Genetic variations in GPSM3 associated with protection from rheumatoid arthritis affect its transcript abundance

BJ Gall1, A Wilson2, AB Schroer1, JD Gross1, P Stoilov3, V Setola1,4, CM Watkins2 and DP Siderovski1

G protein signaling modulator 3 (GPSM3) is a regulator of G protein-coupled receptor signaling, with expression restricted to leukocytes and lymphoid organs. Previous genome-wide association studies have highlighted single-nucleotide polymorphisms (SNPs; rs204989 and rs204991) in a region upstream of the GPSM3 transcription start site as being inversely correlated to the prevalence of rheumatoid arthritis (RA)—this association is supported by the protection afforded to Gpsm3-deficient mice in models of inflammatory arthritis. Here, we assessed the functional consequences of these polymorphisms. We collected biospecimens from 50 volunteers with RA diagnoses, 50 RA-free volunteers matched to the aforementioned group and 100 unmatched healthy young volunteers. We genotyped these individuals for GPSM3 (rs204989, rs204991), CCL21 (rs2812378) and HLA gene region (rs6457620) polymorphisms, and found no significant differences in minor allele frequencies between the RA and disease-free cohorts. However, we identified that individuals homozygous for SNPs rs204989 and rs204991 had decreased GPSM3 transcript abundance relative to individuals homozygous for the major allele. In vitro promoter activity studies suggest that SNP rs204989 is the primary cause of this decrease in transcript levels. Knockdown of GPSM3 in THP-1 cells, a human monocytic cell line, was found to disrupt ex vivo migration to the chemokine MCP-1.

Histologic analyses of affected synovial tissues indicate a relative acellularity of synovial cavities in Gpsm3-deficient mice compared with wild-type mice, which suggests a potential leukocyte migration deficit. These findings suggest a role for GPSM3 in the development of inflammation in rheumatoid arthritis (RA), most likely through a key regulatory role on leukocyte chemotaxis to the synovium. However, a direct demonstration of such a role has yet to be exhibited with human cells.

Genome-wide association studies (GWAS) have identified two single-nucleotide polymorphisms (SNPs), rs204989 and rs204991, to the GPSM3 transcriptional start site that are significantly less prevalent in individuals with RA (and other autoimmune diseases; for example, lupus and multiple sclerosis), implying a potential protective effect of these minor gene variations. Notably, rs204989 and rs204991 both exist within the chromosome 6p21.3 band, a region previously attributed to a large portion of the heritable aspect of RA; specifically, HLA gene region polyallelic haploblocks within the chromosome 6p21.3 region represent some of the greatest risk factors for RA (reviewed in Holoshitz). In particular, the biallelic HLA gene locus polymorphism, rs6457620 [C→G], has been identified as an RA risk factor in a meta-analysis of GWAS studies investigating multiple populations in the Wellcome Trust Case Control Consortium, North American Rheumatoid Arthritis Consortium and the Swedish Epidemiological Investigation of Rheumatoid Arthritis. Thus, the potential exists for linkage disequilibrium between GPSM3 and HLA gene region polymorphisms. In this study, we addressed whether GPSM3 SNPs result in a detectable phenotype that explains their...
inversely associated with RA. Furthermore, we assessed whether linkage disequilibrium with the known RA risk allele in the HLA gene region, rs6457620, may affect the inverse association of GPSM3 SNPs with RA. In addition, another RA risk allele, CCL21 rs2812378 [T > C], located on an unlinked chromosome, was analyzed as both a negative control for linkage and a positive control for RA disease risk. We recruited a group of 50 volunteers with a diagnosis of RA, 50 RA-free volunteers who matched to the aforementioned group by a ‘Bring-a-friend-to-RA’ program, and 50 healthy young volunteers to donate biospecimens for analyses. Based on the location of the polymorphisms and previous reports of protection from inflammatory phenotypes in human GWAS and Gpms3-deficient mouse studies, we hypothesized that individuals homozygous for the minor alleles of rs204989 and rs204991 would exhibit decreased whole blood GPSM3 transcript abundance. In addition, we predicted that knockdown of GPSM3 would result in disruption of chemokine-induced migration in a human monocytic cell line.

**RESULTS**

GPSM3 SNPs rs204989 and rs204991, each previously associated with GWAS with protection from RA, form a haploblock with rs204990. The cohorts recruited for this study included an initial set of 100 unmatched healthy young volunteers, a group of 50 volunteers with a positive diagnosis of RA, and 50 RA-free volunteers matched to the aforementioned group by a ‘Bring-a-friend-to-RA’ program. Upon genotyping all 200 volunteers recruited for this study, we found that GPSM3 SNPs rs204989 and rs204991, originally identified to be independently associated with protection from RA, are in complete linkage disequilibrium within this population. In addition, sequencing a 3.5-kb region 5’ to the GPSM3 transcriptional start site in eight volunteers revealed a total of four polymorphisms in this region: rs204989, rs204990, rs204991, and rs3096688 (Figure 1a). All of these chromosome 6 SNPs have been previously identified in the NCBI Database of Genes and Immunity (2016) 139 – 147 © 2016 Macmillan Publishers Limited
Single-Nucleotide Polymorphisms (dbSNPs). In this study, three of these GPSM3 SNPs (rs204989, rs204990 and rs204991) were seen to be inherited as a haplblock in complete linkage disequilibrium; therefore, we defined all individuals homozygous for the minor (lower frequency) alleles of all three SNPs as having the ‘m/m’ genotype and all individuals homozygous for the major (higher frequency) allele for all three SNPs as having the ‘M/M’ genotype (Figure 1a).

