42]. However the reported coexistence of UC with PBC is low [43] yet, in a mouse model of PBC-like autoimmune cholangitis induced by genetic crippling of regulatory T-cell pathways, there is a regular concurrence of UC [44]. Further interpretation of these coexistences should emerge as GWAS studies in IBD and PBC [38] come to address these overlaps.

1.2.4. Autoimmune diseases for which coexistence is debatable: MS an example?

Some autoimmune diseases cluster only weakly, if at all: examples include MS, peripheral neuropathies, diseases of the kidney, bullous skin diseases, uveitis and cardiomyopathy but, for some at least, adequate case collections may not have been assembled.

Considering MS, in 1972 an initial uncontrolled retrospective study of 326 Australian patients revealed coexistence in nine with some other autoimmune disease, including thyroiditis in four [45]. A later case-controlled study of 117 patients and their relatives with a questionnaire design revealed moderate increases in prevalence of other autoimmune diseases including thyroiditis and RA in patients and/or relatives [46]. Further, a medical record linkage showed a 3.7-fold increase in prevalence of IBD [47]. Autoimmune diabetes was implicated in published case reports, and there was coexistence with diabetes within the relatively ‘closed’ population of Sardinia [48]. Contrariwise, a study from Denmark co-authored by Rose [49] and based on medical record linkage, examined population prevalences, co-morbidities and familial aggregations for 31 different autoimmune diseases; results corroborated known autoimmune clusterings but, for MS, were not informative; perhaps the ‘dominant’ diagnosis of MS eclipsed registration of other diagnoses (Berkson bias). Thus MS may well associate, weakly, with diseases in either of the major clusters i.e. thyroiditis, autoimmune diabetes or rheumatoid arthritis, or with IBD, but further studies are needed. Furthermore the immunogenetic studies so far are not supportive. Thus the strongest effect is exerted by HLA DR15, with epistasis others for the UC-CrD overlap e.g., the TNFSF15 gene encoding tumor necrosis factor superfamily member 15. The reader is referred to refs 40 and 41 for a more detailed overview of this impressive body of work.

Yet to be fully explored genetically are inflammatory disease coexistences affecting liver and/or colon, and other tissues particularly joints and skin (psoriasis). The coexistence of UC with AIH (~10% of cases of AIH) is well recognized [34], and PSC coexists with UC in as many as 70–80% of cases [42]. However the reported coexistence of UC with PBC is low [43] yet, in a mouse model of PBC-like autoimmune cholangitis induced by genetic crippling of regulatory T-cell pathways, there is a regular concurrence of UC [44]. Further interpretation of these coexistences should emerge as GWAS studies in IBD and PBC [38] come to address these overlaps.

1.3. The interpretation of clustering

1.3.1. A look at ‘etiology’

Clustering, coexistences and overlaps of autoimmune diseases, among affected individuals and/or their family members, and among populations, accord with Rose’s notion of ‘common threads’, and two prominent disease clusters, thyrogastric and lupus-associated, are discernible. The immunogenetic data shown in Table 1 clearly support the notion of clustering but one has to look hard to discern.
lead diseases of either major cluster, thyroiditis and SLE, themselves a rationale for allocation into separate disease clusters. Moreover the Chromosomal loci and prominent risk genes in selected autoimmune diseases.
a
Table 1

<table>
<thead>
<tr>
<th>Locus</th>
<th>Gene</th>
<th>Thyrogastric cluster</th>
<th>Lupus associated cluster</th>
<th>Non-aligned diseases</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>AITD</td>
<td>T1D</td>
<td>SLE</td>
</tr>
<tr>
<td>1p13</td>
<td>PTEN 22</td>
<td>$+$[57]</td>
<td>$+$[60]</td>
<td>$+$[73]</td>
</tr>
<tr>
<td>2q32</td>
<td>STAT4</td>
<td>$+$[74]</td>
<td>$+$[64]</td>
<td>$+$[64]</td>
</tr>
<tr>
<td>2q33</td>
<td>C6orf4</td>
<td>$+$[57]</td>
<td>$+$[60]</td>
<td>$+$[64]</td>
</tr>
<tr>
<td>2q37</td>
<td>PDCD1</td>
<td>$+$[57]</td>
<td>$+$[57]</td>
<td>$+$[57]</td>
</tr>
<tr>
<td>6p11-21</td>
<td>HLA complex</td>
<td>DR3-DQB1[57]</td>
<td>DR3-DQB1[57]</td>
<td>DR2, DR3 [62]</td>
</tr>
<tr>
<td>10p15</td>
<td>IL2RA (CD25)</td>
<td>$+$[CD] [60]</td>
<td>$+$[60]</td>
<td>$+$[60]</td>
</tr>
<tr>
<td>16q3</td>
<td>NOD2</td>
<td>$+$[62]</td>
<td>$+$[62]</td>
<td>$+$[62]</td>
</tr>
<tr>
<td>18p11</td>
<td>PTPN2</td>
<td>$+$[60]</td>
<td>$+$[60]</td>
<td>$+$[60]</td>
</tr>
</tbody>
</table>

