

The human *GIMAP5* gene has a common polyadenylation polymorphism increasing risk to systemic lupus erythematosus

Anna Hellquist, Marco Zucchelli, Katja Kivinen, Ulpu Saarialho-Kere, Sari Koskenmies, Elisabeth Widen, Heikki Julkunen, Andrew Wong, Marja-Liisa Karjalainen-Lindsberg, Tiina Skoog, Johanna Vendelin, Deborah S Cunninghame-Graham, Timothy J Vyse, Juha Kere, Cecilia M Lindgren

J Med Genet 2007;44:314–321. doi: 10.1136/jmg.2006.046185



Supplementary methodological information, tables and figures are available at <http://jmg.bmj.com/supplemental>

See end of article for authors' affiliations

Correspondence to: Professor Juha Kere, Karolinska Institutet, Department of Biosciences and Nutrition, S-141 57 Huddinge, Sweden; juha.kere@biosci.ki.se

Received 30 August 2006
Revised 26 November 2006
Accepted 16 December 2006
Published Online First
12 January 2007

Background: Several members of the *GIMAP* gene family have been suggested as being involved in different aspects of the immune system in different species. Recently, a mutation in the *GIMAP5* gene was shown to cause lymphopenia in a rat model of autoimmune insulin-dependent diabetes. Thus it was hypothesised that genetic variation in *GIMAP5* may be involved in susceptibility to other autoimmune disorders where lymphopenia is a key feature, such as systemic lupus erythematosus (SLE).

Material and methods: To investigate this, seven single nucleotide polymorphisms in *GIMAP5* were analysed in five independent sets of family-based SLE collections, containing more than 2000 samples.

Result: A significant increase in SLE risk associated with the most common *GIMAP5* haplotype was found (OR 1.26, 95% CI 1.02 to 1.54, $p=0.0033$). In families with probands diagnosed with thrombocytopenia, the risk was increased (OR 2.11, 95% CI 1.09 to 4.09, $p=0.0153$). The risk haplotype bears a polymorphic polyadenylation signal which alters the 3' part of *GIMAP5* mRNA by producing an inefficient polyadenylation signal. This results in higher proportion of non-terminated mRNA for homozygous individuals ($p<0.005$), a mechanism shown to be causal in thalassaemias. To further assess the functional effect of the polymorphic polyadenylation signal in the risk haplotype, monocytes were treated with several cytokines affecting apoptosis. All the apoptotic cytokines induced *GIMAP5* expression in two monocyte cell lines (1.5–6 times, $p<0.0001$ for all tests).

Conclusion: Taken together, the data suggest the role of *GIMAP5* in the pathogenesis of SLE.

Systemic lupus erythematosus (SLE, OMIM 152700) is a multisystem autoimmune disease with largely unknown aetiology affecting predominantly women. Its prevalence in the Western population is about 0.05%,^{1,2} a number that differs widely in other ethnic groups.³ Although many candidate genes have been suggested as being involved in susceptibility to SLE,^{4,5} few have been replicable.⁶

SLE commonly shows familial co-segregation with other autoimmune disorders: 10–20% of SLE probands have at least one first- or second-degree relative with another autoimmune disorder.³ Disturbances of apoptosis of regulatory T-cells, which normally facilitates maintenance of immune tolerance, has been suggested as causing autoimmunity⁷ and is thought to contribute to lymphopenia in several autoimmune disorders such as autoimmune insulin-dependent diabetes,⁸ Sjogren's syndrome⁹ and SLE.¹⁰

Recently, *GIMAP5* (OMIM 608086; previously called *Immunity-associated nucleotide 4-like1*, *IAN4L1*, fig 1A) was found to cause lymphopenia and autoimmune insulin-dependent diabetes in a rat model.^{11–13} Several members of the *GIMAP* gene family have been suggested as being involved in different aspects of the immune system in different species.^{11,12,14,15} These events appear to be associated with cell proliferation and apoptosis, suggesting a role for *GIMAP* proteins in control of cell death/survival.¹⁶ In line with this, several studies have shown that *GIMAP5* normally protects cells from apoptosis and that loss of *GIMAP5* function results in decreased mitochondrial integrity with lymphocyte apoptosis of primarily T cells.^{11,12,17–20}

Thus, based on current research, *GIMAP5* is a potential candidate gene for autoimmune disorders where apoptosis and lymphopenia are key features, such as SLE. Thus, we set out to test association between SLE and *GIMAP5* and designed our study to control as many common confounders of genetic association studies as possible.⁶

MATERIAL AND METHODS

Subjects and samples

All subjects gave written informed consent for participation in genetic studies on SLE. The study protocols for all samples were reviewed and approved by local ethical committees, for the Finnish samples at University of Helsinki and at Karolinska Institutet, and for the UK samples at the Multi-Centre Research Ethics Committee.

Finnish SLE samples

Collection of this sample has been described in detail previously.¹ All patients with SLE were interviewed by the same physician, and the case records from the hospitals where the patients were treated were reviewed. All patients met the American College of Rheumatology criteria for the diagnosis of SLE.²¹ According to the prevalence of SLE in Finland,²² approximately 1200 of 1500 available patients were contacted

Abbreviations: BSA, bovine serum albumin; PBS, phosphate-buffered solution; RT-PCR, polymerase chain reaction with reverse transcription; SLE, systemic lupus erythematosus

