HHS Public Access

Author manuscript

Genes Immun. Author manuscript; available in PMC 2016 July 28.

Published in final edited form as:

Genes Immun. 2016 March ; 17(2): 139–147. doi:10.1038/gene.2016.3.

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

BJ Gall[1], **A Wilson**[2], **AB Schroer**[1], **JD Gross**[1], **P Stoilov**[3], **V Setola**[1,4], **CM Watkins**[2], and **DP Siderovski**[1]

DP Siderovski: dpsiderovski@hsc.wvu.edu

[1]Department of Physiology & Pharmacology, West Virginia University School of Medicine, Morgantown, WV, USA 26506-9229

[2]Department of Orthopaedics, West Virginia University School of Medicine, Morgantown, WV, USA 26506-9229

[3]Department of Biochemistry, West Virginia University School of Medicine, Morgantown, WV, USA 26506-9229

[4]Department of Behavioral Medicine & Psychiatry, West Virginia University School of Medicine, Morgantown, WV, USA 26506-9229

Abstract

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, 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.

Conflicts of Interest: The authors declare no conflict of interest.

Introduction

Chemokine receptors comprise a subfamily of the G protein-coupled receptor (GPCR) superfamily of transmembrane receptors that are expressed on a number of leukocyte subsets and function predominantly to regulate chemotaxis[1–5]. Upon binding their cognate chemokine agonists, chemokine receptors transduce signals by inducing dissociation of their associated, intracellular G_i protein heterotrimers ($G_i\alpha$·GDP/$G\beta\gamma$). This process is highly regulated through additional intracellular proteins that act upon the $G_i\alpha$ subunit and ultimately affect the rate of signal inactivation[4,6,7]. In particular, proteins containing one or more conserved GoLoco motifs are capable of sequestering inactivated $G_i\alpha$·GDP, preventing its reassociation with $G\beta\gamma$ and GPCRs and thereby disrupting continued $G_i\alpha$-induced signaling without quenching $G\beta\gamma$-mediated signaling[6–10]. The importance of $G\beta\gamma$-associated signaling to chemokine actions has recently been highlighted by reports that specific $G\beta\gamma$-activating compounds are sufficient to induce neutrophil chemotaxis[11] and, conversely, a $G\beta\gamma$ antagonist can inhibit fMLP-induced chemotaxis[12]. GoLoco proteins may directly regulate signaling pathways required for chemotaxis by sequestering $G_i\alpha$·GDP and prolonging $G\beta\gamma$-mediated signaling processes[13,14], thereby exacerbating inflammation.

G protein signaling modulator 3 (GPSM3) contains two functional GoLoco motifs and is restricted in its expression to leukocytes and myeloid-derived cells[15,16]. *Gpsm3*-deficient mice exhibit a profoundly blunted course of joint damage in an acute, collagen antibody-induced model of inflammatory arthritis[16]. Histologic analyses of affected synovial tissues indicate a relative acellularity of synovial cavities in *Gpsm3*-deficient mice compared to 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, most likely through a key regulatory role on leukocyte chemotaxis to the synovium[17]. 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, 5′ to the *GPSM3* transcriptional start site that are significantly less prevalent in individuals with rheumatoid arthritis (and other autoimmune diseases; *e.g.*, lupus and multiple sclerosis), implying a potential protective effect of these minor gene variations[18–20]. 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[21] (reviewed in ref. [22]). 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 (WTCCC), North American Rheumatoid Arthritis Consortium (NARAC), and the Swedish Epidemiological Investigation of Rheumatoid Arthritis (EIRA)[23,24]. 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 inverse association with rheumatoid arthritis. Furthermore, we assessed whether linkage disequilibrium with the known RA risk allele in the *HLA gene* region, rs6457620[23,24], may affect the inverse association of *GPSM3* SNP alleles with RA. Additionally, 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[24]. We recruited a group of 50 volunteers with a diagnosis of RA, 50 RA-free volunteers who were matched to the aforementioned group by a 'Bring-a-friend-to-clinic' program, and 100 unmatched 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[18–20] and *Gpsm3*-deficient mouse studies[16], we hypothesized that individuals homozygous for the minor alleles of rs204989 and rs204991 would exhibit decreased whole-blood *GPSM3* transcript abundance. Additionally, 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 by GWAS with protection from rheumatoid arthritis, 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-clinic' program. Upon genotyping all 200 volunteers recruited for this study, we found that *GPSM3* SNPs rs204989 and rs204991, originally identified to be independently[18–20] associated with protection from RA, are in complete linkage disequilibrium within this population. Additionally, 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 (Fig. 1A). All of these chromosome 6 SNPs have been previously identified in the NCBI Database of Single Nucleotide Polymorphisms (dbSNPs) [25]. In this study, three of these *GPSM3* SNPs (rs204989, rs204990, rs204991) were seen to be inherited as a haploblock 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 (Fig. 1A).

The *GPSM3* rs204989/rs204991 haploblock 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% C.I. of ± 7.6%) of the average *GPSM3* transcript level present in homozygotes for the major alleles (M/M; n = 53; 95% C.I. of ± 5.6%). This was a decrease of 24.1% (95% C.I. of 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 (Fig. 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% C.I. of ± 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[26,27]. 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 haploblock with a minor allele frequency (MAF) of 29% within the unmatched healthy young cohort (n=100), 23% within the rheumatoid arthritis 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 rheumatoid arthritis 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 rheumatoid arthritis 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 (Fig. 1C).

