Molecular Medicine

Genome-Wide Association Study of Late-Onset Myasthenia Gravis: Confirmation of *TNFRSF11A* and Identification of *ZBTB10* and Three Distinct HLA Associations

Michael F Seldin,[1] Omar K Alkhairy,[2] Annette T Lee,[3] Janine A Lamb,[4] Jon Sussman,[5] Ritva Pirskanen-Matell,[6] Fredrik Piehl,[6] Jan J G M Verschuuren,[7] Anna Kostera-Pruszczyk,[8] Piotr Szczudlik,[8] David McKee,[5] Angelina H Maniaol,[9] Hanne F Harbo,[10] Benedicte A Lie,[11] Arthur Melms,[12] Henri-Jean Garchon,[13] Nicholas Willcox,[14] Peter K Gregersen,[3] and Lennart Hammarstrom[2]

[1]Department of Biochemistry and Molecular Medicine, and Department of Medicine, University of California, Davis, California, United States of America; [2]Division of Clinical Immunology, Karolinska Institutet at Karolinska University Hospital Huddinge, Stockholm, Sweden; [3]The Robert S. Boas Center for Genomics and Human Genetics, Feinstein Institute for Medical Research, North Shore-LIJ Health System, Manhasset, New York, United States of America; [4]Centre for Integrated Genomic Medical Research, Manchester Academic Health Science Centre, University of Manchester, Manchester, United Kingdom; [5]Department of Neurology, Greater Manchester Neuroscience Centre, Manchester, United Kingdom; [6]Department of Neurology, Karolinska University Hospital Solna, Stockholm, Sweden; [7]Department of Neurology, Leiden University Medical Center, Leiden, the Netherlands; [8]Department of Neurology, Medical University of Warsaw, Warsaw, Poland; [9]Department of Neurology, Oslo University Hospital, Ullevål, Oslo, Norway; [10]Department of Neurology, Oslo University Hospital and University of Oslo, Oslo, Norway; [11]Department of Medical Genetics, University of Oslo and Oslo University Hospital, Oslo, Norway; [12]Department of Neurology, Tübingen University Medical Center, Tübingen, Germany, and Neurologische Klinik, Universitätsklinikum Erlangen, Erlangen, Germany; [13]INSERM U1173, University of Versailles, Campus Paris-Saclay, France; and [14]Nuffield Department of Clinical Neurosciences, Weatherall Institute for Molecular Medicine, University of Oxford, Oxford, United Kingdom

To investigate the genetics of late-onset myasthenia gravis (LOMG), we conducted a genome-wide association study imputation of >6 million single nucleotide polymorphisms (SNPs) in 532 LOMG cases (anti–acetylcholine receptor (AChR) antibody positive; onset age ≥50 years) and 2,128 controls matched for sex and population substructure. The data confirm reported *TNFRSF11A* associations (rs4574025, $P = 3.9 \times 10^{-7}$, odds ratio (OR) 1.42) and identify a novel candidate gene, *ZBTB10*, achieving genome-wide significance (rs6998967, $P = 8.9 \times 10^{-10}$, OR 0.53). Several other SNPs showed suggestive significance including rs2476601 ($P = 6.5 \times 10^{-6}$, OR 1.62) encoding the PTPN22 R620W variant noted in early-onset myasthenia gravis (EOMG) and other autoimmune diseases. In contrast, EOMG-associated SNPs in *TNIP1* showed no association in LOMG, nor did other loci suggested for EOMG. Many SNPs within the major histocompatibility complex (MHC) region showed strong associations in LOMG, but with smaller effect sizes than in EOMG (highest OR ~2 versus ~6 in EOMG). Moreover, the strongest associations were in opposite directions from EOMG, including an OR of 0.54 for *DQA1*05:01* in LOMG ($P = 5.9 \times 10^{-12}$) versus 2.82 in EOMG ($P = 3.86 \times 10^{-45}$). Association and conditioning studies for the MHC region showed three distinct and largely independent association peaks for LOMG corresponding to (a) MHC class II (highest attenuation when conditioning on *DQA1*), (b) *HLA-A* and (c) MHC class III SNPs. Conditioning studies of human leukocyte antigen (HLA) amino acid residues also suggest potential functional correlates. Together, these findings emphasize the value of subgrouping myasthenia gravis patients for clinical and basic investigations and imply distinct predisposing mechanisms in LOMG.
Online address: http://www.molmed.org
doi: 10.2119/molmed.2015.00232

INTRODUCTION

Myasthenia gravis (MG) is an uncommon idiopathic autoimmune disease characterized by impaired neuromuscular transmission and fatigable muscle weakness. Its overall incidence is ~3/100,000, and its prevalence is ~2 per 10,000. Both incidence and prevalence have apparently increased in recent decades, especially in subjects aged >50 years at onset (1–3). Over 80% of patients with generalized MG have antibodies (Abs) to the skeletal muscle acetylcholine receptor (AChR), which are generally accepted to be pathogenic (2,4). About 10% of them have thymomas, which mostly generate (and export) abundant T cells. The neoplastic thymic epithelial cells usually fail to express both human leukocyte antigen

Table 1. Subgrouping of MG patients in previous studies.

	EOMG	LOMG	Thymoma-MG
MG onset age (years)	10–40, 50 or 60 (puberty to menopause)	>40, >50 or >60	>10
Female:male	≥ 3:1	1:1.7	1:1
Thymic histology	Medullary lymph node-like infiltrates (>80%)	Atrophy/normal-for-age	Epithelial neoplasm
Autoantibodies (at diagnosis)			
AChR	100% by definition	100% by definition	≤100%
Titin	<5%	>70%	>95%
IFN-α and/or IL12	<5%	~60%	>70%
HLA associations (references)	DQ2–DR3–B8–A1: especially in females (10,15)	DR15 (6); DR2/B7 in males (10); DR7 if titin-Ab⁺ (12)	None confirmed

The summary is based on studies of European patients. The specific autoantibodies are rare in other diseases, apart from those against IFN-α values, which are found in ≤100% of patients with autoimmune polyendocrine syndrome type 1 (5).

(HLA) class II and the autoimmune regulator gene (AIRE). The lack of expression of these genes presumably impairs self-tolerance induction in T cells developing in thymomas (which has been implicated in these patients' distinctive clinical and autoantibody profiles) (Table 1) (2,5).

Patients with generalized AChR-Ab–positive MG without thymoma are generally sub-grouped into those with early-onset myasthenia gravis (EOMG) or late-onset myasthenia gravis (LOMG) (Table 1), with a cutoff at ages 40, 45, 50 or even 60 years in different studies (6–10; rev. in 5). As summarized in Table 1, these subgroups consistently show contrasting sex ratios, thymic histology and HLA associations (6,10–13). Moreover, several autoantibodies are almost uniquely shared by patients with LOMG or thymomatous MG but not EOMG, which remains unexplained.

We have previously reported a genome-wide association study (GWAS) in 649 EOMG patients using rigid entry criteria (onset age <40 or <45 years if the thymus showed hyperplasia). We confirmed the association with *PTPN22* (14) and identified a novel association with a *TNIP1* coding single nucleotide polymorphism (SNP) and defined HLA associations (15). The study mapped the strongest HLA association to *HLA-B*08* rather than class II SNPs in linkage disequilibrium (LD) with this class I allele (15).

Recently, another group reported a GWAS in EOMG (n = 235) and LOMG

patients (n = 737), both separately and combined (16). Although not supporting all of the previously reported associations (see Discussion), this study provided strong evidence for an association with *TNFRSF11A* (alias RANK) in the LOMG subset defined by age of onset >40 years (16). The current study was undertaken to further define genetic associations in stringently defined LOMG patients (onset ≥50 and ≥60 years), again focusing on Northern Europeans, using population substructure matching of publicly available control genotypes, and imputation methods extending to specific HLA variants and amino acids (AAs). The results suggest a biological basis for defining a cutoff for age of onset between EOMG and LOMG and again imply distinct genetic and environmental factors in predisposition in these groups.

MATERIALS AND METHODS

Study Subjects

All the LOMG patients included in these studies were European and met the following criteria: (a) clinical diagnostic criteria for MG; (b) positive for anti-AChR Ab; (c) no radiological evidence of thymoma; and (d) onset age ≥ 50 years. Cases were collected from Stockholm, Sweden; Oslo, Norway; Manchester and Oxford, England; Paris, France; Leiden, the Netherlands; Tübingen, Germany; and Warsaw, Poland. Samples for all participants of this study were collected

under institutionally approved inform consent (at the respective institutes) and complied with the Helsinki Declaration as revised in 1983.

Of a combined total of 557 LOMG samples collected, 25 were excluded because they failed checks for quality control (QC) (<90% complete genotyping data), cryptic relationship (ρ̂ >0.15, using PLINK [17]), ancestry and/or SNP heterozygosity (>3 standard deviation above or below mean). The remaining 532 patients were matched 1:4 with controls available from these same populations plus others from European–American populations (Supplementary Table S1).

Quality Control

In addition to the QC for individual samples, the SNP data were carefully reviewed, and exclusion criteria were applied to minimize potential batch effects as described below. The application of these criteria was particularly important given the derivation of genotypes from multiple platforms typed in different laboratories. We included only SNPs with <5% missing data, Hardy-Weinberg equilibrium P values $>1 \times 10^{-4}$ in controls and $>1 \times 10^{-5}$ in combined cases and controls, and minor allele frequencies >0.02. These procedures were applied in a stepwise approach separately for each data set and after combining the data sets. Thus, for each of the separately derived control genotyping sets, SNPs were excluded if they failed the above criteria within the

individual control set (platform and laboratory) or in combination with any of the other control groups, or in the complete data set. The Hardy-Weinberg criteria were applied after exclusion of non-European individuals. Finally, after selection of the control data set, SNPs were excluded if the allele frequencies differed by >10% between different control groups.

