Coherent Somatic Mutation in Autoimmune Disease

Kenneth Andrew Ross*

Department of Computer Science, Columbia University, New York, New York, United States of America

Abstract

Background: Many aspects of autoimmune disease are not well understood, including the specificities of autoimmune targets, and patterns of co-morbidity and cross-heritability across diseases. Prior work has provided evidence that somatic mutation caused by gene conversion and deletion at segmentally duplicated loci is relevant to several diseases. Simple tandem repeat (STR) sequence is highly mutable, both somatically and in the germ-line, and somatic STR mutations are observed under inflammation.

Results: Protein-coding genes spanning STRs having markers of mutability, including germ-line variability, high total length, repeat count and/or repeat similarity, are evaluated in the context of autoimmunity. For the initiation of autoimmune disease, antigens whose autoantibodies are the first observed in a disease, termed primary autoantigens, are informative. Three primary autoantigens, thyroid peroxidase (*TPO*), phogrin (*PTPRN2*) and filaggrin (*FLG*), include STRs that are among the eleven longest STRs spanned by protein-coding genes. This association of primary autoantigens with long STR sequence is highly significant ($p < 3.0 \times 10^{-7}$). Long STRs occur within twenty genes that are associated with sixteen common autoimmune diseases and atherosclerosis. The repeat within the TTC34 gene is an outlier in terms of length and a link with systemic lupus erythematosus is proposed.

Conclusions: The results support the hypothesis that many autoimmune diseases are triggered by immune responses to proteins whose DNA sequence mutates somatically in a coherent, consistent fashion. Other autoimmune diseases may be caused by coherent somatic mutations in immune cells. The coherent somatic mutation hypothesis has the potential to be a comprehensive explanation for the initiation of many autoimmune diseases.

Introduction

I have previously provided evidence that somatic gene conversion and/or deletion in sequence harboring long segmental duplications is correlated with disease [1]. According to this hypothesis, autoimmunity is a response to novel (somatically mutated) antigens. Others have proposed a role for somatic mutation in autoimmunity [2,3]. The remarkable extent of somatic mutation, including copy number variation and somatic mosaicism, has recently been elucidated, with several proposed links to neurological disease [4–9]. The connection between somatic mutation and autoimmunity requires that somatic mutations be coherent [1], i.e., that the same type of mutation occur in many cells, to the point that the somatically mutated protein either disrupts normal function or is noticed by the immune system as non-self. A coherent mutation may be recurrent (occuring independently in many cells) [10] or clonal (occuring once and replicating many times).

Somatic Mutation of Tandem Repeat Sequence

Coherent somatic mutation of the haptoglobin gene (*HP*) has been observed in vivo in humans [11]. Carriers of the *HP2* allele have a segmentally duplicated 1.7kb sequence fragment within the gene that includes two additional exons beyond the shorter *HP1* allele. In an *HP2* homozygote, Asakawa et al [11]. found a shorter

DNA sequence corresponding to an exact excision of one copy of the tandem repeat. In each of several *HP2* homozygotes subsequently tested, a small but measurable concentration of the shorter sequence was identified. Asakawa et al. argued that rare but regular somatic deletion events occur in vivo. In the mouse, a similar kind of somatic mutation has been observed in vivo at a longer 70 kb segmental duplication [12,13]. The mutation frequency was much higher than for *HP* in humans, presumably due to both the longer duplicon and the fact that phenotypic measurement was performed in gene-expressing tissues where mutations would be more common, rather than in blood cells [11,14]. Somatic mutation at additional loci, mediated by inverted repeats [15] or tandem repeats [16], has been observed in vivo in humans.

Long segmental duplications are not the only repetitive sequence subject to high mutation frequencies. Simple tandem repeats (STRs), including microsatellites and minisatellites that are highly mutable in germ-line cells, are also mutable in somatic cells [17,18]. Some STRs encode proteins, and somatic mutations would generate novel, potentially immunogenic proteins. While not strictly an STR, such an effect has been observed at the La antigen associated with Systemic Lupus Erythematosus (SLE) and Sjogren's Syndrome (SJ), where somatic mutations of an 8bp poly-A sequence into a 7 bp mutant have been observed [19]. These mutations correlate with autoimmunity, in that about 30% of La-

reactive SLE/SJ patients respond specifically to the mutant protein [19] and somatic mutant DNA can be detected in such individuals [20].

Other STRs occur within introns, where changes in repeat counts can change splicing behavior [21]. Altered splicing of autoantigens has been proposed as a mechanism for generating immunogenic protein variants [22]. In particular, inflammation can lead to reduced levels of the splicing factor ASF/SF2 [22]. Low levels of ASF/SF2 are associated with DNA double strand breaks and DNA rearrangements triggered by R loops between DNA and transcribed RNA [23]. R loops promote instability in GC-rich trinucleotide repeats [24], suggesting that transcribed repetitive sequence may be particularly vulnerable to somatic mutation induced by ASF/SF2 depletion.

Additionally, repeat mutations are often accompanied by significant changes in methylation [25]. Demethylation can potentially lead to aberrant transcription initiation in the middle of the gene sequence [26]. Repetitive sequence is also an essential factor in cellular mechanisms for methylating nearby sequence [27,28]. Changes to the methylation pattern can also affect splicing [29]. Altered methylation patterns have been observed in several autoimmune diseases [30].

Yet another reason to focus on somatic repeat mutations in autoimmune disease is the observation that somatic tandem repeat mutations can be induced by inflammation typical of an immune or autoimmune response [31,32]. This observation provides the basis for a feedback loop. An initial immune response against a pathogen could, as a side-effect of inflammation, trigger the initial production of aberrant protein. The aberrant protein induces a second immune response, with further inflammation and coherent somatic mutation in nearby cells (or remote cells opsonized by autoantibodies [33,34]) creating a cycle of autoimmunity. Anti-inflammatory medications reduce rates of somatic mutation in some cancers [35], further supporting a link between inflammation and somatic mutation,

Human STR sequence is overabundant near telomeres [18,36]. Nevertheless, the germ-line variability of a minisatellite repeat in a population does not depend on its chromosomal location [37]. Instead, the primary determinants of minisatellite variability are (a) the number of repeat units it contains, and (b) the degree of identity between different repeat units within the sequence [37]. Variability is a nonlinear function of these measures: Doubling the copy number increases the probability of being variable about 15-fold, and adding 10% to the repeat unit similarity increases the probability of being variable about 18-fold [37]. A more recent model also takes into account the size of the repeat unit [38]. The total repeat length (i.e., the product of the repeat unit size and the repeat count) is strongly correlated with variability [38]. For segmental duplications, high sequence identity is most important for structural variability, with high duplicon length and low duplicon separation also playing a role [39].

While somatic and germ-line microsatellite mutation patterns appear similar [18], somatic and germ-line mutation patterns differ for minisatellites [40]. Germ-line minisatellite mutations involve recombination-based repair of double strand breaks (DSBs), while spontaneous somatic minisatellite mutations arise by replication slippage or mitotic recombination [40]. For somatic mutations induced by inflammation [31,32], DNA damage appears to be critical, including DNA strand breaks [41]. The resulting mutation patterns in STRs may therefore more closely resemble germ-line mutations or somatic mutations in cancer [42] than spontaneous somatic mutations. Structural mutations in repetitive sequence are orders of magnitude more frequent than point mutations [43]. Mitotic mutation rates of up to 2% have been observed in the longest human tandem repeat sequences [44].

Autoimmunity

Autoimmune diseases have overlapping features, including shared susceptibility loci [45–48] and cross-heritability [49]. Nevertheless, each autoimmune disease has specific manifestations, causing damage to particular organs or systems. The central enigma of autoimmune disease is why a relatively small set of specific proteins are immunologically targeted [50]. Many, but not all autoantigens in systemic autoimmune diseases are proteins that are cleaved during apoptosis [51,52], but the reason for this association is unclear given that T cell tolerization to such cleaved proteins is expected [52,53]. Autoantigens appear to have longer exons and harbor more SNPs than other genes [3,54], and they are enriched in several biologically relevant categories [3].

The most prominent phenotype of autoimmune disease is the presence of specific antibodies (Tables 1 and 2). While T-cell epitopes are also implicated in autoimmunity, they are more difficult to measure [55]. Mutant protein can induce antibodies to wild-type protein, even when T-cell tolerance to wild-type protein is maintained [56]. Thus, antibodies are likely to provide the most robust signal about autoimmune targets.

A B cell epitope does not have to be from the same protein molecule as the T cell epitope in order for the B cell to be activated by a CD4+ (helper) T cell. A B cell that endocytoses a protein complex by binding to one of its proteins can be activated by a CD4+ T cell specific to another protein in the complex. Such a mechanism has been used to explain anti-TG2 antibodies in celiac disease, where a TG2-specific B cell is activated by a CD4+ T cell specific to gliadin after endocytosis of a TG2-gliadin complex [57].

Thus, a protein is a candidate CD4+ T cell target either if it elicits antibodies itself, or if an in-vivo binding partner of the protein elicits antibodies. B cell specificities (and thus antibodies) to multiple proteins can be supported by a single CD4+ T cell epitope. I use the term *peri-antigen* to mean an in-vivo binding partner of an autoantigen. A peri-antigen can potentially function as a CD4+ T-cell target supporting B cell specificity to the autoantigen.

Testing the Coherent Somatic Mutation Hypothesis

I sought data to test the hypothesis that autoimmune disease is associated with mutable repetitive sequence. Because of its construction from long contigs [58], the reference human genome has reliable sequence for most repetitive regions, although gaps still remain. Because shorter reads were used, the Celera sequence is missing the interiors of many repetitive elements [59]. Most current sequencing technologies use short reads that must be assembled into whole genomes. Both de-novo assembly and alignment-based assembly are unreliable in highly repetitive regions [60–62]. The reference human genome is therefore the primary currently available source of robust repetitive sequence throughout the genome.

Antibodies that develop early in disease progression provide the strongest evidence for a causative role for the corresponding antigen. A *primary* autoantigen is one whose antibodies have been shown, in at least a subset of cases, to be the first disease-associated antibodies to appear. A test of the coherent somatic mutation hypothesis can be formulated as follows: Is there a statistical link between primary autoantigens (and/or their peri-antigens) and genes containing highly mutable sequence?

Once such a statistical link is established, a subsequent test of the comprehensiveness of the coherent somatic mutation hypothesis would consider other mutable (e.g., long STR) sequence. To

Table 1. Twenty-one of the most prevalent human autoimmune diseases, in approximately decreasing order of prevalence [49,275].

Abbrev.	Name	*PTPN22* Assn.	B-cell Autoantigens	Refs.
GD	Graves' Disease	Yes [45]	**TPO**, **TG**, TSHR	[276]
RA	Rheumatoid Arthritis	Yes [143]	**FLG**, **VIM**, **FGA**, **FGB**, **ENO1**, IgG (rheumatoid factor), IFI16, ANXA1, PADI4	[169,178,277–282]
HT	Hashimoto's Thyroiditis	Yes [143]	**TPO**, **TG**	[276,283]
CEL	Celiac disease	Unclear [284–286]	TG2, HP, actin, CALR, TG3, ganglioside, collagen	[57,287,288]
PSO	Psoriasis	No [143]	PALLD, AGAP3, DSP, collagen-21, ATXN3	[289]
VIT	Vitiligo	Yes [290]	TYR, TH, TYRP1, MCHR1, lamin A	[291–293]
SJ	Sjogren syndrome	Unclear [294,295]	**SPTAN1**, SPTBN1, Ro52(TRIM21), Ro60(TROVE2), La(SSB), CHRNA3, IFI16, VIM, CHRM3	[178,296–302]
UC	Ulcerative Colitis	Yes [105]	**HMGB1**, **HMGB2**, **pANCA**, tropomyosin	[303–305]
AS	Ankylosing Spondylitis	No [306]	multispecific	[307]
T1D	Type-1 diabetes	Yes [143]	**PTPRN2**, **PTPRN**, INS, GAD2, SLC30A8, VAMP2, NPY; AMY2A (fulminant T1D)	[308–313]
AA	Alopecia Areata	Yes [314,315]	TH, TCHH, KRT16	[316–318]
JIA	Juvenile Idiopathic Arthritis	Yes [319,320]	DEK, HSP70, citrullinated peptides	[321–324]
PA	Pernicious Anemia	Unknown	ATP4A/ATP4B, pepsinogen A	[325–327]
MS	Multiple Sclerosis	No [143]	**MAG** *, MBP, PLP, MOG, CRYAB, CR1, neuronal antigens	[150,151,328–332]
CD	Crohn's disease	Opposite # [333]	GP2, CUDZ1	[334]
SLE	Systemic Lupus Erythematosus	Yes [143]	**Ro60(TROVE2)**, **SNRPA**, **APOH/cardiolipin-complex**, **ribosomal P**, VIM/cardiolipin-complex, La(SSB), Ro52(TRIM21), ds-DNA, Sm, SNRNP70, SNRPC, chromatin/histones, Ku, CALR, NCL, RF, CR1, IFI16, VIM, lamin B, F2, F2/Phosphatidylserine, ANXA1, ANXA2, ANXA5, NPM1, HMGB1, LTF, SR proteins, others	[178,281,299,328,335–352]
UV	Uveitis	No [353]	CRALBP, CRYAA, CRYAB, CRYBB1	[354,355]
AD	Addison's disease	Yes [356]	**CYP21A2**	[357,358]
MG	Myasthenia Gravis	Yes+ [359–362]	AChR, MUSK, LRP4, AGRN, ColQ, TTN, KCNA1, RYR	[363–365]
DM	Dermatomyositis	Yes [366]	Mi-2-complex, IFIH1, TRIM33, MORC3, Ro52(TRIM21)	[367,368]
SSc	Systemic Sclerosis	Yes [369,370]	RNA Polymerase III, CENPB, CENPA, RNA Polymerase I, RNA Polymerase II, TOP1, PM/Scl-complex, Ro52(TRIM21), SNRNP70, NOR-90, Ku, Th/To, U3RNP/FBL, IFI16, ANXA5, NPM1, HMGB1, HMGB2, Mitochondrial-M2	[211,281,299,371–375]

Autoantibodies to antigens in bold are known to be primary antibodies that occur early in disease progression, often prior to the appearance of symptoms. The tryptophan allele of the Arg620Trp polymorphism at rs2476601 in the *PTPN22* gene is associated with many autoimmune diseases, as indicated in the "*PTPN22* Assn." column. Atherosclerosis (CAD) is not universally considered an autoimmune disease, and is therefore not listed. Nevertheless, CAD does have autoimmune features [376] and an association with *PTPN22* [377–379]. *The initial pathology in some MS lesions is associated with MAG loss [329,380,381].
#The tryptophan PTPN22 allele is protective from CD [333].
+In MG, two studies conflict about whether PTPN22 is specifically associated with the subset of cases having anti-TTN antibodies.
doi:10.1371/journal.pone.0101093.t001

what extent could somatic mutation at these loci explain other autoimmune phenomena?

Results

Genes Containing Long Repeats Include Primary Autoantigens for Common Autoimmune Diseases

Using the Tandem Repeat Finder [63] (TMRF) track of the UCSC Genome Browser [64], I queried the database for protein-coding genes whose DNA sequence spans STR sequence, and filtered the results as described in the Methods. Figure 1 shows all 37 gene-internal repeats longer than 5 kb. *NSUN6*, *TTC34*, and *ANKRD36C* each contain multiple long repeats, and thus appear more than once. As previously discussed, high repeat length, high repeat count, and high repeat identity are markers of repeat mutability. Additionally, for intronic repeats, longer repeats are more likely to induce long mutations that in turn are more likely to

alter methylation and splicing. At this scale, all repeats are minisatellites with intermediate to long repeat units.

Among the eleven genes with longest repeat length are thyroid peroxidase (*TPO*); protein-tyrosine phosphatase, receptor-type, n, polypeptide 2 (*PTPRN2*); and filaggrin (*FLG*). *TPO* encodes a primary autoantigen in both Hashimoto's Thyroiditis (HT) and Graves' Disease (GD); *PTPRN2*, and *FLG* encode primary autoantigens in Type-1 Diabetes (T1D) and Rheumatoid Arthritis (RA) respectively (Table 1). The presence of three primary autoantigens among the top eleven genes is highly significant ($p < 3.0 \times 10^{-7}$, see Methods).

Additionally, the tenth ranked gene, *BRF1*, encodes an RNA-Polymerase-III (RNAP-III) initiation factor that binds to RNAP-III [65]. RNAP-III is an autoantigen specific to Systemic Sclerosis (SSc) (Table 1); *BRF1* thus encodes a peri-antigen for SSc.

Table 2. Autoantigens for selected low-prevalence autoimmune diseases.

Abbrev.	Name	Autoantigens	Refs.
PV	Pemphigus Vulgaris	DSG3, DSG1, HLA-DRA, DSC3, DSC1, ATP2C1, PKP3, CHRM3, COL21A1, ANXA8L1, CD88, CHRNE	[382,383]
RHF	Rheumatic Fever	VIM, MYBPC3, tropomyosin, collagen	[186,384]
LEMS	Lambert-Eaton Myasthenic Syndrome	CACNA1A, CACNB2	[385,386]
AH1	Autoimmune Hepatitis (type 1)	HMGB1, HMGB2	[387]
AH2	Autoimmune Hepatitis (type 2)	CYP2D6, CES1, PDIA3	[388,389]
HA	Autoimmune hemolytic anemia	RHD, GYPA	[390]
AP	Autoimmune pancreatitis	AMY2A, CA2, LTF, HSP10, plasminogen-binding protein, trypsinogens, SPINK1	[313,391]
PBC	Primary Biliary Cirrhosis	Mitochondrial-M2, SP100, PML, NUP210, Ro52(TRIM21), CENPB, SUMO2, SUMO1, CHRM3	[210,392,393]
NMO	Neuromyelitis Optica	AQP4	[329]
GPS	Goodpasture's Syndrome	COL4A3	[394]

doi:10.1371/journal.pone.0101093.t002

Autoimmune Associations of Genes with Long High-Identity Tandem Duplications

Motivated by known somatic mutation of *HP*, I looked for examples of long tandem duplications with at least 96% identity occuring within protein coding genes. Since the tandem repeat finder algorithm limits repeat units to 2000 bp, and its coverage of some longer units appears to be incomplete (e.g., it misses the 1.7 kb repeat in *HP*), such repeats may have been overlooked in the earlier analysis. The segmental duplications track of the UCSC database [64] was used as described in the Methods.

Table 3 shows all tandem duplications of total length at least 3400 bp where at least one duplicon occurs entirely within a protein-coding gene locus, and the tandem duplicons have the same orientation and are separated by at most 100 bp. Several genes appearing in Figure 1 also appear in Table 3, having long segments that are high identity tandem repeats. Of the remaining genes, five are autoantigens: complement component receptor 1 (*CR1*) in SLE and multiple sclerosis (MS); pepsinogen 4, group I (*PGA4*) in pernicious anemia (PA); titin (*TTN*) in myasthenia gravis (MG); interferon-gamma-inducible protein 16 (*IFI16*) in RA, SLE, SSc and SJ; and *HP* in celiac disease (CEL) (Table 1). The presence of five autoantigens among the top 33 genes is statistically significant ($p < 0.0004$, see Methods).

Copy number variations in the 54.7 kb STR of *CR1* (Table 3) have been associated with SLE [66] and Alzheimer's disease (ALZ) [67]. The *CR1-S* allele has three repeats (as in the human reference genome) and has a population frequency of about 15%, while the shorter *CR1-F* allele has two repeats and a frequency of 83% [67]. The repeat length is functionally important, since the repeat includes sequence that codes for complement binding sites [67]. In both SLE and ALZ, the longer *CR1-S* allele is the high-risk variant [66,67]. *CR1* plays an important immunological role in various cell types [68].

PGA4 is one of three genes in the human reference genome coding for highly similar (but not identical) versions of pepsinogen A, an autoantigen in PA. Low levels of pepsinogen A are specific in diagnosing PA [69]. Variant alleles observed in the population contain three, two or one pepsinogen A gene [70]. The other major autoantigen in PA is *ATP4A/ATP4B* (Table 1), which both interacts with and colocalizes with pepsinogen A on the parietal cell surface [71].

The *HP* gene that has been observed in vivo to be somatically mutated [11] also codes for zonulin in individuals carrying the *HP2* allele [72]. The functions of haptoglobulin and zonulin are diverse, including some specific immunological capabilities conferred by the *HP2* allele [72,73]. *HP2* alleles are overrepresented in several autoimmune diseases, coronary artery disease, and mental disorders [73–77].

## Additional Long Repeats Obtained from Self-Chain Alignments

To ensure completeness of the long repeat dataset, I queried the self-chain track of the UCSC database as described in the Methods. These alignments capture tandem repeats that may be slightly imperfect, i.e., there may be gaps between segments in the alignments, as well as repeats whose unit length is above the 2 kb threshold for TMRF. The results, shown in Table 4, are largely in agreement with Figure 1 and Table 3. Table 4 includes the following additional genes with alignments over 13 kb and exhibiting germ-line structural variation (File S1): *LPA*, *DMBT1*, *MGAM*, *KIR3DL1*, *KATNAL2*.

TTC34 is a Candidate CD4+ T Cell Antigen for Systemic Lupus Erythematosus

The gene *TTC34* is an outlier in Figure 1, both in terms of the length of the repetitive segment (an underestimate because the repeat is terminated by a gap in the human reference assembly) as well as the number of repeat units. *TTC34* encodes an uncharacterized protein that binds to PPP4C [78]. In support of a functional role for TTC34/PPP4C binding, RNAi depletion of either protein induces a common elongated cell phenotype [79].

If somatic mutation of *TTC34* induces autoimmunity, then antibodies to binding partners of TTC34/PPP4C would be expected. PPP4C is a ubiquitous serine/threonine phosphatase that regulates a variety of cellular functions [80]. Based on the localization of those cellular functions, I hypothesize that *TTC34* mutation underlies the initial pathenogenesis of SLE. Table 5 shows that many autoantigens in SLE, including known primary SLE autoantigens, associate with PPP4C. Under this hypothesis, the broad array of autoantigens in SLE is a consequence of the many functions of PPP4C, together with secondary immunogenicity caused by the aberrant clearance of apoptotic cells [52,81].