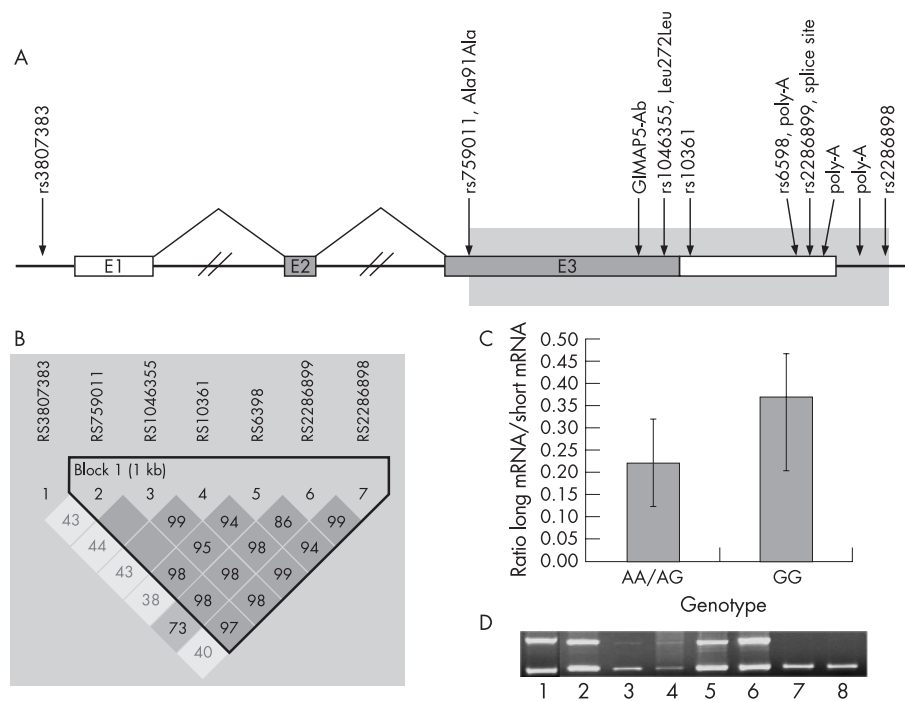


Figure 1 Schematic of the *GIMAP5* gene. A, *GIMAP5* spans 6286 base pairs (bp) and contains three exons; exon 1 is untranslated and exon 3 contains more than 94% of the coding sequence. The deduced 307 amino-acid protein (34.8 kilodaltons) contains five motifs conserved in guanosine triphosphate (GTP)-binding proteins and a putative C-terminal transmembrane region.¹⁹ B, representation of LD structure in the *GIMAP5* gene using D' with 95% CI.²⁸ Each square depicts the magnitude of linkage disequilibrium (LD) for a pair of markers, red colour indicating high D' and white colour low D'. C, The difference in ratio of amounts between the longer (526 bp) and the shorter (457 bp) bands compared between individuals either homozygous or heterozygous for the common AATAAA polyA signal and individuals homozygous for the AATAGA signal in the 3' untranslated region (UTR) of *GIMAP5* ($p < 0.005$). D, Positive complementary DNA (cDNA) controls for the polyA polymerase chain reaction with reverse transcription (RT-PCR) showing a difference in bands independent of genotype (all non-pooled samples are homozygous GG for rs6598 and are thereby homozygous for the rare AATAGA polyA signal) in different tissues, suggesting that the *GIMAP5* polyA signal functions in a cell-type-specific manner; (1) placenta, (2) U937 monocytes, (3) A549 lung epithelial cells, (4) neuroblasts BE (2)-C cells, (5) Jurkat E6-1 T lymphocytes, (6) normal pooled skin, (7) pooled brain and (8) pooled leucocytes.

during two phases of recruitment, accounting for more than 80% of all Finnish patients with SLE requiring hospital-based treatment. Blood samples from patients and their relatives were obtained, comprising a total of 192 families; of which 86 were multiply affected by SLE (multiplex families); the rest were families with sporadic cases. Clinical characteristics of the patients included in this cohort are described in supplementary table 1 and have been described previously.¹ Of these families,

76 parent-affected offspring trios and 84 discordant sib pairs were included in association studies.

British samples

In the UK collection, diagnosis of SLE was established by telephone interview, health questionnaire and details from clinical notes. All collected probands conformed to the American College of Rheumatology criteria for SLE²¹ and

Table 1 Minor allele frequencies and nominal pointwise association in the different samples

	rs3807383	rs759011	rs1046355	rs10361	rs6598	rs2286899	rs2286898
SNP rs#	[A/C]	[A/G]	[C/T]	[C/G]	[A/G]	[C/T]	[A/C]
Position (base pairs)	149871806	149877148	149877691	149877891	149878297	149878362	149878624
FIN_CAU (MAF)	0.25	0.47	0.47	0.46	0.45	0.29	0.47
UK_CAU (MAF)	0.27	0.30	0.28	0.29	0.26	0.12	0.30
UK_ASIA (MAF)	0.22	0.39	0.29	0.29	0.30	0.16	0.29
Pointwise association analysis							
FIN_CAU trios			0.057 (34:20)	0.068 (37:23)			
FIN_CAU sibs		0.025 (n.a.)					0.035 (n.a.)
UK_CAU trios							
UK_CAU sibs							
UK_ASIA trios		0.039 (14:5)			0.033 (11:3)		

All marker positions are in base pairs (bp) according to NT_007914.14 (see electronic information). All assays were in Hardy-Weinberg equilibrium ($p > 0.05$) when tested in 90 unrelated independent healthy controls (Finnish blood donors), genotyped on the Sequenom platform. MAF, minor allele frequency; T:U, transmitted:untransmitted ratio. All p values were calculated using TDT (30).

written consent was obtained from all participants. The study cohort comprised 378 parent-affected offspring trios and 171 discordant sib pairs of European Caucasian descent together with 23 parent-affected offspring trios of Indian-Asian descent. An additional 14 Indian-Asian discordant sib pairs, 16 parent-affected offspring trios and 15 discordant sib pairs of Afro-Caribbean descent and 12 families with mixed ethnicity were excluded from the analysis because of their low information content. The families have all been described in detail previously²³ and the clinical characteristics of the patients in the cohort are described in supplementary table 1.

Genotyping

First, we sequenced (see supplementary material for details) the entire *GIMAP5* gene in one patient homozygous for all of the associated single nucleotide polymorphisms (SNPs) within *GIMAP5* to identify and validate SNPs in the *GIMAP5* region, but only identified previously reported SNPs.²⁴ We chose 10 markers in the *GIMAP5* gene for genotyping: rs3807383, rs759011, rs1046355, rs9657881, rs13229808, rs13229806, rs10361, rs6598, rs2286899 and rs2286898. Primers and probes for genotyping were designed using Primer3 (see electronic information).

The Finnish samples were genotyped using the SnuPe kit according to the manufacturer's instructions (Amersham Pharmacia Biotech, Uppsala, Sweden). Three of the tested assays, rs9657881, rs13229806 and rs13229808, failed to meet our quality criteria (as described below) and were subsequently excluded from further genotyping. We found the remaining seven SNPs to identify all common variation within *GIMAP5* redundantly.

Subsequently, DNA from our UK replication samples were genotyped for the same seven markers using matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry according to the manufacturer's instructions (Sequenom Inc., San Diego, California, USA).²⁵ PCR assays and associated extension reactions were designed using the SpectroDESIGNER software (Sequenom Inc.).

Using either genotyping method, genotypes for each SNP were confirmed by two independent scorers. For all data, PedCheck²⁶ was used to detect Mendelian inconsistencies and markers were excluded if they yielded more than 10% errors. Hardy-Weinberg calculations were performed to ensure that each marker was in equilibrium ($p > 0.05$). Success rate of all assays was required to be at least 85%. To ensure consistency between the two genotyping methods, three markers were re-genotyped in the Finnish material with the matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry method (rs6598, rs2286899 and rs2286898). The concordance

levels between the methods were greater than 98%. The concordance within the UK sample sets (repeated in two different labs on two different Sequenom platforms) was greater than 99%.

Expression analysis of *GIMAP5*

Peripheral blood mononuclear cells were isolated from fresh blood samples from nine patients with SLE and nine healthy controls from Finland using Ficoll-Paque PLUS (Amersham Biosciences, Uppsala, Sweden). Total RNA was extracted using RNeasy Mini-kit (Qiagen, Chatsworth, California, USA) as instructed by the manufacturer and 800 ng were reverse transcribed to cDNA using TaqMan Reverse transcription reagents (Applied Biosystems, Warrington, UK).