We used two different genotyping platforms: the 370K or compatible Illumina platforms for set 1 (n = 214 cases and 856 controls) and the 2.5 million Illumina platform for set 2 (n = 318 cases and 1,272 controls (Supplementary Table S1). For sets 1 and 2, a total of 280,929 and 1,406,133 SNPs, respectively, passed all data filters. Because only ~150,000 SNPs were shared between platforms, imputation (see below) was performed separately for each platform, and the data were combined and resubjected to the same QC metrics described above.

Ancestry and Case-Control Matching

European ancestry in both cases and controls was determined using a panel of 2,440 SNPs (intermarker r^2 <0.05, minor allele frequency >0.05) and analyzed using STRUCTURE v2.3.4 (18,19) and subjects of known European, Amerindian, East Asian, and West African origin as previously described (20). We used STRUCTURE to exclude non-European and admixed study participants, because this Bayesian clustering method allows exclusion/inclusion criteria to be set using reference populations. Unlike principal component analysis (PCA), the clustering algorithm can be run under conditions that are only marginally affected by the inclusion of the unknown samples. Briefly, analyses were performed using >100,000 resamplings and >50,000 burn-in cycles under the admixture model and using the prior population information for the reference populations. Runs were performed under the λ = 1 option, where λ estimates the prior probability of the allele frequency and is based on the Dirichlet distribution of allele frequencies. Subjects with >10% non-European ancestry were excluded from further analyses.

Each case was then matched with four controls based on PCAs performed using 25,914 selected SNPs shared across platforms (both sets 1 and 2) using EIGEN-SOFT v5 (21). SNPs were chosen from overlapping SNPs with minimal inter-marker LD (r^2 < 0.05) in Windows of 2 Mb after exclusion of the major histocompatibility complex (MHC) and genomic regions with large inversions.

Matching (4 controls:1 case) was performed separately for males and females and also for set 1 (370K Illumina platform, 87 female and 127 male cases) and set 2 (HumanOmni 2.5M-8v1-1 chip Illumina platform, 113 female and 205 male cases) using procedures previously described (EOMG) (15). A total of 3,190 female and 2,159 male European control genotypes for set 1 and 5,392 female and 3,940 male European control genotypes for set 2 were available for matching after QC (Supplementary Table S1). Matching used the first five principal components (PCs) (which accounted for >99% of EIGENVALUE variance in Tracy-Widom significant PCs). As shown in Supplementary Table S1, there were a total of 856 controls for set 1 (4× the number of set 1 cases) and 1,272 controls for set 2 (4× the number of set 2 cases).

After matching, the first 7 PCs showed significant Tracy-Widom statistics, but just the first 3 PCs explained over 99% of their total EIGENVALUE variance for these PCs. After matching, λ_{gc} was 1.039 without correction and decreased to 1.029 after controlling for these three PCs. PCA and scree plots are shown in Supplementary Figures S1 and S2.

Genome-Wide Imputation of SNPs

Imputation was performed by using a 1000 Genomes Phase 1 V3 reference set and the University of Michigan Imputation Server (https://imputationserver.sph.umich.edu/start.html#!run/minimac). This method uses a computationally efficient implementation of the Markov chain-based haplotyper (MaCH) algorithm with prephasing using SHAPEIT (22) as previously described (23). The program outputs the most likely geno-

type and provides r^2 metrics for quality measurement. We performed imputation separately for set 1 and set 2 genotypes. A total of 6,175,472 genotypes showed r^2 >0.8 common to both set 1 and set 2, and all QC criteria as described above that were applied to these joined set 1 and set 2 genotypes. (Note: testing of strongly suggestive associated SNPs (P values <10^{-7}) with genotypes that met lower r^2 thresholds in one or both sets showed P values >5×10^{-5} after regional imputation and testing. We therefore restricted further analyses to genotypes with r^2 > 0.8 in both sets 1 and 2.) The combined set 1 and 2 case (532) and control (2,128) imputed genotypes (6,175,472) were used as the data set for initial analyses.

Regional Imputation of SNPs and Insertion/Deletion Polymorphisms

For any SNP that showed suggestive association ($P < 10^{-5}$) in the screening imputation (described above), we performed regional imputation using Impute2 (IMPUTE v2.3.1) (https://mathgen.stats.ox.ac.uk/impute/impute_v2.html) (24) under recommended settings and the afore-noted 1,000-genome phase 1v3 reference set. This software provides genotype probabilities (rather than base assignments), enabling a more accurate assessment of statistical support. Overall errors compared favorably with other approaches (24). The intervals examined (excluding the MHC) included at least a three order of magnitude fall-off of significance from the association peak and a minimum of 250 kb. For the MHC, we examined a 4-Mb interval (chromosome 6, ~29–33 Mb, HG19/HG38). We applied the algorithm with 250-kb buffers (in addition to the interval under examination) and without prephasing (standard IMPUTE2 MCMC algorithm) as recommended by the developers for more accurate, although less computationally efficient, imputation.

Imputation of HLA Antigens

To impute HLA antigens, we used a separate reference data set collected by the Type 1 Diabetes Genetics Consortium
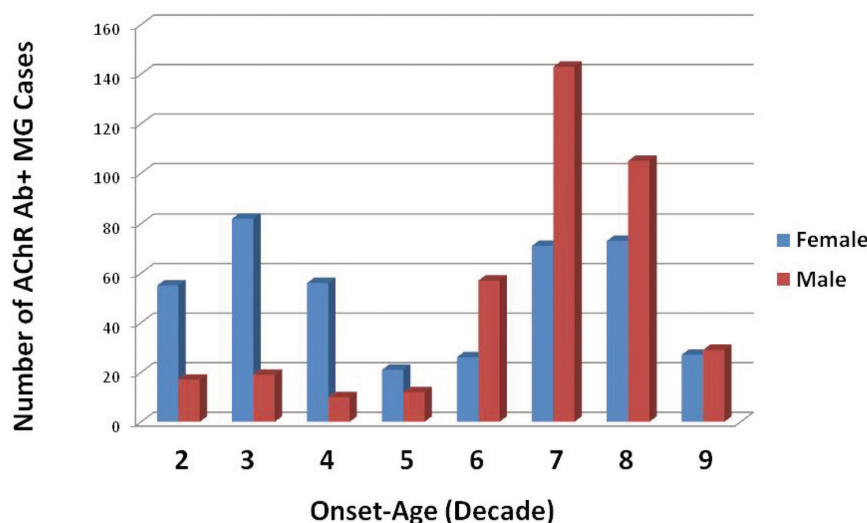
Figure 1. Relationship of sex and age of onset for AChR Ab + MG. The histogram shows the number of female (blue) and male (red) MG cases for 10-year age of onset intervals starting at decade 2 (age of onset 10–19 years). The cases include EOMG (onset age by definition <40 years if no histology or <45 years if histology shows thymic hyperplasia, n = 649) (15), unclassified (onset age 40 to < 50 years, n = 23) and LOMG (onset age ≥50 years, n = 532) (current study).

and the SNP2HLA software, as described by its developers (25). The Beagle software package (26) was used for imputation in this data set (compatible with this software). Here, we also used genotype probabilities for association testing; our final data consisted of only antigens, AAs and SNPs with information scores >0.8 in SNPTEST analyses.

Association Testing Statistics

For all association results, analyses included covariates for sex, genotype set and the first three PCs (as discussed above). All association tests were done by using the combined imputed SNP data set as described above. For autosomal data, we used SNPTEST V2.3 (https://mathgen.stats.ox.ac.uk/genetics_software/snptest/snptest.html) for the primary association analyses for the imputed genotypes, as well as for conditional analyses. For the X chromosome analyses we also used SNPTEST V2.5. Analyses included an X inactivation model and a model that allows for heterogeneity (stratify_on option). With the exception of the screening imputed genotypes, genotype probabilities (and a minimum information score of 0.8) were used in all analyses.

Epigenetic Markers

To assess possible regulatory function, we used RegulomeDB (27). This database (search engine) annotates SNPs for predicted and known regulatory region for DNase hypersensitivity, transcription factor binding sites and promoter regions. It includes the current Encyclopedia of DNA Elements (ENCODE) releases and chromatin states from the Roadmap Epigenome Consortium and uses multiple resources that include DNase footprinting, position weight matrices and DNA methylation information (for further description of RegulomeDB see http://regulomedb.org). For this assessment, we examined all SNPs that were in at least weak LD ($r^2 > 0.25$) to the strongest associated SNP and had annotations that showed possible regulatory function. Conditional analyses were then performed, and the SNP with the strongest likelihood of function was presented in Results. We also tested combinations of putative functional SNPs. However, conditional analyses on combi-

nations of SNPs did not enhance the signal attenuation compared with the SNP presented in Results for *ZBTB10*, *TNFRSF11A* or the class III MHC region.

Other Software

For graphical presentation of association results, we used qqman for quantile–quantile (Q–Q) and Manhattan plots (http://cran.r-project.org/web/packages/qqman/qqman.pdf) and LocusZoom for chromosome region association plots (28).

All supplementary materials are available online at www.molmed.org.

RESULTS

Screening for Association with LOMG

To screen for putative candidate loci, we selected AChR-Ab–positive LOMG cases with no radiologic evidence of thymomas (see Materials and Methods and Supplementary Table S2). We used two age of onset cutoffs (≥50 and ≥60 years) based both on previous studies (see Introduction) and the bimodal distribution of MG with age of onset (Figure 1).