The GPSM3 rs204989/rs204991 haplblock is associated with decreased GPSM3 mRNA expression in whole blood. In disease-free homozygotes for the minor alleles of GPSM3 SNPs, whole blood-derived RNA contained only 75.9% (m/m; n = 11; 95% confidence interval (CI) ± 7.6%) of the average GPSM3 transcript level present in homozygotes for the major alleles (M/M; n = 53; 95% CI ± 5.6%). This was a decrease of 24.1% (95% CI 7.1–41.7%; t = 3.803; df = 62; P = 0.0003) for individuals with the m/m genotype (n = 11) relative to individuals with the M/M genotype (Figures 1b; n = 53). Heterozygous individuals (M/m genotype) exhibited an average of 82.0% of the GPSM3 transcript abundance of the M/M genotype, but had a marked amount of variability (data not shown; n = 6; 95% CI ± 31.1%). In our analyses of GPSM3 mRNA abundance, we excluded RA patient samples to avoid potential confounding effects caused by anti-RA pharmacotherapy, as shown previously by RNA-seq.26 According to the 1000 Genomes Project,28 the minor allele frequencies (MAFs) in the general population of rs204989 and rs204991 are 23.3% and 23.4%, respectively. In our volunteer populations, the chromosome 6 rs204989 and rs204991 SNPs occurred in a haplblock with a MAF of 29% within the unmatched healthy young cohort (n = 100), 23% within the RA cohort (n = 50) and 18% within the matched control (disease-free) cohort (n = 50). The MAF of the chromosome 9 CCL21 gene locus polymorphism rs2812378 (T > C) is 30.1% according to the 1000 Genomes Project.28 In our recruited populations, the CCL21 gene polymorphism MAF was 31.5% within the unmatched healthy young population (n = 100), 30.0% within the RA cohort (n = 50) and 25.0% within the matched control cohort (n = 50). The MAF of the chromosome 6 HLA gene region polymorphism rs6457620 was 56.0% within the unmatched healthy young cohort (n = 100), 35.0% within the RA cohort (n = 50) and 45.0% within the matched control population (n = 50). None of the differences in GPSM3 (P = 0.4839), CCL21 (P = 0.5267) and HLA gene region (P = 0.1938) polymorphism MAFs between the RA and matched, disease-free cohorts was statistically significant (Figure 1c).

The GPSM3 SNP haplblock is in weak linkage disequilibrium with HLA SNP rs6457620, but the latter is unlinked to GPSM3 transcript abundance. Both the gene for GPSM3 and the HLA gene region SNP rs6457620 reside on chromosome 6, presenting the possibility of linkage disequilibrium; however, over 0.5 megabases of DNA sequence separate rs6457620 from the GPSM3 rs204989/rs204991 SNP haplblock.29 In individuals homozygous for the GPSM3 SNPs (M/M), the HLA gene region SNP MAF was 41.3% (Figure 2). In heterozygotes for the GPSM3 SNPs (M/m), the HLA gene region SNP MAF was 56.3%, whereas homozygotes for the minor alleles of GPSM3 SNPs (m/m) exhibited a 60.7% HLA gene region SNP MAF. This distribution was significantly different from chance distribution of unlinked loci (Figure 2a; P = 0.0066), suggesting the presence of weak linkage disequilibrium and consistent with data from the NCBI dbSNP 25 (Figure 2d). Within the RA cohort, there was significant variation from chance distribution of GPSM3 SNPs MAF when stratified by HLA gene region SNP genotype (Figure 2b; P = 0.0123), which was not observed within the matched control group (Figure 2c; P = 0.2739). However, when stratified by HLA SNP rs6457620 genotype, GPSM3 transcript abundance in whole blood was not significantly different (Figure 2e; F = 1.219; P = 0.3050); this result is in contrast to the significant difference in GPSM3 transcript abundance observed when stratified by GPSM3 SNP haplblock genotype (Figure 1b).

rs204989 of the GPSM3 SNP haplblock is solely linked to decreased GPSM3 transcript abundance. A 3.5-kb region 5′ to the human GPSM3 transcription start site, spanning the region of the GPSM3 SNP haplblock under study (Figure 1a), was subcloned into a promoterless luciferase-reporter vector. Transfecting this construct into HEK293T cells, we found the normalized firefly luciferase activity induced by the ‘minor/polymorphic’ GPSM3 promoter (pGL3-m; containing the minor rs204989/rs204990/rs204991 haplblock as well as the minor allele of SNP rs3096688) was 26.5% less than that of the wild-type promoter (pGL3-M; containing the major allele of all four SNPs) (P < 0.0001; Figure 3a). To assess the relative effect(s) of each of the four minor SNP variants within this GPSM3 promoter region, each individual polymorphism was separately introduced to the wild-type promoter by site-directed mutagenesis (that is, rs204989 [C > T], rs204990 [G > T], rs204991 [A > G] and rs3096688 [A > G]). Introduction of the minor allele of rs204990 (P = 0.1435), rs204991 (P = 0.1316) or rs3096688 (P = 0.6618) did not significantly alter the normalized firefly luciferase activity of the wild-type promoter (Figure 3a). In contrast, introduction of the minor allele of rs204989 [C > T] into the wild-type promoter significantly decreased promoter-induced luminescence (P < 0.0001; Figure 3a); conversely, introduction of the major allele of rs204989 [T > C] into the polymorphic promoter (‘pGL3-m-989’) was found to restore wild-type promoter-induced luminescence (P = 0.9384; Figure 3b). Using the JASPAR transcription factor binding profile database,29 two potential transcription factor binding sites were found to overlap with the site of the rs204989 SNP: namely, AP1 and C/EBPβ consensus binding sites (Figure 4).