The long *TTC34* STR appears (with shorter length) in several primate species, but not in more distantly related species whose genomes have been sequenced [64]. Surprisingly, a 12 kb long STR has independently evolved in the mouse (GRCm38) genome, 3.2 kb upstream of the mouse *Ttc34* start site [82]. The mouse
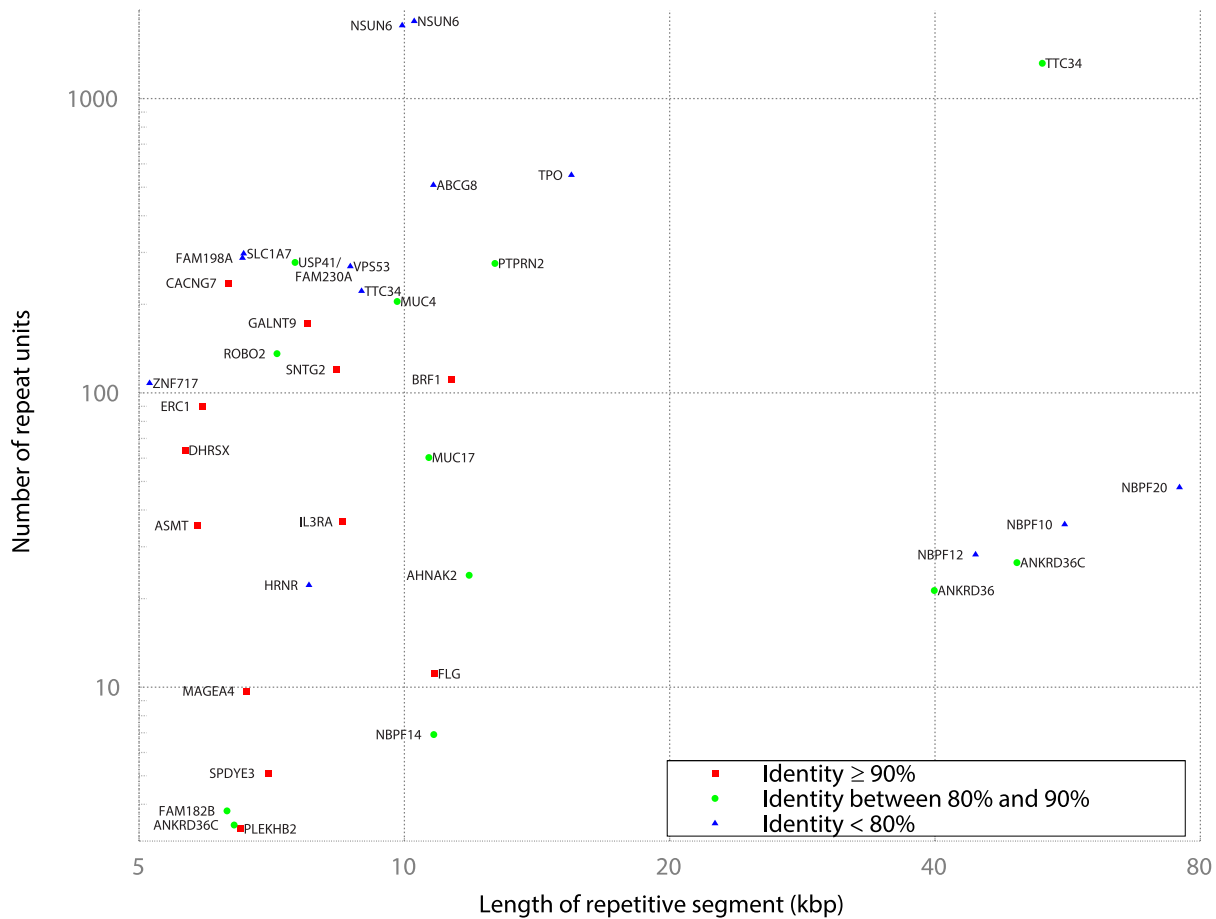
Figure 1. Genes with long internal repeats. The x-axis denotes the total length of the tandem repeat (log-scale), and the y-axis represents the number of repeat units within the tandem repeat (log-scale). The degree of repeat identity reported by TMRF is indicated by the color of the data point. Genes in bold have exonic sequence overlapping the repeat. Genes containing multiple disjoint long repeats appear more than once. doi:10.1371/journal.pone.0101093.g001

repeat unit length is 37, similar to the unit length of 40 in the human repeat. As for humans, the 12 kb mouse repeat is an outlier within the mouse genome: among all STRs that overlap a protein-coding gene locus, including a 5 kb segment upstream of the gene, the *Ttc34* repeat is the fifth longest (Table 6). The independent evolution of such a similar long repeat argues strongly for a functional role.

If the *TTC34* repeat mutates under inflammation [31], then the desired functional role would be one where changes in *TTC34* expression and/or PPP4C activity would be adaptive under inflammation. PPP4C depletion makes T cells resistant to apoptosis [83]. The association of apoptosis reduction with inflammation is biologically plausible, since T cells in inflammatory environments would be expected to receive survival signals during normal immune responses.

LPA in Atherogenesis

LPA encodes a protein that binds to ApoB-100 in LDL particles to form Lp(a) lipoprotein particles containing lipids, phospholipids and cholesterol [84]. In coronary artery disease (CAD) ApoB-100 and LDL are immune targets of T cells and antibodies [85], meaning that *LPA* encodes a peri-antigen for CAD. Under the coherent somatic mutation hypothesis, rare but regular somatic mutation to *LPA* would occur, analogously to that observed for *HP* [11]. Epitopes of the mutant protein would be presented by

immune cells in blood vessels, leading to activation of immune cells in atherosclerotic lesions [85] and autoimmune responses against other components of Lp(a) lipoprotein particles. *LPA* is central to CAD pathenogenesis, since an elevated plasma Lp(a) lipoprotein level predicts stroke and vascular disease, particularly in men [86,87]. SNPs in *LPA* have the largest known effect on CAD risk [88], including an odds ratio of 1.74 for the minor allele of rs3798220.

ABCG8 in Hypercholesterolemia

ABCG8 contains a long (10.8 kb) intronic repeat, part of a larger compound repeat separated by a LINE insertion (Figure 2). *ABCG8* encodes a cholesterol transporter that has been implicated in CAD [88,89] and in gallstone formation [90]. SNPs rs41360247 and rs4245791 in *ABCG8* are associated with both CAD risk and LDL cholesterol levels [89]. Additionally, the SNP rs4952688 was shown to influence the mRNA expression of both *ABCG8* and its co-transporter *ABCG5* in liver cells [91]. rs4952688 is located within the compound repeat sequence (Figure 2), implicating this repeat sequence (or nearby linked sequence) in the expression levels of these two cholesterol transporters.

The normal function of *ABCG8* and *ABCG5* in liver cells is to excrete cholesterol into the bile [92]. Disruption of this process could lead to hypercholesterolemia, the initial manifestation of atherosclerosis. *ABCG8* variants can also influence cholesterol

Table 3. Long tandem duplications with at least 96% identity that occur within a gene locus.

Gene	Length	Copies	Gap	Coding
NBPF20	76181	2[#]	55	Y
NBPF8	65137	2	0	Y
CR1	54708	3	0	Y
ANKRD30A	47663	2	2	Y
RBMY1A1	47081	3	0	Y
NBPF12	44119	2[#]	31	Y
PGA4	37662	2	0	Y
TRPM3	35986	2	0	N
FCGBP	31945	2*	0	Y
NEB	31782	3	0	N
NKG2-E [+]	30864	2	0	Y
TBC1D3C/TBC1D3H	27063	2	0	N
HCAR1	26136	2	14	Y
TTC34	22675	2[#]	22	N
DAZ1	21690	2	0	Y
NBPF1	12620	2	17	Y
NBPF12	12568	2[#]	4	Y
BRF1	11321	2[#]	0	N
C2orf78	10103	2	58	Y
CLEC17A	8924	2	0	Y
TTN	8521	2	0	Y
SNTG2	8383	2[#]	1	N
IFI16	8282	2	0	Y
MUC5B	7627	2[#]	1	Y
SPDYE3	7020	5	0	Y
ERC1	5850	2[#]	23	N
HRNR	5637	2[#]	0	Y
ACRC	4289	2	66	Y
SPRN	4144	2	0	N
TMEM132D	3907	2[#]	12	N
HP/HPR	3431	2	4	Y

Duplications were identified as described in the Methods. The length indicates the total length of the high-identity tandem duplicons. The gap is the separation between the two highest-identity (long) duplicons, which was required to be less than 100 bp. The duplication is "coding" if a duplicon overlaps at least one exon.
*FCGBP has a third duplicon, but with less than 96% identity.
#These genes have duplicons that are themselves STRs of lower fidelity; only the copy number for the high-identity long tandem duplication is reported in this table.
+The segmental duplication containing NKG2-E overlaps the three genes KLRC1, KLRC2 and KLRC3.
doi:10.1371/journal.pone.0101093.t003

levels by modulating cholesterol absorption [93]. Somatic repeat mutations accumulating over time could change expression levels of these proteins, thereby altering the rate of cholesterol excretion/absorption. Germ-line mutations in these genes are associated with premature atherosclerosis [91,94], as are mutations in other cholesterol transporters such as APOE [95,96].

In principle, somatic repeat mutations could induce the production of aberrant ABCG8 protein variants that would be immunogenic, as previously argued for autoimmune disease.

Antibodies to such variants could interfere with cholesterol excretion, but ABCG8-specific antibodies have not been documented in CAD. The molecular mechanisms by which the proteins encoded by ABCG8 and ABCG5 transport cholesterol are not fully understood [97]. If the ABCG5/ABCG8 complex binds to LDL, then ABCG8 would encode a peri-antigen for CAD since oxidized LDL is an autoantigen [85].

DMBT1, FCGBP, and the Mucins MUC4, MUC5B, MUC12 and MUC17

Mucins including MUC4, MUC12 and MUC17 are important for intestinal integrity and have previously been associated with both ulcerative colitis (UC) and Crohn's disease (CD) [98–100]. MUC17 depletion increases epithelial permeability in the face of E. coli exposure [101]. FCGBP is a component of the mucus layer coating of the intestinal tract [102], and expression is higher in several autoimmune diseases [103]. The DMBT1 protein also provides mucosal protection of the intestine, and expression levels correlate with disease activity in CD and UC [104]. Host-microbe interactions appear to be central to the pathogenesis of UC and CD [105]. CD, UC, psoriasis (PSO) and ankylosing spondylitis (AS) have common features [105,106] that suggest a cluster of diseases with related etiology. AS has been associated with the gut microbiome [107], and PSO has been associated with intestinal yeast infections [108].

A critical clue is provided by the PTPN22 rs6679677 C/A polymorphism that is in high linkage disequilibrium with the rs2476601 C/T polymorphism associated with many autoimmune diseases [109]. At rs6679677, the A allele appears to be a risk allele for UC (as for most other autoimmune diseases) but protective for CD [105]. In the context of the coherent somatic mutation hypothesis, one could interpret this opposite PTPN22 association in terms of alternative responses to somatic mutation. UC would be caused by an autoimmune response against the mutant protein, while CD would be caused by the failure of the mutant protein's function, in the absence of a direct immune response against that protein. This interpretation is consistent with a clear role for MHC alleles in UC but not CD [105,110], and with a reduction in mucus quantity and/or goblet cell density specifically in UC [111,112].

CD and UC have opposite risk alleles for NOD2 polymorphisms [105]; NOD2 variation modulates adaptive immune responses to microbial antigens [113], and regulates DMBT1 expression in CD [114]. Significantly, short alleles of the DMBT1 tandem repeat that encode fewer bacterial recognition sites are overrepresented in CD but not UC [104]. DMBT1 has high protein homology with the CD autoantigen CUZD1 [115], potentially leading to cross-reactive antibodies. Further, DMBT1 -coded protein binds to pancreatic amylase [116,117] that in turn binds to the CD autoantigen GP2 [118], meaning that DMBT1 encodes a peri-antigen for CD.

In Sjogren's sydrome, a primary initiating change is the dysregulation of mucins [119], including the aberrant exocytosis of MUC5B [120]. MUC4 is an interesting somatic mutation candidate because its expression pattern in the eye, vagina, ectocervix, trachea, and salivary gland [121] closely aligns with locations where symptoms occur [122]. MUC5B is expressed in many of these tissues [123], but not in the tear fluid [124]. Somatically mutated mucins could induce an immune response against the mutant protein. Alternatively, aberrant mucin protein may offer reduced protection of epithelial cells, making them vulnerable to infection. Apoptosis of the epithelial cell could trigger the induction of antibodies to apoptotically generated proteins in Sjogren's syndrome.

Table 4. Long (>5 kb) regions of self-alignment within protein-coding genes.

Gene	Length	Gene	Length
NBPF10	45133	MTUS2	10090
ANKRD30A	40083	ANKRD36	8739
NBPF20	39623	PTPRN2	8649
DAZ2	38211	FAM153A	8495
DAZ1	36567	TTC34	8343
LPA	35017	FAM153B	7969
NBPF12	32343	FLG	7934
FCGBP	30167	BRF1	7650
DMBT1	26579	ST3GAL4	6583
MGAM	24595	TTN	6447
DAZ4	23181	MUC12	6346
KIR3DL1	22943	MUC5B	6303
NEB	20252	GALNT9	6290
ANKRD30B	18603	TRHDE	6161
NBPF8	14249	ERC1	5794
TBC1D3C	13432	ROBO2	5789
TBC1D3H	13432	TM4SF2	5498
NBPF1	13424	NBPF14	5345
KATNAL2	13368	CACNG7	5304
CR1	12971	SNTG2	5229
HCAR1	12648	TNXB	5227
POTEJ	12480	MAGEA4	5091
DAZ3	12115	ASMT	5021

Sequences with a self-similarity score of 60 or above having both query and target mapped within a protein-coding gene locus were obtained from the self-alignment track [267] of the UCSC database as described in the Methods, and ranked by match-length. In this table, the match length corresponds to the length of identity between the two duplicons. Note that self-aligned duplicons may overlap.
doi:10.1371/journal.pone.0101093.t004

Long Repeats Reside in Genes Expressed in Immune Cells and Implicated in Autoimmunity

KIR3DL1 encodes an inhibitory receptor expressed on natural killer (NK) cells and T cells [125]. There is a high degree of copy number variation of the KIR genes around this locus, and some haplotypes do not possess KIR3DL1 [126]. HLA-Bw4 is the ligand for KIR3DL1, and is protective in MS [125] and primary sclerosing cholangitis [127]. The presence of KIR3DL1 is protective for AS [128], particularly AS with uveitis (UV) [129]. Somatic mutations to KIR3DL1 could reduce inhibition of NK cells and/or T cells, leading to selective activation and clonal expansion.

The segmental duplication at the NKG2-E locus overlaps the genes KLRC1, KLRC2 and KLRC3. Copy number variation at NKG2-E (manifested as a deletion of KLRC2) is associated with psoriasis susceptibility [130]. Reduced KLRC2 expression in T cells is observed in PSO [131], and enhanced expression of KLRC2 on CD4+ T cells is observed in MS [132]. KLRC1 encodes a critical receptor on NK cells, regulating the elimination of autoreactive CD4+ T cells in animal models of MS [133]. KLRC1 plays a critical role in tolerization by regulatory T cells [134], and is downregulated in PSO [135].

KIR3DL1 and KLRC1 encode NK cell receptors. NK cells and their receptors regulate autoimmunity in MS [136], and NK cell populations rise and fall in ways that correlate with the development of lesions in relapsing-remitting MS [137,138]. NK cells are found in psoriatic plaques, and circulating NK cells are reduced in PSO, MS, SLE and T1D [139,140].

The segmental duplication within the long HCAR1 repeat identified in Tables 3 and 4 covers the two genes HCAR2 and HCAR3. HCAR2 codes for a niacin receptor that is expressed on antigen presenting cells and functions in a tolerization pathway for T cells [141]. Niacin administration ameliorates an animal model of MS through this pathway [141].

Summary: Long Simple Tandem Repeats in Autoimmunity

Table 7 summarizes the autoimmune associations of genes with long STRs. This key table shows that long STRs within twenty genes are associated with sixteen common autoimmune diseases and atherosclerosis. Each of these putatively mutable STRs exhibits germ-line structural variation (File S1), consistent with a somatically mutable locus. The coherent somatic mutation hypothesis thus has the potential to be a comprehensive explanation for many autoimmune diseases.

With the exception of MS and possibly PA and SJ, each of the diseases associated with an autoantigen or peri-antigen in Table 7 is influenced by the functional rs2476601 single-nucleotide polymorphism in the PTPN22 gene (Table 1). This polymorphism specifically influences T cell signaling [142,143], B cell signaling [144,145], autoreactive B cell generation [144], and T cell and dendritic cell hyper-responsiveness [146]. The role of PTPN22 in some but not all autoimmune diseases suggests a common

Table 5. Correspondence of PPP4C localization with many known SLE autoantigens.

Autoantigen(s)	Putative PPP4C function/localization
SNRPA, SNRPC, SNRNP70, Sm	These are spliceosome proteins. PPP4C is involved in spliceosome assembly [80,395].
SR proteins	SR proteins associate with the spliceosome [396]. See SNRPA, SNRPC, SNRNP70, Sm.
chromatin/histones	PPP4C binds to HDAC3, a histone deacetylase [397]. A PPP4C complex dephosphorylates γ-H2AX histones [398].
ds-DNA	Stabilisation of stalled replication forks [80], histone deacetylation [397], histone dephosphorylation [398].
Ku70, Ku80	Ku70 and Ku80 associate with γ-H2AX histones during double strand break repair, mediated by DNA [399,400]. A PPP4C complex dephosphorylates γ-H2AX histones [398].
PARP1	PARP1 binds with Ku [401]. See Ku.
ribosomal P	dephosphorylated during apoptosis by a caspase-induced phosphatase [402].
La(SSB)	La is dephosphorylated during apoptosis by a caspase-induced PP2A-like phosphatase [403].
Ro60(TROVE2)	See La; Ro60 and La are components of a common protein/RNA complex [404].
APOH/cardio-lipin-complex	APOH (coding for beta 2 glycoprotein I) associates with ANXA2/TLR4/CALR/NCL complexes [405]. Anti-APOH antibodies target bound APOH, triggering NF-Kappa B activation in a TRAF6/MyD88 dependent fashion in endothelial cells [405-407]. PPP4C physically interacts with TRAF6, and is recruited to the TLR4 complex on lipopolysaccharide (LPS) stimulation [408]. Further, LPS stimulation induces expression of PPP4C [408].
VIM/cardiolipin-complex, VIM	Vimentin is also observed in analysis of APOH/ANXA2/TLR4/CALR/NCL complexes [405]
NPM1	NPM1 binds cardiolipin [350]. See APOH/cardiolipin-complex.
CALR	CALR may be dephosphorylated by an okadaic-acid-sensitive protein phosphatase [409]. See La; CALR interacts with the Ro60/La/RNA complex [410]. See also APOH/cardiolipin-complex.
Ro52(TRIM21)	See CALR; Ro52 and CALR are binding partners [410].
NCL	See APOH/cardiolipin-complex.
ANXA2	See APOH/cardiolipin-complex.
F2/Phosphatidyl-serine, F2	Phosphatidylserine bound by ANXA2 [411]. See ANXA2.
ANXA1	Binds ANXA2 [412], phosphatidylserine [413], and colocalizes with ANXA5 [413]. See ANXA2, F2/Phosphatidylserine.
ANXA5	Binds phosphatidylserine as a monomer or dimer [414] and colocalizes with ANXA1 [413]. See ANXA1, F2/Phosphatidylserine.
HMGB1	Binds phosphatidylserine [415]. See F2/Phosphatidylserine.
LTF	Binds to TLR4 and activates the TRAF6/MyD88 pathway [416]. See APOH/cardiolipin-complex.

Primary autoantigens are in bold.
doi:10.1371/journal.pone.0101093.t005

underlying pathway for this subset of diseases [45,143] that may be related to STR length and/or mutability.

Table 8 shows that the conditions associated with autoantigens/peri-antigens above have a high degree of co-morbidity and/or familial association. Taken together, the data support the following model for this subset of diseases:

Table 6. Murine long (>8 kb) STRs overlapping protein-coding RefSeq gene loci, including 5 kb upstream of the gene start site.

Gene	Largest Repeat length (bp)	Smallest unit size
Ulk4	19480	10
St3gal4*	19106	1236
Dmd	18426	734
Flg2	16228	234
Ttc34	12289	37
Hrnr	8023	513

The smallest repeat unit for each region is given together with the total STR length. The Ttc34 repeat ends 3.2 kb upstream of the gene start site.
*Two adjacent repeats reported by TMRF have similar repeat structure, and have been combined.
doi:10.1371/journal.pone.0101093.t006

- For each gene containing a mutable repeat locus, individuals have a small population of somatically mutant cells.
- Under normal conditions, these mutant cells either induce peripheral tolerance or are too rare to trigger an immune response.
- Under inflammatory conditions (e.g., during an infection) the population of mutant cells increases, concurrently with immune system stimulation.
- In individuals with impaired tolerance or with sensitive B-cell or T-cell activation thresholds, reactions against mutant cells occur.
- Inflammation caused by immune reponses induces new coherent mutation in neighboring cells, and creates a cycle of autoimmunity.

A disjoint subset of diseases, including MS, PSO, UV, and AS have no association with the *PTPN22* gene polymorphism (Table 1). All four of these conditions are associated with immune-cell expressed genes spanning long repeats. Somatic mutation in those genes, rather than in antigenic genes, may be the critical step for such diseases.

A Repeat Constituting 97% of the Intron Sequence within an Autoantigen for Pemphigus Vulgaris

Somatic repeat mutations in introns could be particularly disruptive when the intron is almost exclusively tandem repeat
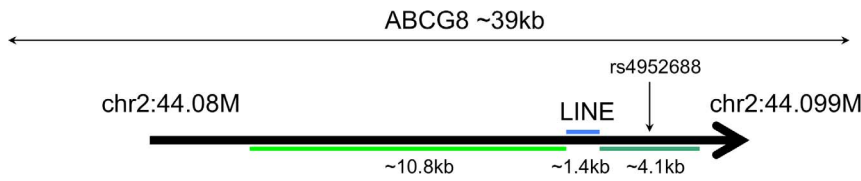
Figure 2. Structure of the long *ABCG8* repeat in the human reference genome. A 10.8 kb repeat and a 4.1 kb repeat have closely related repeat unit sequence, and are separated by a 1.4 kb LINE insertion. The SNP rs4952688 occurs in the middle of the 4.1 kb repeat.
doi:10.1371/journal.pone.0101093.g002

sequence. I therefore queried the reference genome for genes containing introns where a single tandem repeat occupies a large fraction of the intron (Table 9). The top-ranked gene in this analysis is *PKP3*, containing a 2310 bp repeat occupying over 97% of the eighth intron. There is germ-line structural variation at this locus in the HapMap population, with deletion variants encompassing almost the entire STR sequence [147].

PKP3 encodes an autoantigen in pemphigus vulgaris (Table 2). Furthermore, PKP3 binds in vivo to several other primary pemphigus vulgaris autoantigens including DSG3, DSG1, DSC1, and DSC3 [148]. Aberrant PKP3 could therefore serve as a CD4+ T cell antigen in the induction of antibodies to these other proteins. The *p* value for the top gene being an autoantigen is 0.02 (see Methods).