Since two disparate SNP GWAS chipsets were used for genotyping, we melded the data using a genome-wide imputation strategy (see Materials and Methods for additional details). We tested for associations with variants imputed with high confidence ($r^2 \geq 0.8$, required for each imputed variant from both GWAS panels); the analyses included over 6 million variants with minor allele frequencies ≥0.02. The association tests included covariates to control for residual population substructure, and sex.

As expected in any autoimmune disease, we found major contributions from the MHC region on chromosome 6 in patients with cutoff ages ≥50 and ≥60 years, using Q–Q analyses and Manhattan plots (Figure 2). HLA associations were further refined with more precise imputation methods, as detailed later.

The results also suggest other loci on chromosomes 1, 5, 8, 12, 18 and 21 (Figure 2, Supplementary Table S3). Even
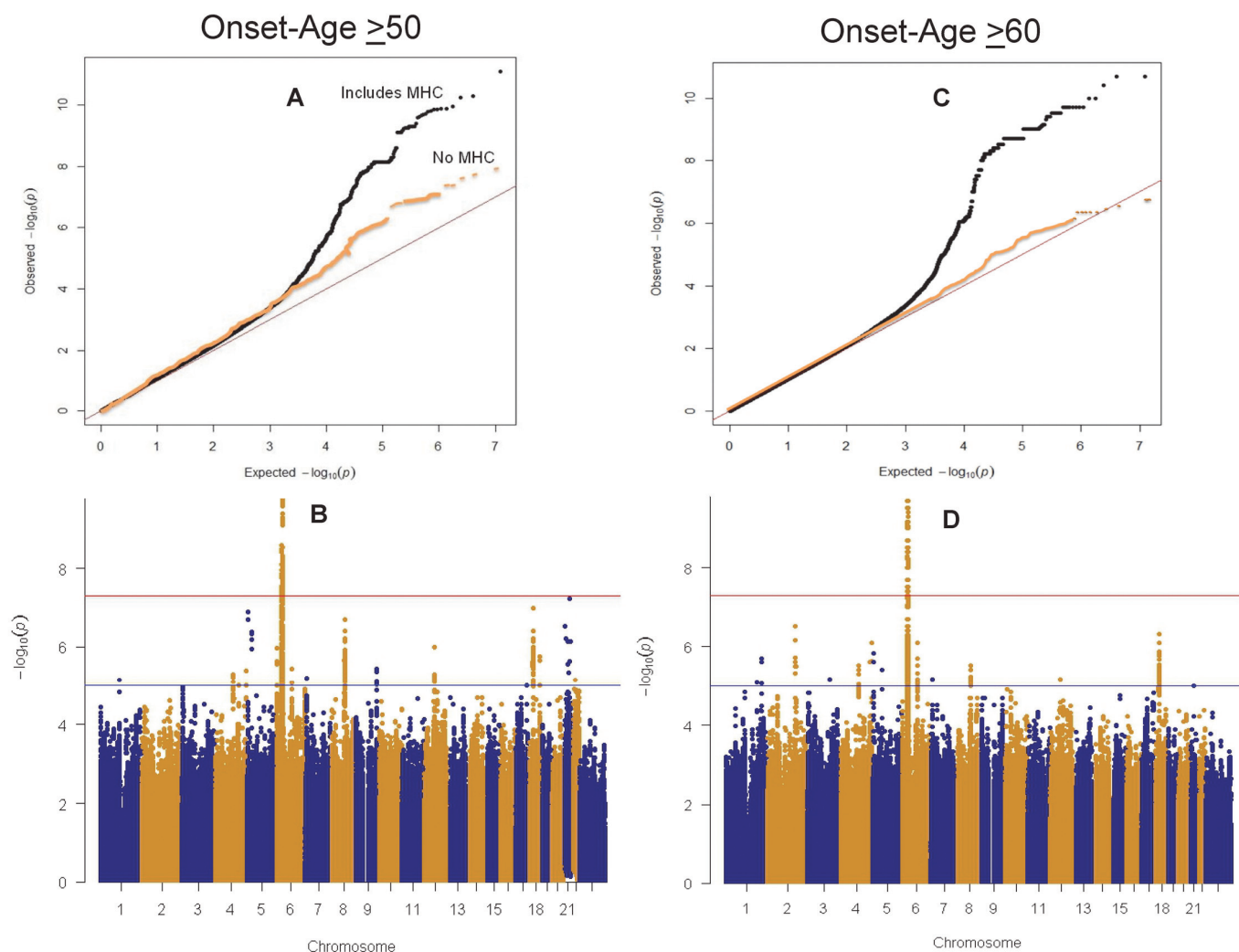
Figure 2. Quantile–quantile (Q–Q) and Manhattan plots for LOMG GWAS results. (A) Results are from screening study (for Minimac imputation, see Materials and Methods). (B) Age of onset ≥50 years (n = 532). (C, D) Age of onset ≥60 years (n = 449). (A, C) The observed deviation above the diagonal (expected Q–Q distribution) indicates evidence for genetic association. (B, D) Blue and orange show alternate chromosomes.

after removal of the extended MHC region, the Q–Q plots still show deviations from the expected P value distribution (Figures 2A, C), especially in the stronger-powered subgroup with onset at ≥50 years of age. We found suggestive evidence for association at the *PTPN22* locus, as previously noted in EOMG (14,15) (Supplementary Table S3). In contrast, there was only marginal evidence for any association signal for *TNIP1* (all P values >0.001 within 100 kb flanking this gene or in r^2 > 0.25), the strongest non-HLA signal we observed in EOMG (15).

More Precise Imputation of Potentially Associated Loci

To further define the loci thus imputed ($P < 10^{-5}$), we next applied a more accurate but computationally less efficient imputation method without prephasing or applying pruning procedures (see Materials and Methods). Several of the above candidates maintained either suggestive or significant associations (Table 2, Supplementary Tables S4 and S5).

Outside the MHC region, only the chromosome 8 SNP, rs6998967, 5′ of *ZBTB10*, reached conservative genome-

wide criteria (5×10^{-8}) for association using either cutoff age. We also found suggestive evidence for associations in eight disparate regions (Table 2); they include the same R620W variant of PTPN22 noted previously, and one for the tumor necrosis factor receptor superfamily, member 11a, NFKB activator (*TNFRSF11A*, *alias RANK* and its particular SNP, rs4574025), reported recently in LOMG (16). For each of the loci achieving at least suggestive significance ($P < 10^{-5}$), odds ratios (ORs) were similar for the more stringent cutoff age (≥60 years) (Table 2 and Supplementary Table S4).

Table 2. Strongest association signals for LOMG.

Variant	Chromosome	Position (HG19)	Minor Allele	Major Allele	Case MAF	Control MAF	Additive P value	OR (95% CI)	OR age ≥60 years	Gene/location
rs111945767	6	32517045	C	G	0.46	0.62	3.13×10^{-17}	0.52 (0.44–0.61)	0.49	MHC region
rs6998967	8	81364205	A	G	0.10	0.17	8.86×10^{-10}	0.53 (0.42–0.65)	0.53	ZBTB10/34 kb 5′
rs4574025	18	60009814	C	T	0.55	0.47	3.91×10^{-7}	1.42 (1.24–1.63)	1.31	TNFRSF11/intron
rs9963862	18	20587911	G	A	0.42	0.34	4.19×10^{-7}	1.44 (1.25–1.66)	1.47	RBBP8/intron
rs12653117	5	9386585	T	C	0.13	0.08	5.34×10^{-7}	1.76 (1.42–2.17)	1.75	SEMA5A/intron
rs6914704	6	1759880	G	C	0.42	0.49	2.32×10^{-6}	0.72 (0.63–0.83)	0.74	GMDS/intron
rs4128527	8	72682693	T	C	0.14	0.19	3.68×10^{-6}	0.62 (0.50–0.77)	0.62	MSC/71 kb 5′
5:31213962:D	5	31213962	TC	T	0.09	0.05	3.84×10^{-6}	1.84 (1.44–2.37)	1.68	CDH6/intron
rs4518467	6	92196398	A	G	0.11	0.07	4.31×10^{-6}	1.77 (1.40–2.23)	1.94	MIR4643/35 kb 5′
rs2476601	1	1143777568	A	G	0.16	0.11	6.50×10^{-6}	1.62 (1.31–1.99)	1.66	PTPN22/coding

Data are shown for onset age ≥50 years, except for the OR for onset age ≥60 years as comparison. The strongest P values (if $P < 10^{-5}$) for putative associations of each region are shown after re-analyses using IMPUTE2 including relevant covariates (see Materials and Methods). Additional SNPs for each region are shown in Supplementary Table S4. The OR and 95% CI are shown for the minor allele. The minor allele and minor allele frequency (MAF) was determined based on the entire data set (controls matched 4:1 with cases).

Conditional Analyses of the *ZBTB10* and *TNFRSF11A* Locus SNPs

The details of the *ZBTB10* association are shown in a LocusZoom plot (28) graph (Figure 2). The rs6998967 SNP is located ~34 kb 5′ of the zinc finger– and BTB domain–containing 10 (*ZBTB10*) gene, and there are multiple SNPs in strong LD showing suggestive associations (Figure 3A). Conditioning for the rs6998967 SNP implicated a single haplotype marked by this variant (Figure 3B). Conditioning on rs1065238, a potential strong regulatory SNP (RegulomeDB score = 1f, eQTL + transcription factor [TF] binding site/DNase peak; see Materials and Methods) (27), also attenuated most of the signal from this region (Figure 3C).

Similarly, the *TNFRSF11A* association was primarily due to a single haplotype (Figure 4). It includes rs8086340, also with potential regulatory function (RegulomeDB score = 4, possible regulatory site), in LD with the strongest associated SNP, rs4574025.