Decreased GPSM3 abundance leads to a deficit in migration toward monocyte chemoattractant protein-1 (MCP-1) by the human monocytic THP-1 cell line. The human monocytic THP-1 cell line has previously been shown to endogenously express GPSM3;36 we confirmed the presence of the GPSM3 ancestral (‘major’) alleles of rs204989 and rs204991 within this cell line (data not shown). We established stable GPSM3 knockdown and scrambled control THP-1 cell lines (Figure 5a), using previously validated short hairpin RNA (shRNA)-expressing lentiviral vectors directed against GPSM3 (that is, shRNA19 and shRNA20).30 Using these stable THP-1 cell lines in real-time Transwell plate migration assays, we measured vehicle-normalized relative fluorescence units over a 30-min timeframe to quantify MCP-1/CCL2-induced cell transmigration. After normalizing to vehicle-induced nonspecific migration, the GPSM3 knockdown cell lines shRNA19 and shRNA20 exhibited disrupted migration to the chemokine MCP-1 (a.k.a. CCL2) relative to scrambled shRNA control lines (Figure 5b); these migration differences were independent of any changes in transcript abundance of the MCP-1 receptor (that is, CCR2) between cell lines assayed (Figure 5c; F = 2.40; P = 0.119). This migratory deficit upon GPSM3 knockdown was observed both by a decrease in initial rate of migration (Figure 5d) and by a decrease in maximum transmigration-induced fluorescence achieved (Figure 5e).

**DISCUSSION**

GWAS have empowered researchers to investigate complex non-Mendelian genetic disorders, including RA; however, these studies can be hampered by the identification of large numbers of noncoding variants with unknown effects, reproducibility across multiple populations being uncertain and linkage with other risk
alleles obscuring results. In multiple GWAS reports, noncoding variants 5' to the transcription start site of GPSM3 (Figure 1a) have been identified as occurring less frequently in individuals with RA. However, like many identified variants, no functional consequence has been elucidated to explain this association. In this study, we addressed (1) genotype-dependent phenotypes of the identified chromosome 6 GPSM3 SNPs, (2) potential linkage with rs6457620, a biallelic RA risk allele with a large effect size in the HLA gene region on chromosome 6 (refs. 23,24) and (3) a functional effect of altered GPSM3 expression levels in a human monocytic cell line.

We recruited 50 volunteers with confirmed RA diagnoses, 50 RA-free volunteers matched to the arthritis group and 100 younger healthy individuals to analyze genotype-dependent functional consequences of GPSM3 SNP alleles. Our finding that homozygotes for the GPSM3 SNPs (m/m) have lower GPSM3 transcript abundance in whole blood relative to homozygotes for the ancestral allele (M/M) (Figure 1b) supports the hypothesis first put forth from GWAS findings that there is a potential functional consequence of the identified GPSM3 SNPs rs204989 and rs204991, namely protection from development of RA; we have previously demonstrated that lowered Gpsm3 expression in mice, as generated by gene knockout technology, results in protection from collagen autoantibody-induced inflammatory arthritis.

Although our present finding of lowered GPSM3 mRNA levels in the whole blood of homozygotes bearing the GPSM3 SNP haplblock is consistent with GWAS and mouse gene knockout studies, the observed frequencies of the rs204989/rs204991 haplblock within the various genotyped groups are seemingly discrepant, given the trend toward decreased GPSM3 MAF, albeit not significantly different, in the disease-free matched control

| a | HLA SNP (rs6457620) | Homozygous Ancestral GPSM3 (M/M; n = 155) | Heterozygous GPSM3 (M/m; n = 71) | Homozygous SNP GPSM3 (m/m; n = 14) | Significance levels of observed differences |
|---|-------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| C/C | 41 | 53 | 21 | 10 | 42 | 19 | 3 | 5 | 6 |
| C/G | 14 | 16 | 6 | 2 | 13 | 1 | 1 | 2 | 1 |
| G/G | 7 | 7 | 3 | 0 | 1 | 2 | 1 | 2 | 1 |

Figure 2. Linkage of the GPSM3 SNP haplblock and the known RA risk allele in the HLA gene region, rs6457620 [C > G], is more pronounced in RA samples than controls, but this linkage does not result in a rs6457620 genotype-dependent effect on GPSM3 transcript abundance in whole blood. (a) Descriptive statistics of all 200 volunteers, including the RA (n = 50), disease-free matched control (n = 50) and young healthy control volunteers recruited for this study. These data are stratified by GPSM3 SNP haplotype, showing that the rs6457620 [G] risk allele frequency amongst homozygous ancestral GPSM3 haplblock individuals (M/M) is 41.3%, heterozygous individuals (M/m) is 56.3%, and homozygous minor GPSM3 haplblock (m/m) is 60.7%. These data are significantly different from chance distribution (Fisher–Freeman–Halton exact test; P = 0.0066). (b) Analyses of GPSM3 SNP haplblock linkage within RA samples stratified by rs6457620 genotype, again exhibiting statistically significant difference from chance distribution (Fisher–Freeman–Halton exact test; P = 0.0123). (c) Within the disease-free matched control samples, GPSM3 SNP minor allele (m) frequency, as stratified by rs6457620 genotype, is not significantly different from chance distribution as determined by Fisher–Freeman–Halton exact test (P = 0.2739). (d) Graphical representation of MAFs (of European [eu] population), physical distance (in basepairs), and linkage disequilibrium (LD) value (as quantified by heatmap) between the three linked SNPs of the GPSM3 SNP haplblock (rs204989, rs204990 and rs204991) and the HLA gene region SNP rs6547620, as obtained using the NIHES SNPinfo webserver (http://snpinfo.nih.gov/) from NCBI dbSNP data. (e) Despite modest linkage between GPSM3 SNP minor alleles (m) and the rs6457620 risk allele (G), there is no effect of the genotype at HLA gene region SNP rs6457620 on GPSM3 transcript abundance. Significance determined by one-way analysis of variance (ANOVA) with Bonferroni post-hoc.
isons with Dunnett analysis of variance (ANOVA) after controlling for multiple comparisons. Significance was determined using one-way ANOVA, and p values are reported as means ± s.e.m. (that is, ‘′p<0.0001′′). Introduction of independent minor alleles to the average level measured from the wild-type pGL3-M vector (set to 1.00; dotted line). Introduction of independent minor alleles (denoted ‘′<last 3 digits of SNP rs#>′′) into the wild-type pGL3-M vector is seen to differentially affect resultant firefly luciferase activity. (b) Restoration of the major allele at rs204989 in the pGL3-m vector (that is, ‘′pGL-m-989′′) restores wild-type pGL3-M promoter activity. All data are compiled from three independent experiments. Error measure is s.e.m. Significance was determined using one-way analysis of variance (ANOVA) after controlling for multiple comparisons with Dunnett’s test using pGL3-M as the defined control.