Genes with High Copy-Number Internal Repeats Include Autoantigens for Multiple Sclerosis and Myasthenia Gravis

Figure 3 shows repeats of length up to 5 kb with repeat counts of at least 700 units. At this scale, all repeats are microsatellites with short repeat units. The genes with the eleventh and twelfth highest repeat counts genomewide are *MUSK* and *MAG* respectively. *MUSK* encodes an autoantigen in myasthenia gravis (Table 1). *MAG* encodes a multiple sclerosis autoantigen that binds in vivo to MBP and PLP [149], two other MS autoantigens (Table 1). Anti-MAG antibodies have also been observed in various polyneuropathies [150–152]. The presence of two autoantigens among the top twelve is statistically significant ($p<0.022$, see Methods). On the other hand, the STRs in *MAG* and *MUSK* do not exhibit germ-line structural variation at 50 bp resolution (File S1); germ-line variation would be expected for a somatically mutable locus.

Discussion

Somatic mutation has been overlooked or discounted as a cause of autoimmunity, primarily because "random" mutation would not lead to consistent and specific disease characteristics [153]. However, many kinds of somatic mutation are nonrandom, caused by mechanisms that yield coherent mutation patterns both within and across individuals. Coherent somatic mutation is a unifying and biologically plausible hypothesis to explain the specific targets of autoimmune disease.

Longer-Range Segmental Duplications

Long high-identity segmental duplications that are not strict tandem repeats may still lead to somatic protein changes via deletion or duplication if they partially overlap genes. Examples of this pattern include: RHD and GYPA, autoantigens in autoimmune hemolytic anemia; AMY2A, an autoantigen in autoimmune pancreatitis and fulminant T1D, and a binding partner of the CD autoantigen GP2 [118]; CES1 and PDIA3, autoantigens in type-2

autoimmune hepatitis; TYR, an autoantigen in vitiligo; and CHRNA7, an autoantigen observed in schizophrenia (Tables 1 and 2, [154]). The genomic structure of *TYR* makes it particularly susceptible to gene conversion and deletion (Figure 4).

The human genome contains segmental duplications that span whole genes, and copy number variation in these tandem repeats is likely to affect gene dosage [155]. These duplications are not considered in the primary anaylsis since repeat-dependent somatic mutation via deletion and/or duplication is less likely to induce altered protein. Nevertheless, the potential for altered protein exists through gene conversion or other processes that combine sequence from multiple instances of the gene. The primary autoantigen in Addison's disease is encoded by *CYP21A2* (Table 1), which resides within a segmentally duplicated region and is a known locus of germ-line gene conversion [156].

A five gene cluster (*GH1*, *GH2*, *CSH1*, *CSHL1*, *CSH2*) on chromosome 17 resides in a region characterized by complex segmental duplications with identity ranging from 92% to 96%. This cluster is a hot-spot for germ-line gene conversion [157]. Variations in these genes are associated with metabolic syndrome later in life [158]. Anti-pituitary antibodies are observed in conjunction with type-2 diabetes [159,160] and GH1 is one of the autoantigens [161]. *GH1* codes for human growth hormone, and growth impairment is observed in celiac disease in conjunction with anti-pituitary antibodies [162].

Mechanisms of Coherent Somatic Mutation

PTPRN2 is an outlier not just in the length of its repetitive sequence; it has the most predicted sites of R loop formation in the whole genome [163]. The R loop sites do not overlap the 12 kb repeat in *PTPRN2*, but several long R loop sites occur about 20 kb upstream of this repeat. These R loops may contribute to the instability of the repeat region, and implicate mis-splicing [22] of *PTPRN2* in T1D.

Coherent somatic mutation can occur through a variety of mechanisms besides repeat instability and gene conversion, discussed below and summarized in Table 10.

RAG-mediated Somatic Recombination and Rheumatoid Factor

Cancer studies provide valuable information about coherent somatic mutation in vivo. Many cancers elicit antibodies that are also found in autoimmune disease [164], further supporting a role for somatic mutation in autoimmunity. A striking example of coherent somatic mutation in cancer is the gene *IKZF1*. Internal *IKZF1* deletions occur in over 80% of cases of BCR-ABL1 acute lymphoblastic leukemia (ALL) [165]. Consistent breakpoints suggest aberrant RAG-mediated recombination [165]. The mutations coincide with a transition in the cancer from Chronic Lymphocytic Leukemia (CLL) to ALL.

CD5 expression on B cells is a common feature of both RA and CLL [166], CD5 expression correlates with RAG activity in B cells

Table 7. Known links between genes with long STRs and human autoimmune diseases.

Gene(s)	Disease	Antigen type	CNV	Expr. changes
FLG	RA	**Autoantigen**		
TPO	HT	**Autoantigen**		
TPO	GD	**Autoantigen**		
PTPRN2	T1D	**Autoantigen**		
CR1	SLE	Autoantigen	Yes	Yes
CR1	MS	Autoantigen		
PGA4	PA	Autoantigen		Yes
TTN	MG	Autoantigen		
IFI16	SLE, SSc, RA, SJ	Autoantigen		Yes
HP	CEL	Autoantigen	Yes	Yes
BRF1	SSc	Peri-antigen		
TTC34	SLE	Peri-antigen*		
LPA	CAD	Peri-antigen	Yes	Yes
ABCG8	CAD	Peri-antigen?		Yes
DMBT1	CD	Peri-antigen	Yes	Yes
DMBT1	UC			Yes
MUC4, MUC12, MUC17	CD, UC			Yes
MUC5B	SJ			Yes
HP	RA, SLE, CD, CAD, SSc		Yes	
HP	T1D		Yes	Yes
FCGBP	several			Yes
KLRC2	PSO		Yes	Yes
KLRC2	MS			Yes
KIR3DL1	AS, UV		Yes	

Genes with long STRs come from Figure 1, Table 3 and Table 4. A bold autoantigen label corresponds to a known primary autoantigen. The CNV column indicates whether a germ-line STR length variant is associated with the disease. Gene expression changes during disease are also shown.
*While many genes qualify as encoding peri-antigens in SLE, TTC34 encodes a peri-antigen for many autoantigens (Table 5).
doi:10.1371/journal.pone.0101093.t007

of people with autoimmune disease [167], and RAG is expressed in B cells in the RA synovium [168]. In RA, the appearance of rheumatoid factor (RF, an antibody to Fc-IgG) correlates with the hypogalactosylation of IgG, occuring roughly two years after the appearance of antibodies to citrullinated proteins, but two years before RA diagnosis [169]. RF is detected in several other autoimmune and infectious diseases [170].

If the RAG-dependent *IKZF1* mutations that consistently occur in ALL also occur in RA B cells, possibly followed by clonal expansion, then aberrant glycosylation would be explained because *IKZF1* appears to be critical for proper IgG glycosylation [171]. The improperly glycosylated IgG would be immunogenic. In the context of a normal immune response to a pathogen, a somatic mutation to *IKZF1* could be adaptive, because it would lead to RF production and potentially enhanced clearance of

Table 8. Co-morbidity and/or familial associations between six autoimmune diseases and atherosclerosis.

	GD	RA	T1D	SLE	SSc	CAD
HT	[417]	[49,418,419]	[49,418]	[49,419,420]	[419]	[421]
GD		[417]	[422]	[417]	[423]	[421]*
RA			[49,418,424]	[49,424]	[49]	[421,425]
T1D				[424]		[426]
SLE					[49]	[421,425]
SSc						[421,425]

Comorbidity may reflect common susceptibility factors or secondary disease effects, such as inflammation in RA contributing to CAD risk [427]. Comorbidities with some of these diseases exist for alopecia areata [428,429], vitiligo [430,430,431], juvenile idiopathic arthritis [432], myasthenia gravis [433,434], and Addison's disease [435], five additional *PTPN22* -associated diseases, as well as celiac disease [436,437] and pernicious anemia [438,439].
*A link between GD and CAD is potentially confounded by the anti-atherogenic properties of thyroid hormones [440].
doi:10.1371/journal.pone.0101093.t008

Table 9. Genes with intronic tandem repeats occupying more than 90% of an intron.

Gene	Chrom.	Intron start	Intron end	Int. length	Rpt. Length	Occupancy	Copies
PKP3	chr11	400706	403076	2371	2310	0.974	155.9
NMRK2	chr19	3937287	3938599	1313	1274	0.970	27.8
HSD17B14	chr19	49337616	49339060	1445	1401	0.970	38.1
PPP1R12C	chr19	55604602	55605711	1110	1066	0.960	29.4
ASMT	chrX	1755454	1761694	6241	5827	0.934	35.4
TCF25	chr16	89973731	89975372	1642	1526	0.929	43.8
NSMF	chr9	140344708	140346815	2108	1948	0.924	33.4
SCNN1D	chr1	1223418	1225649	2232	2055	0.921	110.9
AHNAK2	chr14	105407316	105420216	12901	11844	0.918	24
BRF1	chr14	105695251	105707600	12350	11322	0.917	111
TOP1MT	chr8	144403557	144406167	2611	2367	0.907	164.1

immune complexes [172]. However, in the context of an autoimmune response, RF production could increase the severity of disease [172]. RF is also found in SLE [173], and reduced *IKZF1* expression has been associated with SLE [174,175].

Mutagens and Oxidative Stress

Cigarette smoking is mutagenic, and appears to be selectively associated with antibodies to the primary autoantigens encoded by *ENO1* [176], *VIM* [177], and *FGB* [177] in RA. *VIM* mutations induced by oxidative stress influence antigenicity [178]. The association of RA with smoking is strong only among individuals with particular HLA alleles. A similar phenomenon occurs in MS [179]. This interaction of mutagen, autoantigen and HLA suggests that mutation is pathogenic primarily when the mutant epitope is well-presented by the corresponding antigen presenting molecules.

Clonal Expansion Following Somatic Mutation

Somatic mutations in the *TSHR* gene are relatively common [180] and can induce activation and clonal expansion in thyroid tissue [181,182], potentially explaining TSHR-antigenicity in GD. Paraneoplastic autoimmunity [164,183,184] is a related phenomenon in which an immune response to a tumor expressing mutant antigens also affects normal tissues expressing wild-type proteins.

## Pathogen-Induced Protein Binding and Modification

A pathogen-expressed protein that binds with an endogenous protein complex could serve as a CD4+ T cell target, providing help to B cells generating antibodies to proteins in the protein complex. A pathogen-modified endogenous protein could behave in a similar fashion

Rheumatic Fever (RHF) is a condition characterized by autoimmune attack against cardiac muscle, usually associated with group A streptococcal infections [185]. There is some in-vitro evidence of cross-reactivity of antibodies to streptococcal proteins and autoantigens in RHF [186]. Nevertheless, there is also evidence that mimicry may not be an important feature of RHF [187]. Autoreactivity to collagen in RHF has been proposed to result from collagen binding to streptococcal proteins [187].

The RHF autoantigens vimentin, myosin, and tropomyosin (Table 2) form part of the calcium-bound sarcomere protein complex [188]. Two lines of evidence implicate vimentin as an initiating autoimmune target (and peri-antigen) in RHF. First, vimentin is modified (ADP-ribosylated) by the group A streptococcal protein SpyA in a way that alters both its sequence and its organization [189]. Second, group A streptococci are known to bind to vimentin, particularly at sites of muscle injury [190].

Apoptotic Cleavage

Adaptive immune reponses require the joint participation and mutual activation of CD4+ T cells and antigen-presenting cells such as B cells. B cells become anergic under chronic low-level exposure to antigen with limited costimulation [191]. Nevertheless, even anergic B cells can be activated with sufficient stimulation [191]. Protein that is post-translationally modified only upon apoptosis would presumably generate only low-level exposure to B cells. A post-translationally modified protein that forms part of a protein complex containing a somatic mutant is liable to trigger B cell/T cell co-activation. In such a case, a CD4+ T cell specific to the mutant peri-antigen could activate a previously anergic B cell clone. Such a mechanism could explain why post-translationally modified proteins, particularly those geneated during apoptosis, would be over-represented among B cell autoantigens [51,52,192].
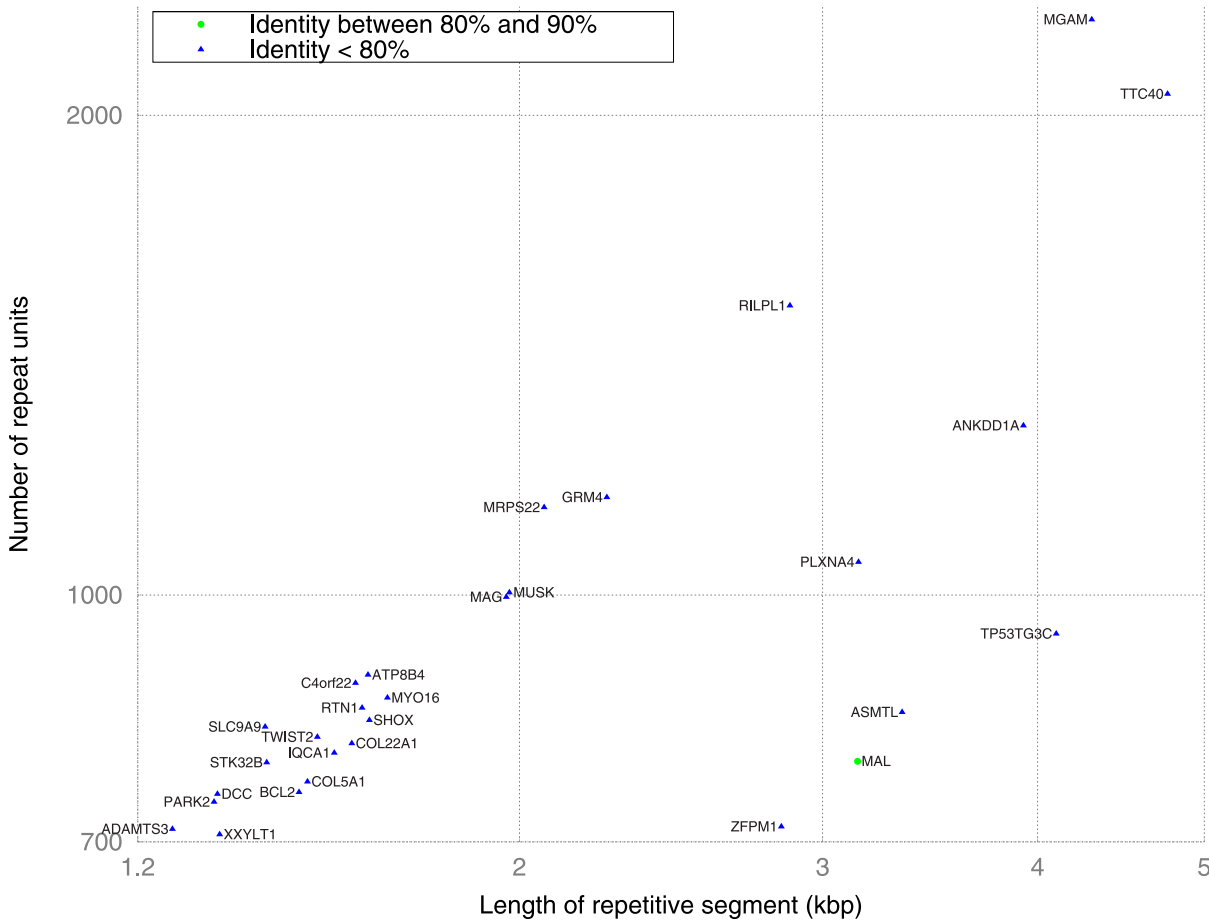
Figure 3. Genes with high copy number internal repeats. The x-axis denotes the total length of the tandem repeat (log-scale), and the y-axis represents the number of repeat units within the tandem repeat (log-scale). The degree of homology between repeat units is indicated by the color of the data point. All repeats in this diagram reside in introns. Genes containing multiple disjoint repeats appear more than once.
doi:10.1371/journal.pone.0101093.g003

Retrotransposition

An additional potential mechanism of coherent somatic mutation is retrotransposition. Retrovirus [193,194] and retrotransposon [195] integration hotspots exist, independent of selective pressure for cell growth/survival. This form of mutation could be relevant to Bout Onset Multiple Sclerosis (BOMS) in which an endogenous retrovirus has been implicated [196,197], as well as schizophrenia [198] and amyotrophic lateral sclerosis [199]. Alternatively, retroviral expression could be a driver of neuroinflammation [200], leading to somatic mutation at other mutable repeat sequence.
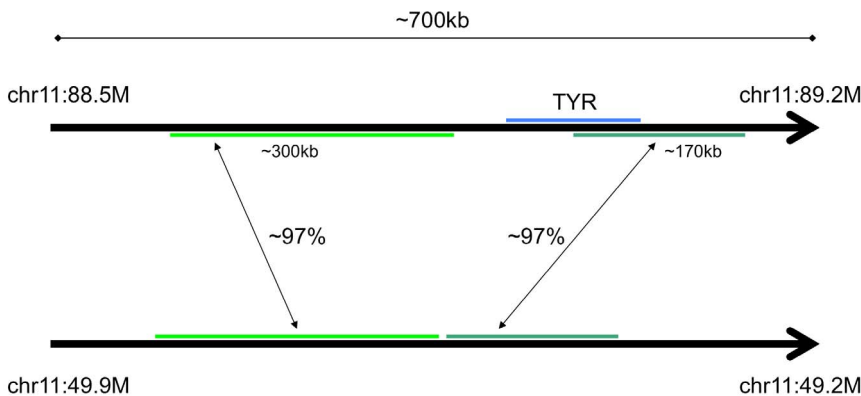
Figure 4. Structure of the *TYR*-related tandem duplications in the human reference genome. The long, high-identity duplicons make the region susceptible to gene conversion [274].
doi:10.1371/journal.pone.0101093.g004

Table 10. Multiple mechanisms generating coherent somatic mutation, and possible examples where autoimmunity results.

Mechanism	Possible Examples
Mutations at long tandem repeats	T1D, HT, RA, SLE, …
Gene conversion at segmental duplications	AD
Clonal expansion	Paraneoplastic autoimmune diseases, GD
Oxidative stress	*VIM* mutation in RA
RAG-dependent somatic mutation	*IKZF1* in RA
Pathogen Binding/Modification	*VIM* in RHF
Retrotransposition	BOMS
Apoptotic protein cleavage	Many cleaved proteins are autoantigens
Dysregulation of protein modification	Anti-TOP1 SSc
Environmental mutagens	*ENO1, VIM, FGB* in RA

doi:10.1371/journal.pone.0101093.t010

Dysregulation of Protein Modification Pathways

In SSc, the presence of one antibody type is generally exclusive of the others [201,202], suggesting several subtypes of SSc with different mechanisms of induction. Chromosomal abnormalities are found at high frequency in the lymphocytes of patients with anti-centromere or anti-TOP1 antibodies, but at normal frequency in patients with anti-RNAPIII antibodies [203]. In SSc fibroblasts, increased sumoylation of TOP1 induces deficits in TOP1-mediated supercoiled-DNA relaxation [204] and disruption of TOP1 is known to cause chromosomal aberrations [203]. Inhibition of sumoylation improves TOP1 function in fibroblasts [203] and reduces fibrosis [205]. One interpretation of this data is that anti-TOP1 SSc is a sumoylation disorder. Hyper-sumoylated TOP1 could induce cell death via chromosomal aberrations, and at the same time trigger an immune response. Because the post-translationally modified protein would not be normally presented to immature B or T cells, tolerization to modified TOP1 would not occur. The centromere protein and SSc autoantigen CENPB is also a sumoylation target [206–209].

A similar neoantigen-creating role for sumoylation in a subset of patients with primary biliary cirrhosis (PBC) has previously been proposed [210]. In patients with antibodies to PML or SP100, two sumoylation target proteins [206–209], antibodies to SUMO2 and SUMO1 have been observed [210]. CENPB is also an autoantigen in PBC (Table 2). SSc and PBC are comorbid, with anti-CENPB as a common risk factor [211,212], suggesting a shared etiology.

Schizophrenia and Autism

Schizophrenia and autism have prominent immunological features, including HLA associations, comorbidity with autoimmune diseases, and associations with viral triggers and maternal infections during pregnancy (Table 11). Immunological theories of schizophrenia have been proposed [213].

A clue that somatic repeat mutation may contribute to schizophrenia comes from a twin study in which a genomewide measure of somatic trinucleotide repeat mutation was obtained [214]. A high somatic trinucleotide mutation rate associated selectively with the schizophrenic proband in monozygotic twins discordant for disease [214].

Four *NBPF* family genes are among the top twelve in Figure 1, including the two longest STR sequences. The four *NBPF* genes in Figure 1 are located between positions 145.2 M and 148.3 M on chromosome 1, overlapping the 1q21.1 region. *NBPF* genes contain many copies of the *DUF1220* element; *DUF1220* copy number is closely related to brain size, and humans have many more copies than other primate species [215,216]. In humans, high *DUF1220* copy number correlates with macrocephaly, and low copy number correlates with microcephaly [217,218]. Germline deletions within the 1q21.1 region are associated with schizophrenia [219,220], while duplications are associated with autism [217]. Somatic genomic instability is likely in such highly repetitive regions [217]. Somatic mutations early in embryonic development [221], suggested by the link to maternal infections during pregnancy, could lead to effects that mirror those of germline mutations. Early somatic mutation also creates the possibility that the thymus and brain express different haplotypes, preventing thymic deletion of T cells reactive to proteins coded by a brain-specific haplotype.

Other schizophrenia-associated genes among those in Figure 1 include *IL3RA* [222] and *CACNG7* [223]. *IL3RA* encodes a receptor for IL3 that is expressed in neurons, and IL3 expression is correlated with brain volume [222]. CACNG7 modulates neurite growth [224] and regulates AMPA receptor gating [225].

Several autism-related genes appear in Figure 1 and Table 4. SNTG2 binds to neuroligins 3 and 4, genes that have been associated with autism, and known autism-related mutations in those neuroligins weaken the binding with SNTG2 [226]. ROBO2 is an axon-guidance protein with significantly reduced expression in autistic brains [227]. *ASMT* encodes the last enzyme in the melatonin biosynthesis pathway, low melatonin expression is observed in autism spectrum disorders, and rare *ASMT* mutations are associated with autism [228–230]. *MGAM* is a gene involved in starch metabolism, with dysregulated mRNA expression in autism [231]. Germ line loss-of-function mutations in *KATNAL2* have been associated with autism [232].

Additional autism related genes appear in Figure 3 and exhibit structural variation in their STR sequence (File S1). Like ROBO2, PLXNA4 is an axon-guidance protein with significantly reduced expression in autistic brains [227]. ASMTL binds with TDO2 [233]; TDO2 is the rate-limiting enzyme in the catabolism of tryptophan, the precursor of serotonin, which is known to be elevated in 30% of autism cases [234].