HLA Associations

To further examine associations within the MHC region, we also imputed both classical *HLA* alleles and AA variants in them. Three peaks of association were observed in the MHC region (Figure 5A). These corresponded to the *HLA-A*, MHC class III and MHC class II genes.

For classical HLA class II antigens, multiple genes showed evidence for associations (Table 3 and Supplementary Table S6). However, most of these show protective ORs <1, in stark contrast with those observed in EOMG. Furthermore, the LOMG risk (ORs >1) associations were weaker, and none of them achieved genome-wide significance in the series as a whole. The strongest association we found in LOMG was with *DQA1*05:01* ($P = 5.9 \times 10^{-12}$, OR 0.54). Similarly, *DRB1*03:01*, which is in strong LD, showed an OR of 0.5 that is consistent with a previous study showing negative association in anti-titin-Ab–positive LOMG patients (12). These alleles belong to the B8.1 haplotype that predisposes strongly to EOMG (OR 2.82); *HLA-B*08* (OR 6.41) also belongs, which gave a contrasting OR of 0.69 in LOMG. Evidently, these alleles favor early onset of MG in susceptible subjects, rather than protecting against it at all ages.

Although none of the *HLA-A* antigens reached genome-wide significance, the [70]Q present in HLA-A*68:01, *03:01 and *11:01 (and other rare alleles) showed a significant risk association (OR 1.5, $P = 2.6 \times 10^{-8}$). For MHC class III, the minor alleles of multiple SNPs showed a mixed pattern of significant positive and negative associations (Supplementary Table S5).

Conditional Analyses of the MHC Region

To localize the key variants more precisely, we performed extensive conditional analyses (Figure 5 and Table 4). Conditioning on all nominally significant classical *HLA-DQA1* antigen genes ($P < 0.05$) almost completely eliminated the strongest peak of association, which was located in the MHC class II region (Figure 5B and Table 4). Conditioning instead on *DRB1* antigens had similar but less complete effects on the MHC class II region (residual $P = 5.58 \times 10^{-5}$ versus $P = 1.11 \times 10^{-3}$; Table 4).

Nearly all the signal in the HLA-A region resisted conditioning on MHC class II genes or the strongest associated SNP there, rs111945767 (which attenuated the MHC class II signal nearly as well as *HLA-DQA1*) (Table 4). Conversely, conditioning on all classic *HLA-A* antigen genes ($P < 0.05$) left strong residual association signals in the MHC class II region (P value still $<10^{-15}$; Table 4).

The MHC class III region signals were largely unaffected, even by conditioning on the combination of *HLA-A* and *DQA1* (Figure 5C), although this step almost completely attenuated the *HLA-A* and the MHC class II region peaks. Conditioning on *HLA-A*, *DQA1* and *HLA-C* also left residual signals in class III (Table 4). We also note that the combination of *HLA-A*,

DRB1 and *HLA-C*, and the combination of *HLA-A*, *DQB1* and *HLA-C*, also did not attenuate the class III signal (data not shown). Finally, conditioning on *HLA-A*, *DQA1* and an MHC class III SNP effectively eliminated almost all association signals in the entire region (residual $P > 0.002$) (Figure 5D). We cannot locate the precise locus in MHC class III because of the strong LD within this region (Figure 5C); three of its SNPs each showed nearly equivalent effects in combination with the *HLA-A* and *DQA1* conditioning: rs2256974, $P = 0.0022$; rs6929796, $P = 0.0018$; and rs2071596, $P = 0.0015$.

The results were similar whether we conditioned on HLA alleles or selected AAs within them (Table 4). Conditioning just on the [11]S (present in HLA-DR1*03:01 and most DRB1*11, *13 and *14 subtypes) plus [70]Q in HLA-A (in HLA-A*68:01, 03:01, 11:01), plus the rs2256974 class III SNP, attenuated most of the association signal (residual $P = 2.3 \times 10^{-4}$). Conditioning on [130]A and [75]I in DQA1 (present in DQA1*01:03 and DQA1*05:01, respectively) plus the same HLA-A AA and class III SNPs showed marginally better attenuation (residual $P = 5.4 \times 10^{-4}$), which became almost complete after adding [152]A (includes HLA-A*11:01 and 01:01) and/or [152]E (includes HLA-A*03:01, 25:01 and 26:01) in HLA-A (residual $P = 1.95 \times 10^{-3}$).

Comparing Associations in Males and Females

Because we found significantly stronger associations and effect sizes with *HLA-B8* in EOMG females than in males (10,15), we checked for analogous differences in LOMG at all the strongest associated loci within and outside HLA (Supplementary Table S7). Although none of these differences were significant (based on *Z* score calculations), effect sizes (whether risk or protective alleles) were larger in the males for each MHC sub-region. Among the other loci, only the *PTPN22* SNP showed any significant gender difference (unlike in EOMG), the association being confined to males (OR 1.83, 95% confidence interval [CI] 1.4–2.4), with no significant association in females (OR 0.75, 95% CI 0.54–1.04).
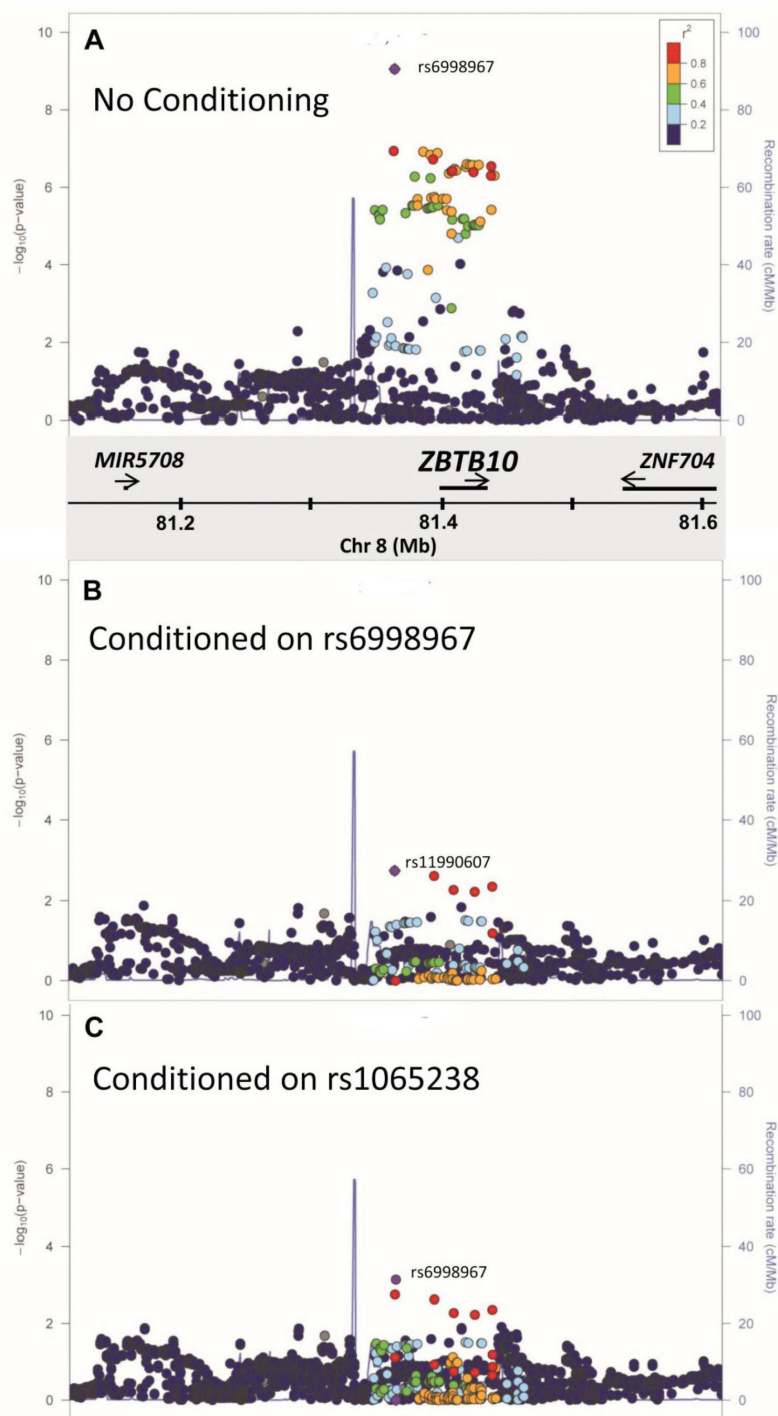