(MAF = 18.0%; n = 50) compared with the RA cohort (MAF = 23.0%; n = 50). Considering the small sample size, the observed variability is not remarkable in light of the reported MAF in the general population being 23.3%. Small sample size likely also affected our ability to assign statistical significance to the trend of increased frequency of the positive RA risk allele (Figure 1c). CCL21 gene rs2812378 [T>C] on chromosome 9 (that is, unrelated to the human GPSM3 locus on chromosome 6).24

To address whether the GPSM3 SNP haploblock on chromosome 6 was serving as a surrogate marker for a linked and causal (yet unrelated) chromosome 6 variation affecting GPSM3 mRNA abundance and/or RA disease status, additional genotyping was performed on another chromosome 6 biallelic allele, specifically near the well-established risk region in the HLA gene region (reviewed by Holoshitz22 and Yuta et al.23). This RA risk allele, rs6457620 [C>G],23,24 is reported by RegulomeDB23 (based on Hapmap2 metadata) to exist in complete linkage disequilibrium (D′ = 1.0; r2 = 1.0) with another HLA region RA risk allele rs6457617 [C>T]24,25 in a population of Europeans with American ancestry (CEPH collection)26 that is very similar to the demographics of our West Virginia-based sample population. Both polymorphisms (rs6457617 and rs6457620) exist ~500 kb from the GPSM3 SNP haploblock (for example, Figure 2d). When stratifying the HLA region-associated SNP genotype by GPSM3 SNP genotype, a preferential coexistence is seen between the GPSM3 SNP haploblock minor allele and the minor allele (G) of rs6457620 (Figure 2). Upon further stratification by disease state, we discovered the GPSM3 SNP alleles were more strongly associated with the minor allele of rs6457620 (G) in our RA cohort when compared with the matched control cohort (Figures 2b vs c). These findings suggest that the protective effect of the GPSM3 SNP haploblock minor allele is most likely masked in our small-sized cohort by the large effect size of the HLA gene region RA risk allele at rs6457620 (RA odds ratio of 2.55)27 or other RA risk-related phenotypes and/or a founder effect among the local recruited volunteers.37 Regardless, our finding of a genotype-dependent decrease in GPSM3 mRNA abundance is supportive of the prior GWAS findings,18–20 especially in light of the protection from inflammatory arthritis exhibited by Gpsm3-deficient mice.15

To probe the molecular mechanism whereby presence of the minor allele of the rs204989/rs204991 haploblock leads to decreased GPSM3 transcript levels in whole blood, a 3.5-kb

(C)
promoter region containing the haplotype region was tested for its promoter activity upon transfection into HEK293T cells (used given low transient transfection efficiency of the THP-1 cell line). There was a significant decrease in promoter activity in the reporter construct containing all four 5’ SNPs versus the ancestral sequence, in a direct parallel to decreased GPSM3 transcript abundance seen in whole blood samples from individuals homozygous for the minor rs204989/rs204991 haplotype versus the ancestral (‘major’) haplotype. The presence of rs204989 (C>T) was found to be the sole determinant responsible for the reduction in GPSM3 promoter activity. We did observe slight variability in the decrease in GPSM3 promoter-driven luciferase activity between experiments, most likely due to experimental variability, environmental factors (for example, temperature differentially affecting firefly and Renilla luciferase activity), and elapsed time between experiments. Regardless, all constructs containing the minor allele of rs204989 (C>T) were not significantly different from the construct with all 4 GPSM3 SNP alleles. To our knowledge, this is the first reported functional connection between the GPSM3 rs204989 polymorphism and GPSM3 transcript abundance in human cells and human samples. The singular effect of the rs204989 cytosine-to-thymine transition on GPSM3 promoter activity may reside in the disruption of the consensus DNA-binding motif of AP1 and/or C/EBPβ; this hypothesis is now the topic of future work.

A potential functional consequence of this genotype-dependent decrease in GPSM3 abundance is dysregulated GPCR signaling via Gia6-10—a common signaling system used by chemokine receptors. Considering the restricted expression of GPSM3 in leukocytes and lymphoid organs15,16 and the importance of of Gia-coupled chemokine receptors in leukocyte migration during RA pathogenesis38, we hypothesized that a reduction in GPSM3 expression in human monocytic THP-1 cells would result in perturbed chemokine-dependent chemotaxis. Given the role of MCP-1/CCR2 signaling in the recruitment of monocytes and macrophages during inflammatory bouts in RA patients,39,40 we elected to examine MCP-1-induced chemotaxis. shRNA-induced knockdown of GPSM3 in THP-1 cells disrupted Transwell migration (both initial rate and maximal migration) toward MCP-1, but did not affect expression of its cognate receptor, C-C chemokine receptor 2 (CCR2) (Figure 5). Taken together, these data support a role for GPSM3 in supporting MCP-1-induced migration of human THP-1 cells. We note that the knockdown efficiency of the successful GPSM3-targeted shRNAs was significantly greater than the 24.1% decrease found in individuals homozygous for the minor alleles of the GPSM3 SNP haplotype. Therefore, we hypothesize that the functional effect of the rs204989 GPSM3 SNP in vivo is most likely more subtle, yet it likely has a role in the pathogenesis of RA as a component of a complex genetic framework that is, to date, still being elucidated.