Primers were designed to cover a 80 bp fragment, spanning from exon 2 to exon 3, specific for the coding domain (359–438 bp in *GIMAP5* mRNA; NM_018384) using Primer3 (electronic information). Quantitative real-time PCR (TaqMan) was performed on an ABI PRISM 7700 Sequence Detection System according to the manufacturer's instructions (Applied Biosystems, see supplementary material for detailed information).

Allelic expression analysis of rs759011, rs1046355, rs10361, rs6598 and rs2286899

Genomic DNA from blood and cDNA from the same individuals as described above were sequenced (see supplementary material) for SNP detection and investigation of allelic expression levels.²⁷

Northern Blot analysis of *GIMAP5* and β actin

An 80 bp RT-PCR product from brain (Clontech, Palo Alto, California, USA), specific for the coding domain of *GIMAP5* (359–438 bp in *GIMAP5* mRNA; accession number NM_018384) was labelled with P³²-dCTP (Amersham Biosciences) by random priming (High Prime, Roche, Penzberg, Germany) and used as probe on Human Multiple Tissue Northern Blots (Ambion, Austin, Texas, USA) and Human Immune System MTN[®] Blot II (Clontech) with β actin as control probe.

RT-PCR on *GIMAP5* and *GAPDH*

The same primers used for the *GIMAP5* expression (TaqMan) analysis were used to study the expression of *GIMAP5* at mRNA level. Total RNA (0.5 μ g) from neuroblast BE (2)-C cells, Jurkat E6-1 T-cells, 15 HL-60 eosinophils, A549 lung carcinoma epithelial cells, U937 histiocytic lymphoma monocytes, normal pooled human skin (BD Biosciences, Palo Alto, California, USA) and human placenta (Ambion) were converted to cDNA using

Table 2 Analysis of the common six marker haplotypes spanning *GIMAP5* in five independent sample sets

	T	U	OR	Low			High			Haplotypes					
				95% CI	95% CI	p value	95% CI	95% CI	p value	rs759011	rs1046355	rs10361	rs6598	rs2286899	rs2286898
Finnish Caucasian trios (global p 0.046)	42	26	1.52	0.99	2.63	0.05	C	C	G	G	T	C			
Finnish Caucasian sibships (global p n.s.)	37	34	1.09	0.68	1.73	0.13	C	C	G	G	T	C			
UK Caucasian trios (global p not significant)	119	95	1.25	0.96	1.64	0.1	C	C	G	G	T	C			
UK Asian (global p < 0.07)	11	3	3.67	1.02	13.2	0.03	C	C	G	G	T	C			
UK Caucasian sibships (global p not significant)	85	92	0.92	0.69	1.24	0.27	C	C	G	G	T	C			

All six marker haplotypes with frequency greater than 5% in the control chromosomes in the populations were analysed. rs3807383 is not in linkage disequilibrium with the other markers and falls outside the haplotype block as defined by Gabriel et al.²⁸ Two-sided p values are reported, except in the Asian sample, which is considered a replication and thus reported as one-sided. A global p value was calculated for the distribution of haplotypes using 50 000 permutations in each population. T, transmitted; U, undertransmitted.

SuperScriptTM III Reverse Transcriptase reagents (Invitrogen). Amplification reactions were run in a total volume of 25 μ l with 2 μ l of cDNA, 0.8 μ M of each primer, 0.2 mM dNTP, 1.5 mM MgCl₂ and 0.03 U/ μ l of HotStarTaq (Qiagen) using the following programme; 95°C for 15 min followed by 40 cycles (45 s at 95°C, 30 s at 55°C, and 1 min 15 s at 72°C) and 10 min extension at 72°C and run on an agarose gel.

Regulation of *GIMAP5* expression by SLE-associated cytokines

Two monocytic cell lines, THP-1 and U937, were plated on six-well-plates (1.4 \times 10⁶ cells per well) and grown overnight. The cells were depleted of serum overnight before stimulation with tumor necrosis factor alpha (TNF- α , 10 ng/ml, Sigma, St Louis, Missouri, USA), transforming growth factor beta 1 (TGF- β 1, 10 ng/ml, Sigma), basic fibroblast growth factor (FGF2, 10 ng/ml, Sigma), interleukin-1 beta (IL-1 β , 0.5 and 5 U/ml, Roche Molecular Biochemicals, Indianapolis, Indiana, USA) and gamma-interferon (IFN- γ , 1 and 10 ng/ml, Sigma). Stimulation was allowed to proceed for 24 h after which the cells were lysed and total RNA was extracted using RNeasy Mini-kit (Qiagen). Cells grown in serum-free medium were used as controls. All experiments were performed at least twice, in triplicate wells.

Total RNA (0.5 μ g) from cultured cells was reverse transcribed to cDNA (Invitrogen) and used as a template for TaqMan PCR. Reactions were performed with the 7500 Fast Real-Time PCR system using standard protocols (Applied Biosystems).

Differential polyA RT-PCR on *GIMAP5*

One forward primer located 412 bp upstream (F) and two reverse primers located 132 bp upstream (R1) or 42 bp downstream (R2) of rs6598 were designed. Fresh peripheral blood mononuclear cells were collected from 17 individuals, and lymphocytes were immortalised by Epstein-Barr virus transformation. RNA extraction and cDNA transcription were performed as above.

As positive control, cDNA from placenta (Ambion), U937 monocytes, A549 lung epithelial cells, neuroblasts BE (2)-C cells, Jurkat E6-1 T lymphocytes, normal pooled skin (BD Biosciences), pooled human fetal brain (Clontech) and pooled human leucocytes (Clontech) were used. Genotypes for rs6598 in all non-pooled samples were obtained by genotyping, sequencing or both.

Amplification reactions were run in duplicates in 25 μ l with 20 ng of cDNA, 1.6 μ M of forward primer (Proligo, France) and 0.8 μ M each reverse primer (Proligo), 0.2 mM dNTP, 1.5 mM MgCl₂ and 0.075 U/ μ l of HotStarTaq (Qiagen). Reactions were first heated to 95°C for 15 min, and then subjected to 39 cycles (45 s at 95°C, 30 s at 59°C and 1 min 15 s at 72°C) followed by a 10 min extension at 72°C and run on an agarose gel.

The signal of the long and short band for the different genotypes was quantified by measuring the area of the band and the number of pixels in the area after maximising the intensity on the gel. A mean of the ratio of the long band over the short band from two measurements on each individual was used to adjust for laboratory variation in the signal amounts. The individuals homozygous for the rs6598 canonical polyA signal (n = 2) were combined with the individuals heterozygous for the rs6598 canonical polyA signal (n = 6) for statistical analysis and compared with the individuals homozygous for the rare AATAGA polyA signal (n = 9).