Figure 3. *ZBTB10* association signals in LOMG (age of onset ≥50 years). The ordinate shows the strengths of the association signals, with the position on chromosome 8 shown in megabases (HG19 map) along the abscissa. The *P* value for each SNP is shown before (A) and after (B) conditioning with rs6998967, the strongest associated SNP, and rs1065238 (C), a SNP in strong LD, with probable regulatory effects (RegulomeDB score = 1f). The color-coded symbols correspond to the strength of linkage disequilibrium (based on HG19/1000 genome, November 2014, European population) with the most significantly associated SNP (purple filled circle) in each panel.
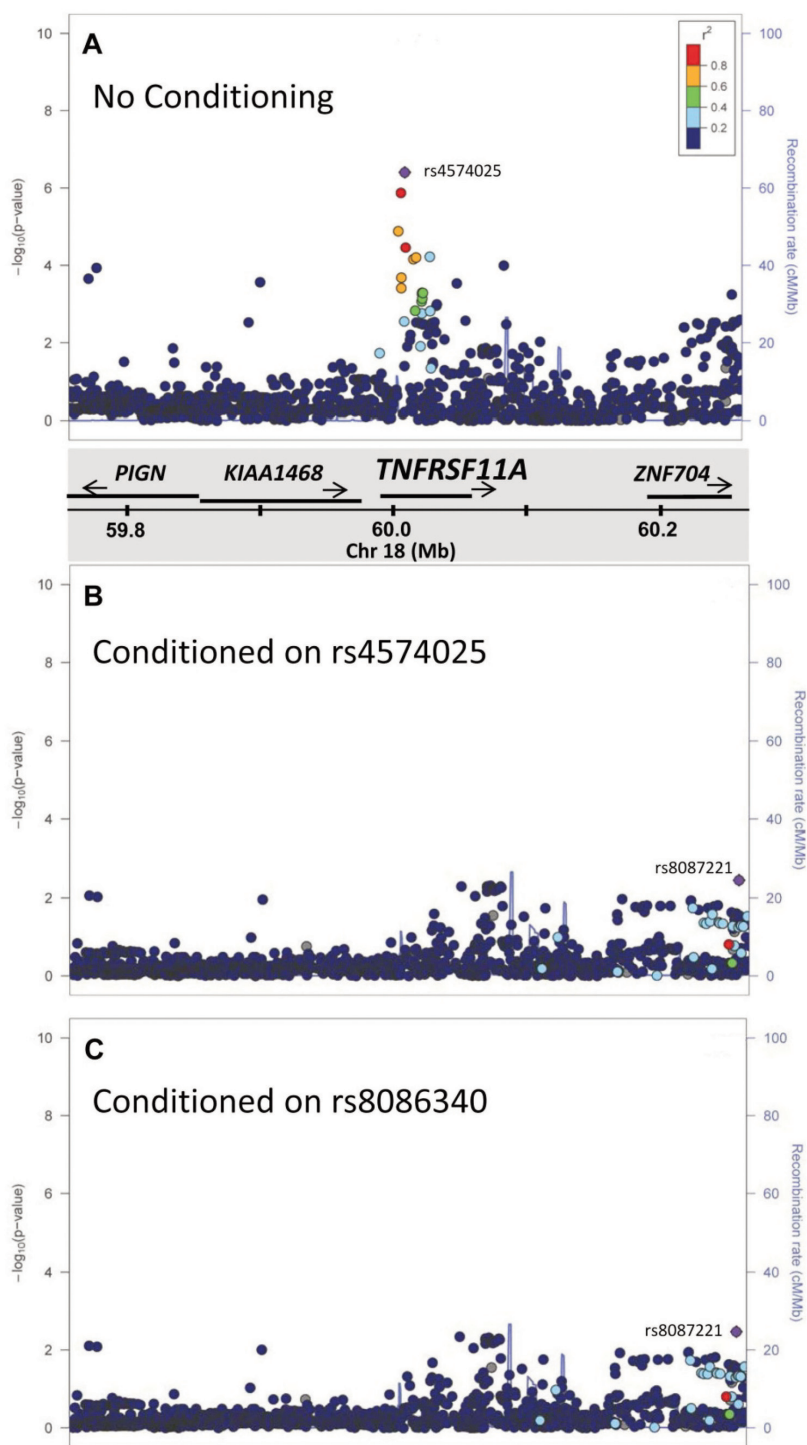
Figure 4. *TNFRSF11A* association signals in LOMG. The ordinate shows the strength of the association signals, with the position on chromosome 18 shown in megabases (HG19 map) along the abscissa. The *P* value for each SNP is shown before (A) and after (B) conditioning with rs4574025, the strongest associated SNP, and rs8086340 (C), a SNP with possible regulatory effects (RegulomeDB score = 4). The color-coded symbols correspond to the strength of linkage disequilibrium, with the most significantly associated SNP (purple-filled circle) in each panel.

Comparing Associations in Generalized and Ocular LOMG

Lastly, we evaluated whether stratifying AChR-Ab–positive LOMG for clinical features of generalized versus ocular disease would affect our results. The ORs were similar among the entire LOMG study grou*p* (n = 532) and those classified as generalized (n = 408) for both MHC (for example, *HLA-DQA1*05:01*, OR 0.52 versus OR 0.51) and non-MHC loci (Supplementary Table S8). The ocular group showed more fluctuation in the ORs, as expected from the small sample size (n = 78), but the pattern was broadly consistent with that observed for the generalized subgroup (Supplementary Table S8).

DISCUSSION

This genetic study included only patients strictly defined by MG with age of onset ≥50 years. Importantly, our GWAS findings show sharp contrasts between EOMG and LOMG (whether with onset ≥60 years or even ≥50 years) and implicate regions both within and outside the MHC complex; only *PTPN22* showed similar associations in both groups. By providing a more biological basis for the cutoff between EOMG and LOMG, our findings should provoke new hypotheses about the distinct genetic and environmental factors influencing susceptibility.

For non-MHC variants, only a single locus with a peak at rs6998967, 34 kb proximal to *ZBTB10*, met conventional GWAS criteria for a significant association. It is in LD with a variant, rs1065238, that could account for most of the signal and is located in a DNase-sensitive quantitative trait locus; in fact, it is located within a binding site for POU2F1, Zfp187 (ZSCAN26), POU3F1 and POU4F3 (29). This gene previously met genome-wide association criteria in studies of sex hormone–binding globulin levels (linked to sex steroid regulation) (30) and asthma with hay fever (31). It has also shown suggestive levels of association in hypothyroidism (32), atopic dermatitis (33) and self-reported allergy (34). ZBTB10 is thought to regulate specificity proteins Sp1, Sp3 and Sp4 (35,36); in cell culture
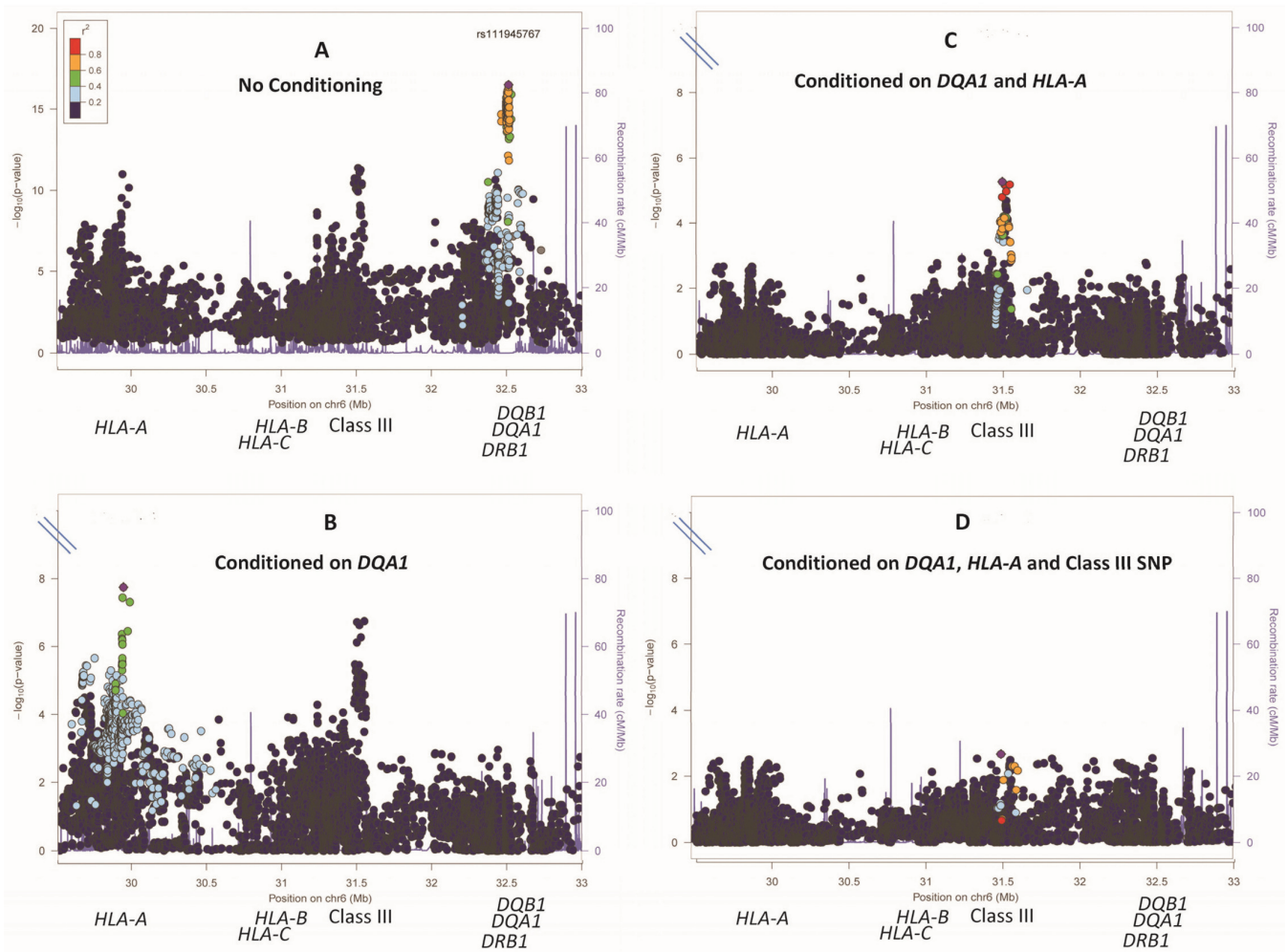
Figure 5. Analysis of the HLA region association signals in LOMG. In each panel, the symbols show the strength of the association signal (ordinate) for this region of chromosome 6 shown in Mb (HG19) along the abscissa; the SNPs, HLA antigens and HLA AAs are given color-coded symbols that correspond to the strength of LD with the most significantly associated SNP (purple filled circle). The *P* values for each variant are plotted: (A) no conditioning; (B) conditioning on *HLA DQA1*; (C) *HLA DQA1* and *HLA-A*; and (D) *HLA DQA1, HLA-A* and class III SN*P* (rs2071596). Additional conditioning results are summarized in Table 4.

studies, it is suppressed by ROS-microRNA27a, thereby enhancing estrogen receptor α expression and mediating estrogen effects (36). Potential immunoregulation could either be through Sp1 (critical for interleukin [IL]-10 and CD40 responses) (37,38) or indirectly via estrogen effects on autoimmunity (39,40).