Based on our collective results, the GPSM3 SNPs rs204989 and rs204991, identified as protective against RA in prior GWAS studies, likely exhibit this protective effect via linkage to rs204989 and its alteration of GPSM3 transcription. An rs204989-mediated
decrease in GPSM3 transcript level likely decreases GPSM3 protein levels and thereby decreases chemokine-dependent leukocyte migration in inflammation. RNA-seq ‘transcriptomic’ data sets report the GPSM3 mRNA as being expressed at an appreciable level in all human leukocyte subsets examined ( GEO accession GDS62408), suggesting that decreased GPSM3 mRNA transcript caused by SNP rs204989 may affect both the innate and adaptive immune responses that act sequentially in the pathogenesis of RA.28 Future work will initially address whether the presence of rs204989 is functionally relevant in primary leukocytes isolated from patient samples or in CRISPR-mediated introduction of rs204989 in human leukocyte cell lines. In addition, we seek to establish whether one or both of the candidate transcription factors (that is, AP1 or C/EBPβ) are indeed involved in GPSM3 promoter regulation and whether such activity is directly affected by the major or minor allele of SNP rs204989.

MATERIALS AND METHODS

Subjects

A WVU IRB-approved case–control study (IRB protocol #1304033165) recruited a total of 200 consenting local volunteers for buccal swabs and blood samples. The volunteer population was subdivided into three groups of 100 young, healthy control volunteers from among the WVU Health Sciences Center student population, 50 RA patients (I-9 diagnosed and anti-RF-positive) and 50 demographically matched controls. Inclusion criteria for the 100 healthy control volunteers included (1) at least 18 years of age, (2) self-reported good health and (3) no family history of autoimmune disease. Inclusion criteria for the 50 RA individuals included (1) at least 18 years of age, (2) positive ICD-9 diagnosis and (3) rheumatoid autoimmune disease. Inclusion criteria for the 50 RA individuals included (1) at least 18 years of age, (2) positive ICD-9 diagnosis and (3) rheumatoid autoimmune disease. Inclusion criteria for the 50 RA individuals included (1) at least 18 years of age, (2) positive ICD-9 diagnosis and (3) rheumatoid autoimmune disease. Inclusion criteria for the 50 RA individuals included (1) at least 18 years of age, (2) positive ICD-9 diagnosis and (3) rheumatoid autoimmune disease.

Isolation of genomic DNA (gDNA) from buccal swab samples was performed using a QiAamp DNA Blood Mini Kit (Qiagen, Germantown, MD, USA) on a QiAcube automated nucleic acid extraction/purification instrument (Qiagen). Briefly, buccal swabs were immersed in 400 μl of Buffer ATL (Qiagen) and treated with proteinase K by incubation at 56 °C for 10 min. Following incubation, buccal swabs were removed, placed into QiAamp spin columns and centrifuged at 13,000 r.p.m. for 1 min on a tabletop microfuge. Buccal swabs were then discarded. Cell lysates were then loaded into separate QiAamp spin columns as per the manufacturer’s suggested protocols; purified gDNA was automatically collected using the custom QiAcube ‘buccal swab’ spin protocol (Qiagen). Quantification of human gDNA was performed with the Investigator Quantiplex Kit (Qiagen) using the manufacturer’s suggested protocols. Briefly, buccal swab eluate was diluted 1:1 with nuclease-free water. Quantitative PCR to measure buccal swab eluate gDNA content was performed using the Investigator Quantiplex Kit on a RotorGene Q-400 thermocycler (Qiagen) with an initial denaturation step of 95 °C for 10 min, followed by 40 cycles of 95 °C for 1 s and 60 °C for 10 s. DNA quantity was recorded as a mean of technical replicates.

SNP genotyping of subjects

SNP genotyping were determined using the Type-it Fast Probe PCR Kit (Qiagen), TaqMan SNP genotyping probes, and primers for rs204989 (GPSM3), rs204991 (GPSM3), rs2182137 (CCL21) and rs6457620 (HFA gene region) (Life Technologies, Carlsbad, CA, USA; cat. #C_3293828_20, C_2412452_20, C_1611356_10 and C_29315329_10, respectively) using the manufacturer’s suggested protocols. Briefly, equal masses of human gDNA were added to each reaction and the manufacturer’s recommended cycling parameters were used to detect both major and minor alleles in multiple quantitative PCR reactions performed in technical replicates on a RotorGene Q-400 thermocycler (Qiagen). Coded identifiers were used to blind investigators of the disease state of volunteers until all samples had been genotyped.

Isolation of whole blood RNA

To maintain homogeneity of sample treatment, all RNA was extracted from archived whole blood collected into PAXgene Blood RNA tubes with the PAXGene Blood RNA kit (Qiagen) using a modification of the manufacturer’s suggested protocol. Briefly, when possible genotypes (M/M and m/m) were prepared concurrently. Upon removal from cryostorage, PAXGene tubes were incubated on a rocker at room temperature for 15–18 h, centrifuged at 5000 × g for 10 min at 4 °C and the resulting cell pellet was resuspended in 4 ml of nuclease-free water. The solution was then centrifuged at 5000 × g for 10 min at 4 °C and the cell pellet was resuspended in 350 μl of the manufacturer’s resuspension buffer (BRI). Isolation of whole blood RNA was performed using an automated QIAcube workflow according to the manufacturer’s suggested protocol to minimize variability between preparations. Eluted nucleic acids were quantified by absorbance at 260 nm using a Nanodrop 2000 (Thermo Scientific, Pittsburgh, PA, USA) and assessed for purity using 260/280 nm ratio of > 1.8. Immediately following isolation, a sample (250 ng) of eluted nucleic acids from each subject was used to synthesize complementary DNA with Thermoscript Reverse Transcriptase Kit (Thermo Scientific). Complementary DNA was then diluted 1:5 and used immediately for quantitative reverse transcriptase-PCR analyses.

Quantitative reverse transcriptase-PCR

A sample of complementary DNA, derived from equal masses of RNA from each volunteer, was amplified in duplicate (with an additional independent ‘no reverse transcriptase-treated’ control) using the USB VeriFast Quick SYBR Green qPCR Master Mix (Affymetrix, Cleveland, OH, USA) and custom-designed, intron-spanning PCR primers (Supplementary Table S1 and Supplementary Figure S1). Cycling conditions were modified from the manufacturer’s suggested protocol to establish linearity of transcript detection (Supplementary Figures S2A and B). Relative quantification of gene expression was performed using the ΔΔCT method,38 and the specificity of each amplicon was verified by gel electrophoresis and melt curve tests (for example, Supplementary Figure S2C).