There is a high concentration of autism-related genes among a relatively small set of putatively mutable genes. In light of the autoimmune features of autism (Table 11), this concentration suggests that somatic repeat mutation may contribute to the etiology of autism.

Explaining Autoimmunity

A satisfying feature of the coherent somatic mutation hypothesis is that it provides a parsimonious yet comprehensive account of autoimmunity. The initiation of most diseases is attributed to a single mutable locus. A handful of diseases having several known subtypes include more than one corresponding mutable locus. Only four of the top sixteen genes in Figure 1 (*ANKRD36C*, *ANKRD36*, *AHNAK2*, *NSUN6*) do not have a link with an autoimmune disease, an autoimmune-associated mental disorder, or atherosclerosis. These relatively uncharacterized genes are promising candidates for future study.

The most prominent prior theory of autoimmunity is molecular mimicry, the hypothesis that peptides similar to host proteins are expressed by host-resident microbes, sometimes inducing an autoimmune reaction against the host proteins. The attractive feature of molecular mimicry has been that it provides a plausible explanation for the known link between infection and autoimmunity [235,236]. However, despite decades of research, no human autoimmune diseases have been clearly attributed to molecular mimicry [235,237,238].

Autoimmune diseases have historically been categorized as organ-specific or systemic, with some diseases hard to categorize [239]. Under the coherent somatic mutation hypothesis, both kinds of disease have a common etiology, with the phenotype dependent on the expression patterns of the autoantigen. A narrow expression pattern (such as *PTPRN2*) leads to an organ-specific disease (T1D), while a widely expressed protein complex (TTC34/PPP4C as proposed in this report) leads to a systemic disease (SLE).

The incidence of each of several autoimmune diseases has been rising in recent years [240], as has the apparent incidence of autism [241]. The "hygiene hypothesis" states that autoimmune disease is linked to the absence of infections, through one of several possible immunoregulatory mechanisms [240]. Some infections that are protective if they occur early in development are possible triggers of autoimmunity if they occur later [240]. The present theory is consistent with a variant of the hygiene hypothesis in which tolerance to coherently mutated antigens is dependent on the early generation of such mutants. Infections or other inflammatory stimuli would increase the rate of somatic mutation, allowing for more efficient induction of peripheral tolerance. In the absence of peripheral tolerance, late generation of somatic mutants could induce autoimmunity. Alternative hypotheses based on increasing exposure to environmental mutagens [242,243] are also consistent with an etiology dependent on somatic mutation.

Autoinflammatory Disease

Several non-autoimmune diseases may also be caused by somatic mutation of highly mutable repeat sequence in the context of inflammation. Atopic dermatitis and icthyosis vulgaris are inflammatory skin conditions caused by inactivating germ-line mutations of the *FLG* gene in some cases [244,245]. Somatic inactivating mutations of the 10.8 kb coding tandem repeat in *FLG*, reinforced by local inflammation, could contribute to the pathogenesis of these conditions. An accumulation of somatic mutations in *PTPRN2* (without autoimmunity) could lead to glucose intolerance [246]. Similar mechanisms could underlie various autoinflammatory conditions [247].

Genetics

Our study is limited by its reliance on a single human genome for long repetitive sequence. Some reference alleles are much shorter than those typically observed in the population (e.g., *MUC1* [248,249]). It is likely that long repetitive sequence is highly variable in the population [37,38,250], and that variations in germ-line sequence would modulate disease risk as seen for *CR1*, *LPA*, *HP* and *DMBT1*. Nevertheless, primary autoantigens whose genes contain long repeats were identified in a presumably healthy random individual, suggesting that, at least for those genes, all humans have some degree of somatic mutation and risk for disease.

Linkage based analysis of sequence variation in a population would not identify mutable repetitive regions because the high germ-line mutation rate would rapidly eliminate any linkage disequilibrium with adjacent sequence [157]. In contrast, there are likely to be few germ-line mutations within a pedigree, meaning that estimates of heritability [251] will include any effects of commonly inherited mutable sequence. Together, these effects could explain at least some of the missing heritability observed in many genomewide association studies [252–254].

Immunological Aspects

Not all somatic mutation is likely to be immunogenic, even in protein-coding sequence. Somatic mosaicism observed in triplet repeat expansion diseases [255] would not generate immunogenic protein if the repeat length is longer than the fragment expressible in MHC molecules (8–10 amino acids for MHC-I, 15–24 amino acids for MHC-II). On the other hand, a long triplet repeat could be vulnerable to somatic deletions, yielding a short, potentially immunogenic peptide repeat.

Keratinocytes express *FLG* [256] and are non-professional antigen presenting cells (APCs) [257]. Pancreatic beta cells express *PTPRN2*, and thyroid epithelial cells express *TPO*; both of these cell types are also non-professional APCs. The purpose of antigen presentation by such cells is assumed to be tolerization in the absence of costimulatory molecules [258], which seems appropriate in the case of three primary (and putatively mutable) autoantigens. The presence of antigen presentation on these cell types may have allowed the evolution of mutable genes without significant risk of abrogating tolerance. Alternatively, antigen presentation within these cell types may have evolved as a response to selective pressure for longer repeat sequences in these genes.

While T cell tolerance can be induced by the administration of peptides [259,260], attempts to induce tolerance in humans suffering from autoimmune disease have been largely unsuccessful [261]. Nevertheless, the success of these attempts is critically dependent on the peptide sequence used. The coherent somatic mutation hypothesis suggests that for intronic repeats, the initial immunogenic proteins may be mis-spliced or truncated forms of a native protein. Peptides covering the splice or truncation boundaries of putative mutant protein would be natural candidates for tolerance induction.

Validation

Many of the high prevalence diseases in Table 1 have been specifically associated with mutable antigens or peri-antigens in the present report. Some more speculative hypotheses for the involvement of somatic mutation in other diseases are presented in File S1. The proposed associations should be considered tentative, and subject to experimental validation. For reasons described previously and below, experimental validation may be technically difficult.

Recent sequencing advances have the potential to accurately sequence long repetitive regions [250]. Accurately sequencing many cells in search of rare somatic mutants will require significantly more effort, although new technologies will help [6]. Obtaining putatively mutated cells from sites of autoimmune

Table 11. Immunological features of autism and schizophrenia.

Feature	Autism	Schizophrenia
HLA Association	Yes [441,442]	Yes [442,443]
Co-morbidity with autoimmune disease	Yes [441,444–446]	Yes [447–449]
Viral triggers for disease	Yes [450]	Yes [450,451]
Association with maternal infection during pregnancy.	Yes [441]	Yes [452]
Autoantibodies	Brain-specific antibodies in mothers and probands [441,446,453,454]; Anti-nuclear antibodies [446,455]	Yes [456]
Other	Gene expression changes reminiscent of autoimmunity [457]; NK cell dysregulation [458]; Amelioration of aberrant behaviors during fever [459]	Various immunological abnormalities [460–462]; differentially expressed genes involve immune pathways [463]

doi:10.1371/journal.pone.0101093.t011

damage is challenging, since such cells would be subject to immunological destruction as soon as the mutation occurs.

Conclusions

The coherent somatic mutation hypothesis states that recurrent or clonal somatic mutation underlies the initiation of autoimmune disease. Long STR sequence is likely to be somatically mutable in vivo, motivating the present study. A highly significant association between three primary autoantigens (covering four autoimmune diseases) and long STR sequence was established. Additional autoantigens and peri-antigens were identified among genes spanning long STR sequence, and among genes with other known markers of somatic mutation. The work presented here could lead to a partial resolution of the mystery of why particular proteins are targets of autoimmune destruction [50]. Experimental validation of the specific predictions made here is the next step.

Materials and Methods

Genome coordinates use the GRCh37 (hg19) sequence. Gene names use HGNC approved nomenclature. Queries were submitted to the UCSC MySQL database server [64] and processed as described below. The SQL queries can be found in File S1. Gene transcripts were required to be protein-coding according to GENCODE version 17 [262] or (for Queries 2 and 6) RefSeq [263].

Identifying Genes with Intragenic Repeats

Query 1 was submitted to obtain genes containing long or frequent repeats. The output from this query was edited as follows:

- Genes not on the reference chromosomes were removed. Only one such gene (MGC39584/AC018692.2 on chr4_gl000193_random) had length over 5 kb and none had a repeat count over 100.
- For genes occurring on both the X and Y chromosomes, only the X chromosome instance was retained.
- TMRF often generated multiple repeat candidates for a region with the periods of the candidates being multiples of the shortest period. In such cases, only the shortest-period candidate with the highest repeat-unit count was kept, even if it spanned a slightly smaller region.
- When TMRF generated a consensus repeat unit that was itself repetitive (e.g., AGTTAGTTAGTT) the TMRF entry was replaced by one with a shorter repeat unit (e.g., AGTT) and a

higher repeat-unit count, retaining the degree of identity from the longer sequence. Examples include VPS53 (in which a 96 bp repeat is itself made of 3 instances of a 32 bp repeat), MUC4 (in which a 96 bp repeat consists of two consecutive instances of a 48 bp repeat), and MAL (with an 8 bp AGTGAGTG repeat).

- In a small number of cases, TMRF generated multiple essentially contiguous repeats with the same period and consensus sequence. The only such case where the repeat was either more than 5 kb long or contained more than 600 repeat units was PTPRN2 (chr7:158122660–158135328) where the contiguous repeat records were combined into a single longer 12.6 kb repeat.

To see whether the output was dependent on the source of the gene annotations, I reformulated the query as Query 2 using RefSeq [263]. The following differences were noted for repeats longer than 5 kb:

- There was some discrepant labeling of the NBPF genes. The NBPF repeat sequences were the same, with the exception of one NBPF10 repeat (see below).
- The following genes/repeat-lengths were identified by GEN-CODE but not RefSeq: ANKRD36C/49539; FAM230A, USP1/7516; PLEKHB2/6521; ANKRD36C/6410; FAM182B/6292.
- The following genes/repeat-lengths were identified by RefSeq but not GENCODE: NBPF10/15997; ANKRD36B/25486; MUC19/8607.
- A large majority of repeats were common to the two annotations, with the differences mentioned above largely due to differences in the labeling of a gene transcript as protein coding.

The differences between the two annotations appear to be small. The MUC19 transcript identified by RefSeq may have immunological significance given the association of MUC19 with Crohn's disease and ulcerative colitis [105,264,265].

Genes that span gaps in the human assembly where the gaps are presumed to include repetitive sequence (e.g., MUC5AC [250]) are absent from the query result. Applying the tandem repeat finder algorithm [63] to the MUC5AC exon 31 sequence reported by Guo et al [250] revealed a longest tandem repeat of 1.6 kb.

Identifying Genes Spanning Long Segmental Duplications

Query 3 was used to identify a preliminary set of segmental duplications occuring within protein-coding genes, using the segmental duplication track [266] of the UCSC MySQL database server [64]. At least one duplicon was required to occur entirely within the gene sequence. The structure of the identified segmental duplications was examined using the UCSC genome browser. Where more than two contiguous tandem duplications exist (CR1, NEB, SPDYE3), the records for the gene were combined into a single record for the longer compound tandem repeat. When multiple segmental duplications overlapped (TTC34) only the longer duplication was retained.

Additional Queries

Query 4 was used to identify long self-alignments (score at least 60) within protein-coding genes, using the self-alignment track [267] of the UCSC MySQL database server [64]. Query 5 was used to identify repeats constituting almost an entire intron within a gene. Query 6 was used to identify long repeats in the mouse genome; repeats are required to overlap a protein-coding RefSeq gene, including 5 kb of sequence upstream of the gene start site.

Query 7 was used to identify pairs of long repeats where the second repeat unit is the reverse complement of the first. The purpose of this analysis is to understand the genomewide significance of this feature of the NSUN6 repeats (File S1). The output of this query was filtered to remove sequences on unplaced chromosomes and rows in which the two repeat sequences are not reverse complements. Queries 8 through 12 identify structural variation at STR loci utilizing information from the DGV database [268–270] (File S1).

Significance of Autoantigen Over-Representation in Gene Lists

Primary Autoantigens. To determine the statistical significance of a set of primary autoantigens within a gene list, an estimate of the number of known primary autoantigens for common autoimmune diseases is required. Based on Table 1, there are nineteen known primary autoantigens for those diseases. This number includes pANCA, a category covering five proteins in UC [271], and ribosomal P (3 proteins), so a more precise

estimate of the number of genes is 25. The null hypothesis H_0 states that each gene associated with a primary autoantigen is equally likely to appear anywhere in the ranked list of genes. There are 20,330 protein-coding genes in GENCODE V17 [272]. Choosing the top eleven genes is therefore well approximated by a binomial process, where a selected gene has a $25/20330 \approx 0.0012$ probability of being a primary autoantigen under the null hypothesis.

I apply an exact one-sided binomial test of goodness of fit. The p-value for 3 or more of the top 11 genes being primary autoantigens under the null hypothesis is 3.0×10^{-7}. The significance is robust to the size of the prefix of the gene list. For example, taking the top 35 genes rather than the top 11 yields $p < 1.2 \times 10^{-5}$. One can therefore reject the null hypothesis and conclude that the overrepresentation of primary autoantigens near the top of the list is highly significant.

Autoantigens. Determining the significance of a set of autoantigens within a gene list requires an estimate of the total number of autoantigens. Stadler et al. [54] tabulate 348 known autoantigens, but this list is incomplete (e.g., it does not include FLG or PKP3). For the purposes of determining a p value, 400 autoantigens and 20,330 protein-coding genes [272] are assumed for a one-sided binomial goodness of fit test. All p values calculated above remain significant at $p < 0.05$ even if an estimate of 600 autoantigens was used.

Supporting Information

File S1
(PDF)

Acknowledgments

Thanks to Judy Cho, Itsik Pe'er and Zoe Ross for helpful discussions. Several of the protein interactions mentioned in this report were identified with the aid of the STRING protein interaction database [273].

Author Contributions

Conceived and designed the experiments: KR. Performed the experiments: KR. Analyzed the data: KR. Contributed reagents/materials/analysis tools: KR. Wrote the paper: KR.

References

1. Ross KA (2011) Evidence for somatic gene conversion and deletion in bipolar disorder, Crohn's disease, coronary artery disease, hypertension, rheumatoid arthritis, type-1 diabetes, and type-2 diabetes. BMC Med 9: 12.
2. Engelhorn ME, Guevara-Patino JA, Noffz G, Hooper AT, Lou O, et al. (2006) Autoimmunity and tumor immunity induced by immune responses to mutations in self. Nat Med 12: 198–206.
3. Backes C, Ludwig N, Leidinger P, Harz C, Hoffmann J, et al. (2011) Immunogenicity of autoantigens. BMC Genomics 12: 340.
4. Lupski JR (2013) Genetics. Genome mosaicism–one human, multiple genomes. Science 341: 358–359.
5. Abyzov A, Mariani J, Palejev D, Zhang Y, Haney MS, et al. (2012) Somatic copy number mosaicism in human skin revealed by induced pluripotent stem cells. Nature 492: 438–442.
6. Poduri A, Evrony GD, Cai X, Walsh CA (2013) Somatic mutation, genomic variation, and neurological disease. Science 341: 1237758.
7. Iourov IY, Vorsanova SG, Yurov YB (2013) Somatic cell genomics of brain disorders: a new opportunity to clarify genetic-environmental interactions. Cytogenet Genome Res 139: 181–188.
8. Forsberg LA, Absher D, Dumanski JP (2013) Non-heritable genetics of human disease: spotlight on post-zygotic genetic variation acquired during lifetime. J Med Genet 50: 1–10.
9. Bruder CE, Piotrowski A, Gijsbers AA, Andersson R, Erickson S, et al. (2008) Phenotypically concordant and discordant monozygotic twins display different DNA copy-number-variation profiles. Am J Hum Genet 82: 763–771.
10. Liu P, Carvalho CM, Hastings PJ, Lupski JR (2012) Mechanisms for recurrent and complex human genomic rearrangements. Curr Opin Genet Dev 22: 211–220.
11. Asakawa J, Kodaira M, Nakamura N, Satoh C, Fujita M (1999) Chimerism in humans after intragenic recombination at the haptoglobin locus during early embryogenesis. Proc Natl Acad Sci USA 96: 10314–10319.
12. Melvold RW (1971) Spontaneous somatic reversion in mice. Effects of parental genotype on stability at the p-locus. Mutat Res 12: 171–174.
13. Schiestl RH, Khogali F, Carls N (1994) Reversion of the mouse pink-eyed unstable mutation induced by low doses of x-rays. Science 266: 1573–1576.
14. Lobachev KS, Rattray A, Narayanan V (2007) Hairpin- and cruciform-mediated chromosome breakage: causes and consequences in eukaryotic cells. Front Biosci 12: 4208–4220.
15. Flores M, Morales L, Gonzaga-Jauregui C, Dominguez-Vidana R, Zepeda C, et al. (2007) Recurrent DNA inversion rearrangements in the human genome. Proc Natl Acad Sci USA 104: 6099–6106.
16. Lam KW, Jeffreys AJ (2006) Processes of copy-number change in human DNA: the dynamics of alpha-globin gene deletion. Proc Natl Acad Sci USA 103: 8921–8927.
17. Bois PR (2003) Hypermutable minisatellites, a human affair? Genomics 81: 349–355.
18. Ellegren H (2004) Microsatellites: simple sequences with complex evolution. Nat Rev Genet 5: 435–445.
19. Bachmann MP, Bartsch H, Gross JK, Maier SM, Gross TF, et al. (2006) Autoimmunity as a result of escape from RNA surveillance. J Immunol 177: 1698–1707.

20. Semsei I, Maier S, Workman-Azbill J, Urban L, Moser K, et al. (2007) Detection of a rare oligo(A) repeat tract mutation (8As–>7As) in the sequence encoding the La/SS-B autoantigen. Anal Biochem 370: 47–53.

21. Tazi J, Bakkour N, Stamm S (2009) Alternative splicing and disease. Biochim Biophys Acta 1792: 14–26.

22. Yang F, Chen IH, Xiong Z, Yan Y, Wang H, et al. (2006) Model of stimulation-responsive splicing and strategies in identification of immunogenic isoforms of tumor antigens and autoantigens. Clin Immunol 121: 121–133.

23. Li X, Manley JL (2005) Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. Cell 122: 365–378.

24. Lin Y, Dent SY, Wilson JH, Wells RD, Napierala M (2010) R loops stimulate genetic instability of CTG.CAG repeats. Proc Natl Acad Sci USA 107: 692–697.

25. Yauk CL, Polyzos A, Rowan-Carroll A, Kortubash I, Williams A, et al. (2008) Tandem repeat mutation, global DNA methylation, and regulation of DNA methyltransferases in cultured mouse embryonic fibroblast cells chronically exposed to chemicals with different modes of action. Environ Mol Mutagen 49: 26–35.

26. Ball MP, Li JB, Gao Y, Lee JH, LeProust EM, et al. (2009) Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. Nat Biotechnol 27: 361–368.

27. Holmes R, Chang Y, Soloway PD (2006) Timing and sequence requirements defined for embryonic maintenance of imprinted DNA methylation at Rasgrf1. Mol Cell Biol 26: 9564–9570.

28. Brideau CM, Kauppinen KP, Holmes R, Soloway PD (2010) A non-coding RNA within the Rasgrf1 locus in mouse is imprinted and regulated by its homologous chromosome in trans. PLoS ONE 5: e13784.

29. Shukla S, Kavak E, Gregory M, Imashimizu M, Shutinoski B, et al. (2011) CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. Nature 479: 74–79.

30. Renaudineau Y, Youinou P (2011) Epigenetics and autoimmunity, with special emphasis on methylation. Keio J Med 60: 10–16.

31. Okada F, Nakai K, Kobayashi T, Shibata T, Tagami S, et al. (1999) Inflammatory cell-mediated tumour progression and minisatellite mutation correlate with the decrease of antioxidative enzymes in murine fibrosarcoma cells. Br J Cancer 79: 377–385.

32. Lee SH, Chang DK, Goel A, Boland CR, Bugbee W, et al. (2003) Microsatellite instability and suppressed DNA repair enzyme expression in rheumatoid arthritis. J Immunol 170: 2214–2220.

33. Bondanza A, Manfredi AA, Zimmermann VS, Iannacone M, Tincani A, et al. (2001) Anti-beta2 glycoprotein I antibodies cause inflammation and recruit dendritic cells in platelet clearance. Thromb Haemost 86: 1257–1263.

34. Janeway C, Travers P, Walport M, Shlomchik M (2001) Immunobiology: The Immune System in Health and Disease, New York: Garland Science, chapter 9: The destruction of antibody-coated pathogens via Fc receptors. 5th edition edition.

35. Kostadinov RL, Kuhner MK, Li X, Sanchez CA, Galipeau PC, et al. (2013) NSAIDs modulate clonal evolution in Barrett's esophagus. PLoS Genet 9: e1003553.

36. Royle NJ, Clarkson RE, Wong Z, Jeffreys AJ (1988) Clustering of hypervariable minisatellites in the proterminal regions of human autosomes. Genomics 3: 352–360.

37. Naslund K, Saetre P, von Salome J, Bergstrom TF, Jareborg N, et al. (2005) Genome-wide prediction of human VNTRs. Genomics 85: 24–35.

38. Legendre M, Pochet N, Pak T, Verstrepen KJ (2007) Sequence-based estimation of minisatellite and microsatellite repeat variability. Genome Res 17: 1787–1796.

39. Dittwald P, Gambin T, Szafranski P, Li J, Amato S, et al. (2013) NAHR-mediated copy-number variants in a clinical population: mechanistic insights into both genomic disorders and Mendelizing traits. Genome Res 23: 1395–1409.

40. Buard J, Collick A, Brown J, Jeffreys AJ (2000) Somatic versus germline mutation processes at minisatellite CEB1 (D2S90) in humans and transgenic mice. Genomics 65: 95–103.

41. Firestein GS, Echeverri F, Yeo M, Zvaifler NJ, Green DR (1997) Somatic mutations in the p53 tumor suppressor gene in rheumatoid arthritis synovium. Proc Natl Acad Sci USA 94: 10895–10900.

42. Armour JA, Patel I, Thein SL, Fey MF, Jeffreys AJ (1989) Analysis of somatic mutations at human minisatellite loci in tumors and cell lines. Genomics 4: 328–334.

43. Lupski JR (2007) Genomic rearrangements and sporadic disease. Nat Genet 39: S43–47.

44. Schaap M, Lemmers RJ, Maassen R, van der Vliet PJ, Hoogerheide LF, et al. (2013) Genome-wide analysis of macrosatellite repeat copy number variation in worldwide populations: evidence for differences and commonalities in size distributions and size restrictions. BMC Genomics 14: 143.