Genes particular to one disease among some disease examples

<table>
<thead>
<tr>
<th>Locus</th>
<th>Gene</th>
<th>Thyrogastric cluster</th>
<th>Lupus associated cluster</th>
<th>Non-aligned diseases</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>AITD</td>
<td>T1D</td>
<td>SLE</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As a possible outcome, a deeper understanding of clustering should dispense with the worn cliché that this or that disease ‘is autoimmune in nature but of unknown etiology’, insofar as autoimmunity itself becomes part of ‘etiology’. ‘Etiology’ becomes comprehensible thus as inheritance of an ‘at risk’ genotype comprising female gender, MHRC risk alleles, and combinations of mutant alleles of multiple ‘low-penetrance’ tolerance/immunity (and other) genes, so creating an environment-susceptible phenotype comprising aberrant responsiveness to any one of numerous normally banal extrinsic triggers, overlaid upon which is an immeasurable quantum of random chance (‘bad luck’).

1.3.2. Inheritance of an ‘at risk’ genotype

Female susceptibility to autoimmunity is much discussed, and likely depends particularly on estrogenic hormones [53]. The possibility exists of immuoregulatory genes on the X chromosome; of interest the gene for interleukin-1 receptor associated protein (IRAK1, Xq28) is one of the many risk genes for SLE [54].

MHRC (HLA) alleles often as a haplotype confer genetic risk for most autoimmune diseases in humans (HLA) [55], and in many autoimmune strains of mice (H-2) [8]. Particular but far from exclusive risk attaches to the ‘ancestral’ HLA haplotype B8-DR3 (DRB1*0301) that is highly expressed among autoimmune-prone populations of Northern European origin [56]; a wide variety of other HLA risk alleles have been identified. Some at least can account for tissue specificity because the encoded molecules have (or lack) particular binding affinities for critical autoantigenic peptides, and so are likely also to contribute to disease clustering. However considering the B8-DR3-associated diseases overall, many including adrenalitis, thyrotoxicosis, autoimmune diabetes and CeD (together with other HLA risk alleles) distribute in the thyrogastric cluster, whereas others including SLE and SJ5s distribute in the lupus cluster, and still others including ‘thymitis-type’ MG and autoimmune hepatitis, distribute with either of the two clusters. Then there are further inconsistencies in that, for various of the constituent diseases in either cluster, the risk alleles can differ and, in overlap diseases such as hashtoxicity and rhusus (see above), there are differing HLA-associated risk alleles for each of the two partners of the overlap. For the sake of simplicity there is no need to consider the two clusters, but there are differences in linkage disequilibrium among alleles of the MHRC [57] but, clearly, there is far more to clustering than HLA alone.

Combinations of low-penetrance ‘tolerance-autoimmunity’ genes and others are being investigated by highly powered multi-center collaborations based on molecular genetics and GWA studies [58] in an increasing range of autoimmune diseases, notably MS [59], autoimmune diabetes [60], RA [60,61], SLE [62], IBD [40,41] and PBC [38]. A (non-all-inclusive) listing of genes of interest in these diseases, shown in Table 1, points to a contribution of effects of multiple low-risk alleles, although the gap between the known and still unknown genetic risk elements remains sizeable.