Cell lines and culture conditions

The human embryonic kidney 293 T cell line (HEK293T) was obtained from the American Type Tissue Collection (ATCC, Manassas, VA, USA). Cell lines were maintained at 37 °C in a humidified atmosphere containing 5% CO2. The HEK293T cell line was maintained in Dulbecco’s modified Eagle’s medium supplemented with 10% fetal bovine serum, 100 IU penicillin and 100 μg ml−1 streptomycin. Transient transfections of HEK293T cells were performed as previously described.42,43 Briefly, transfection of monolayers grown to 60–70% confluence was performed by incubating cells 15–18 h with Ca3(PO4)2 and plasmid DNA solution. The cell monolayers were then washed twice with phosphate-buffered saline (pH 7.4), then fresh medium was added. THP-1 cells, from the ATCC, were maintained in an ATCC modification of RPMI medium supplemented with 10% heat-inactivated fetal bovine serum, 100 IU penicillin, 100 μg ml−1 streptomycin and 2 μg ml−1 puromycin. GPSM3 expression was stably knocked-down in THP-1 cells via lentiviral transduction (under the WVU IBC protocol #15-04-08) of Mission shRNA in the pLKO.1 vector (clones: TRCN0000036810 and TRCN0000036820). As a control, cells were established stably expressing scrambled control (a gift from Dr David Sabatini,44 Addgene plasmid # 1864; Addgene, Cambridge, MA, USA). Briefly, transduction was performed at a multiplicity of infection of 10 using the spinoculation protocol (described in O’Doherty et al.,45 transduced cells were selected with 2 μg ml−1 puromycin, and monolonal populations were established by limiting dilution. All THP-1 mononal cell lines and HEK293 T cells were confirmed free of Mycoplasma spp. and Acholeplasma spp. contamination (Supplementary Figure S3) using Lookout Mycoplasma PCR Detection kit following the manufacturer’s protocols (Sigma-Aldrich, St Louis, MO, USA).

Cloning and mutagenesis of plasmids

DNA from plasmids pGL3-basic-fuc2P (Promega, Madison, WI, USA) and pcDNA3.1(+)- puromycin, GPSM3 expression was stably knocked-down in THP-1 cells via lentiviral transduction (under the WVU IBC protocol #15-04-08) of Mission shRNA in the pLKO.1 vector (clones: TRCN0000036810 and TRCN0000036820). As a control, cells were established stably expressing scrambled control (a gift from Dr David Sabatini,44 Addgene plasmid # 1864; Addgene, Cambridge, MA, USA). Briefly, transduction was performed at a multiplicity of infection of 10 using the spinoculation protocol (described in O’Doherty et al.,45 transduced cells were selected with 2 μg ml−1 puromycin, and monolonal populations were established by limiting dilution. All THP-1 mononal cell lines and HEK293 T cells were confirmed free of Mycoplasma spp. and Acholeplasma spp. contamination (Supplementary Figure S3) using Lookout Mycoplasma PCR Detection kit following the manufacturer’s protocols (Sigma-Aldrich, St Louis, MO, USA).
enzymes and buffers (Thermo Scientific), and restricted DNA ends were phosphatase treated with calf intestinal alkaline phosphatase (Promega). Vectors were purified via extraction from a 1% agarose gel (QiAquick, Qiagen). Sequences of the GPSM3 gene locus were retrieved from the NCBI reference sequence (NC_000006.12 c3219553-32197066). A 3510-bp region of the GPSM3 promoter was amplified from gDNA (GPSM3 amplification primers; Supplementary Table S2) from one individual with the major GPSM3 alleles (‘pG3L-M’); another individual with all minor polymorphic alleles (‘pG3L- M’); nested primers (Supplementary Table S2) were used to amplify a 3512-bp span of this promoter region (NC_000006.12 32192523–32196035) and thereby introduce flanking HindIII and XhoI cut sites (GPSM3 cloning primers). The resultant amplicons were digested with HindIII and XhoI, and ligated into both pGL3-luc2P and pcDNA3.1(+) vectors (Rapid DNA ligation kit Roche, Basel, Switzerland). Mutagenesis was performed using the QuickChange Lightning Mutagenesis kit (Agilent Technologies, Santa Clara, CA, USA) and manufacturer’s suggested protocols with primers (Supplementary Table S3). Briefly, mutagenesis of GPSM3 promoter inserts was performed within the pcDNA3.1(+) vector, sequenced by the Sanger method to confirm mutagenesis, and then subcloned into pGL3-luc2P vectors.

Promoter-driven luciferase assay
HEK293T cells were transiently transfected using the calcium phosphate method, as described previously,5,46 with 5 μg of pGL3 firefly luciferase-based reporter vector and 1 μg of pRL-TK Renilla luciferase-based reporter control vector (Promega). Forty-eight hours following transfection, the cells were assayed for both firefly and Renilla luciferase activities using the manufacturer’s suggested protocol for the Dual-Luciferase Reporter Assay System (Promega). Briefly, cells were harvested, washed in phosphate-buffered saline at pH 7.4, and lysed using the manufacturer’s ‘passive lysis’ buffer. Using automatic injectors within a Flexstation-3 (Molecular Devices, Sunnyvale, CA, USA), firefly luciferase substrate was added to the cellular lysate and luminescence was recorded in a white, clear-bottom 96-well plate for 10 s. Firefly luminescence was then quenched and Renilla luciferase substrate was added to the lysate and luminescence was recorded for 10 s. Data were normalized to luminescence induced by Renilla luciferase activity and reported as fold-change from pGL3-M transfected HEK 293 T cells. All data are representative of biologic triplicates.

Transcription factor bioinformatics
The 3.5 kb of sequence 5’ to the GPSM3 transcription start site was acquired from the Genome Reference Consortium human genome 38 (GRCh38) and imported into JASPAR.29 Predictions of consensus binding motifs for transcription factors were cut off at 20% dissimilarity.