45. Burn GL, Svensson L, Sanchez-Blanco C, Saini M, Cope AP (2011) Why is PTPN22 a good candidate susceptibility gene for autoimmune disease? FEBS Lett 585: 3689–3698.

46. Gough SC, Walker LS, Sansom DM (2005) CTLA4 gene polymorphism and autoimmunity. Immunol Rev 204: 102–115.

47. Zenewicz LA, Abraham C, Flavell RA, Cho JH (2010) Unraveling the genetics of autoimmunity. Cell 140: 791–797.

48. Cho JH, Gregersen PK (2011) Genomics and the multifactorial nature of human autoimmune disease. N Engl J Med 365: 1612–1623.

49. Cardenas-Roldan J, Rojas-Villarraga A, Anaya JM (2013) How do autoimmune diseases cluster in families? A systematic review and meta-analysis. BMC Med 11: 73.

50. Plotz PH (2003) The autoantibody repertoire: searching for order. Nat Rev Immunol 3: 73–78.

51. Casciola-Rosen L, Andrade F, Ulanet D, Wong WB, Rosen A (1999) Cleavage by granzyme B is strongly predictive of autoantigen status: implications for initiation of autoimmunity. J Exp Med 190: 815–826.

52. Utz PJ, Gensler TJ, Anderson P (2000) Death, autoantigen modifications, and tolerance. Arthritis Res 2: 101–114.

53. Izquierdo M, Grandien A, Criado LM, Robles S, Leonardo E, et al. (1999) Blocked negative selection of developing T cells in mice expressing the baculovirus p35 caspase inhibitor. EMBO J 18: 156–166.

54. Stadler MB, Arnold D, Frieden S, Luginbuhl S, Stadler BM (2005) Single nucleotide polymorphisms as a prerequisite for autoantigens. Eur J Immunol 35: 371–378.

55. Quarantino S, Ruf J, Osman M, Guo J, McLachlan S, et al. (2005) Human autoantibodies modulate the T cell epitope repertoire but fail to unmask a pathogenic cryptic epitope. J Immunol 174: 557–563.

56. Gauba V, Grunewald J, Gorney V, Deaton LM, Kang M, et al. (2011) Loss of CD4 T-cell-dependent tolerance to proteins with modified amino acids. Proc Natl Acad Sci USA 108: 12821–12826.

57. Alaedini A, Green PH (2008) Autoantibodies in celiac disease. Autoimmunity 41: 19–26.

58. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, et al. (2001) Initial sequencing and analysis of the human genome. Nature 409: 860–921.

59. Myers EW, Sutton GG, Smith HO, Adams MD, Venter JC (2002) On the sequencing and assembly of the human genome. Proc Natl Acad Sci USA 99: 4145–4146.

60. Alkan C, Sajjadian S, Eichler EE (2011) Limitations of next-generation genome sequence assembly. Nat Methods 8: 61–65.

61. Gnerre S, Maccallum I, Przybylski D, Ribeiro FJ, Burton JN, et al. (2011) High-quality draft assemblies of mammalian genomes from massively parallel sequence data. Proc Natl Acad Sci USA 108: 1513–1518.

62. Treangen TJ, Salzberg SL (2012) Repetitive DNA and next-generation sequencing: computational challenges and solutions. Nat Rev Genet 13: 36–46.

63. Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res 27: 573–580.

64. Meyer LR, Zweig AS, Hinrichs AS, Karolchik D, Kuhn RM, et al. (2013) The UCSC Genome Browser database: extensions and updates 2013. Nucleic Acids Res 41: D64–69.

65. White RJ (2011) Transcription by RNA polymerase III: more complex than we thought. Nat Rev Genet 12: 459–463.

66. Nath SK, Harley JB, Lee YH (2005) Polymorphisms of complement receptor 1 and interleukin-10 genes and systemic lupus erythematosus: a meta-analysis. Hum Genet 118: 225–234.

67. Brouwers N, Van Cauwenberghe C, Engelborghs S, Lambert JC, Bettens K, et al. (2012) Alzheimer risk associated with a copy number variation in the complement receptor 1 increasing C3b/C4b binding sites. Mol Psychiatry 17: 223–233.

68. Erdei A, Isaak A, Torok K, Sandor N, Kremlitzka M, et al. (2009) Expression and role of CR1 and CR2 on B and T lymphocytes under physiological and autoimmune conditions. Mol Immunol 46: 2767–2773.

69. Annibale B, Lahner E, Fave GD (2011) Diagnosis and management of pernicious anemia. Curr Gastroenterol Rep 13: 518–524.

70. Taggart RT, Samloff IM, Raffel LJ, Graham A, Cass C, et al. (1986) Relationships between the human pepsinogen DNA and protein polymorphisms. Am J Hum Genet 38: 848–854.

71. Zavros Y, Waghray M, Tessier A, Bai L, Todisco A, et al. (2007) Reduced pepsin A processing of sonic hedgehog in parietal cells precedes gastric atrophy and transformation. J Biol Chem 282: 33265–33274.

72. Tripathi A, Lammers KM, Goldblum S, Shea-Donohue T, Netzel-Arnett S, et al. (2009) Identification of human zonulin, a physiological modulator of tight junctions, as prehaptoglobin-2. Proc Natl Acad Sci USA 106: 16799–16804.

73. Langlois MR, Delanghe JR (1996) Biological and clinical significance of haptoglobin polymorphism in humans. Clin Chem 42: 1589–1600.

74. Guerranti R, Bertocci E, Fioravanti A, Papakostas P, Montella A, et al. (2010) Serum proteome of patients with systemic sclerosis: molecular analysis of expression and prevalence of haptoglobin alpha chain isoforms. Int J Immunopathol Pharmacol 23: 901–909.

75. Adams JN, Cox AJ, Freedman BI, Langefeld CD, Carr JJ, et al. (2013) Genetic analysis of haptoglobin polymorphisms with cardiovascular disease and type 2 diabetes in the Diabetes Heart Study. Cardiovasc Diabetol 12: 31.

76. Fasano A, Shea-Donohue T (2005) Mechanisms of disease: the role of intestinal barrier function in the pathogenesis of gastrointestinal autoimmune diseases. Nat Clin Pract Gastroenterol Hepatol 2: 416–422.

77. Fasano A (2012) Zonulin, regulation of tight junctions, and autoimmune diseases. Ann N Y Acad Sci 1258: 25–33.

78. Ewing RM, Chu P, Elisma F, Li H, Taylor P, et al. (2007) Large-scale mapping of human protein-protein interactions by mass spectrometry. Mol Syst Biol 3: 89.

79. Fuchs F, Pau G, Kranz D, Sklyar O, Budjan C, et al. (2010) Clustering phenotype populations by genome-wide RNAi and multiparametric imaging. Mol Syst Biol 6: 370.

80. Cohen PT, Philp A, Vazquez-Martin C (2005) Protein phosphatase 4–from obscurity to vital functions. FEBS Lett 579: 3278–3286.

81. Salmon M, Gordon C (1999) The role of apoptosis in systemic lupus erythematosus. Rheumatology (Oxford) 38: 1177–1183.

82. Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, et al. (2002) Initial sequencing and comparative analysis of the mouse genome. Nature 420: 520–562.

83. Mourtada-Maarabouni M, Williams GT (2009) Protein phosphatase 4 regulates apoptosis in leukemic and primary human T-cells. Leuk Res 33: 1539–1551.

84. Scanu AM (2003) Lp(a) lipoprotein–coping with heterogeneity. N Engl J Med 349: 2089–2090.

85. Hansson GK, Hermansson A (2011) The immune system in atherosclerosis. Nat Immunol 12: 204–212.

86. Ariyo AA, Thach C, Tracy R (2003) Lp(a) lipoprotein, vascular disease, and mortality in the elderly. N Engl J Med 349: 2108–2115.

87. Dangas G, Ambrose JA, D'Agate DJ, Shao JH, Chockalingham S, et al. (1999) Correlation of serum lipoprotein(a) with the angiographic and clinical presentation of coronary artery disease. Am J Cardiol 83: 583–585.

88. Butterworth AS, Braund PS, Farrall M, Hardwick RJ, Saleheen D, et al. (2011) Large-scale genecentric analysis identifies novel variants for coronary artery disease. PLoS Genet 7: e1002260.

89. Teupser D, Baber R, Ceglarek U, Scholz M, Illig T, et al. (2010) Genetic regulation of serum phytosterol levels and risk of coronary artery disease. Circ Cardiovasc Genet 3: 331–339.

90. Hirschfield GM, Chapman RW, Karlsen TH, Lammert F, Lazaridis KN, et al. (2013) The genetics of complex cholestatic disorders. Gastroenterology 144: 1357–1374.

91. Fitzgerald ML, Mujawar Z, Tamehiro N (2010) ABC transporters, atherosclerosis and inflammation. Atherosclerosis 211: 361–370.

92. Graf GA, Yu L, Li WP, Gerard R, Tuma PL, et al. (2003) ABCG5 and ABCG8 are obligate heterodimers for protein trafficking and biliary cholesterol excretion. J Biol Chem 278: 48275–48282.

93. Silbernagel G, Chapman MJ, Genser B, Kleber ME, Fauler G, et al. (2013) High intestinal cholesterol absorption is associated with cardiovascular disease and risk alleles in ABCG8 and ABO: evidence from the LURIC and YFS cohorts and from a meta-analysis. J Am Coll Cardiol 62: 291–299.

94. Kenny EE, Gusev A, Riegel K, Lutjohann D, Lowe JK, et al. (2009) Systematic haplotype analysis resolves a complex plasma plant sterol locus on the Micronesian Island of Kosrae. Proc Natl Acad Sci USA 106: 13886–13891.

95. Ghiselli G, Schaefer EJ, Gascon P, Breser HB (1981) Type III hyperlipoproteinemia associated with apolipoprotein E deficiency. Science 214: 1239–1241.

96. Greenow K, Pearce NJ, Ramji DP (2005) The key role of apolipoprotein E in atherosclerosis. J Mol Med 83: 329–342.

97. Brown JM, Yu L (2009) Opposing Gatekeepers of Apical Sterol Transport: Niemann-Pick C1-Like 1 (NPC1L1) and ATP-Binding Cassette Transporters G5 and G8 (ABCG5/ABCG8). Immunol Endocr Metab Agents Med Chem 9: 18–29.

98. Moehle C, Ackermann N, Langmann T, Aslanidis C, Kel A, et al. (2006) Aberrant intestinal expression and allelic variants of mucin genes associated with inflammatory bowel disease. J Mol Med 84: 1055–1066.

99. Luu Y, Junker W, Rachagani S, Das S, Batra SK, et al. (2010) Human intestinal MUC17 mucin augments intestinal cell restitution and enhances healing of experimental colitis. Int J Biochem Cell Biol 42: 996–1006.

100. Buisine MP, Desreumaux P, Debailleul V, Gambiez L, Geboes K, et al. (1999) Abnormalities in mucin gene expression in Crohn's disease. Inflamm Bowel Dis 5: 24–32.

101. Resta-Lenert S, Das S, Batra SK, Ho SB (2011) Muc17 protects intestinal epithelial cells from enteroinvasive E. coli infection by promoting epithelial barrier integrity. Am J Physiol Gastrointest Liver Physiol 300: G1144–1155.

102. Kim YS, Ho SB (2010) Intestinal goblet cells and mucins in health and disease: recent insights and progress. Curr Gastroenterol Rep 12: 319–330.

103. Kobayashi K, Yagasaki M, Harada N, Chichibu K, Hibi T, et al. (2001) Detection of Fcgamma binding protein antigen in human sera and its relation with autoimmune diseases. Immunol Lett 79: 229–235.

104. Renner M, Bergmann G, Krebs I, End C, Lyer S, et al. (2007) DMBT1 confers mucosal protection in vivo and a deletion variant is associated with Crohn's disease. Gastroenterology 133: 1499–1509.

105. Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, et al. (2012) Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature 491: 119–124.

106. Najarian DJ, Gottlieb AB (2003) Connections between psoriasis and Crohn's disease. J Am Acad Dermatol 48: 805–821.

107. Schaeverbeke T, Truchetet ME, Richez C (2013) Gut metagenome and spondyloarthritis. Joint Bone Spine 80: 349–352.

108. Waldman A, Gilhar A, Duek L, Berdicevsky I (2001) Incidence of Candida in psoriasis–a study on the fungal flora of psoriatic patients. Mycoses 44: 77–81.

109. Smyth DJ, Cooper JD, Howson JM, Walker NM, Plagnol V, et al. (2008) PTPN22 Trp620 explains the association of chromosome 1p13 with type 1 diabetes and shows a statistical interaction with HLA class II genotypes. Diabetes 57: 1730–1737.

110. Parkes M (2012) The genetics universe of Crohn's disease and ulcerative colitis. Dig Dis 30 Suppl 1: 78–81.

111. McCormick DA, Horton LW, Mee AS (1990) Mucin depletion in inflammatory bowel disease. J Clin Pathol 43: 143–146.

112. Strugala V, Dettmar PW, Pearson JP (2008) Thickness and continuity of the adherent colonic mucus barrier in active and quiescent ulcerative colitis and Crohn's disease. Int J Clin Pract 62: 762–769.

113. Devlin SM, Yang H, Ippoliti A, Taylor KD, Landers CJ, et al. (2007) NOD2 variants and antibody response to microbial antigens in Crohn's disease patients and their unaffected relatives. Gastroenterology 132: 576–586.

114. Rosenstiel P, Sina C, End C, Renner M, Lyer S, et al. (2007) Regulation of DMBT1 via NOD2 and TLR4 in intestinal epithelial cells modulates bacterial recognition and invasion. J Immunol 178: 8203–8211.

115. Liaskos C, Rigopoulou EI, Orfanidou T, Bogdanos DP, Papandreou CN (2013) CUZD1 and anti-CUZD1 antibodies as markers of cancer and inflammatory bowel diseases. Clin Dev Immunol 2013: 968041.

116. Boulatnikov I, De Lisle RC (2004) Binding of the Golgi sorting receptor muclin to pancreatic zymogens through sulfated O-linked oligosaccharides. J Biol Chem 279: 40918–40926.

117. De Lisle RC (2002) Role of sulfated O-linked glycoproteins in zymogen granule formation. J Cell Sci 115: 2941–2952.

118. Jacob M, Laine J, LeBel D (1992) Specific interactions of pancreatic amylase at acidic pH. Amylase and the major protein of the zymogen granule membrane (GP-2) bind to immobilized or polymerized amylase. Biochem Cell Biol 70: 1105–1114.

119. Castro I, Sepulveda D, Cortes J, Quest AF, Barrera MJ, et al. (2013) Oral dryness in Sjgren's syndrome patients. Not just a question of water. Autoimmun Rev 12: 567–574.

120. Barrera MJ, Sanchez M, Aguilera S, Alliende C, Bahamondes V, et al. (2012) Aberrant localization of fusion receptors involved in regulated exocytosis in salivary glands of Sjgren's syndrome patients is linked to ectopic mucin secretion. J Autoimmun 39: 83–92.

121. Chaturvedi P, Singh AP, Batra SK (2008) Structure, evolution, and biology of the MUC4 mucin. FASEB J 22: 966–981.

122. Tincani A, Andreoli L, Cavazzana I, Doria A, Favero M, et al. (2013) Novel aspects of Sjgren's syndrome in 2012. BMC Med 11: 93.

123. Wickstrom C, Davies JR, Eriksen GV, Veerman EC, Carlstedt I (1998) MUC5B is a major gel-forming, oligomeric mucin from human salivary gland, respiratory tract and endocervix: identification of glycoforms and C-terminal cleavage. Biochem J 334 (Pt 3): 685–693.

124. Spurr-Michaud S, Argueso P, Gipson I (2007) Assay of mucins in human tear fluid. Exp Eye Res 84: 939–950.

125. Lorentzen AR, Karlsen TH, Olsson M, Smestad C, Mero IL, et al. (2009) Killer immunoglobulin-like receptor ligand HLA-Bw4 protects against multiple sclerosis. Ann Neurol 65: 658–666.

126. Vierra-Green C, Roe D, Hou L, Hurley CK, Rajalingam R, et al. (2012) Allele-level haplotype frequencies and pairwise linkage disequilibrium for 14 KIR loci in 506 European-American individuals. PLoS ONE 7: e47491.

127. Karlsen TH, Boberg KM, Olsson M, Sun JY, Senitzer D, et al. (2007) Particular genetic variants of ligands for natural killer cell receptors may contribute to the HLA associated risk of primary sclerosing cholangitis. J Hepatol 46: 899–906.

128. Zvyagin IV, Mamedov IZ, Britanova OV, Staroverov DB, Nasonov EL, et al. (2010) Contribution of functional KIR3DL1 to ankylosing spondylitis. Cell Mol Immunol 7: 471–476.

129. Moon SJ, Oh EJ, Kim Y, Kim KS, Kwok SK, et al. (2013) Diversity of killer cell immunoglobulin-like receptor genes in uveitis associated with autoimmune diseases: ankylosing spondylitis and Behet disease. Ocul Immunol Inflamm 21: 135–143.

130. Zeng X, Chen H, Gupta R, Paz-Altschul O, Bowcock AM, et al. (2013) Deletion of the activating NKG2C receptor and a functional polymorphism in its ligand HLA-E in psoriasis susceptibility. Exp Dermatol 22: 679–681.

131. Li X, Li J, Yang Y, Hou R, Liu R, et al. (2013) Differential gene expression in peripheral blood T cells from patients with psoriasis, lichen planus, and atopic dermatitis. J Am Acad Dermatol.

132. Zaguia F, Saikali P, Ludwin S, Newcombe J, Beauseigle D, et al. (2013) Cytotoxic NKG2C+ CD4 T cells target oligodendrocytes in multiple sclerosis. J Immunol 190: 2510–2518.

133. Nielsen N, Ødum N, Ursø B, Lanier LL, Spee P (2012) Cytotoxicity of CD56(bright) NK cells towards autologous activated CD4+ T cells is mediated through NKG2D, LFA-1 and TRAIL and dampened via CD94/NKG2A. PLoS ONE 7: e31959.

134. Lu L, Kim HJ, Werneck MB, Cantor H (2008) Regulation of CD8+ regulatory T cells: Interruption of the NKG2A-Qa-1 interaction allows robust suppressive activity and resolution of autoimmune disease. Proc Natl Acad Sci USA 105: 19420–19425.

135. Son SW, Kim EO, Ryu ES, Kim TJ, Kim JN, et al. (2009) Upregulation of Fas and downregulation of CD94/NKG2A inhibitory receptors on circulating natural killer cells in patients with new-onset psoriasis. Br J Dermatol 161: 281–288.

136. Kaur G, Trowsdale J, Fugger L (2013) Natural killer cells and their receptors in multiple sclerosis. Brain 136: 2657–2676.

137. Kastrukoff LF, Morgan NG, Zecchini D, White R, Petkau AJ, et al. (1999) Natural killer cells in relapsing-remitting MS: effect of treatment with interferon beta-1B. Neurology 52: 351–359.

138. Kastrukoff LF, Lau A, Wee R, Zecchini D, White R, et al. (2003) Clinical relapses of multiple sclerosis are associated with 'novel' valleys in natural killer cell functional activity. J Neuroimmunol 145: 103–114.

139. Cameron AL, Kirby B, Griffiths CE (2003) Circulating natural killer cells in psoriasis. Br J Dermatol 149: 160–164.

140. Baxter AG, Smyth MJ (2002) The role of NK cells in autoimmune disease. Autoimmunity 35: 1–14.

141. Penberthy WT (2009) Nicotinic acid-mediated activation of both membrane and nuclear receptors towards therapeutic glucocorticoid mimetics for treating multiple sclerosis. PPAR Res 2009: 853707.

142. Fiorillo E, Orru V, Stanford SM, Liu Y, Salek M, et al. (2010) Autoimmune-associated PTPN22 R620W variation reduces phosphorylation of lymphoid phosphatase on an inhibitory tyrosine residue. J Biol Chem 285: 26506–26518.

143. Criswell LA, Pfeiffer KA, Lum RF, Gonzales B, Novitzke J, et al. (2005) Analysis of families in the multiple autoimmune disease genetics consortium (MADGC) collection: the PTPN22 620W allele associates with multiple autoimmune phenotypes. Am J Hum Genet 76: 561–571.

144. Menard L, Saadoun D, Isnardi I, Ng YS, Meyers G, et al. (2011) The PTPN22 allele encoding an R620W variant interferes with the removal of developing autoreactive B cells in humans. J Clin Invest 121: 3635–3644.

145. Arechiga AF, Habib T, He Y, Zhang X, Zhang ZY, et al. (2009) Cutting edge: the PTPN22 allelic variant associated with autoimmunity impairs B cell signaling. J Immunol 182: 3343–3347.

146. Zhang J, Zahir N, Jiang Q, Miliotis H, Heyraud S, et al. (2011) The autoimmune disease-associated PTPN22 variant promotes calpain-mediated Lyp/Pep degradation associated with lymphocyte and dendritic cell hyperresponsiveness. Nat Genet 43: 902–907.

147. Conrad DF, Pinto D, Redon R, Feuk L, Gokcumen O, et al. (2010) Origins and functional impact of copy number variation in the human genome. Nature 464: 704–712.

148. Bonne S, Gilbert B, Hatzfeld M, Chen X, Green KJ, et al. (2003) Defining desmosomal plakophilin-3 interactions. J Cell Biol 161: 403–416.

149. Arvanitis DN, Yang W, Boggs JM (2002) Myelin proteolipid protein, basic protein, the small isoform of myelin-associated glycoprotein, and p42MAPK are associated in the Triton X-100 extract of central nervous system myelin. J Neurosci Res 70: 8–23.

150. Langkamp M, Hornig SC, Hornig JB, Kirschner M, Pridzun L, et al. (2009) Detection of myelin autoantibodies: evaluation of an assay system for diagnosis of multiple sclerosis in differentiation from other central nervous system diseases. Clin Chem Lab Med 47: 1395–1400.

151. Andersson M, Yu M, Soderstrom M, Weerth S, Baig S, et al. (2002) Multiple MAG peptides are recognized by circulating T and B lymphocytes in polyneuropathy and multiple sclerosis. Eur J Neurol 9: 243–251.

152. Lunn MP, Crawford TO, Hughes RA, Griffin JW, Sheikh KA (2002) Anti-myelin-associated gly-coprotein antibodies alter neurofilament spacing. Brain 125: 904–911.

153. Leslie RD, Hawa M (1994) Twin studies in auto-immune disease. Acta Genet Med Gemellol (Roma) 43: 71–81.

154. Chandley MJ, Miller MN, Kwasigroch CN, Wilson TD, Miller BE (2009) Increased antibodies for the alpha7 subunit of the nicotinic receptor in schizophrenia. Schizophr Res 109: 98–101.