In a theoretical estimate of germ-line genetic loading for autoimmunity based on one new deleterious mutation per average protein-coding gene per 100,000 generations, together with the predicted existence of some several hundred immune response genes, every ‘normal’ individual could carry in the germ-line multiple potentially deleterious mutations of such genes in a heterozygous state [63]. Chance would determine the frequency in offspring of homozygosity for culprit genes.
The modest number of confirmed examples of ‘autoimmune-risk’ genes/promoters notwithstanding, those already known adequately explain the general clustering of autoimmune diseases, given their impact on critical functions of dendritic cells (DCs), T cells and B cells, but do not sufficiently explain the specific clusterings described herein. Among the examples cited in Table 1 are genes that encode molecules that activate B cells (CD40, [57]), down-regulate activated T cells (CTLA4), inhibit T-cell signaling (PTEN22), promote differentiation of Th 17 T cells (STAT4, [64]), induce interferon release in response to injury (IFIH1,[65]) or potentiate immune (and autoimmune) responses (subunits of IL-12 [38]).

1.3.3. Environment-susceptible phenotype

An environmental contribution to occurrence of autoimmunity, if not clustering, is generally accepted, for good reasons, but culprit agents, whether infectious, toxic, dietary, or products of tissue damage, remain difficult to identify. Presumably these are multiple, ubiquitous and relatively banal, and need to interact with a particular genotype such that the combination serves as the actual etiological factor. Known examples would be dietary gluten and specific HLA risk alleles for CeD [20], and insufficient sunlight and genetic variants of the vitamin D receptor resulting in loss of immunoregulatory (or other) effects of vitamin D for MS [66]; this latter factor could pertain too for other autoimmune diseases that increase in prevalence with distance from the equator.

All-in-all, the major environmental suspect is infection by effects that range from cell injury causing autoantigen exposure, epitope mimicry, and pro-inflammatory activation of antigen-presenting cells via Toll-like receptors. Synergistic effects of HLA alleles are also relevant here. Numerous infectious agents have come under suspicion, including viruses (Coxsackie B virus in autoimmune diabetic insulinitis and Epstein–Barr virus infection or an aberration of the host response to it in SLE, RA, Sjogren’s syndrome and MS), bacteria (Streptococci in rheumatic carditis, Campylobacter in peripheral neuropathy, intestinal microflora in CrD) or parasites (Trypanosoma cruzi in Chargas-type cardiomyopathy). However, conceding that an infectious agent may serve as the initial trigger, these environmental agents per se seem hardly sufficient to explain patterns of clustering unless infection of multiple tissues were to occur.

1.3.4. Random chance

The contribution of chance [67] is suggested by modest concordances for various autoimmune diseases in monozygous twins and differing rates of progression to diabetes in similarly-housed inbred NOD mice. Chance could operate in human autoimmunity at various levels: a) individual, depending on degrees of parental transmission to offspring of defective tolerance genes, randomness of individual exposure to infections or toxins, or somatic mutations in hemopoietic cell lineages that inactivate a single functional copy of an immune response gene; b) cellular, depending on the numerous critical cell–cell interactions needed for orderly progression of innate to adaptive immunity and on to an effective immune response; and c) molecular, depending on ligand–receptor interactions, intracellular signalling and transcriptional regulation of gene activation, and here could be included disordered processes of apoptosis.

1.4. Determinants of clustering and overlap among autoimmune diseases

First, to recapitulate: many but by no means all of the autoimmune diseases distribute in two discernible clusters, among which exist discrete disease associations or overlaps. The organ-specific thyrogastric diseases exhibit autoimmunity to one or a few intracellularly sequestered autoantigens (often functional enzymes) to which responses are usually experimentally inducible by immunization with tissue extracts. The multisystem lupus-associated diseases exhibit autoimmunity to ubiquitous non-sequestered autoantigens (particularly prevalent nuclear and cytoplasmic moieties derived from products of cellular degradation) to which responses are poorly inducible by immunization (with the exception here of arthritis induced by immunization with type II collagen).