Transwell migration assay
THP-1 cells stably expressing either GPSM3 knockdown shRNA or scrambled shRNA control were serum-starved in Hank’s balanced salt solution+1% bovine serum albumin at 37 °C for 1 h and labeled with calcein-AM cell permeant fluorescent dye (494/517 nm). Cells were washed and suspended in Hank’s balanced salt solution+1% bovine serum albumin+10 μM HEPES, and 5 × 107 cells ml⁻¹ were loaded into the upper well of a Falcon HTS FluoroBlok 96-Multiwell Insert System (Corning, NY, USA). Cells were allowed to migrate toward vehicle or 100 ng ml⁻¹ MCP-1/CCL2 (R&D Systems, Minneapolis, MN, USA) in the lower chamber. Fluorescence was measured at 2-min intervals for the course of 60 min at 37 °C using a Flexstation-3 microplate reader (Molecular Devices, Sunnyvale, CA, USA). All data are representative of biological triplicates with technical replicates (that is, n = 6).

Statistical analyses
Descriptive data from the case–control cohort were analyzed using SAS University Edition 3.1 (SAS Institute Inc., Cary, NC, USA). Differences in 2 × 2 discrete variables were determined using Fisher’s exact test, whereas 2 × 3 discrete variables were compared using the Fisher–Freeman–Halton exact test.46 Regression analyses in Transwell migration assays were performed using four parameter logistic nonlinear regression analyses over the period of migration with post-hoc comparison by one-way analysis of variance of the unshared maximum and time to 50% migration parameters. All nonlinear models were confirmed to have r-squared values >0.99, and accepted the null hypothesis of the Shapiro–Wilks normality test (P > 0.05). All other analyses were performed in Prism 6.0 (La Jolla, CA, USA) using a two-tailed independent t-test or one-way analysis of variance with Bonferroni or Dunnett’s post-hoc tests where indicated.

CONFLICT OF INTEREST
The authors declare no conflict of interest.

ACKNOWLEDGEMENTS
This work was supported in part by the National Institutes of Health under Award Number U54GM10492. BJG and DPS thank Di Teresa Tarrant (UNC-Chapel Hill TARC) for helpful discussions on this topic.

REFERENCES
1. White GE, Iqbal AJ, Greaves DR. CC chemokine receptors and chronic inflammation—therapeutic opportunities and pharmacological challenges. Pharmacol Rev 2013; 65: 47–89.
2. Griffith JW, Sokol CL, Luster AD. Chemokines and chemokine receptors: positioning cells for host defense and immunity. Annu Rev Immunol 2014; 32: 659–702.
3. Kehr JH. Chemoinstructant receptor signaling and the control of lymphocyte migration. Immunol Res 2006; 34: 211–227.
4. Cho H, Kehr JH. Regulation of immune function by G protein-coupled receptors, trimeric G proteins, and RGS proteins. Prog Mol Biol Transl Sci 2009; 86: 249–298.
5. Genger W, Biro M, Jain R. Leukocyte migration in the intestinal epithelium of non-lymphoid organs. Nat Rev Immunol 2014; 14: 232–246.
6. Willard FS, Kittle RJ, Siderovski DP. Return of the GDI: the GoLoco motif in cell signaling. Mol Pharmacol 2004; 64: 925–951.
7. Lambert NA, Johnston CA, Cappell SD, Kuravi S, Kittle AJ, Willard FS et al. Regulators of G-protein signaling accelerate GPCR signaling kinetics and govern sensitivity solely by accelerating GTPase activity. Proc Natl Acad Sci USA 2010; 107: 7066–7071.
8. Takesono A, Cisowski MJ, Bernard M, Chung P, Hazard S, Duzic E et al. Receptor-independent activators of heterotrimeric G-protein signaling pathways. J Biol Chem 1999; 274: 33202–33205.
9. Siderovski DP, Diverse-Pierluissi MA, De Vries L. The GoLoco motif: a Gα/i binding motif and potential guanine-nucleotide exchange factor. Trends Biochem Sci 1999; 24: 340–341.
10. Kittle RJ, Kittle ME, Betts L, Sondek J, Siderovski DP. Structural determinants for GDI-induced inhibition of nucleotide release by Galphab subunits. Nature 2002; 416: 876–881.
11. Surve CR, Lehmann D, Smrcka AV. A chemical biology approach demonstrates G protein βγ subunits are sufficient to mediate directional neutrophil chemotaxis. J Biol Chem 2014; 289: 30705–30711.
12. Lehmann DM, AMP8 Serevatintric, Smrcka AV, York N. Small molecule disruption of G protein βγ subunit signaling inhibits neutrophil chemotaxis and inflammation. Mol Pharmacol 2008; 73: 410–418.
13. Kinoshita-Kawada M, Oberdick J, Xi Zhu M, A Purkinje cell specific GocLod domain protein, L7/Pcp-2, modulates receptor-mediated inhibition of Cav2.1 Ca2+ channels in a dose-dependent manner. Mol Brain Res 2004; 132: 73–86.
14. Dennis EA, Bradshaw RA. Signal transduction by G proteins: basic principles, molecular diversity, and structural basis of their actions. In: Transduction Mechanisms in Cellular Signaling. Academic Press: Waltham, MA, 2011, pp 499–505.
15. Wu C, Orozco C, Boyer J, Leglise M, Goodyle J, Batalov S et al. BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. Genome Biol 2009; 10: R130.
16. Gigoue P, Billard M, Laresco G, Buckley B, Timoshchko R, McGinnis M et al. G-protein signaling modulator-3, a gene linked to autoimmune diseases, regulates monocyte function and its deficiency protects from inflammatory arthritis. Mol Immunol 2013; 54: 193–198.
17. Billard MJ, Gall BJ, Richards KL, Siderovski DP, Tarrant TK. G protein signaling modulator-3: a leukocyte regulator of inflammation in health and disease. Am J Clin Exp Immunol 2014; 3: 97–106.
18. Sigot M, Schaub M, Batzoglou S, Robinson WH, Butte AJ. Autoimmune disease classification by inverse association with SNP alleles. PLoS Genet 2009; 5: e1000792.
19. Corona E, Dudley JT, Butte AJ. Extreme evolutionary disparities seen in positive selection across seven complex diseases. PLoS One 2010; 5: e12326.
SNP rs204989 decreases GPSM3 transcription
BJ Gall et al

20 Plenge RM, Seielstad M, Padyukov L, Lee T, Remmers EF, Ding B et al. TRAF1-C5 as a risk locus for rheumatoid arthritis — a genomewide study. Nature 2007; 447: 661–678.