Non-MHC loci meeting suggestive criteria for association notably include *PTPN22* and *TNFRSF11A*. *PTPN22* was reported in several studies in MG (14,15,41), reaching GWAS criteria in EOMG (15). This association is with the same coding SNP, rs2476601, as in EOMG, with similar

effect sizes (ORs 1.62–1.71), and in multiple other autoimmune diseases including type 1 diabetes mellitus, rheumatoid arthritis (RA) and systemic lupus erythematosus (42–44). The predisposing [620]W variant is thought to restrain T-cell responses less than its [620]R counterpart, as well as increase B-cell activation on antigen binding (45). The higher prevalence of this variant in our males together with the lack of association with *TNIP1* in LOMG and the starkly contrasting HLA associations (discussed below) argue strongly that the *PTPN22* association is not due to EOMG contaminants in our LOMG series.

Our findings with *TNFRSF11A* confirm a recent report where one of the same SNPs (rs4574025) showed one of the strongest associations, again only in MG patients with onset after age 40 years (16). This result might reflect the potential regulatory function of rs8086340 SNP, which is in such strong LD with the rs4574025 that conditioning on either SNP gave almost identical residual *P* values (<0.004). Variants of *TNFRSF11A* (RANK) underlie a familial form of Paget disease of the bone (46) and a recessive form of osteopetrosis with hypogammaglobulinemia (47). *TNFRSF11A* variants are also im-

Table 3. Classic HLA associations in LOMG and comparison with EOMG.

	LOMG				EOMG			
Allele	Case frequency	Control frequency	P	OR (95% CI)	Case frequency	Control frequency	P	OR (95% CI)
*HLA-A*01:01*	0.13	0.15	1.09×10^{-1}	0.85 (0.70–1.04)	0.35	0.16	1.99×10^{-44}	2.99 (2.57–3.49)
*HLA-A*03:01*	0.20	0.15	9.20×10^{-5}	1.42 (1.19–1.69)	0.08	0.14	5.92×10^{-9}	0.56 (0.46–0.68)
*HLA-C*07:01*	0.11	0.15	1.80×10^{-4}	0.67 (0.54–0.83)	0.52	0.34	5.67×10^{-31}	2.18 (1.91–2.48)
*HLA-C*05:01*	0.05	0.08	1.88×10^{-3}	0.64 (0.47–0.86)	0.03	0.10	7.84×10^{-13}	0.44 (0.35–0.55)
*HLA-B*07:02*	0.18	0.15	1.20×10^{-2}	1.26 (1.05–1.50)	0.07	0.14	3.54×10^{-11}	0.53 (0.44–0.64)
*HLA-B*08:01*	0.08	0.11	1.75×10^{-3}	0.69 (0.54–0.88)	0.41	0.13	2.87×10^{-113}	6.41 (5.46–7.53)
*HLA-B*44:02*	0.07	0.09	2.01×10^{-2}	0.73 (0.56–0.96)	0.04	0.09	3.27×10^{-11}	0.44 (0.35–0.56)
*HLA-DRB1*03:01*	0.06	0.12	6.70×10^{-8}	0.50 (0.39–0.66)	0.35	0.13	9.69×10^{-74}	4.70 (3.98–5.56)
*HLA-DRB1*07:01*	0.15	0.12	2.43×10^{-4}	1.45 (1.19–1.76)	0.04	0.12	3.74×10^{-16}	0.42 (0.34–0.52)
*HLA-DRB1*13:01*	0.03	0.07	1.90×10^{-7}	0.39 (0.26–0.58)	0.04	0.06	3.18×10^{-04}	0.60 (0.46–0.79)
*HLA-DRB1*15:01*	0.18	0.14	3.10×10^{-3}	1.31 (1.10–1.57)	0.10	0.14	8.92×10^{-06}	0.66 (0.55–0.79)
*HLA-DQA1*05:01*	0.15	0.24	5.92×10^{-12}	0.54 (0.45–0.65)	0.43	0.24	3.63×10^{-45}	2.82 (2.44–3.26)
*HLA-DQA1*02:01*	0.15	0.12	2.96×10^{-4}	1.44 (1.19–1.75)	0.04	0.12	3.19×10^{-16}	0.42 (0.35–0.52)
*HLA-DQA1*01:03*	0.04	0.07	4.31×10^{-6}	0.48 (0.34–0.68)	0.04	0.07	7.34×10^{-04}	0.64 (0.49–0.83)
*HLA-DQB1*02*	0.19	0.20	5.37×10^{-1}	0.95 (0.80–1.12)	0.37	0.22	1.63×10^{-33}	2.50 (2.16–2.90)
*HLA-DQB1*02:01*	0.06	0.12	1.23×10^{-7}	0.52 (0.40–0.67)	NA	NA	NA	NA
*HLA-DQB1*02:02*	0.12	0.08	8.74×10^{-6}	1.66 (1.33–2.06)	NA	NA	NA	NA

Results of LOMG with onset age ≥50 years are compared with results from our previous study of EOMG (15). Case and control antigen frequencies are shown. The ORs and 95% CIs are shown for each HLA antigen except where not available (NA). Data were selected to show the 10 most significant associations in the MG subgroups and also included *DRB1*15:01* and *HLA-B*07:01* for comparison with previous studies (see Supplementary Table S6 for complete HLA analyses).

plicated in GWAS both in other forms of Paget disease and of bone mineral density (48,49). With respect to immunologic functions, *TNFRSF11A* appears to be important in lymph node development and thymic selection (50), and its expression on dendritic cells may also be critical to specific interactions with T cells (51,52).

In the MHC, we found strong associations with multiple different loci. For HLA alleles, the strongest association was with *HLA-DQA1*05:01* ($P = 5.3 \times 10^{-12}$, OR 0.54) and minimal with *HLA-B*08:01* ($P = 1.8 \times 10^{-3}$, OR 0.69), unlike in EOMG, where *HLA-B*08* predisposes very strongly (OR ~6) (15) and evidently biases toward earlier onset of MG. In addition, the sharp contrast between these risk associations in EOMG (15) and the protective associations in LOMG (and their much greater strength for this -DQ allele than for -B8 or -A1 [Table 3]) implies an extra protective role for DQA1*0501 or linked variants (for example, *DRB1*03:01*) in LOMG. In both subgroups, the findings also implicate each variant's function in presenting antigens to specific T cells (see below). We postulate that the striking dif-

ferences in MHC region associations between LOMG and EOMG may in turn be due to differences in the environmental stimulus that presumably induces the break in tolerance to AChR.

In some of our LOMG patients, we previously noted risk associations with *DRB1*15:01* ($P = 7.4 \times 10^{-5}$, OR 2.38) in Norwegians (6) or with -DR2 and/or -B7 in British males (10). Although still evident in our larger present series, these associations are substantially weaker (for example, *DRB1*15:01*, $P = 3.1 \times 10^{-3}$, OR 1.32), even just in the ~200 combined Scandinavian and British males ($P = 8.9 \times 10^{-3}$, OR 1.44). This difference may reflect underlying heterogeneity within LOMG and might best be clarified in further studies with larger patient numbers with onset before or after 1990, focusing separately on males and females with or without anti-titin or other autoantibodies listed in Table 1 and with onset before or after ages 40, 50, or 60 years, or even higher.

In addition to the *HLA-DQA1/DRB1* associations, there were independent signals from the HLA-A and the MHC class III re-

gion. The HLA-A signal was due to two risk-associated AAs as discussed below. In class III, the associated SNPs are located within introns of three genes: the read-through transcript of the Dead box polypeptide 39B and V-type proton ATPase subunit G2 (*DDX39B-ATP6V1G2*); the nuclear factor of κ light polypeptide gene enhancer in B-cells inhibitor-like 1 (*NFKBIL1*); and leukocyte-specific transcript 1 (*LST1*). *NFKBIL1* and *LST1* are potential modulators of immune and autoimmune responses. Moreover, the 50-kb interval between these SNPs also contains both tumor necrosis factor (*TNF*) and lymphotoxin α (*LTA*), each of which are strong candidate genes. Notably, differences in macrophage expression of these class III genes after stimulation reportedly associate with different *LST1* haplotypes (53). Of the three SNPs, only rs2071596 has a high predicted regulatory function (RegulomeDB score = 2b indicating TF binding + any TF binding motif + DNase Footprint + DNase peak). While another SNP (rs45457097; RegulomeDB score = 2b) located in the 3′UTR of *NFKBIL1* may be another functional candidate, it was not imputed with

Table 4. Highlights of conditional analyses of the MHC region in LOMG.