GIMAP5 protein analysis

We examined the expression of *GIMAP5* protein (accession number Q96F15) in formalin-fixed paraffin embedded tissue sections from healthy individuals, using exon-3-specific

polyclonal antibodies (fig 1). The antibodies were generated against the peptide LGRREGSFHSDNLF in two rabbits (Sigma Genosys, Suffolk, UK) and affinity purified to final concentrations of 0.4 mg/ml and 1.0 mg/ml, respectively.

The tissue sections were stained using the avidin-biotin-peroxidase complex technique (DakoCytomation Norden, Solna, Sweden). Diaminobenzidine (Vector Laboratories, Immunkemi F&D AB, Jarfalla, Sweden) was used as the chromogenic substrate and Mayer's haematoxylin as the counterstain (Merck, Darmstadt, Germany). Incubation with primary antibody, diluted to a final concentration of 0.02 mg/ml in phosphate-buffered solution (PBS) containing 3% bovine serum albumin (BSA, Sigma-Aldrich), was performed overnight at 4°C. Corresponding preimmune serum was used as a negative control.

To test the specificity of the polyclonal antibodies, a Western blot analysis was performed. U937 monocytes were grown to a density of 1 \times 10⁶ cells per millilitre, washed twice in ice-cold PBS and harvested in 1 ml RIPA buffer. Aliquots from cleared lysates containing 10, 20 or 30 μ g protein were subjected to electrophoresis on a 10% SoftGel precast polyacrylamide gel (Mirador DNA Design INC, Montreal, Canada) and transferred to a Hybond-P PVDF transfer membrane (Amersham Biosciences). Protein was detected using an ECL Advance Western Blotting Detection Kit (Amersham Biosciences), and then visualised on an ECL HyperfilmTM (Amersham Biosciences).

Statistics

To evaluate haplotype block patterns over the *GIMAP5* locus, linkage disequilibrium between the SNPs was analysed using Haploview 2.1²⁸. Data for 119 SNPs (between rs952825 and rs717658) from the International HapMap project spanning the *GIMAP5* locus was used as a reference. Two of our seven SNPs (rs759011 and rs2286898) were also genotyped in the Centre d'Etude du Polymorphisme Humain (CEPH) reference material from the HapMap data (greater than 99% concordance with genotyping in our laboratory).

Haplotypes were reconstructed using maximum-likelihood estimation in nuclear families by FAMHAP.²⁹ Subsequent family-based, nominal pointwise and haplotype associations were first assessed by the transmission disequilibrium test³⁰ as implemented in Genehunter 2.1 R2.³¹ Global significance of the distribution of haplotypes was assessed by a randomisation test on the transmitted alleles/haplotypes using 50 000 iterations. For association analysis using discordant sib pairs, we used the discordant allele test.³² Gender and age of onset as covariates were non-significant and therefore disregarded from further analyses.

Meta-analysis³³ of the independent sample sets was performed in two steps, first for SLE and then for thrombocytopenia; first the Breslow-Day test for non-compatibility of ORs was performed (p>0.2) and then the pooled estimate of ORs was obtained using the Mantel-Haenzel method as implemented in the "R" software (see electronic information).

The differences in expression levels between individuals with SLE compared with healthy controls as well as by haplotype, independent of affection status, were calculated as Δ Ct values between (*GIMAP5*-*GAPDH*) and subsequently compared using a *t* test as implemented in StatsDirect software version 2.3.4.

The differences in expression levels between untreated and cytokine treated monocytes were calculated as Δ Ct values between (*GIMAP5*-*GAPDH*) and subsequently compared using a *t* test as implemented in StatsDirect version 2.3.4. Fold changes in *GIMAP5* expression levels were calculated according to the manufacturer's instructions (Applied Biosystems).

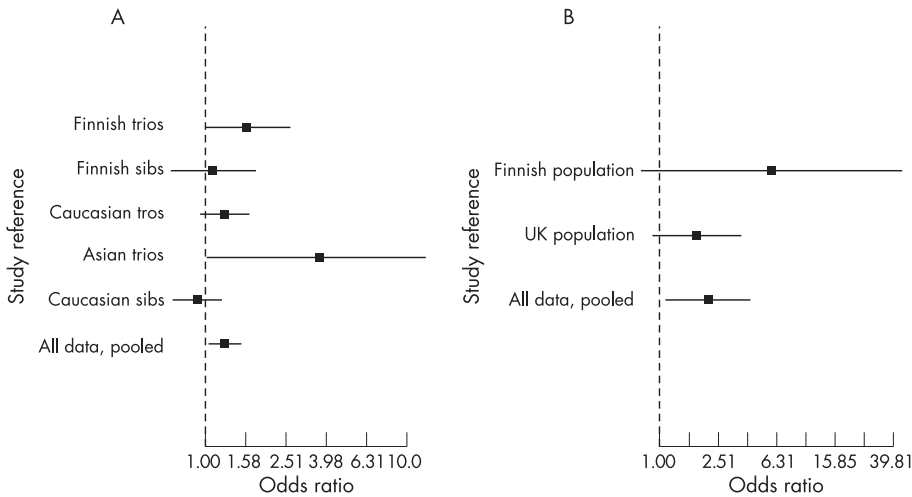


Figure 2 Meta-analysis of the association between the *GIMAP5* risk haplotype and SLE (A) and the association of the *GIMAP5* risk haplotype thrombocytopenia in patients with SLE (B). In both cases the OR of each analysis surrounded by the 95% CI for the six-marker haplotype of *GIMAP5* is given. For each study, • represents the estimated OR for the risk haplotype. Bars indicate the 95% CIs.

The differences in the signal amount of the long and short band between the different genotypes in the polyA differential RT-PCR was assessed by *t* test as implemented in StatsDirect version 2.3.4.

All *p* values are given as nominal, two sided *p* values if not stated otherwise.

RESULTS

To capture all common variation in *GIMAP5*, we genotyped all individuals using seven SNPs distributed over the gene (fig 1A). All markers except one (rs3807383) belong to one haplotype