The specificity of tissue targeting in the first place, given a genetically permissive host, depends primarily on the site and nature of some incidental initiating tissue damage, with spillage of provocative autoantigens in an inflammatory milieu. However, our particular concern is seeking explanations for disease clustering since, even allowing for possible intermolecular epitope spreading, the structural diversity of the various autoantigenic molecules, within either cluster, demands additional processes, very likely genetic. Relevant genes can be arbitrarily allocated to four groups: those contributing to i) tissue targeting, ii) tolerance maintenance, iii) apoptosis, and iv) ‘end-organ’ vulnerability. Group i) includes insulin gene INS polymorphisms and mode of thymic expression of the gene product as a contributor to autoimmune diabetes [57]; thyroglobulin gene Tg polymorphisms that result in pro-autoimmunogenic sequence variations in Tg [57], the PADI4 gene that influences production of citrullinated TG proteins as important autoimmune antigens in RA [61], or genes that could influence propensity to develop a particular autoantibody e.g. sI I in mice encoding ANA as a first step in the evolution of SLE [62]. Additionally, of course, particular HLA risk alleles will influence which, and the means whereby, tissue-specific autoantigens can be presented to the immune system, in thymus or periphery.

Group ii) includes the numerous genes required for orderly progression of adaptive immune responses or tolerance, already described herein as ‘tolerance-autoimmunity’ genes. Among these is AIRE which projects exposure in the thymus of tissue-specific autoantigens, with homozygous deficiency resulting in APS-1; however the intuitive likelihood of heterozygous mutant AIRE contributing to thyrogastric tissue-specific autoimmunity is not established. Notably, particular combinations of tolerance-autoimmunity genes may identify subsets within disease clusters, e.g. APS3v [57]; thus clinical recognition of a particular disease subset should prompt a search for the corresponding genetic variant, and vice versa.

Group iii), apoptosis genes, impact on numerous points on apoptosis pathways, and hence various mutant apoptosis genes enhance risk for autoimmunity. Although these have not come up prominently in GWA studies in humans, there are mouse models involving the Bcl or Fas-FasL gene families that are amply illustrative. Apoptosis contributes in two ways: by inhibition of apoptotic disposal of negatively selected autoimmune lymphocytes, or by promotion after tissue damage of cell surface blebs that harbour autoimmunogenic cellular particles, termed apotopes [68,69]. It was earlier postulated that lupus-associated autoimmunity depended on presentation of ‘dynamic particles’ containing aggregates of subcellular immunogenic moieties [70], and a later comment made in 2003 was that ‘structural features of autoantigens, their location and catabolism during cell death, and their translocation to cells that can present antigens to the immune system must contribute to the selection of the repertoire of autoantigens’ [71]. Still, however, there are unanswered questions, one being on the relatively unique nuclear or cytoplasmic autoantibody specificities that are demonstrable in many of the individual lupus-associated diseases, some examples being anti-topoisomerase I in SSc, anti-centromere in limited CREST-type SSC, anti-RNA synthetases in polymyositis, anti-Ro (SS-A) and anti-La (SS-B) RNPs in Sjs, anti-PM/Sc in the SSC-polymyositis overlap, and anti-U1-RNP in MCTD, among various others.
Group iv) takes in genetically dependent single- or multi-end-organ vulnerability to damage or capability for cell survival [72], perhaps reflected by a developmental defect in hox11-derived tissues in NOD mice claimed to enhance susceptibility to diabetes [72].

1.5. Concluding remarks

Noel Rose's dictum of 'common threads' among the autoimmune diseases encompasses the characteristic clustering of these disorders into two major groups, the thyrogastric and lupus-associated clusters and other aggregations. Association in the one individual (or direct relatives) of two independent autoimmune diseases is referred to as coexistence, and features of different autoimmune diseases expressed in the one organ or tissue is referred to as overlap. Autoimmunity is determined by genetics, environment and chance, but clustering of autoimmune diseases points to the major role of genes. The intuitive expectation of many different 'pro-autoimmune' genetic influences in autoimmunity is being verified by highly powered genome-wide association studies that are identifying candidate gene loci, and single nucleotide polymorphisms at these risk-associated loci. Considered herein are genes required for tissue targeting; complex cellular interactions related to innate and adaptive immune responses and tolerance; the molecular assemblies of autoantigens in apoptotic blebs at the cell surface; and 'end-organ' vulnerability to immune-mediated tissue injury. Identification of pro-autoimmune genetic polymorphisms, in the context of clinically well-defined disease clusters/overlaps, can become 'the light on the hill' that characterizes outstanding immunologists [78–80].

Acknowledgements

The author is grateful to Elaine Pearson for her skilled assistance in the preparation of the manuscript, to Tamara Bailey for help with the graphics, and Senga Whittingham for critique of the text.

References