Conditioning	HLA-A: 29510630-30465978			Class III: 30466879-31763744			Class II: 31765860-33005932		
	Lowest residual P value	Variant	Position	Lowest residual P value	Variant	Position	Lowest residual P value	Variant	Position
None	1.00×10^{-11}	rs150881176	29947740	4.11×10^{-12}	rs2071591	31515799	$\mathbf{3.13 \times 10^{-17}}$	rs111945767	32517045
rs111945767	$\mathbf{2.24 \times 10^{-8}}$	rs150881176	29947740	1.17×10^{-7}	6:31534010:D	31534010	7.91×10^{-4}	rs204887	32029226
DRB1	$\mathbf{3.26 \times 10^{-8}}$	rs150881176	29947740	8.33×10^{-7}	rs2071596	31506691	5.58×10^{-5}	rs9501624	32399286
DQA1	$\mathbf{1.78 \times 10^{-8}}$	rs150881176	29947740	1.77×10^{-7}	rs2256974	31555392	1.11×10^{-3}	rs147008997	32447598
DQB1	$\mathbf{1.84 \times 10^{-9}}$	rs115207654	29942293	7.19×10^{-8}	rs2071591	31515799	9.52×10^{-8}	DQA1_AA107_I	32609806
HLA-A	1.16×10^{-4}	rs1632958	29688643	2.40×10^{-10}	rs2071591	31515799	$\mathbf{7.02 \times 10^{-16}}$	rs111945767	32517045
HLA-B	2.82×10^{-8}	rs150881176	29947740	3.17×10^{-8}	rs116706902	31529335	$\mathbf{6.85 \times 10^{-13}}$	rs111938555	32509707
HLA-C	6.99×10^{-8}	rs150881176	29947740	2.68×10^{-6}	rs2071591	31515799	$\mathbf{3.75 \times 10^{-12}}$	rs112780983	32511348
DRB1, HLA-A, Class III SNP	2.55×10^{-3}	rs1632958	29688643	$\mathbf{4.15 \times 10^{-4}}$	rs138758805	31533722	$\mathbf{1.57 \times 10^{-5}}$	rs9501624	32399286
DQA1, HLA-A, Class III SNP	2.18×10^{-3}	6:30014294:I	30014294	$\mathbf{2.16 \times 10^{-3}}$	rs12665489	31507987	2.96×10^{-3}	chr6:31834138	31834138
DQA1, HLA-A, HLA-C	2.52×10^{-3}	6:30014294:I	30014294	$\mathbf{4.20 \times 10^{-5}}$	rs2071596	31506691	4.08×10^{-3}	rs28732173	32071010
DRB1_AA11, A_AA70, Class III SNP	$\mathbf{2.34 \times 10^{-4}}$	rs150881176	29947740	7.05×10^{-4}	rs12665489	31507987	2.19×10^{-4}	rs114053271	32535948
DQA1_AA75, DQA1_AA130, A_AA70, Class III SNP	$\mathbf{5.42 \times 10^{-4}}$	rs114705891	29990708	1.57×10^{-3}	rs12665489	31507987	3.12×10^{-3}	rs6907185	31844684
DQA1 AA75, DQA1 AA130, A_AA70, A_AA152, Class III SNP	$\mathbf{1.95 \times 10^{-3}}$	rs150881176	29947740	2.06×10^{-3}	rs12665489	31507987	3.40×10^{-3}	rs143832718	32287072

The most significant P value for each condition is shown for each of the three regions with signal peaks (Figure 5). The most significant P value for each condition (row) is shown in bold. Gray shading indicates where the signal has been attenuated above $P = 5 \times 10^{-5}$. The class III SNP included in the conditioning is rs2071596, but nearly equivalent results are obtained with rs6929796 or rs2256974. For some of the AAs indicated in this table, others gave nearly identical results: [130]A was in complete LD with [41]K, and [75]I was in nearly complete LD with AAs at multiple positions (40, 47, 50, 51, 53, 107, 156, 161, 163, and 175). For these studies, conditioning was performed using all HLA antigens for each gene that had nominally significant P values ($P < 0.05$). For AAs, we individually conditioned on each AA (in DQA1, DRB1 and HLA-A) that had nominal P values <0.001. Results are for those AAs with the strongest attenuation of the association signals.

high confidence and could not be evaluated by conditioning studies.

Previous studies have shown several independent signals within the MHC region in multiple autoimmune diseases, including RA and systemic lupus erythematosus (54–57), and specifically with *HLA-A* in type 1 diabetes (54). The data are less clear for MHC class III because of the strong LD there (Figure 4C). One study has suggested a possible independent role for *DDX39B* (formerly *BAT1*) in RA (58).

Conditional association tests using specific HLA AA variants may also provide insights into possible functional cor-

relates. For *DQA1*, [130]A (DQA1*05:01) and [75]I (DQA1*01:03) gave equivalent results to conditioning on all *DQA1* antigen genes. Structural modeling studies suggested a critical role for AA 130 in DQA1/DM interactions in peptide loading (59). Because [75]I protrudes into the peptide-binding groove, it might affect recognition of the bound peptides. Although the data modestly favor *DQA1* over *DRB1*, we also noted a strong effect at AA position 11 in *DRB1* (Table 4), which is also critical in peptide binding.

For *HLA-A*, [70]Q (present in HLA-A*68:01, *03:01 and *11:01 and other

lower-frequency HLA-A AAs) and [152]A or [152]E (present in HLA-A*01:01, *03:01, *11:01, *25:01 and *26:01) could account for the signal and were risk variants. Interestingly, these AAs have been implicated in cytotoxic T-cell recognition of influenza A epitopes (60) and may thus implicate environmental provoking factors. The high prevalence in LOMG of oligoclonal expansions in both CD8+ and CD4+ T cells (61) is another tantalizing hint that viruses might be involved. Changes in exposure could be relevant to the recent increase in incidence of LOMG. In this regard, we did not find any significant difference in associa-

tion signals when we subdivided LOMG cases with onset before or after 1990 (data not shown).

When we compared associations in males and females, none of the differences proved significant (Z scores with *P* values >0.4), but the effect sizes for the HLA associations were larger in the LOMG males, whereas they were significantly higher in the females with EOMG (15). Outside of the MHC, we observed a significant association with *PTPN22* only in males with LOMG. By contrast, effect sizes for the *ZBTB10*-linked SNP were almost identical in males and females. Confirmation of these potential differences might imply stronger genetic and/or hormonal influences on MG susceptibility in younger females and older males. Sexual dimorphism and the potential role of X chromosome genes have been emphasized in other studies of autoimmunity (40,62).

In our patients with generalized and pure ocular AChR-Ab–positive LOMG, we also found broadly similar ORs for the associated SNPs, in both *HLA* and other regions (Supplementary Table S8). This finding may not be surprising, since neurotransmission is impaired in peripheral muscles in nearly all patients with AChR-Ab–positive pure ocular LOMG when tested by sensitive single-fiber electromyography, just as in generalized AChR-Ab–positive LOMG, but unlike in AChR-Ab–negative ocular MG (63–65).

Finally, this study has not detected associations with several genes previously implicated in EOMG and MG. These include *TNIP1* and such possible candidates as *STAT4*, *IKZF1*, *IRF5*, *NKX2-3*, *ORMDL3*, *CD226* and *PPG1* (15). We also found no firm evidence for the *CTLA4* associations previously reported in MG (5,16). Similarly, for the LOMG-associated genes, *ZBTB10* and *TNFRSF11A*, there was no suggestion of association in our previous EOMG study (15).

Our study has several limitations. First, the sample size limited power to identify loci with relatively modest effect sizes. This finding was also partially compounded by the incomplete ability to impute all variants in the 1,000-genome resource at a high

confidence ($r^2 > 0.8$). Second, control genotypes were obtained from disparate sources. This result was mitigated by our assessment and matching strategy based on PCA as well as including appropriate covariates to account for differences in population substructure and genotyping platform. Finally, an independent replication was precluded by the small numbers of these uncommon patients.

CONCLUSION

The data presented here provide compelling evidence that LOMG has a different genetic basis from EOMG. The bimodal distribution of onset ages of MG, and the similar results with cutoffs ≥50 and ≥60 years for LOMG (except for diminished power), together with previous studies, suggest that a cutoff age of onset of ≥50 years is a useful criterion for future investigative and clinical studies. The study confirms that a variant of *TNFRSF11A* is a risk factor for LOMG and defines independent associations with specific MHC class 1, class 2 and class 3 polymorphisms. Additional replication studies are warranted to further support the novel identification of *ZBTB10* as an LOMG risk factor, and larger sample sets using strict phenotypic definitions, as well as directed functional studies, will be necessary to further unravel the complex genetics of different forms of MG.

DISCLOSURE

The authors declare that they have no competing interests as defined by *Molecular Medicine*, or other interests that might be perceived to influence the results and discussion reported in this paper.

REFERENCES

1. McGrogan A, Sneddon S, de Vries CS. (2010) The incidence of myasthenia gravis: a systematic literature review. *Neuroepidemiology.* 34:171–83.
2. Meriggioli MN, Sanders DB. (2009) Autoimmune myasthenia gravis: emerging clinical and biological heterogeneity. *Lancet.* 8:475–90.
3. Carr AS, Cardwell CR, McCarron PO, McConville J. (2010) A systematic review of population based epidemiological studies in Myasthenia Gravis. *BMC Neurol.* 10:46.
4. Lindstrom JM, Seybold ME, Lennon VA, Whittingham S, Duane DD. (1976) Antibody to acetylcholine receptor in myasthenia gravis. Prevalence, clinical correlates, and diagnostic value. *Neurology.* 26:1054–9.
5. Chuang WY, *et al.* (2014) Late-onset myasthenia gravis-CTLA4(low) genotype association and low-for-age thymic output of naive T cells. *J. Autoimmun.* 52:122–9.
6. Maniaol AH, *et al.* (2012) Late onset myasthenia gravis is associated with HLA DRB1*15:01 in the Norwegian population. *PLoS One.* 7:e36603.
7. Compston DA, Vincent A, Newsom-Davis J, Batchelor JR. (1980) Clinical, pathological, HLA antigen and immunological evidence for disease heterogeneity in myasthenia gravis. *Brain.* 103:579–601.
8. Evoli A, Batocchi AP, Minisci C, Di Schino C, Tonali P. (2000) Clinical characteristics and prognosis of myasthenia gravis in older people. *J. Am. Geriatr. Soc.* 48:1442–8.
9. Akaishi T, *et al.* (2014) Insights into the classification of myasthenia gravis. *PLoS One.* 9:e106757.
10. Janer M, *et al.* (1999) A susceptibility region for myasthenia gravis extending into the HLA-class I sector telomeric to HLA-C. *Hum. Immunol.* 60:909–17.
11. Avidan N, Le Panse R, Berrih-Aknin S, Miller A. (2014) Genetic basis of myasthenia gravis: a comprehensive review. *J. Autoimmun.* 52:146–53.
12. Giraud M, *et al.* (2001) Linkage of HLA to myasthenia gravis and genetic heterogeneity depending on anti-titin antibodies. *Neurology.* 57:1555–60.
13. Grob D, Brunner N, Namba T, Pagala M. (2008) Lifetime course of myasthenia gravis. *Muscle Nerve.* 37:141–9.