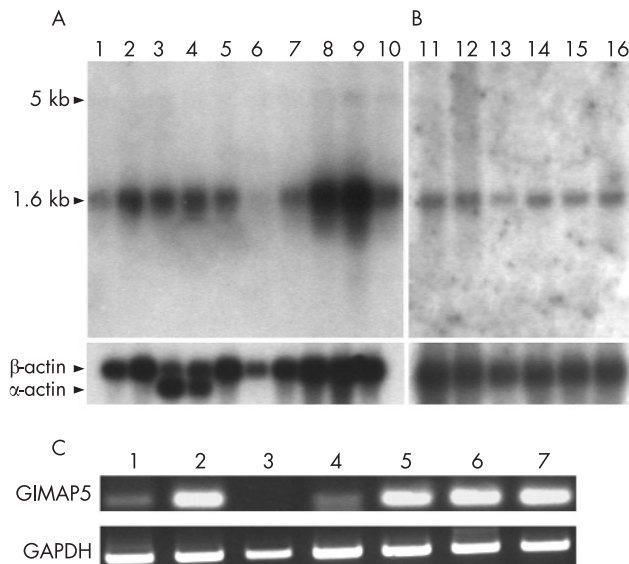


Figure 3 Multiple tissues (A) and immune-related tissue northern blot analysis (B). Commercial northern blots containing 2 μ g polyA⁺ RNA from brain (1), placenta (2), skeletal muscle (3), heart (4), kidney (5), pancreas (6), liver (7), lung (8), spleen (9), colon (10), spleen (11), lymph node (12), thymus (13), peripheral blood leucocytes (14), bone marrow (15) and fetal liver (16) were used. After exposing the blot with a ³²P-random labelled 80 bp probe for 72 hours, we detected a 1.6 kilobase (kb) transcript and after additional 72 h a weak band of approximately 5.5 kb appears, as reported previously.²⁵ For control, the northern blots were probed with β -actin. C, RT-PCR on *GIMAP5* and *GAPDH* on cDNA from neuroblasts BE (2)-C cells (1), Jurkat E6-1 T lymphocytes (2), 15 HL-60 eosinophils (3), A549 lung epithelial cells (4), U937 monocytes (5), normal pooled skin (6), placenta (7).

block (as defined by Gabriel et al²⁸), covering the protein coding parts as well as the 3' untranslated region (UTR) (fig 1B). *GIMAP5* spans a 360 kilobase (kb) region on chromosome 7q36.1, also containing a cluster of 10 additional *GIMAP* gene family members. To ensure that the genetic effect was specific to *GIMAP5*, we evaluated the linkage disequilibrium structure in the region, and *GIMAP5* is clearly separated from the nearby genes.

In Finnish trios we found suggestive association of SLE to a single SNP (table 1; rs1046355 $p \leq 0.06$) and to the most common haplotype (table 2; rs759011–rs2286898, CCGGTC, $p \leq 0.05$). To evaluate these findings more extensively,⁶ we genotyped the same seven SNPs in four additional family-based sample sets and tested for *GIMAP5* association with SLE. In 84 discordant Finnish sib pairs, we found nominal significance for two markers (rs759011 $p < 0.03$ and rs2286898 $p < 0.04$). The suggested risk haplotype showed a trend in the same direction as the initial finding (OR 1.09, 95% CI 0.68 to 1.73, $p = 0.13$). In 378 UK Caucasian trios, the most common haplotype showed a trend towards an association with SLE (OR 1.25, 95% CI 0.96 to 1.64, $p = 0.1$, fig 2).

In 21 UK Asian trios, two SNPs suggested pointwise association (rs759011; $p = 0.039$ and rs6598 $p = 0.033$) and the common risk-haplotype was trending towards an association with SLE (OR 3.67, 95% CI 1.02 to 13.2, $p = 0.03$). However, the 95% CI for the genetic effect in this population was broad, reflecting the small sample size (fig 2). In an additional UK sample, 171 discordant sib pairs were tested for association but no point wise or haplotype associations were detected (OR 0.92, 95% CI 0.69 to 1.24, $p = 0.27$).

The most exact estimate of the suggested haplotype association was obtained by combining the ORs in a Mantel-Haenzel test. This analysis reveals a significant association between the most common *GIMAP5* haplotype and SLE (OR 1.26, 95% CI 1.02 to 1.54, $p = 0.0033$, one sided, fig 2).

GIMAP5 is a small protein, suggested to reside in the mitochondrial membrane^{19 34 35} (fig 1A). None of the genotyped SNPs of the risk haplotype changed the amino-acid sequence of *GIMAP5*; two were silent polymorphisms, one altered the first canonical polyadenylation (polyA) signal (AATAAA) and one was a putative splice variant 22 bp from the end of the mRNA occurring after the canonical polyA signal (fig 1A). We therefore compared total *GIMAP5* mRNA expression level between patients with SLE ($n = 8$) and controls ($n = 9$), or by their individual haplotype, independent of affection status. We saw a nearly significant difference in expression between

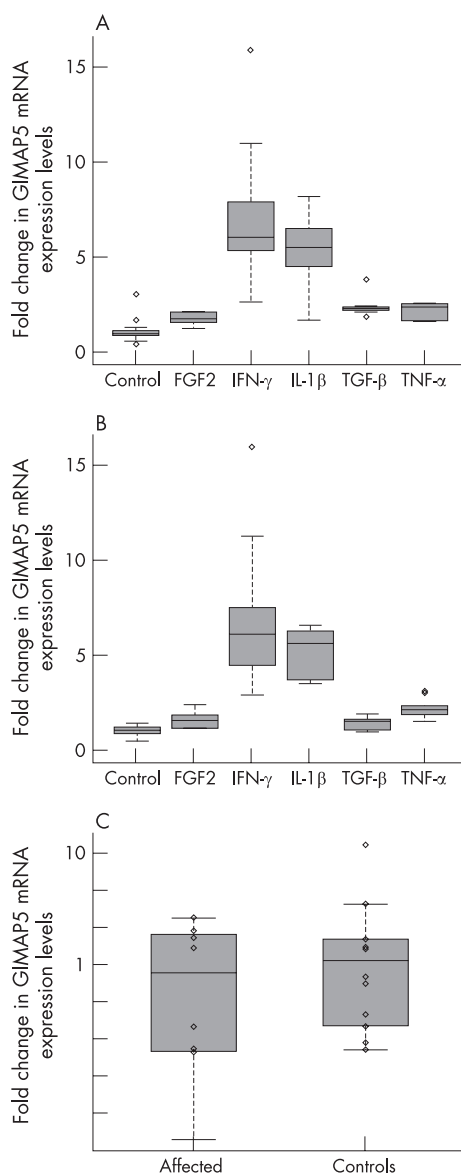


Figure 4 Cytokine effect on *GIMAP5* mRNA expression in (A) THP-1 monocytes and (B) U937 monocytes. Data were collected for untreated cells (controls) and cells treated with five different cytokines. All cytokines significantly upregulated *GIMAP5* in a similar fashion in both tested cell lines (IFN- γ 5.7/6.7-fold, p values $10^{-8}/10^{-5}$; IL-1 β 4.9/3.9-fold, p values $10^{-7}/10^{-6}$, TNF- α 2.1/2.2-fold, p values $10^{-4}/10^{-4}$, TGF- β 1.4/2.4-fold, p values 0.05/ 10^{-7} and FGF2 1.5/1.8-fold, p values 0.02/ 10^{-6}). C, Differences in expression between patients with SLE and controls shown in log scale ($p=0.074$, power for $\alpha<0.05$ significance is 76%), showing a trend for lower expression in patients. Expression levels are given as relative fold changes of *GIMAP5* (y axis).

patients with SLE and controls (fig 4C, $p=0.074$, power for $\alpha<0.05$ significance is 76%), with lower expression in patients. No significant difference was seen between non-risk homozygous + risk-heterozygous ($n=9$) and risk-homozygous ($n=10$) ($p=0.7$, data not shown). We also assessed the allele-specific expression of all genotyped SNPs in exon 3 (rs759011, rs1046355, rs10361, rs6598, rs2286899), but did not find any significant differences between patients and controls.