155. Warburton PE, Hasson D, Guillem F, Lescale C, Jin X, et al. (2008) Analysis of the largest tandemly repeated DNA families in the human genome. BMC Genomics 9: 533.

156. Lee HH (2013) Variants of the CYP21A2 and CYP21A1P genes in congenital adrenal hyperplasia. Clin Chim Acta 418: 37–44.

157. Sedman L, Padhukasahasram B, Kelgo P, Laan M (2008) Complex signatures of locus-specific selective pressures and gene conversion on Human Growth Hormone/Chorionic Somatomam-motropin genes. Hum Mutat 29: 1181–1193.

158. Day IN, Chen XH, Gaunt TR, King TH, Voropanov A, et al. (2004) Late life metabolic syndrome, early growth, and common polymorphism in the growth hormone and placental lactogen gene cluster. J Clin Endocrinol Metab 89: 5569–5576.

159. Kobayashi T, Yabe S, Kikuchi T, Kanda T, Kobayashi I (1997) Presence of anti-pituitary anti-bodies and GAD antibodies in NIDDM and IDDM. Diabetes Care 20: 864–866.

160. Bellastella G, Maiorino MI, Olita L, De Bellis A, Giugliano D, et al. (2013) Anti-pituitary antibodies and hypogonadotropic hypogonadism in type 2 diabetes: in search of a role. Diabetes Care 36: e116–117.

161. Yabe S, Murakami M, Maruyama K, Miwa H, Fukumura Y, et al. (1995) Western blot analysis of rat pituitary antigens recognized by human antipituitary antibodies. Endocr J 42: 115–119.

162. Delvecchio M, De Bellis A, Francavilla R, Rutigliano V, Predieri B, et al. (2010) Anti-pituitary antibodies in children with newly diagnosed celiac disease: a novel finding contributing to linear-growth impairment. Am J Gastroenterol 105: 691–696.

163. Wongsurawat T, Jenjaroenpun P, Kwoh CK, Kuznetsov V (2012) Quantitative model of R-loop forming structures reveals a novel level of RNA-DNA interactome complexity. Nucleic Acids Res 40: e16.

164. Bei R, Masuelli L, Palumbo C, Modesti M, Modesti A (2009) A common repertoire of autoantibodies is shared by cancer and autoimmune disease patients: Inflammation in their induction and impact on tumor growth. Cancer Lett 281: 8–23.

165. Mullighan CG, Miller CB, Radtke I, Phillips LA, Dalton J, et al. (2008) BCR-ABL1 lymphoblastic leukaemia is characterized by the deletion of Ikaros. Nature 453: 110–114.

166. Plater-Zyberk C, Maini RN, Lam K, Kennedy TD, Janossy G (1985) A rheumatoid arthritis B cell subset expresses a phenotype similar to that in chronic lymphocytic leukemia. Arthritis Rheum 28: 971–976.

167. Doster A, Ziegler S, Foermer S, Rieker RJ, Heeg K, et al. (2013) Phosphorothioate-modified CpG oligodeoxynucleotides mimic autoantigens and reveal a potential role for Toll-like receptor 9 in receptor revision. Immunology 139: 166–178.

168. Zhang Z, Wu X, Limbaugh BH, Bridges SL (2001) Expression of recombination-activating genes and terminal deoxynucleotidyl transferase and secondary rearrangement of immunoglobulin kappa light chains in rheumatoid arthritis synovial tissue. Arthritis Rheum 44: 2275–2284.

169. van de Stadt LA, de Koning MH, van de Stadt RJ, Wolbink G, Dijkmans BA, et al. (2011) Development of the anti-citrullinated protein antibody repertoire prior to the onset of rheumatoid arthritis. Arthritis Rheum 63: 3226–3233.

170. Shmerling RH, Delbanco TL (1991) The rheumatoid factor: an analysis of clinical utility. Am J Med 91: 528–534.

171. Lauc G, Huffman JE, Pucic M, Zgaga L, Adamczyk B, et al. (2013) Loci associated with N-glycosylation of human immunoglobulin G show pleiotropy with autoimmune diseases and haema-tological cancers. PLoS Genet 9: e1003225.

172. Newkirk MM (2002) Rheumatoid factors: host resistance or autoimmunity? Clin Immunol 104: 1–13.

173. Witte T, Hartung K, Sachse C, Matthias T, Fricke M, et al. (2000) Rheumatoid factors in systemic lupus erythematosus: association with clinical and laboratory parameters. SLE study group. Rheumatol Int 19: 107–111.

174. Westra HJ, Peters MJ, Esko T, Yaghootkar H, Schurmann C, et al. (2013) Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat Genet 45: 1238–1243.

175. Hu W, Sun L, Gao J, Li Y, Wang P, et al. (2011) Down-regulated expression of IKZF1 mRNA in peripheral blood mononuclear cells from patients with systemic lupus erythematosus. Rheumatol Int 31: 819–822.

176. Mahdi H, Fisher BA, Kallberg H, Plant D, Malmstrom V, et al. (2009) Specific interaction between genotype, smoking and autoimmunity to citrullinated alpha-enolase in the etiology of rheumatoid arthritis. Nat Genet 41: 1319–1324.

177. van der Woude D, Alemayehu WG, Verduijn W, de Vries RR, Houwing-Duistermaat JJ, et al. (2010) Gene-environment interaction influences the reactivity of autoantibodies to citrullinated antigens in rheumatoid arthritis. Nat Genet 42: 814–816.

178. Bang H, Egerer K, Gauliard A, Luthke K, Rudolph PE, et al (2007) Mutation and citrullination modifies vimentin to a novel autoantigen for rheumatoid arthritis. Arthritis Rheum 56: 2503–2511.

179. Hedstrom AK, Sundqvist E, Baarnhielm M, Nordin N, Hillert J, et al. (2011) Smoking and two human leukocyte antigen genes interact to increase the risk for multiple sclerosis. Brain 134: 653–664.

180. Farid NR, Kascur V, Balazs C (2000) The human thyrotropin receptor is highly mutable: a review of gain-of-function mutations. Eur J Endocrinol 143: 25–30.

181. Fuhrer D, Holzapfel HP, Wonerow P, Scherbaum WA, Paschke R (1997) Somatic mutations in the thyrotropin receptor gene and not in the Gs alpha protein gene in 31 toxic thyroid nodules. J Clin Endocrinol Metab 82: 3885–3891.

182. Russo D, Arturi F, Suarez HG, Schlumberger M, Du Villard JA, et al. (1996) Thyrotropin receptor gene alterations in thyroid hyperfunctioning adenomas. J Clin Endocrinol Metab 81: 1548–1551.

183. Sharp L, Vernino S (2012) Paraneoplastic neuromuscular disorders. Muscle Nerve 46: 841–850.

184. Dalmau J, Tuzun E, Wu HY, Masjuan J, Rossi JE, et al. (2007) Paraneoplastic anti-N-methyl-D-aspartate receptor encephalitis associated with ovarian teratoma. Ann Neurol 61: 25–36.

185. Cunningham MW (2000) Pathogenesis of group A streptococcal infections. Clin Microbiol Rev 13: 470–511.

186. Fae KC, da Silva DD, Oshiro SE, Tanaka AC, Pomerantzeff PM, et al. (2006) Mimicry in recognition of cardiac myosin peptides by heart-intralesional T cell clones from rheumatic heart disease. J Immunol 176: 5662–5670.

187. Tandon R, Sharma M, Chandrashekhar Y, Kotb M, Yacoub MH, et al. (2013) Revisiting the pathogenesis of rheumatic fever and carditis. Nat Rev Cardiol 10: 171–177.

188. (2013). Reactome. http://www.reactome.org/.

189. Icenogle LM, Hengel SM, Coye LH, Streifel A, Collins CM, et al. (2012) Molecular and biological characterization of Streptococcal SpyA-mediated ADP-ribosylation of intermediate filament protein vimentin. J Biol Chem 287: 21481–21491.

190. Bryant AE, Bayer CR, Huntington JD, Stevens DL (2006) Group A streptococcal myonecrosis: increased vimentin expression after skeletal-muscle injury mediates the binding of Streptococcus pyogenes. J Infect Dis 193: 1685–1692.

191. Andrews SF, Wilson PC (2010) The anergic B cell. Blood 115: 4976–4978.
192. Doyle HA, Mamula MJ (2012) Autoantigenesis: the evolution of protein modifications in autoimmune disease. Curr Opin Immunol 24: 112–118.
193. Zheng W, Wang Y, Chang T, Huang H, Yee JK (2013) Significant differences in genotoxicity induced by retrovirus integration in human T cells and induced pluripotent stem cells. Gene 519: 142–149.
194. Santoni FA, Hartley O, Luban J (2010) Deciphering the code for retroviral integration target site selection. PLoS Comput Biol 6: e1001008.
195. Baillie JK, Barnett MW, Upton KR, Gerhardt DJ, Richmond TA, et al. (2011) Somatic retro-transposition alters the genetic landscape of the human brain. Nature 479: 534–537.
196. Hansen B, Oturai AB, Harbo HF, Celius EG, Nissen KK, et al. (2011) Genetic association of multiple sclerosis with the marker rs391745 near the endogenous retroviral locus HERV-Fc1: analysis of disease subtypes. PLoS ONE 6: e26438.
197. Perron H, Bernard C, Bertrand JB, Lang AB, Popa I, et al. (2009) Endogenous retroviral genes, Herpesviruses and gender in Multiple Sclerosis. J Neurol Sci 286: 65–72.
198. Karlsson H, Bachmann S, Schroder J, McArthur J, Torrey EF, et al. (2001) Retroviral RNA identified in the cerebrospinal fluids and brains of individuals with schizophrenia. Proc Natl Acad Sci USA 98: 4634–4639.
199. Douville R, Liu J, Rothstein J, Nath A (2011) Identification of active loci of a human endogenous retrovirus in neurons of patients with amyotrophic lateral sclerosis. Ann Neurol 69: 141–151.
200. Rolland A, Jouvin-Marche E, Viret C, Faure M, Perron H, et al. (2006) The envelope protein of a human endogenous retrovirus-W family activates innate immunity through CD14/TLR4 and promotes Th1-like responses. J Immunol 176: 7636–7644.
201. Nihtyanova SI, Denton CP (2010) Autoantibodies as predictive tools in systemic sclerosis. Nat Rev Rheumatol 6: 112–116.
202. Mahler M, You D, Baron M, Taillefer SS, Hudson M, et al. (2011) Anti-centromere antibodies in a large cohort of systemic sclerosis patients: comparison between immunofluorescence, CENP-A and CENP-B ELISA. Clin Chim Acta 412: 1937–1943.
203. Majone F, Cozzi F, Tonello M, Olivieri S, Montaldi A, et al. (2007) Unstabilized DNA breaks in lymphocytes of patients with different subsets of systemic sclerosis. Ann N Y Acad Sci 1108: 240–248.
204. Zhou X, Lin W, Tan FK, Assassi S, Fritzler MJ, et al. (2011) Decreased catalytic function with altered sumoylation of DNA topoisomerase I in the nuclei of scleroderma fibroblasts. Arthritis Res Ther 13: R128.
205. Khodzhigorova A, Distler A, Lang V, Dees C, Schneider H, et al. (2012) Inhibition of sumoylation prevents experimental fibrosis. Ann Rheum Dis 71: 1904–1908.
206. Golebiowski F, Matic I, Tatham MH, Cole C, Yin Y, et al. (2009) System-wide changes to SUMO modifications in response to heat shock. Sci Signal 2: ra24.
207. Tatham MH, Matic I, Mann M, Hay RT (2011) Comparative proteomic analysis identifies a role for SUMO in protein quality control. Sci Signal 4: rs4.
208. Bruderer R, Tatham MH, Plechanovova A, Matic I, Garg AK, et al. (2011) Purification and identification of endogenous polySUMO conjugates. EMBO Rep 12: 142–148.
209. Becker J, Barysch SV, Karaca S, Dittner C, Hsiao HH, et al. (2013) Detecting endogenous SUMO targets in mammalian cells and tissues. Nat Struct Mol Biol 20: 525–531.
210. Janka C, Selmi C, Gershwin ME, Will H, Sternsdorf T (2005) Small ubiquitin-related modifiers: A novel and independent class of autoantigens in primary biliary cirrhosis. Hepatology 41: 609–616.
211. Imura-Kumada S, Hasegawa M, Matsushita T, Hamaguchi Y, Encabo S, et al. (2012) High prevalence of primary biliary cirrhosis and disease-associated autoantibodies in Japanese patients with systemic sclerosis. Mod Rheumatol 22: 892–898.
212. Cavazzana I, Ceribelli A, Taraborelli M, Fredi M, Norman G, et al. (2011) Primary biliary cirrhosis-related autoantibodies in a large cohort of italian patients with systemic sclerosis. J Rheumatol 38: 2180–2185.
213. Carpenter WT, Buchanan RW (1994) Schizophrenia. N Engl J Med 330: 681–690.
214. Nguyen GH, Bouchard J, Boselli MG, Tolstoi LG, Keith L, et al. (2003) DNA stability and schizophrenia in twins. Am J Med Genet B Neuropsychiatr Genet 120B: 1–10.
215. O'Bleness MS, Dickens CM, Dumas LJ, Kehrer-Sawatzki H, Wyckoff GJ, et al. (2012) Evolutionary history and genome organization of DUF1220 protein domains. G3 (Bethesda) 2: 977–986.
216. Vandepoele K, Van Roy N, Staes K, Speleman F, van Roy F (2005) A novel gene family NBPF: intricate structure generated by gene duplications during primate evolution. Mol Biol Evol 22: 2265–2274.
217. Dumas L, Sikela JM (2009) DUF1220 domains, cognitive disease, and human brain evolution. Cold Spring Harb Symp Quant Biol 74: 375–382.
218. Dumas LJ, O'Bleness MS, Davis JM, Dickens CM, Anderson N, et al. (2012) DUF1220-domain copy number implicated in human brain-size pathology and evolution. Am J Hum Genet 91: 444–454.
219. Stone JL, O'Donovan MC, Gurling H, Kirov GK, Blackwood DH, et al. (2008) Rare chromosomal deletions and duplications increase risk of schizophrenia. Nature 455: 237–241.
220. Stefansson H, Rujescu D, Cichon S, Pietilainen OP, Ingason A, et al. (2008) Large recurrent microdeletions associated with schizophrenia. Nature 455: 232–236.
221. Lupski JR (2012) Brain copy number variants and neuropsychiatric traits. Biol Psychiatry 72: 617–619.
222. Luo XJ, Li M, Huang L, Nho K, Deng M, et al. (2012) The interleukin 3 gene (IL3) contributes to human brain volume variation by regulating proliferation and survival of neural progenitors. PLoS ONE 7: e50375.
223. Drummond JB, Tucholski J, Haroutunian V, Meador-Woodruff JH (2013) Transmembrane AMPA receptor regulatory protein (TARP) dysregulation in anterior cingulate cortex in schizophrenia. Schizophr Res 147: 32–38.
224. Waithe D, Ferron L, Dolphin AC (2011) Stargazin-related protein γ_7 is associated with signalling endosomes in superior cervical ganglion neurons and modulates neurite outgrowth. J Cell Sci 124: 2049–2057.
225. Kato AS, Gill MB, Ho MT, Yu H, Tu Y, et al. (2010) Hippocampal AMPA receptor gating controlled by both TARP and cornichon proteins. Neuron 68: 1082–1096.
226. Yamakawa H, Oyama S, Mitsuhashi H, Sasagawa N, Uchino S, et al. (2007) Neuroligins 3 and 4X interact with syntrophin-gamma2, and the interactions are affected by autism-related mutations. Biochem Biophys Res Commun 355: 41–46.
227. Suda S, Iwata K, Shimmura C, Kameno Y, Anitha A, et al. (2011) Decreased expression of axon-guidance receptors in the anterior cingulate cortex in autism. Mol Autism 2: 14.
228. Melke J, Goubran Botros H, Chaste P, Betancur C, Nygren G, et al. (2008) Abnormal melatonin synthesis in autism spectrum disorders. Mol Psychiatry 13: 90–98.
229. Wang L, Li J, Ruan Y, Lu T, Liu C, et al. (2013) Sequencing ASMT identifies rare mutations in Chinese Han patients with autism. PLoS ONE 8: e53727.
230. Jonsson L, Ljunggren E, Bremer A, Pedersen C, Landen M, et al. (2010) Mutation screening of melatonin-related genes in patients with autism spectrum disorders. BMC Med Genomics 3: 10.
231. Williams BL, Hornig M, Buie T, Bauman ML, Cho Paik M, et al. (2011) Impaired carbohydrate digestion and transport and mucosal dysbiosis in the intestines of children with autism and gastrointestinal disturbances. PLoS ONE 6: e24585.
232. Neale BM, Kou Y, Liu L, Ma'ayan A, Samocha KE, et al. (2012) Patterns and rates of exonic de novo mutations in autism spectrum disorders. Nature 485: 242–245.
233. Rual JF, Venkatesan K, Hao T, Hirozane-Kishikawa T, Dricot A, et al. (2005) Towards a proteomescale map of the human protein-protein interaction network. Nature 437: 1173–1178.
234. Nabi R, Serajee FJ, Chugani DC, Zhong H, Huq AH (2004) Association of tryptophan 2,3 dioxy-genase gene polymorphism with autism. Am J Med Genet B Neuropsychiatr Genet 125B: 63–68.
235. Benoist C, Mathis D (2001) Autoimmunity provoked by infection: how good is the case for T cell epitope mimicry? Nat Immunol 2: 797–801.
236. Blander JM, Torchinsky MB, Campisi L (2012) Revisiting the old link between infection and autoimmune disease with commensals and T helper 17 cells. Immunol Res 54: 50–68.
237. Rose NR, Mackay IR (2000) Molecular mimicry: a critical look at exemplary instances in human diseases. Cell Mol Life Sci 57: 542–551.
238. Albert LJ, Inman RD (1999) Molecular mimicry and autoimmunity. N Engl J Med 341: 2068–2074.
239. Gershwin ME, Shoenfeld Y (2011) Cutting-edge issues in organ-specific autoimmunity. Clin Rev Allergy Immunol 41: 123–125.
240. Rook GA (2012) Hygiene hypothesis and autoimmune diseases. Clin Rev Allergy Immunol 42: 5–15.
241. Duchan E, Patel DR (2012) Epidemiology of autism spectrum disorders. Pediatr Clin North Am 59: 27–43.
242. Mostafalou S, Abdollahi M (2013) Pesticides and human chronic diseases: evidences, mechanisms, and perspectives. Toxicol Appl Pharmacol 268: 157–177.
243. Farhat SC, Silva CA, Orione MA, Campos LM, Sallum AM, et al. (2011) Air pollution in autoimmune rheumatic diseases: a review. Autoimmun Rev 11: 14–21.
244. Smith FJ, Irvine AD, Terron-Kwiatkowski A, Sandilands A, Campbell LE, et al. (2006) Loss-of-function mutations in the gene encoding filaggrin cause ichthyosis vulgaris. Nat Genet 38: 337–342.
245. Nomura T, Sandilands A, Akiyama M, Liao H, Evans AT, et al. (2007) Unique mutations in the filaggrin gene in Japanese patients with ichthyosis vulgaris and atopic dermatitis. J Allergy Clin Immunol 119: 434–440.
246. Kubosaki A, Gross S, Miura J, Saeki K, Zhu M, et al. (2004) Targeted disruption of the IA-2beta gene causes glucose intolerance and impairs insulin secretion but does not prevent the development of diabetes in NOD mice. Diabetes 53: 1684–1691.
247. Kastner DL, Aksentijevich I, Goldbach-Mansky R (2010) Autoinflammatory disease reloaded: a clinical perspective. Cell 140: 784–790.
248. Fowler JC, Teixeira AS, Vinall LE, Swallow DM (2003) Hypervariability of the membrane-associated mucin and cancer marker MUC1. Hum Genet 113: 473–479.
249. Kirby A, Gnirke A, Jaffe DB, Baresova V, Pochet N, et al. (2013) Mutations causing medullary cystic kidney disease type 1 lie in a large VNTR in MUC1 missed by massively parallel sequencing. Nat Genet 45: 299–303.
250. Guo X, Zheng S, Dang H, Pace RG, Stonebraker JR, et al. (2013) Genome Reference and Sequence Variation in the Large Repetitive Central Exon of Human MUC5AC. Am J Respir Cell Mol Biol.

251. Selmi C, Lu Q, Humble MC (2012) Heritability versus the role of the environment in autoimmunity. J Autoimmun 39: 249–252.

252. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, et al. (2009) Finding the missing heritability of complex diseases. Nature 461: 747–753.

253. Hannan AJ (2010) TRPing up the genome: Tandem repeat polymorphisms as dynamic sources of genetic variability in health and disease. Discov Med 10: 314–321.

254. Lupski JR, Belmont JW, Boerwinkle E, Gibbs RA (2011) Clan genomics and the complex architecture of human disease. Cell 147: 32–43.

255. Ito Y, Tanaka F, Yamamoto M, Doyu M, Nagamatsu M, et al. (1998) Somatic mosaicism of the expanded CAG trinucleotide repeat in mRNAs for the responsible gene of Machado-Joseph disease (MJD), dentatorubral-pallidoluysian atrophy (DRPLA), and spinal and bulbar muscular atrophy (SBMA). Neurochem Res 23: 25–32.

256. Sakabe J, Yamamoto M, Hirakawa S, Motoyama A, Ohta I, et al. (2013) Kallikrein-related peptidase 5 functions in proteolytic processing of profilaggrin in cultured human keratinocytes. J Biol Chem 288: 17179–17189.

257. Laning JC, Deluca JE, Isaacs And CM, Hardin-Young J (2001) In vitro analysis of CD40-CD154 and CD28-CD80/86 interactions in the primary T-cell response to allogeneic "nonprofessional" antigen presenting cells. Transplantation 71: 1467–1474.

258. Vlad G, Cortesini R, Suciu-Foca N (2005) License to heal: bidirectional interaction of antigen-specific regulatory T cells and tolerogenic APC. J Immunol 174: 5907–5914.

259. Aichele P, Kyburz D, Ohashi PS, Odermatt B, Zinkernagel RM, et al. (1994) Peptide-induced T-cell tolerance to prevent autoimmune diabetes in a transgenic mouse model. Proc Natl Acad Sci USA 91: 444–448.

260. Toes RE, Offringa R, Blom RJ, Melief CJ, Kast WM (1996) Peptide vaccination can lead to enhanced tumor growth through specific T-cell tolerance induction. Proc Natl Acad Sci USA 93: 7855–7860.