21 Gregersen PK, Silver J, Winchester RJ. The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. Arthritis Rheum 1987; 30: 1205–1213.

22 Holoshitz J. The rheumatoid arthritis HLA-DRB1 shared epitope. Curr Opin Rheumatol 2010; 22: 293–298.

23 Wellcome T, Case T, Consortium C. Genomewide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007; 447: 661–678.

24 Raychaudhuri S, Remmers EF, Lee AT, Hackett R, Burtt NP, Gianniny L et al. Common variants at CD40 and other loci confer risk of rheumatoid arthritis. Nat Genet 2008; 40: 1216–1223.

25 Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM et al. dbSNP: the NCBI database of genetic variation. Nucleic Acids Res 2001; 29: 308–311.

26 Hobl E, Mader R, Erfacher L, Duhm B, Mustak M, Broll H et al. The influence of methotrexate on the gene expression of the pro-inflammatory cytokine IL-12 A in the therapy of rheumatoid arthritis. Clin Exp Rheumatol 2011; 29: 963–969.

27 Goldminz AM, Suárez-Fariñas M, Wang AC, Dumont N, Krueger JG, Gottlieb AB. The shared epitope hypothesis. Arthritis Rheum 2006; 53: 1–14.

28 The 1000 Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. Nature 2012; 491: 56–65.

29 Mathelier A, Zhao X, Zhang AW, Parcy F, Worsley-Hunt R, Arenillas DJ et al. JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. Nucleic Acids Res 2014; 42: D124–D147.

30 Giguère PM, Laroche G, Oestreich Ea, Siderovski DP. G-protein signaling modulator-3 regulates heterotrimeric G-protein dynamics through dual association with Gβ and Gαi protein subunits. J Biol Chem 2012; 287: 4863–4874.

31 Ward LD, Kellis M. Interpreting noncoding genetic variation in complex traits and human disease. Nat Biotechnol 2012; 30: 1095–1106.

32 Kochi Y, Suzuki A, Yamamoto K. Genetic basis of rheumatoid arthritis: a current review. Biochem Biophys Res Commun 2014; 452: 254–262.

33 Boyle AP, Hong EL, Harirhan M, Cheng Y, Schaub MA, Kasowski M et al. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res 2012; 22: 1790–1797.

34 Wellcome Trust Case Control Consortium. Genomewide-association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007; 447: 661–678.

35 Han T-U, Bang S-Y, Kang C, Bae S-C. TRAF1 polymorphisms associated with rheumatoid arthritis susceptibility in Asians and in Caucasians. Arthritis Rheum 2009; 60: 2577–2584.

36 Dausset J, Cann H, Cohen D, Latrop M, Lalouel JM, White R. Centre d'étude du polymorphisme humain (CEPH): collaborative genetic mapping of the human genome. Genomics 1990; 6: 575–577.

37 Grisu RA, Dilger RJ, Hammock AS, Plein C. West Virginia Politics and Government 2nd edn. University of Nebraska Press: Lincoln, NE, 2009.

38 Tarrant TK, Patel DD. Chemokines and leukocyte trafficking in rheumatoid arthritis. Pathophysiology 2006; 13: 1–14.

39 Brühl H, Ciak J, Schneider Ma, Plachý J, Rupp T, Wenzel I et al. Dual role of CCR2 during initiation and progression of collagen-induced arthritis: evidence for regulatory activity of CCR2+ T cells. J Immunol 2004; 172: 890–898.

40 Pavkova Goldbergova M, Lipkova J, Pavek N, Gatterova J, Vasku A, Soucek M et al. RANTES, MCP-1 chemokines and factors describing rheumatoid arthritis. Mol Immunol 2012; 52: 273–278.

41 Hrdlickova B, Kumar V, Kandun K, Zhemakova DV, Tripathi S, Karjalainen J et al. Expression profiles of long non-coding RNAs located in autoimmune disease-associated regions reveal immune cell type specificity. Genome Med 2014; 6: 88–101.

42 Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 2001; 25: 402–408.

43 Kingston RE, Chen CA, Rose JK. Calcium phosphate transfection. In: Bonifacino JS, Harford JB, Lippincott-Schwartz J, Yamada KM (eds). Current Protocols in Cell Biology. John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2009. pp 10.13.9–10.13.10.

44 Kingston RE, Chen CA, Rose JK. Calcium phosphate transfection. In: Bonifacino JS, Harford JB, Lippincott-Schwartz J, Yamada KM (eds). Current Protocols in Molecular Biology. John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2001. pp 10.13.1–10.13.9.

45 O'Doherty U, Swiggard WJ, Malim MH. Human immunodeficiency virus type 1 spinoculation enhances infection through virus binding. J Virol 2000; 74: 10074–10080.

46 Kingston RE, Chen CA, Rose JK. Transfection of DNA. In: Bonifacino JS, Harford JB, Lippincott-Schwartz J, Yamada KM (eds). Current Protocols in Cell Biology. John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2003. pp 10.13.1–10.13.9.

47 Freeman GH, Halton JH. Note on an exact treatment of contingency, goodness of fit and other problems of significance. Biometrika 1951; 38: 141–149.

Supplementary Information accompanies this paper on Genes and Immunity website (http://www.nature.com/gene)