GIMAP5 contains two consecutive polyA signals, and one SNP (rs6598) that marks the risk haplotype resides at the fifth base of the first canonical polyA signal (fig 1A). This SNP causes a switch from the canonical AATAAA polyA signal to a rare polyA

signal (AATAGA), present in only 0.7% of mature mRNA species.³⁶ We hypothesised that the alternative polyA signal might cause a differential termination of the transcription of *GIMAP5*, an mechanism that has previously been described for the globulin gene causing α - and β thalassaemia.^{37, 38} Thus, we investigated differences in mRNA lengths between individuals with different genotypes at rs6598. Individuals homozygous for the rarer AATAGA signal have a higher amount of longer mRNA transcripts than individuals heterozygous or homozygous for the conservative polyA signal AATAAA (comparing the ratio of long bands/short bands in individuals, $n=9$, with either AA/AG genotypes = 0.22 ± 0.10 or individuals, $n=8$, with GG genotypes = 0.37 ± 0.16 , $p<0.005$, power for $\alpha<0.05$ significance is 76%). The AATAGA polyA signal efficiency has been suggested as cellular-context dependent,³⁶ and there is emerging evidence that different polyA signals in the same gene could react differently on environmental factors in a tissue, differentiation or development specific way.^{36, 39} Supporting this, we observed either one (suggesting predominant use of the first polyA signal) or two bands (suggesting leaky recognition and use of both polyA signals) in our polyA RT-PCR when using cDNA from placenta, U937 monocytes, A549 lung epithelial cells, neuroblasts BE (2)-C cells, Jurkat E6-1 T lymphocytes, normal pooled skin, pooled human fetal brain and pooled human leucocytes (fig 1D), with all non-pooled cDNA samples being homozygous GG for rs6598, and thereby also homozygous for the rarer AATAGA polyA signal. The genotype for the pooled samples could not be determined because of mixed genotypes in these samples.

Next, we investigated whether the identified *GIMAP5* risk haplotype might affect the haematological profile of patients with SLE, as could be expected from the initial results from the animal model of autoimmune type 1 diabetes.^{11, 12} As a systemic, autoimmune disorder, SLE typically displays several different haematological disturbances, such as thrombocytopenia or lymphopenia, in a large proportion of patients (20–40%).^{40, 41} Thrombocytopenia was available as the only haematological phenotype in all our patient cohorts. We analysed our data stratified by the presence of thrombocytopenia and found evidence that in patients with SLE, carriage of the risk haplotype increased the likelihood of developing thrombocytopenia (OR 2.11, 95% CI 1.09 to 4.09, $p=0.0153$). *GIMAP5* mRNA was strongly expressed in the spleen (the major site of thrombocyte clearance), lymph nodes and lung, but less abundantly in other tissues (fig 3).

To characterise the human *GIMAP5* protein expression in humans more precisely, we studied relevant tissues using immunohistochemistry. Interestingly, *GIMAP5* was detected in endothelial cells in blood vessels and kidney, lung septal cells, parenchymal cells in the thymus, and polymorphonuclear leucocytes, in agreement with mRNA expression patterns (supplementary fig 2).

To investigate what effect several cytokines (IFN- γ , IL-1 β , TNF- α , TGF- β and FGF2), well-known to affect apoptosis and inflammation, would have on *GIMAP5* expression, we treated two different monocyte cell lines (U937 and THP-1) with these cytokines and measured changes in mRNA levels. We found consistent and significant upregulation of *GIMAP5* expression in both monocyte cell lines (fig 4A,B; IFN- γ 5.7/6.7-fold, p values $10^{-8}/10^{-5}$; IL-1 β 3.9/4.9-fold, p values $10^{-7}/10^{-6}$, TNF- α 2.1/2.2-fold, p values $10^{-4}/10^{-4}$, TGF- β 1.4/2.4-fold, p values 0.05/ 10^{-7} and FGF2 1.5/1.8-fold, p values 0.02/ 10^{-6} ; power for $\alpha<0.05$ significance is 80%) after cytokine treatment.

DISCUSSION

An association between SLE and *GIMAP5* is consistent in all independent family-based samples (combined OR 1.26, 95% CI

1.02 to .54, $p=0.0033$). The suggested risk haplotype is common (about 50% in control chromosomes) as seen in several associations to common complex disorders.^{42–43} This may explain why no genome-wide scans have identified this region as a significant candidate for harbouring an SLE-susceptibility gene,⁴ because most genome-wide studies are underpowered to detect common risk variants.⁴⁴ However, *GIMAP5* is located about 2.5 centimorgans away from the suggestive scores of linkage in the combined mapping results for Minnesota SLE cohorts 1 and 2,⁴⁵ and the region on chromosome 7q34–36.3 is ranked as one of the most interesting regions in a recent meta-analysis of all SLE genome-wide linkage studies.⁵

GIMAP5 was associated to thrombocytopenia in patients with SLE with the risk haplotype. Because thrombocytopenia predicts severe disease progression and premature death in patients with SLE, this finding may potentially have clinical relevance. It would be important to investigate the *GIMAP5* association with the other haematological features of SLE, but larger collaborations will be needed for sufficient power for the less common subphenotypes.

The expression of *GIMAP5* was not significantly altered, but the detected trend towards lower levels of mRNA in patients is in line with findings of increased leucocyte apoptosis in patients with SLE and findings of reduced T-lymphocyte survival in a mouse knockdown for *GIMAP5*.⁴⁶ The two consecutive polyA signals in *GIMAP5* are a rather unusual feature in human-expressed sequence tags.³⁶ A similar A→G replacement of the fifth base of the polyA signal as rs6598 is observed in the human *globin* mRNA and is known to cause α - and β thalassaemia.^{37–38} The generated *globin* polyA signal is inefficient and allows transcription to continue past it, producing a high proportion of non-terminated, long *globin* mRNA causing thalassaemia.^{37–38} Our data suggested similarly that the shorter *GIMAP5* transcript is the predominant mRNA species, but that the A→G substitution causes an inefficient polyA site and results in a longer transcript, as hypothesised and shown previously.^{37–38} Our data also support the idea that different polyA sites in the same gene might react differently on environmental factors in a tissue-, differentiation- or development-specific way.^{36–39} Thus, taken together, we suggest that the alternative polyA sites created by the disease-associated, risk haplotype-tagging SNP may be functionally relevant for disease susceptibility in SLE.