261. Han B, Serra P, Amrani A, Yamanouchi J, Maree AF, et al. (2005) Prevention of diabetes by manipulation of anti-IGRP autoimmunity: high efficiency of a low-affinity peptide. Nat Med 11: 645–652.

262. Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, et al. (2012) GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res 22: 1760–1774.

263. Pruitt KD, Tatusova T, Brown GR, Maglott DR (2012) NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy. Nucleic Acids Res 40: D130–135.

264. Kumar V, Mack DR, Marcil V, Israel D, Krupoves A, et al. (2013) Genome-wide association study signal at the 12q12 locus for Crohn's disease may represent associations with the MUC19 gene. Inflamm Bowel Dis 19: 1254–1259.

265. Phillips AM, Nimmo ER, Van Limbergen J, Drummond HE, Smith L, et al. (2010) Detailed haplotype-tagging study of germline variation of MUC19 in inflammatory bowel disease. Inflamm Bowel Dis 16: 557–558.

266. Bailey JA, Yavor AM, Massa HF, Trask BJ, Eichler EE (2001) Segmental duplications: organization and impact within the current human genome project assembly. Genome Res 11: 1005–1017.

267. Kent WJ, Baertsch R, Hinrichs A, Miller W, Haussler D (2003) Evolution's cauldron: duplication, deletion, and rearrangement in the mouse and human genomes. Proc Natl Acad Sci USA 100: 11484–11489.

268. Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, et al. (2004) Detection of large-scale variation in the human genome. Nat Genet 36: 949–951.

269. Zhang J, Feuk L, Duggan GE, Khaja R, Scherer SW (2006) Development of bioinformatics resources for display and analysis of copy number and other structural variants in the human genome. Cytogenet Genome Res 115: 205–214.

270. Wong LP, Ong RT, Poh WT, Liu X, Chen P, et al. (2013) Deep whole-genome sequencing of 100 southeast Asian Malays. Am J Hum Genet 92: 52–66.

271. Locht H, Skogh T, Wiik A (2000) Characterisation of autoantibodies to neutrophil granule constituents among patients with reactive arthritis, rheumatoid arthritis, and ulcerative colitis. Ann Rheum Dis 59: 898–903.

272. (2013). Encode project statistics. http://www.gencodegenes.org/archive stats.html.

273. Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, et al. (2013) STRING v9.1: protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Res 41: D808–815.

274. Chen JM, Cooper DN, Chuzhanova N, Ferec C, Patrinos GP (2007) Gene conversion: mechanisms, evolution and human disease. Nat Rev Genet 8: 762–775.

275. Cooper GS, Bynum ML, Somers EC (2009) Recent insights in the epidemiology of autoimmune diseases: improved prevalence estimates and understanding of clustering of diseases. J Autoimmun 33: 197–207.

276. Morshed SA, Latif R, Davies TF (2012) Delineating the autoimmune mechanisms in Graves' disease. Immunol Res 54: 191–203.

277. Choi KH, Lee EB, Yoo CD, Baek HJ, Kang SW, et al. (2005) Clinical significance of anti-filaggrin antibody recognizing uncitrullinated filaggrin in rheumatoid arthritis. Exp Mol Med 37: 546–552.

278. Aho K, Palosuo T, Heliovaara M, Knekt P, Alha P, et al. (2000) Antifilaggrin antibodies within "normal" range predict rheumatoid arthritis in a linear fashion. J Rheumatol 27: 2743–2746.

279. Huffmeier U, Boiers U, Lascorz J, Reis A, Burkhardt H (2008) Loss-of-function mutations in the filaggrin gene: no contribution to disease susceptibility, but to autoantibody formation against citrullinated peptides in early rheumatoid arthritis. Ann Rheum Dis 67: 131–133.

280. Uchida K, Akita Y, Matsuo K, Fujiwara S, Nakagawa A, et al. (2005) Identification of specific autoantigens in Sjgren's syndrome by SEREX. Immunology 116: 53–63.

281. Iaccarino L, Ghirardello A, Canova M, Zen M, Bettio S, et al. (2011) Anti-annexins autoantibodies: their role as biomarkers of autoimmune diseases. Autoimmun Rev 10: 553–558.

282. Halvorsen EH, Pollmann S, Gilboe IM, van der Heijde D, Landewe R, et al. (2008) Serum IgG antibodies to peptidylarginine deiminase 4 in rheumatoid arthritis and associations with disease severity. Ann Rheum Dis 67: 414–417.

283. Hutfless S, Matos P, Talor MV, Caturegli P, Rose NR (2011) Significance of prediagnostic thyroid antibodies in women with autoimmune thyroid disease. J Clin Endocrinol Metab 96: E1466–1471.

284. Rueda B, Nunez C, Orozco G, Lopez-Nevot MA, de la Concha EG, et al. (2005) C1858T functional variant of PTPN22 gene is not associated with celiac disease genetic predisposition. Hum Immunol 66: 848–852.

285. Santin I, Castellanos-Rubio A, Aransay AM, Castano L, Vitoria JC, et al. (2008) The functional R620W variant of the PTPN22 gene is associated with celiac disease. Tissue Antigens 71: 247–249.

286. Smyth DJ, Plagnol V, Walker NM, Cooper JD, Downes K, et al. (2008) Shared and distinct genetic variants in type 1 diabetes and celiac disease. N Engl J Med 359: 2767–2777.

287. Fasano A, Not T, Wang W, Uzzau S, Berti I, et al. (2000) Zonulin, a newly discovered modulator of intestinal permeability, and its expression in coeliac disease. Lancet 355: 1518–1519.

288. Clemente MG, Musu MP, Frau F, Brusco G, Sole G, et al. (2000) Immune reaction against the cytoskeleton in coeliac disease. Gut 47: 520–526.

289. Hagforsen E, Sunnerberg K, Michaelsson G, Kampe O, Hedstrand H (2007) Psoriasis autoantigens in normal scalp skin–identification by expression cloning. J Invest Dermatol 127: 2276–2280.

290. Song GG, Kim JH, Lee YH (2013) The CTLA-4 +49 A/G, CT60 A/G and PTPN22 1858 C/T polymorphisms and susceptibility to vitiligo: a meta-analysis. Mol Biol Rep 40: 2985–2993.

291. Jin Y, Birlea SA, Fain PR, Gowan K, Riccardi SL, et al. (2010) Variant of TYR and autoimmunity susceptibility loci in generalized vitiligo. N Engl J Med 362: 1686–1697.

292. Kemp EH, Emhemad S, Akhtar S, Watson PF, Gawkrodger DJ, et al. (2011) Autoantibodies against tyrosine hydroxylase in patients with non-segmental (generalised) vitiligo. Exp Dermatol 20: 35–40.

293. Li Q, Lv Y, Li C, Yi X, Long HA, et al. (2011) Vitiligo autoantigen VIT75 is identified as lamin A in vitiligo by serological proteome analysis based on mass spectrometry. J Invest Dermatol 131: 727–734.

294. Gomez LM, Anaya JM, Gonzalez CI, Pineda-Tamayo R, Otero W, et al. (2005) PTPN22 C1858T polymorphism in Colombian patients with autoimmune diseases. Genes Immun 6: 628–631.

295. Ittah M, Gottenberg JE, Proust A, Hachulla E, Puechal X, et al. (2005) No evidence for association between 1858 C/T single-nucleotide polymorphism of PTPN22 gene and primary Sjgren's syndrome. Genes Immun 6: 457–458.

296. Shiari R, Kobayashi I, Toita N, Hatano N, Kawamura N, et al. (2006) Epitope mapping of antialpha-fodrin autoantibody in juvenile Sjgren's syndrome: difference in major epitopes between primary and secondary cases. J Rheumatol 33: 1395–1400.

297. Ulbricht KU, Schmidt RE, Witte T (2003) Antibodies against alpha-fodrin in Sjgren's syndrome. Autoimmun Rev 2: 109–113.

298. Nguyen CQ, Peck AB (2009) Unraveling the pathophysiology of Sjogren syndrome-associated dry eye disease. Ocul Surf 7: 11–27.

299. Gugliesi F, De Andrea M, Mondini M, Cappello P, Giovarelli M, et al. (2010) The proapoptotic activity of the Interferon-inducible gene IFI16 provides new insights into its etiopathogenetic role in autoimmunity. J Autoimmun 35: 114–123.

300. Kovacs L, Feher E, Bodnar I, Marczinovits I, Nagy GM, et al. (2008) Demonstration of autoantibody binding to muscarinic acetylcholine receptors in the salivary gland in primary Sjgren's syndrome. Clin Immunol 128: 269–276.

301. Ghillani P, Andre C, Toly C, Rouquette AM, Bengoufa D, et al. (2011) Clinical significance of anti-Ro52 (TRIM21) antibodies non-associated with anti-SSA 60 kDa antibodies: results of a multicentric study. Autoimmun Rev 10: 509–513.

302. Pollock W, Toh BH (1999) Routine immunofluorescence detection of Ro/SS-A autoantibody using HEp-2 cells transfected with human 60 kDa Ro/SS-A. J Clin Pathol 52: 684–687.

303. van Schaik FD, Oldenburg B, Hart AR, Siersema PD, Lindgren S, et al. (2013) Serological markers predict inflammatory bowel disease years before the diagnosis. Gut 62: 683–688.

304. Takaishi H, Kanai T, Nakazawa A, Sugata F, Nikai A, et al. (2012) Anti-high mobility group box 1 and box 2 non-histone chromosomal proteins (HMGB1/HMGB2) antibodies and anti-Saccharomyces cerevisiae antibodies (ASCA): accuracy in differentially diagnosing UC and CD and correlation with inflammatory bowel disease phenotype. J Gastroenterol 47: 969–977.

305. Biancone L, Monteleone G, Marasco R, Pallone F (1998) Autoimmunity to tropomyosin isoforms in ulcerative colitis (UC) patients and unaffected relatives. Clin Exp Immunol 113: 198–205.

306. Orozco G, Garcia-Porrua C, Lopez-Nevot MA, Raya E, Gonzalez-Gay MA, et al. (2006) Lack of association between ankylosing spondylitis and a functional polymorphism of PTPN22 proposed as a general susceptibility marker for autoimmunity. Ann Rheum Dis 65: 687–688.

307. Wright C, Sibani S, Trudgian D, Fischer R, Kessler B, et al. (2012) Detection of multiple autoantibodies in patients with ankylosing spondylitis using nucleic acid programmable protein arrays. Mol Cell Proteomics 11: M9.00384.

308. Hoppu S, Harkonen T, Ronkainen MS, Simell S, Hekkala A, et al. (2006) IA-2 antibody isotypes and epitope specificity during the prediabetic process in children with HLA-conferred susceptibility to type I diabetes. Clin Exp Immunol 144: 59–66.

309. Torii S (2009) Expression and function of IA-2 family proteins, unique neuroendocrine-specific protein-tyrosine phosphatases. Endocr J 56: 639–648.

310. Tsirogianni A, Pipi E, Soufleros K (2009) Specificity of islet cell autoantibodies and coexistence with other organ specific autoantibodies in type 1 diabetes mellitus. Autoimmun Rev 8: 687–691.

311. Wenzlau JM, Juhl K, Yu L, Moua O, Sarkar SA, et al. (2007) The cation efflux transporter ZnT8 (Slc30a8) is a major autoantigen in human type 1 diabetes. Proc Natl Acad Sci USA 104: 17040–17045.

312. Hirai H, Miura J, Hu Y, Larsson H, Larsson K, et al. (2008) Selective screening of secretory vesicle-associated proteins for autoantigens in type 1 diabetes: VAMP2 and NPY are new minor autoantigens. Clin Immunol 127: 366–374.

313. Endo T, Takizawa S, Tanaka S, Takahashi M, Fujii H, et al. (2009) Amylase alpha-2A autoantibodies: novel marker of autoimmune pancreatitis and fulminant type 1 diabetes. Diabetes 58: 732–737.

314. Kemp EH, McDonagh AJ, Wengraf DA, Messenger AG, Gawkrodger DJ, et al. (2006) The nonsynonymous C1858T substitution in the PTPN22 gene is associated with susceptibility to the severe forms of alopecia areata. Hum Immunol 67: 535–539.

315. Betz RC, Konig K, Flaquer A, Redler S, Eigelshoven S, et al. (2008) The R620W polymorphism in PTPN22 confers general susceptibility for the development of alopecia areata. Br J Dermatol 158: 389–391.

316. Kemp EH, Sandhu HK, Weetman AP, McDonagh AJ (2011) Demonstration of autoantibodies against tyrosine hydroxylase in patients with alopecia areata. Br J Dermatol 165: 1236–1243.

317. Hedstrand H, Ekwall O, Haavik J, Landgren E, Betterle C, et al. (2000) Identification of tyrosine hydroxylase as an autoantigen in autoimmune polyendocrine syndrome type I. Biochem Biophys Res Commun 267: 456–461.

318. Leung MC, Sutton CW, Fenton DA, Tobin DJ (2010) Trichohyalin is a potential major autoantigen in human alopecia areata. J Proteome Res 9: 5153–5163.

319. Lee YH, Bae SC, Song GG (2012) The association between the functional PTPN22 1858 C/T and MIF -173 C/G polymorphisms and juvenile idiopathic arthritis: a meta-analysis. Inflamm Res 61: 411–415.

320. Hinks A, Cobb J, Marion MC, Prahalad S, Sudman M, et al. (2013) Dense genotyping of immune-related disease regions identifies 14 new susceptibility loci for juvenile idiopathic arthritis. Nat Genet 45: 664–669.

321. Mor-Vaknin N, Kappes F, Dick AE, Legendre M, Damoc C, et al. (2011) DEK in the synovium of patients with juvenile idiopathic arthritis: characterization of DEK antibodies and posttrans-lational modification of the DEK autoantigen. Arthritis Rheum 63: 556–567.

322. Zlacka D, Vavrincova P, Hien Nguyen TT, Hromadnikova I (2006) Frequency of anti-hsp60, -65 and -70 antibodies in sera of patients with juvenile idiopathic arthritis. J Autoimmun 27: 81–88.

323. Tebo AE, Jaskowski T, Davis KW, Whiting A, Clifford B, et al. (2012) Profiling anti-cyclic citrullinated peptide antibodies in patients with juvenile idiopathic arthritis. Pediatr Rheumatol Online J 10: 29.

324. Lipinska J, Brozik H, Stanczyk J, Smolewska E (2012) Anticitrullinated protein antibodies and radiological progression in juvenile idiopathic arthritis. J Rheumatol 39: 1078–1087.

325. Toh BH, Sentry JW, Alderuccio F (2000) The causative H+/K+ ATPase antigen in the pathogenesis of autoimmune gastritis. Immunol Today 21: 348–354.

326. Mardh S, Song YH (1990) The occurrence of auto-antibodies in patients with gastro-duodenal lesions. J Intern Med Suppl 732: 77–82.

327. Mardh S, Song YH (1989) Characterization of antigenic structures in auto-immune atrophic gastritis with pernicious anaemia. The parietal cell H,K-ATPase and the chief cell pepsinogen are the two major antigens. Acta Physiol Scand 136: 581–587.

328. Sadallah S, Hess C, Trendelenburg M, Vedeler C, Lopez-Trascasa M, et al. (2003) Autoantibodies against complement receptor 1 (CD35) in SLE, liver cirrhosis and HIV-infected patients. Clin Exp Immunol 131: 174–181.

329. McLaughlin KA, Wucherpfennig KW (2008) B cells and autoantibodies in the pathogenesis of multiple sclerosis and related inflammatory demyelinating diseases. Adv Immunol 98: 121–149.

330. Vyshkina T, Kalman B (2008) Autoantibodies and neurodegeneration in multiple sclerosis. Lab Invest 88: 796–807.

331. Ousman SS, Tomooka BH, van Noort JM, Wawrousek EF, O'Connor KC, et al. (2007) Protective and therapeutic role for alphaB-crystallin in autoimmune demyelination. Nature 448: 474–479.

332. Menge T, Lalive PH, von Budingen HC, Genain CP (2011) Conformational epitopes of myelin oligodendrocyte glycoprotein are targets of potentially pathogenic antibody responses in multiple sclerosis. J Neuroinflammation 8: 161.

333. Prescott NJ, Fisher SA, Onnie C, Pattni R, Steer S, et al. (2005) A general autoimmunity gene (PTPN22) is not associated with inflammatory bowel disease in a British population. Tissue Antigens 66: 318–320.

334. Komorowski L, Teegen B, Probst C, Aulinger-Stocker K, Sina C, et al. (2013) Autoantibodies against exocrine pancreas in Crohn's disease are directed against two antigens: The glycoproteins CUZD1 and GP2. J Crohns Colitis 7: 780–790.

335. Sherer Y, Gorstein A, Fritzler MJ, Shoenfeld Y (2004) Autoantibody explosion in systemic lupus erythematosus: more than 100 different antibodies found in SLE patients. Semin Arthritis Rheum 34: 501–537.

336. Arbuckle MR, McClain MT, Rubertone MV, Scofield RH, Dennis GJ, et al. (2003) Development of autoantibodies before the clinical onset of systemic lupus erythematosus. N Engl J Med 349: 1526–1533.

337. Chamley LW, Duncalf AM, Konarkowska B, Mitchell MD, Johnson PM (1999) Conformationally altered beta 2-glycoprotein I is the antigen for anti-cardiolipin autoantibodies. Clin Exp Immunol 115: 571–576.

338. Senecal JL, Rauch J, Grodzicky T, Raynauld JP, Uthman I, et al. (1999) Strong association of autoantibodies to human nuclear lamin B1 with lupus anticoagulant antibodies in systemic lupus erythematosus. Arthritis Rheum 42: 1347–1353.

339. Gomez-Puerta JA, Burlingame RW, Cervera R (2008) Anti-chromatin (anti-nucleosome) antibodies: diagnostic and clinical value. Autoimmun Rev 7: 606–611.

340. Bonfa E, Golombek SJ, Kaufman LD, Skelly S, Weissbach H, et al. (1987) Association between lupus psychosis and anti-ribosomal P protein antibodies. N Engl J Med 317: 265–271.

341. Isshi K, Hirohata S (1996) Association of anti-ribosomal P protein antibodies with neuropsychiatric systemic lupus erythematosus. Arthritis Rheum 39: 1483–1490.

342. Ortona E, Capozzi A, Colasanti T, Conti F, Alessandri C, et al. (2010) Vimentin/cardiolipin complex as a new antigenic target of the antiphospho-lipid syndrome. Blood 116: 2960–2967.

343. Heinlen LD, Ritterhouse LL, McClain MT, Keith MP, Neas BR, et al. (2010) Ribosomal P autoantibodies are present before SLE onset and are directed against non-C-terminal peptides. J Mol Med 88: 719–727.

344. McClain MT, Arbuckle MR, Heinlen LD, Dennis GJ, Roebuck J, et al. (2004) The prevalence, onset, and clinical significance of antiphospholipid antibodies prior to diagnosis of systemic lupus erythematosus. Arthritis Rheum 50: 1226–1232.

345. Boehm J, Orth T, Van Nguyen P, Soling HD (1994) Systemic lupus erythematosus is associated with increased auto-antibody titers against calreticulin and grp94, but calreticulin is not the Ro/SS-A antigen. Eur J Clin Invest 24: 248–257.

346. Eggleton P, Ward FJ, Johnson S, Khamashta MA, Hughes GR, et al. (2000) Fine specificity of autoantibodies to calreticulin: epitope mapping and characterization. Clin Exp Immunol 120: 384–391.

347. Minota S, Jarjour WN, Suzuki N, Nojima Y, Roubey RA, et al. (1991) Autoantibodies to nucleolin in systemic lupus erythematosus and other diseases. J Immunol 146: 2249–2252.

348. Bertolaccini ML, Sciascia S, Murru V, Garcia-Fernandez C, Sanna G, et al. (2013) Prevalence of antibodies to prothrombin in solid phase (aPT) and to phosphatidylserine-prothrombin complex (aPS/PT) in patients with and without lupus anticoagulant. Thromb Haemost 109: 207–213.

349. Reeves WH, Pierani A, Chou CH, Ng T, Nicastri C, et al. (1991) Epitopes of the p70 and p80 (Ku) lupus autoantigens. J Immunol 146: 2678–2686.

350. Lartigue A, Drouot L, Jouen F, Charlionet R, Tron F, et al. (2005) Association between antinucleophosmin and anti-cardiolipin antibodies in (NZW x BXSB)F1 mice and human systemic lupus erythematosus. Arthritis Res Ther 7: R1394–1403.

351. Abdulahad DA, Westra J, Bijzet J, Limburg PC, Kallenberg CG, et al. (2011) High mobility group box 1 (HMGB1) and anti-HMGB1 antibodies and their relation to disease characteristics in systemic lupus erythematosus. Arthritis Res Ther 13: R71.

352. Caccavo D, Rigon A, Picardi A, Galluzzo S, Vadacca M, et al. (2005) Anti-lactoferrin antibodies in systemic lupus erythematosus: isotypes and clinical correlates. Clin Rheumatol 24: 381–387.

353. Martin TM, Bye L, Modi N, Stanford MR, Vaughan R, et al. (2009) Genotype analysis of polymorphisms in autoimmune susceptibility genes, CTLA-4 and PTPN22, in an acute anterior uveitis cohort. Mol Vis 15: 208–212.

354. Deeg CA, Raith AJ, Amann B, Crabb JW, Thurau SR, et al. (2007) CRALBP is a highly prevalent autoantigen for human autoimmune uveitis. Clin Dev Immunol 2007: 39245.

355. Chen L, Holland GN, Yu F, Levinson RD, Lampi KJ, et al. (2008) Associations of seroreactivity against crystallin proteins with disease activity and cataract in patients with uveitis. Invest Ophthalmol Vis Sci 49: 4476–4481.

356. Roycroft M, Fichna M, McDonald D, Owen K, Zurawek M, et al. (2009) The tryptophan 620 allele of the lymphoid tyrosine phosphatase (PTPN22) gene predisposes to autoimmune Addison's disease. Clin Endocrinol (Oxf) 70: 358–362.

357. Winqvist O, Karlsson FA, Kampe O (1992) 21-Hydroxylase, a major autoantigen in idiopathic Addison's disease. Lancet 339: 1559–1562.

358. Myhre AG, Undlien DE, Løvås K, Uhlving S, Nedrebø BG, et al. (2002) Autoimmune adrenocortical failure in Norway autoantibodies and human leukocyte antigen class II associations related to clinical features. J Clin Endocrinol Metab 87: 618–623.

359. Vandiedonck C, Capdevielle C, Giraud M, Krumeich S, Jais JP, et al. (2006) Association of the PTPN22*R620W polymorphism with autoimmune myasthenia gravis. Ann Neurol 59: 404–407.