14. Vandiedonck C, *et al.* (2006) Association of the PTPN22*R620W polymorphism with autoimmune myasthenia gravis. *Ann. Neurol.* 59:404–7.

15. Gregersen PK, *et al.* (2012) Risk for myasthenia gravis maps to a (151) Pro—>Ala change in TNIP1 and to human leukocyte antigen-B*08. *Ann Neurol* 72:927–935.

16. Renton AE, *et al.* (2015) A genome-wide association study of myasthenia gravis. *JAMA Neurol.* 72:396–404.

17. Purcell S, *et al.* (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81:559–75.

18. Pritchard JK, Stephens M, Donnelly P. (2000) Inference of population structure using multilocus genotype data. *Genetics.* 155:945–59.

19. Falush D, Stephens M, Pritchard JK. (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics.* 164:1567–87.

20. Kosoy R, *et al.* (2009) Ancestry informative marker sets for determining continental origin and admixture proportions in common populations in America. *Hum. Mutat.* 30:69–78.

21. Price AL, *et al.* (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* 38:904–9.

22. O'Connell J, *et al.* (2014) A general approach for haplotype phasing across the full spectrum of relatedness. *PLoS Genet.* 10:e1004234.

23. Fuchsberger C, Abecasis GR, Hinds DA. (2014) minimac2: faster genotype imputation. *Bioinformatics.* 31:782–4.

24. Howie BN, Donnelly P, Marchini J. (2009) A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* 5:e1000529.

25. Jia X, *et al.* (2013) Imputing amino acid polymorphisms in human leukocyte antigens. *PLoS One.* 8:e64683.

26. Browning BL, Browning SR. (2009) A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am. J. Hum. Genet.* 84:210–23.

27. Boyle AP, *et al.* (2012) Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res.* 22:1790–7.

28. Pruim RJ, *et al.* (2010) LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics.* 26:2336–7.

29. Degner JF, *et al.* (2012) DNase I sensitivity QTLs are a major determinant of human expression variation. *Nature.* 482:390–4.

30. Coviello AD, *et al.* (2012) A genome-wide association meta-analysis of circulating sex hormone-binding globulin reveals multiple Loci implicated in sex steroid hormone regulation. *PLoS Genet.* 8:e1002805.

31. Ferreira MA, *et al.* (2014) Genome-wide association analysis identifies 11 risk variants associated with the asthma with hay fever phenotype. *J. Allergy Clin. Immunol.* 133:1564–71.

32. Eriksson N, *et al.* (2012) Novel associations for hypothyroidism include known autoimmune risk loci. *PLoS One.* 7:e34442.

33. Paternoster L, *et al.* (2011) Meta-analysis of genome-wide association studies identifies three new risk loci for atopic dermatitis. *Nat. Genet.* 44:187–92.

34. Hinds DA, *et al.* (2013) A genome-wide association meta-analysis of self-reported allergy identifies shared and allergy-specific susceptibility loci. *Nat. Genet.* 45:907–11.

35. Tillotson LG. (1999) RIN ZF, a novel zinc finger gene, encodes proteins that bind to the CACC element of the gastrin promoter. *J. Biol. Chem.* 274:8123–8.

36. Li X, *et al.* (2010) MicroRNA-27a indirectly regulates estrogen receptor α expression and hormone responsiveness in MCF-7 breast cancer cells. *Endocrinology.* 151:2462–73.

37. Tone M, Powell MJ, Tone Y, Thompson SA, Waldmann H. (2000) IL-10 gene expression is controlled by the transcription factors Sp1 and Sp3. *J. Immunol.* 165:286–91.

38. Tone M, Tone Y, Babik JM, Lin CY, Waldmann H. (2002) The role of Sp1 and NF-kappa B in regulating CD40 gene expression. *J. Biol. Chem.* 277:8890–7.

39. Bynote KK, *et al.* (2008) Estrogen receptor-alpha deficiency attenuates autoimmune disease in (NZB x NZW)F1 mice. *Genes Immun.* 9:137–52.

40. Rubtsova K, Marrack P, Rubtsov AV. (2015) Sexual dimorphism in autoimmunity. *J. Clin. Invest.* 125:2187–93.

41. Kaya GA, *et al.* (2014) The association of PTPN22 R620W polymorphism is stronger with late-onset AChR-myasthenia gravis in Turkey. *PLoS One.* 9:e104760.

42. Bottini N, *et al.* (2004) A functional variant of lymphoid tyrosine phosphatase is associated with type I diabetes. *Nat. Genet.* 36:337–8.

43. Begovich AB, *et al.* (2004) A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. *Am. J. Hum. Genet.* 75:330–7.

44. Kyogoku C, *et al.* (2004) Genetic association of the R620W polymorphism of protein tyrosine phosphatase PTPN22 with human SLE. *Am. J. Hum. Genet.* 75:504–7.

45. Bottini N, Peterson EJ. (2014) Tyrosine phosphatase PTPN22: multifunctional regulator of immune signaling, development, and disease. *Ann. Rev. Immunol.* 32:83–119.

46. Hughes AE, *et al.* (2000) Mutations in TNFRSF11A, affecting the signal peptide of RANK, cause familial expansile osteolysis. *Nat. Genet.* 24:45–8.

47. Guerrini MM, *et al.* (2008) Human osteoclast-poor osteopetrosis with hypogammaglobulinemia due to TNFRSF11A (RANK) mutations. *Am. J. Hum. Genet.* 83:64–76.

48. Albagha OM, *et al.* (2010) Genome-wide association study identifies variants at CSF1, OPTN and TNFRSF11A as genetic risk factors for Paget's disease of bone. *Nat. Genet.* 42:520–4.

49. Estrada K, *et al.* (2012) Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nat. Genet.* 44:491–501.

50. Roberts NA, *et al.* (2012) Rank signaling links the development of invariant gammadelta T cell progenitors and Aire(+) medullary epithelium. *Immunity.* 36:427–37.

51. Anderson DM, *et al.* (1997) A homologue of the TNF receptor and its ligand enhance T-cell growth and dendritic-cell function. *Nature.* 390:175–9.

52. Dougall WC, *et al.* (1999) RANK is essential for osteoclast and lymph node development. *Genes Dev.* 13:2412–24.

53. Mewar D, *et al.* (2006) Haplotype-specific gene expression profiles in a telomeric major histocompatibility complex gene cluster and susceptibility to autoimmune diseases. *Genes Immun.* 7:625–31.

54. Noble JA, *et al.* (2010) HLA class I and genetic susceptibility to type 1 diabetes: results from the type 1 Diabetes Genetics Consortium. *Diabetes.* 59:2972–9.

55. Raychaudhuri S, *et al.* (2012) Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. *Nat. Genet.* 44:291–6.

56. Jawaheer D, *et al.* (2002) Dissecting the genetic complexity of the association between human leukocyte antigens and rheumatoid arthritis. *Am. J. Hum. Genet.* 71:585–94.

57. Morris DL, *et al.* (2012) Unraveling multiple MHC gene associations with systemic lupus erythematosus: model choice indicates a role for HLA alleles and non-HLA genes in Europeans. *Am. J. Hum. Genet.* 91:778–93.

58. Quinones-Lombrana A, *et al.* (2008) BAT1 promoter polymorphism is associated with rheumatoid arthritis susceptibility. *J. Rheumatol.* 35:741–4.

59. Gockel I, *et al.* (2014) Common variants in the HLA-DQ region confer susceptibility to idiopathic achalasia. *Nat. Genet.* 46:901–4.

60. McMichael AJ, Gotch FM, Santos-Aguado J, Strominger JL. (1988) Effect of mutations and variations of HLA-A2 on recognition of a virus peptide epitope by cytotoxic T lymphocytes. *Proc. Natl. Acad. Sci. U. S. A.* 85:9194–8.

61. Tackenberg B, *et al.* (2007) Clonal expansions of CD4+ B helper T cells in autoimmune myasthenia gravis. *Eur. J. Immunol.* 37:849–63.

62. Selmi C, Brunetta E, Raimondo MG, Meroni PL. (2012) The X chromosome and the sex ratio of autoimmunity. *Autoimmun. Rev.* 11:A531–7.

63. Padua L, *et al.* (2000) SFEMG in ocular myasthenia gravis diagnosis. *Clin. Neurophysiol.* 111:1203–7.

64. Farrugia ME, Jacob S, Sarrigiannis PG, Kennett RP. (2009) Correlating extent of neuromuscular instability with acetylcholine receptor antibodies. *Muscle Nerve.* 39:489–93.

65. Witoonpanich R, Dejthevaporn C, Sriphrapradang A, Pulkes T. (2011) Electrophysiological and immunological study in myasthenia gravis: diagnostic sensitivity and correlation. *Clin. Neurophysiol.* 122:1873–7.