An *in vivo* role of *GIMAP5* is supported by recent observations in the BB rat. In this animal model, a frameshift mutation creates a stop codon that contributes to autoimmune diabetes type 1 with lymphopenia.^{11–12} The actual effect of the mutation on the immune cells of the BB rat remains unsolved, but the strongly decreased expression of mRNA for *rGimap5*¹¹ suggest that the cells either lack the rGimap5 protein completely or produce a truncated protein. Also, using transgenic rescue in lymphopenic BB-derived congenic F344.lyp/lyp rats, Michalkiewicz et al have demonstrated that the Gimap5 protein is essential for maintaining normal T-cell levels.¹³

Sandal et al suggested that the antiapoptotic action of rGimap5 is executed through its central coiled-coil-containing domain, responsible for centrosomal anchoring and resistance towards death, and that if expressed in the BB rat, the mutated rGimap5 will be nonfunctional for this apoptosis antagonism.²⁰ They show that Jurkat T-cells transfected with antisense-*GIMAP5* cDNA became hypersensitive to general apoptosis inducers (γ -radiation and okadaic acid), suggesting that the endogenous level of *GIMAP5* in Jurkat cells normally protect against cell death.²⁰ The members of the GIMAP family of antiapoptotic proteins have been reported to protect against a broad spectrum of apoptogens (including anti-CD95 (Fas/

Apo-1), TNF- α , staurosporine and etoposide),⁴⁷ but the findings of Sandal et al contrast this.²⁰

We examined what effect several pleiotropic cytokines commonly implicated in SLE pathogenesis had on *GIMAP5* expression and saw a consistent, significant upregulation of *GIMAP5* in two monocyte cell lines. IFN- γ , IL-1 β and TNF- α are all proinflammatory cytokines that are suggested to play a part in apoptosis.^{48–49} TGF- β and FGF2 are growth factors and disturbances in their activation and signalling may result in apoptosis. This supports the notion that *GIMAP5* is affected by cytokines important in the immune system and for the development of autoimmune disease, and is upregulated in response to cytokines known to be elevated in SLE, perhaps to protect against apoptosis. Our identified, potentially functional polyA variation is more common, not fully penetrant and would subsequently be expected to have a milder effect on the *GIMAP5* antiapoptotic effects than the mutation described in the BB rat. It is also possible that the genetic variation in *GIMAP5* predisposing to SLE susceptibility is regulating its action in a tissue- or cell-specific manner.

In summary, by analysing five independent family base samples with over 2000 individuals, we have found a significant (OR 1.26, $p=0.0033$) increase in the risk of SLE associated with a common *GIMAP5* risk haplotype. We also confirm a functional variant on the SLE risk haplotype, which introduces a non-optimal polyA site as a possible susceptibility mechanism. Our data imply that inherited variation in *GIMAP5* is a susceptibility factor for SLE.

ELECTRONIC DATABASE INFORMATION

HapMap, public data release #11 September 2004 (<http://www.hapmap.org/>); NCBI (<http://www.ncbi.nlm.nih.gov/>); Online Mendelian Inheritance in Man (OMIM) (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM>); Primer3 (http://frodo.wi.mit.edu/primer3/primer3_code.html); Saccharomyces Genome Database (<http://seq.yeastgenome.org/cgi-bin/SGD/web-primer>); "R" (<http://www.r-project.org/>).

ACKNOWLEDGEMENTS

We thank Linda Berglind and Ville-Veikko Mäkelä for their contributions in genotyping the samples. We thank all patients and family members participating in the Finnish and UK SLE studies. This study was supported by the Sigrid Jusélius Foundation, the Academy of Finland and the Swedish Research Council (to JK and US-K); the Ulla & Gustaf af Ugglas Foundation and the Alex & Eva Wallströms Foundation and travel grants from the Royal Physiographic Society and the Novartis Foundation (to CML); the Edvard Welanders Foundation and the Finsen Foundation (to TS); and a Wellcome Trust Senior Fellowship (to TJJ, supporting AW and DCG).

Authors' affiliations

Anna Hellquist*, **Marco Zucchelli***, **Katja Kivinen**, **Tiina Skoog**, **Juha Kere**, **Cecilia M Lindgren**, Department of Biosciences at Novum, Karolinska Institute, Stockholm, Sweden

Anna Hellquist, **Tiina Skoog**, **Juha Kere**, Clinical Research Centre, Karolinska University Hospital, Huddinge, Sweden

Ulpu Saarialho-Kere, **Sari Koskenmies**, Department of Dermatology, University of Helsinki, Helsinki, Finland

Saarialho-Kere, Department of Dermatology, Karolinska Institutet at Stockholm Söder Hospital, Stockholm, Sweden

Sari Koskenmies, **Johanna Vendelin**, **Juha Kere**, Department of Medical Genetics, University of Helsinki, Finland

Elisabeth Widén, Finnish Genome Center, University of Helsinki, Helsinki, Finland

Heikki Julkunen, Department of Medicine, Peijas Hospital and Helsinki University Hospital, Helsinki, Finland

Andrew Wong, **Deborah S Cunninghame-Graham**, **Timothy J Vyse**, Rheumatology Section, Imperial College, Hammersmith Hospital, London, UK

Marja-Liisa Karjalainen-Lindsberg, Department of Pathology, Helsinki University Central Hospital, Helsinki, Finland

Competing interests: None.