360. Lefvert AK, Zhao Y, Ramanujam R, Yu S, Pirskanen R, et al. (2008) PTPN22 R620W promotes production of anti-AChR autoantibodies and IL-2 in myasthenia gravis. J Neuroimmunol 197: 110–113.

361. Greve B, Hoffmann P, Illes Z, Rozsa C, Berger K, et al. (2009) The autoimmunity-related polymorphism PTPN22 1858C/T is associated with anti-titin antibody-positive myasthenia gravis. Hum Immunol 70: 540–542.

362. Gregersen PK, Kosoy R, Lee AT, Lamb J, Sussman J, et al. (2012) Risk for myasthenia gravis maps to a (151) ProAla change in TNIP1 and to human leukocyte antigen-B*08. Ann Neurol 72: 927–935.

363. Cossins J, Belaya K, Zoltowska K, Koneczny I, Maxwell S, et al. (2012) The search for new antigenic targets in myasthenia gravis. Ann N Y Acad Sci 1275: 123–128.

364. Suzuki S, Utsugisawa K, Nagane Y, Satoh T, Terayama Y, et al. (2007) Classification of myasthenia gravis based on autoantibody status. Arch Neurol 64: 1121–1124.

365. Cavalcante P, Bernasconi P, Mantegazza R (2012) Autoimmune mechanisms in myasthenia gravis. Curr Opin Neurol 25: 621–629.

366. Chinoy H, Platt H, Lamb JA, Betteridge Z, Gunawardena H, et al. (2008) The protein tyrosine phosphatase N22 gene is associated with juvenile and adult idiopathic inflammatory myopathy independent of the HLA 8.1 haplotype in British Caucasian patients. Arthritis Rheum 58: 3247–3254.

367. Casciola-Rosen L, Mammen AL (2012) Myositis autoantibodies. Curr Opin Rheumatol 24: 602–608.

368. Kubo M, Ihn H, Asano Y, Yamane K, Yazawa N, et al. (2002) Prevalence of 52-kd and 60-kd Ro/SS-A autoantibodies in Japanese patients with polymyositis/dermatomyositis. J Am Acad Dermatol 47: 148–151.

369. Diaz-Gallo LM, Gourh P, Broen J, Simeon C, Fonollosa V, et al. (2011) Analysis of the influence of PTPN22 gene polymorphisms in systemic sclerosis. Ann Rheum Dis 70: 454–462.

370. Lee YH, Choi SJ, Ji JD, Song GG (2012) The association between the PTPN22 C1858T polymorphism and systemic sclerosis: a meta-analysis. Mol Biol Rep 39: 3103–3108.

371. Villalta D, Imbastaro T, Di Giovanni S, Lauriti C, Gabini M, et al. (2012) Diagnostic accuracy and predictive value of extended autoantibody profile in systemic sclerosis. Autoimmun Rev 12: 114–120.

372. Mehra S, Walker J, Patterson K, Fritzler MJ (2013) Autoantibodies in systemic sclerosis. Autoimmun Rev 12: 340–354.

373. Hamaguchi Y (2010) Autoantibody profiles in systemic sclerosis: predictive value for clinical evaluation and prognosis. J Dermatol 37: 42–53.

374. Ceribelli A, Krzyszczak ME, Li Y, Ross SJ, Chan JY, et al. (2011) Atypical clinical presentation of a subset of patients with anti-RNA polymerase III–non-scleroderma cases associated with dominant RNA polymerase I reactivity and nucleolar staining. Arthritis Res Ther 13: R119.

375. Ayer LM, Senecal JL, Martin L, Dixon GH, Fritzler MJ (1994) Antibodies to high mobility group proteins in systemic sclerosis. J Rheumatol 21: 2071–2075.

376. Hansson GK (1999) Inflammation and immune response in atherosclerosis. Curr Atheroscler Rep 1: 150–155.

377. Saccucci P, Banci M, Cozzoli E, Neri A, Magrini A, et al. (2011) Atherosclerosis and PTPN22: a study in coronary artery disease. Cardiology 119: 54–56.

378. Pertovaara M, Raitala A, Juonala M, Kahonen M, Lehtimaki T, et al. (2007) Autoimmunity and atherosclerosis: functional polymorphism of PTPN22 is associated with phenotypes related to the risk of atherosclerosis. The Cardiovascular Risk in Young Finns Study. Clin Exp Immunol 147: 265–269.

379. Saccucci P, Banci M, Amante A, Bottini E, Gloria-Bottini F (2011) Coronary artery disease: evidence of interaction between PTPN22 and p53 genetic polymorphisms. Cardiology 120: 166–168.

380. Quarles RH (2002) Myelin sheaths: glycoproteins involved in their formation, maintenance and degeneration. Cell Mol Life Sci 59: 1851–1871.

381. Sato S, Baba H, Inuzuka T, Miyatake T (1986) Anti-myelin-associated glycoprotein antibody in sera from patients with demyelinating diseases. Acta Neurol Scand 74: 115–120.

382. Di Zenzo G, Di Lullo G, Corti D, Calabresi V, Sinistro A, et al. (2012) Pemphigus autoantibodies generated through somatic mutations target the desmoglein-3 cis-interface. J Clin Invest 122: 3781–3790.

383. Kalantari-Dehaghi M, Anhalt GJ, Camilleri MJ, Chernyavsky AI, Chun S, et al. (2013) Pemphigus vulgaris autoantibody profiling by proteomic technique. PLoS ONE 8: e57587.

384. Martins TB, Hoffman JL, Augustine NH, Phansalkar AR, Fischetti VA, et al. (2008) Comprehensive analysis of antibody responses to streptococcal and tissue antigens in patients with acute rheumatic fever. Int Immunol 20: 445–452.

385. Takamori M, Iwasa K, Komai K (1997) Antibodies to synthetic peptides of the alpha1A subunit of the voltage-gated calcium channel in Lambert-Eaton myasthenic syndrome. Neurology 48: 1261–1265.

386. Rosenfeld MR, Wong E, Dalmau J, Manley G, Posner JB, et al. (1993) Cloning and characterization of a Lambert-Eaton myasthenic syndrome antigen. Ann Neurol 33: 113–120.

387. Sobajima J, Ozaki S, Uesugi H, Osakada F, Inoue M, et al. (1999) High mobility group (HMG) non-histone chromosomal proteins HMG1 and HMG2 are significant target antigens of perinuclear anti-neutrophil cytoplasmic antibodies in autoimmune hepatitis. Gut 44: 867–873.

388. Lleo A, Shimoda S, Ishibashi H, Gershwin ME (2011) Primary biliary cirrhosis and autoimmune hepatitis: apotopes and epitopes. J Gastroenterol 46 Suppl 1: 29–38.

389. Komurasaki R, Imaoka S, Tada N, Okada K, Nishiguchi S, et al. (2010) LKM-1 sera from autoimmune hepatitis patients that recognize ERp57, carboxylesterase 1 and CYP2D6. Drug Metab Pharmacokinet 25: 84–92.

390. Barker RN, Vickers MA, Ward FJ (2007) Controlling autoimmunity–Lessons from the study of red blood cells as model antigens. Immunol Lett 108: 20–26.

391. Smyk DS, Rigopoulou EI, Koutsoumpas AL, Kriese S, Burroughs AK, et al. (2012) Autoantibodies in autoimmune pancreatitis. Int J Rheumatol 2012: 940831.

392. Saito H, Takahashi A, Abe K, Okai K, Katsushima F, et al. (2012) Autoantibodies by line immunoassay in patients with primary biliary cirrhosis. Fukushima J Med Sci 58: 107–116.

393. Berg CP, Blume K, Lauber K, Gregor M, Berg PA, et al. (2010) Autoantibodies to muscarinic acetylcholine receptors found in patients with primary biliary cirrhosis. BMC Gastroenterol 10: 120.

394. Abreu-Velez AM, Howard MS (2012) Collagen IV in Normal Skin and in Pathological Processes. N Am J Med Sci 4: 1–8.

395. Shi Y, Reddy B, Manley JL (2006) PP1/PP2A phosphatases are required for the second step of Pre-mRNA splicing and target specific snRNP proteins. Mol Cell 23: 819–829.

396. Neugebauer KM, Merrill JT, Wener MH, Lahita RG, Roth MB (2000) SR proteins are autoantigens in patients with systemic lupus erythematosus. Importance of phosphoepitopes. Arthritis Rheum 43: 1768–1778.

397. Zhang X, Ozawa Y, Lee H, Wen YD, Tan TH, et al. (2005) Histone deacetylase 3 (HDAC3) activity is regulated by interaction with protein serine/threonine phosphatase 4. Genes Dev 19: 827–839.

398. Chowdhury D, Xu X, Zhong X, Ahmed F, Zhong J, et al. (2008) A PP4-phosphatase complex dephosphorylates gamma-H2AX generated during DNA replication. Mol Cell 31: 33–46.

399. Mischo HE, Hemmerich P, Grosse F, Zhang S (2005) Actinomycin D induces histone gamma-H2AX foci and complex formation of gamma-H2AX with Ku70 and nuclear DNA helicase II. J Biol Chem 280: 9586–9594.

400. Drouet J, Delteil C, Lefrancois J, Concannon P, Salles B, et al. (2005) DNA-dependent protein kinase and XRCC4-DNA ligase IV mobilization in the cell in response to DNA double strand breaks. J Biol Chem 280: 7060–7069.

401. Galande S, Kohwi-Shigematsu T (1999) Poly(ADP-ribose) polymerase and Ku autoantigen form a complex and synergistically bind to matrix attachment sequences. J Biol Chem 274: 20521–20528.

402. Zampieri S, Degen W, Ghiradello A, Doria A, van Venrooij WJ (2001) Dephosphorylation of autoantigenic ribosomal P proteins during Fas-L induced apoptosis: a possible trigger for the development of the autoimmune response in patients with systemic lupus erythematosus. Ann Rheum Dis 60: 72–76.

403. Rutjes SA, Utz PJ, van der Heijden A, Broekhuis C, van Venrooij WJ, et al. (1999) The La (SS-B) autoantigen, a key protein in RNA biogenesis, is dephosphorylated and cleaved early during apoptosis. Cell Death Differ 6: 976–986.

404. Boire G, Craft J (1990) Human Ro ribonucleoprotein particles: characterization of native structure and stable association with the La polypeptide. J Clin Invest 85: 1182–1190.

405. Allen KL, Fonseca FV, Betapudi V, Willard B, Zhang J, et al. (2012) A novel pathway for human endothelial cell activation by antiphospholipid/anti-β2 glycoprotein I antibodies. Blood 119: 884–893.

406. Raschi E, Testoni C, Bosisio D, Borghi MO, Koike T, et al. (2003) Role of the MyD88 transduction signaling pathway in endothelial activation by antiphospholipid antibodies. Blood 101: 3495–3500.

407. Meroni PL, Raschi E, Testoni C, Parisio A, Borghi MO (2004) Innate immunity in the antiphospholipid syndrome: role of toll-like receptors in endothelial cell activation by antiphospholipid antibodies. Autoimmun Rev 3: 510–515.

408. Chen L, Dong W, Zou T, Ouyang L, He G, et al. (2008) Protein phosphatase 4 negatively regulates LPS cascade by inhibiting ubiquitination of TRAF6. FEBS Lett 582: 2843–2849.

409. Coppolino MG, Dedhar S (1999) Ligand-specific, transient interaction between integrins and calreticulin during cell adhesion to extracellular matrix proteins is dependent upon phosphorylation/dephosphorylation events. Biochem J 340 (Pt 1): 41–50.

410. Cheng ST, Nguyen TQ, Yang YS, Capra JD, Sontheimer RD (1996) Calreticulin binds hYRNA and the 52-kDa polypeptide component of the Ro/SS-A ribonucleoprotein autoantigen. J Immunol 156: 4484–4491.

411. Illien F, Piao HR, Coue M, di Marco C, Ayala-Sanmartin J (2012) Lipid organization regulates annexin A2 Ca(2+)-sensitivity for membrane bridging and its modulator effects on membrane fluidity. Biochim Biophys Acta 1818: 2892–2900.

412. Lee KH, Na DS, Kim JW (1999) Calcium-dependent interaction of annexin I with annexin II and mapping of the interaction sites. FEBS Lett 442: 143–146.

413. Arur S, Uche UE, Rezaul K, Fong M, Scranton V, et al. (2003) Annexin I is an endogenous ligand that mediates apoptotic cell engulfment. Dev Cell 4: 587–598.

414. Rand ML, Wang H, Pluthero FG, Stafford AR, Ni R, et al. (2012) Diannexin, an annexin A5 homodimer, binds phosphatidylserine with high affinity and is a potent inhibitor of platelet-mediated events during thrombus formation. J Thromb Haemost 10: 1109–1119.

415. Liu G, Wang J, Park YJ, Tsuruta Y, Lorne EF, et al. (2008) High mobility group protein-1 inhibits phagocytosis of apoptotic neutrophils through binding to phosphatidylserine. J Immunol 181: 4240–4246.

416. Ando K, Hasegawa K, Shindo K, Furusawa T, Fujino T, et al. (2010) Human lactoferrin activates NF-kappaB through the Toll-like receptor 4 pathway while it interferes with the lipopolysaccharide-stimulated TLR4 signaling. FEBS J 277: 2051–2066.

417. Boelaert K, Newby PR, Simmonds MJ, Holder RL, Carr-Smith JD, et al. (2010) Prevalence and relative risk of other autoimmune diseases in subjects with autoimmune thyroid disease. Am J Med 123: 1–9.

418. Somers EC, Thomas SL, Smeeth L, Hall AJ (2009) Are individuals with an autoimmune disease at higher risk of a second autoimmune disorder? Am J Epidemiol 169: 749–755.

419. Robazzi TC, Adan LF (2012) Autoimmune thyroid disease in patients with rheumatic diseases. Rev Bras Reumatol 52: 417–430.

420. Park DJ, Cho CS, Lee SH, Park SH, Kim HY (1995) Thyroid disorders in Korean patients with systemic lupus erythematosus. Scand J Rheumatol 24: 13–17.

421. Zoller B, Li X, Sundquist J, Sundquist K (2012) Risk of subsequent coronary heart disease in patients hospitalized for immune-mediated diseases: a nationwide follow-up study from Sweden. PLoS ONE 7: e33442.

422. Greco D, Pisciotta M, Gambina F, Maggio F (2011) Graves' disease in subjects with type 1 diabetes mellitus: a prevalence study in western Sicily (Italy). Prim Care Diabetes 5: 241–244.

423. Antonelli A, Ferri C, Fallahi P, Cazzato M, Ferrari SM, et al. (2007) Clinical and subclinical autoimmune thyroid disorders in systemic sclerosis. Eur J Endocrinol 156: 431–437.

424. Hemminki K, Li X, Sundquist K, Sundquist J (2009) Shared familial aggregation of susceptibility to autoimmune diseases. Arthritis Rheum 60: 2845–2847.

425. Agmon-Levin N, Selmi C (2013) The autoimmune side of heart and lung diseases. Clin Rev Allergy Immunol 44: 1–5.

426. Margeirsdottir HD, Larsen JR, Brunborg C, Overby NC, Dahlo gensen K (2008) High prevalence of cardiovascular risk factors in children and adolescents with type 1 diabetes: a population-based study. Diabetologia 51: 554–561.

427. Holmqvist ME, Wedren S, Jacobsson LT, Klareskog L, Nyberg F, et al. (2009) No increased occurrence of ischemic heart disease prior to the onset of rheumatoid arthritis: results from two Swedish population-based rheumatoid arthritis cohorts. Arthritis Rheum 60: 2861–2869.

428. Chu SY, Chen YJ, Tseng WC, Lin MW, Chen TJ, et al. (2011) Comorbidity profiles among patients with alopecia areata: the importance of onset age, a nationwide population-based study. J Am Acad Dermatol 65: 949–956.

429. Wang SJ, Shohat T, Vadheim C, Shellow W, Edwards J, et al. (1994) Increased risk for type I (insulin-dependent) diabetes in relatives of patients with alopecia areata (AA). Am J Med Genet 51: 234–239.

430. Narita T, Oiso N, Fukai K, Kabashima K, Kawada A, et al. (2011) Generalized vitiligo and associated autoimmune diseases in Japanese patients and their families. Allergol Int 60: 505–508.

431. Kakourou T, Kanaka-Gantenbein C, Papadopoulou A, Kaloumenou E, Chrousos GP (2005) Increased prevalence of chronic autoimmune (Hashimoto's) thyroiditis in children and adolescents with vitiligo. J Am Acad Dermatol 53: 220–223.

432. Robazzi TC, Adan LF, Pimentel K, Guimaraes I, Magalhaes Filho J, et al. (2013) Autoimmune endocrine disorders and coeliac disease in children and adolescents with juvenile idiopathic arthritis and rheumatic fever. Clin Exp Rheumatol 31: 310–317.

433. Nakamura H, Usa T, Motomura M, Ichikawa T, Nakao K, et al. (2008) Prevalence of interrelated autoantibodies in thyroid diseases and autoimmune disorders. J Endocrinol Invest 31: 861–865.

434. Mao ZF, Yang LX, Mo XA, Qin C, Lai YR, et al. (2011) Frequency of autoimmune diseases in myasthenia gravis: a systematic review. Int J Neurosci 121: 121–129.

435. Fichna M, Fichna P, Gryczynska M, Walkowiak J, Zurawek M, et al. (2010) Screening for associated autoimmune disorders in Polish patients with Addison's disease. Endocrine 37: 349–360.

436. Triolo TM, Armstrong TK, McFann K, Yu L, Rewers MJ, et al. (2011) Additional autoimmune disease found in 33% of patients at type 1 diabetes onset. Diabetes Care 34: 1211–1213.

437. Godfrey JD, Brantner TL, Brinjikji W, Christensen KN, Brogan DL, et al. (2010) Morbidity and mortality among older individuals with undiagnosed celiac disease. Gastroenterology 139: 763–769.

438. Banka S, Ryan K, Thomson W, Newman WG (2011) Pernicious anemia - genetic insights. Auoimmun Rev 10: 455–459.

439. Hemminki K, Li X, Sundquist J, Sundquist K (2010) The epidemiology of Graves' disease: evidence of a genetic and an environmental contribution. J Autoimmun 34: J307–313.

440. Ichiki T (2010) Thyroid hormone and atherosclerosis. Vascul Pharmacol 52: 151–156.

441. Onore C, Careaga M, Ashwood P (2012) The role of immune dysfunction in the pathophysiology of autism. Brain Behav Immun 26: 383–392.

442. Crespi BJ, Thiselton DL (2011) Comparative immunogenetics of autism and schizophrenia. Genes Brain Behav 10: 689–701.

443. Shi J, Levinson DF, Duan J, Sanders AR, Zheng Y, et al. (2009) Common variants on chromosome 6p22.1 are associated with schizophrenia. Nature 460: 753–757.

444. Atladottir HO, Pedersen MG, Thorsen P, Mortensen PB, Deleuran B, et al. (2009) Association of family history of autoimmune diseases and autism spectrum disorders. Pediatrics 124: 687–694.

445. Keil A, Daniels JL, Forssen U, Hultman C, Cnattingius S, et al. (2010) Parental autoimmune diseases associated with autism spectrum disorders in offspring. Epidemiology 21: 805–808.

446. Brimberg L, Sadiq A, Gregersen PK, Diamond B (2013) Brain-reactive IgG correlates with autoimmunity in mothers of a child with an autism spectrum disorder. Mol Psychiatry.

447. Chen SJ, Chao YL, Chen CY, Chang CM, Wu EC, et al. (2012) Prevalence of autoimmune diseases in in-patients with schizophrenia: nationwide population-based study. Br J Psychiatry 200: 374–380.

448. Eaton WW, Byrne M, Ewald H, Mors O, Chen CY, et al. (2006) Association of schizophrenia and autoimmune diseases: linkage of Danish national registers. Am J Psychiatry 163: 521–528.

449. Eaton WW, Pedersen MG, Nielsen PR, Mortensen PB (2010) Autoimmune diseases, bipolar disorder, and non-affective psychosis. Bipolar Disord 12: 638–646.

450. Brown AS (2012) Epidemiologic studies of exposure to prenatal infection and risk of schizophrenia and autism. Dev Neurobiol 72: 1272–1276.

451. Park MH, Kwon YJ, Jeong HY, Lee HY, Hwangbo Y, et al. (2012) Association between Intracellular Infectious Agents and Schizophrenia. Clin Psychopharmacol Neurosci 10: 117–123.

452. Boksa P (2008) Maternal infection during pregnancy and schizophrenia. J Psychiatry Neurosci 33: 183–185.

453. Rosenspire A, Yoo W, Menard S, Torres AR (2011) Autism spectrum disorders are associated with an elevated autoantibody response to tissue transglutaminase-2. Autism Res 4: 242–249.

454. Mazur-Kolecka B, Cohen IL, Gonzalez M, Jenkins EC, Kaczmarski W, et al. (2013) Autoanti-bodies against neuronal progenitors in sera from children with autism. Brain Dev.

455. Mostafa GA, Kitchener N (2009) Serum anti-nuclear antibodies as a marker of autoimmunity in Egyptian autistic children. Pediatr Neurol 40: 107–112.

456. Gilat Y, Shoenfeld Y, Kotler M, Iancu I (2011) Anti-ribosomal P antibody in schizophrenia. Isr J Psychiatry Relat Sci 48: 275–279.

457. Garbett K, Ebert PJ, Mitchell A, Lintas C, Manzi B, et al. (2008) Immune transcriptome alterations in the temporal cortex of subjects with autism. Neurobiol Dis 30: 303–311.

458. Enstrom AM, Lit L, Onore CE, Gregg JP, Hansen RL, et al. (2009) Altered gene expression and function of peripheral blood natural killer cells in children with autism. Brain Behav Immun 23: 124–133.

459. Curran LK, Newschaffer CJ, Lee LC, Crawford SO, Johnston MV, et al. (2007) Behaviors associated with fever in children with autism spectrum disorders. Pediatrics 120: e1386–1392.

460. Schwartz M, Silver H (2000) Lymphocytes, autoantibodies and psychosis–coincidence versus etiological factor: an update. Isr J Psychiatry Relat Sci 37: 32–36.

461. Catts VS, Catts SV, Jablensky A, Chandler D, Weickert CS, et al. (2012) Evidence of aberrant DNA damage response signalling but normal rates of DNA repair in dividing lymphoblasts from patients with schizophrenia. World J Biol Psychiatry 13: 114–125.

462. Printz DJ, Strauss DH, Goetz R, Sadiq S, Malaspina D, et al. (1999) Elevation of CD5+ B lymphocytes in schizophrenia. Biol Psychiatry 46: 110–118.

463. Xu J, Sun J, Chen J, Wang L, Li A, et al. (2012) RNA-Seq analysis implicates dysregulation of the immune system in schizophrenia. BMC Genomics 13 Suppl 8: S2.