*These authors contributed equally to the manuscript

REFERENCES

- Koskenmies S, Widen E, Kere J, et al. Familial systemic lupus erythematosus in Finland. *J Rheumatol* 2001;**28**:758–60.
- Prokunina L, Castillejo-Lopez C, Oberg F, et al. A regulatory polymorphism in PDCD1 is associated with susceptibility to systemic lupus erythematosus in humans. *Nat Genet* 2002;**32**:666–9.
- Nath SK, Kelly JA, Harley JB, et al. Mapping the systematic lupus erythematosus susceptibility genes. *Methods Mol Med* 2004;**102**:11–29.
- Tsao BP. Update on human systemic lupus erythematosus genetics. *Curr Opin Rheumatol* 2004;**16**:513–21.
- Lee YH, Nath SK. Systemic lupus erythematosus susceptibility loci defined by genome scan meta-analysis. *Hum Genet* 2005;**118**:434–43.
- Hirschhorn JN, Altschuler D. Once and again—issues surrounding replication in genetic association studies. *J Clin Endocrinol Metab* 2002;**87**:4438–41.
- Marleau AM, Sarvetnick N. T cell homeostasis in tolerance and immunity. *J Leukoc Biol* 2005;**78**:575–84.
- Kaaba SA, Al-Harbi SA. Abnormal lymphocyte subsets in Kuwaiti patients with type-1 insulin-dependent diabetes mellitus and their first-degree relatives. *Immunol Lett* 1995;**47**:209–13.
- Mandl T, Bredberg A, Jacobsson LT, et al. CD4+ T-lymphocytopenia—a frequent finding in anti-SSA antibody seropositive patients with primary Sjogren's syndrome. *J Rheumatol* 2004;**31**:726–8.
- Mirzayan MJ, Schmidt RE, Witte T. Prognostic parameters for flare in systemic lupus erythematosus. *Rheumatology (Oxford)* 2000;**39**:1316–9.
- Hornum L, Romer J, Markholst H. The diabetes-prone BB rat carries a frameshift mutation in *Ian4*, a positional candidate of *Iddm1*. *Diabetes* 2002;**51**:1972–9.
- MacMurray AJ, Moralejo DH, Kwitek AE, et al. Lymphopenia in the BB rat model of type 1 diabetes is due to a mutation in a novel immune-associated nucleotide (*Ian*)-related gene. *Genome Res* 2002;**12**:1029–39.
- Michalkiewicz M, Michalkiewicz T, Ettinger RA, et al. Transgenic rescue demonstrates involvement of the *Ian5* gene in T cell development in the rat. *Physiol Genomics* 2004;**19**:228–32.
- Reuber TL, Ausubel FM. Isolation of *Arabidopsis* genes that differentiate between resistance responses mediated by the RPS2 and RPM1 disease resistance genes. *Plant Cell* 1996;**8**:241–9.
- Krucken J, Schmitt-Wrede HP, Markmann-Mulisch U, et al. Novel gene expressed in spleen cells mediating acquired testosterone-resistant immunity to *Plasmodium chabaudi* malaria. *Biochem Biophys Res Commun* 1997;**230**:167–70.
- Stamm O, Krucken J, Schmitt-Wrede HP, et al. Human ortholog to mouse gene *imap38* encoding an ER-localizable G-protein belongs to a gene family clustered on chromosome 7q32–36. *Gene* 2002;**282**:159–67.
- Cambot M, Aresta S, Kahn-Perles B, et al. Human immune associated nucleotide 1: a member of a new guanosine triphosphatase family expressed in resting T and B cells. *Blood* 2002;**99**:3293–301.
- Zenz T, Roessner A, Thomas A, et al. *Ilan5*: the human ortholog to the rat *Ian4/Iddm1/lyp* is a new member of the *Ian* family that is overexpressed in B-cell lymphoid malignancies. *Genes Immun* 2004;**5**:109–16.
- Pandarpurkar M, Wilson-Fritch L, Corvera S, et al. *Ian4* is required for mitochondrial integrity and T cell survival. *Proc Natl Acad Sci U S A* 2003;**100**:10382–7.
- Sandal T, Aumo L, Hedin L, et al. *Irod/Ian5*: an inhibitor of gamma-radiation- and okadaic acid-induced apoptosis. *Mol Biol Cell* 2003;**14**:3292–304.
- Tan EM, Cohen AS, Fries JF, et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1982;**25**:1271–7.
- Helve T. Prevalence and mortality rates of systemic lupus erythematosus and causes of death in SLE patients in Finland. *Scand J Rheumatol* 1985;**14**:43–6.
- Russell AI, Cunninghame Graham DS, Shepherd C, et al. Polymorphism at the C-reactive protein locus influences gene expression and predisposes to systemic lupus erythematosus. *Hum Mol Genet* 2004;**13**:137–47.
- Payne F, Smyth DJ, Pask R, et al. Haplotype tag single nucleotide polymorphism analysis of the human orthologues of the rat type 1 diabetes genes *Ian4* (*lyp/Iddm1*) and *Cblb*. *Diabetes* 2004;**53**:505–9.
- Jurinke C, van den Boom D, Cantor CR, et al. Automated genotyping using the DNA MassArray technology. *Methods Mol Biol* 2002;**187**:179–92.
- O'Connell JR, Weeks DE. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet* 1998;**63**:259–66.
- Pastinen T, Sladek R, Gurd S, et al. A survey of genetic and epigenetic variation affecting human gene expression. *Physiol Genomics* 2004;**16**:184–93.
- Gabriel SB, Schaffner SF, Nguyen H, et al. The structure of haplotype blocks in the human genome. *Science* 2002;**296**:2225–9.
- Becker T, Knapp M. Maximum-likelihood estimation of haplotype frequencies in nuclear families. *Genet Epidemiol* 2004;**27**:21–32.
- Spielman RS, McGinnis RE, Ewens WJ. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 1993;**52**:506–16.
- Kruglyak L, Daly MJ, Reeve-Daly MP, et al. Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 1996;**58**:1347–63.
- Boehnke M, Langeveld CD. Genetic association mapping based on discordant sib pairs: the discordant-alleles test. *Am J Hum Genet* 1998;**62**:950–61.
- Lohmueller KE, Pearce CL, Pike M, et al. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet* 2003;**33**:177–82.
- Andersen UN, Markholst H, Hornum L. The antiapoptotic gene *Ian41* in the rat: genomic organization and promoter characterization. *Gene* 2004;**341**:141–8.
- Krucken J, Schroetel RM, Muller IU, et al. Comparative analysis of the human *gimap* gene cluster encoding a novel GTPase family. *Gene* 2004;**341**:291–304.
- Beaudouin E, Gautheret D. Identification of alternate polyadenylation sites and analysis of their tissue distribution using EST data. *Genome Res* 2001;**11**:1520–6.
- van Solinge WW, Lind B, van Wijk R, et al. Clinical expression of a rare beta-globin gene mutation co-inherited with haemoglobin E-disease. *Eur J Clin Chem Clin Biochem* 1996;**34**:949–54.
- Jankovic L, Eremov GD, Petkov G, et al. Two novel polyadenylation mutations leading to beta(+)-thalassaemia. *Br J Haematol* 1990;**75**:122–6.
- Erdal G, Veraldi KL, Milcarek C. Alternative poly(A) site selection in complex transcription units: means to an end? *Nucleic Acids Res* 1997;**25**:2547–61.
- Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1997;**40**:1725.
- Kao AH, Manzi S, Ramsey-Goldman R. Review of ACR hematologic criteria in systemic lupus erythematosus. *Lupus* 2004;**13**:865–8.
- Corder EH, Saunders AM, Strittmatter WJ, et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* 1993;**261**:921–3.
- Altschuler D, Hirschhorn JN, Klannemark M, et al. The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat Genet* 2000;**26**:76–80.
- Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 2005;**6**:95–108.
- Gaffney PM, Ortmann WA, Selby SA, et al. Genome screening in human systemic lupus erythematosus: results from a second Minnesota cohort and combined analyses of 187 sib-pair families. *Am J Hum Genet* 2000;**66**:547–56.
- Baechler EC, Gregersen PK, Behrens TW. The emerging role of interferon in human systemic lupus erythematosus. *Curr Opin Immunol* 2004;**16**:801–7.
- Deveraux QL, Reed JC. IAP family proteins—suppressors of apoptosis. *Genes Dev* 1999;**13**:239–52.
- Firme L, Bush AB. FGF signaling inhibits the proliferation of human myeloma cells and reduces c-myc expression. *BMC Cell Biol* 2003;**4**:17.
- Davidson A, Diamond B. Autoimmune diseases. *N Engl J Med* 2001;**345**:340–50.

Keep up to date: sign up for our alerting services

Find out automatically when an article is published on a specific topic or by a particular author. We can also alert you when an article is cited or if an eletter or correction is published. You can also choose to be alerted when a new issue is published online [and when we post articles Online First]. Check out the New Content Alerts and Citation tracker from the Online tools section on the home page.