The *GPSM3* SNP haploblock 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 haploblock[25]. 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%, while 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 (Fig. 2A; p = 0.0066), suggesting the presence of weak linkage disequilibrium and consistent with data from the NCBI dbSNP [25] (Fig. 2D). Within the RA cohort, there was significant variation from chance distribution of *GPSM3* SNPs MAF when stratified by *HLA* gene region SNP genotype (Fig. 2B; p = 0.0123), which was not observed within the matched control group (Fig. 2C; p = 0.2739). However, when stratified by *HLA* SNP rs6457620 genotype, *GPSM3* transcript abundance in whole blood was not significantly different (Fig. 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 haploblock genotype (Fig. 1B).

rs204989 of the *GPSM3* SNP haploblock 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 haploblock 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 haploblock 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; Fig. 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 (*i.e.*, 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 = 0.2136), or rs3096688 (p = 0.6618) did not significantly alter the normalized firefly luciferase activity of the wild type promoter (Fig. 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; Fig. 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; Fig. 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 (Fig. 4).

Decreased GPSM3 abundance leads to a deficit in migration towards 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 [16]; 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 (Fig. 5A), using previously validated shRNA-expressing lentiviral vectors directed against *GPSM3* (*i.e.*, shRNA19 and shRNA20; ref. [47]). Using these stable THP-1 cell lines in real-time Transwell plate migration assays, we measured vehicle-normalized relative fluorescence units (RFU) over a 30-minute 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 (Fig. 5B); these migration differences were independent of any changes in transcript abundance of the MCP-1 receptor (*i.e.*, *CCR2*) between cell lines assayed (Fig. 5C; F = 2.40; p = 0.119). This migratory deficit upon GPSM3 knockdown was observed both by a decrease in initial rate of migration (Fig. 5D) and by a decrease in maximum transmigration-induced fluorescence achieved (Fig. 5E).

Discussion

Genome-wide association studies (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[30]. In multiple GWAS reports, noncoding variants 5′ to the transcription start site of *GPSM3* (Fig. 1A) have been identified as occurring *less* frequently in individuals with RA[18–20]. However, like many identified variants, no functional consequence has been elucidated to explain this association. In the current 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) (Fig. 1B) supports the hypothesis first put forth from GWAS[18–20] that there is a potential functional consequence of the identified *GPSM3* SNPs rs204989 and rs204991, namely protection from development of rheumatoid arthritis; we have previously demonstrated[16] that lowered *Gpsm3* expression in mice, as generated by gene knockout technology, results in protection from collagen-autoantibody induced inflammatory arthritis.

While our present finding of lowered *GPSM3* mRNA levels in the whole blood of homozygotes bearing the *GPSM3* SNP haploblock is consistent with GWAS and mouse gene knockout studies, the observed frequencies of the rs204989/rs204991 haploblock within the various genotyped groups are seemingly discrepant, given the trend toward decreased *GPSM3* minor allele frequency, albeit not significantly different, in the disease-free matched control (MAF = 18.0%; n = 50) compared to the RA cohort (MAF = 23.0%; n = 50). Considering the small sample size, the observed variability is not remarkable in light of the reported minor allele frequency 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 (Fig. 1C), *CCL21* gene rs2812378 [T>C] on chromosome 9 (*i.e.*, unlinked to the human *GPSM3* locus on chromosome 6; ref. [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 Holoshitz[22] and Yuta, et al.[33]). This RA risk allele, rs6457620 [C>G][23,24] is reported by RegulomeDB[34] (based on Hapmap2 metadata) to exist in complete linkage disequilibrium ($D' = 1.0$; $r^2 = 1.0$) with another *HLA* region RA risk allele rs6457617 [C>T][31,32] in a population of Americans with European ancestry (CEPH collection)[35] that is very similar to the demographics of our West Virginia-based sample population. Both polymorphisms (rs6457617 & rs6457620) exist ~500 kb from the *GPSM3* SNP haploblock (*e.g.*, Fig. 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 (Fig. 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 to the matched control cohort (Fig. 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[24]) or other RA risk-related *HLA* gene haploblocks in linkage disequilibrium[21,33]. Previous GWAS studies have genotyped tens of thousands of ethnically-diverse volunteers to identify the two *GPSM3* SNPs rs204989 and rs204991 as inversely associated with RA [18–20]. Therefore, we hypothesize that our inability to reproduce previous GWAS findings is most likely a consequence of the relatively small sample size we recruited for this

functional study of *GPSM3* phenotypes and/or a founder effect among the local recruited volunteers[36]. 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[16].

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 promoter region containing the haploblock 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 haploblock *versus* the ancestral ("major") haploblock. 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 (e.g., 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 AP-1 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 G protein-coupled receptor signaling via $G_i\alpha$[6–10] – a common signaling system used by chemokine receptors. Considering the restricted expression of GPSM3 in leukocytes and lymphoid organs[15,16] and the importance of of $G_i\alpha$-coupled chemokine receptors in leukocyte migration during RA pathogenesis[37], 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 monocyte chemoattractant protein-1 (MCP-1/CCR2) signaling in the recruitment of monocytes and macrophages during inflammatory bouts in RA patients[38,39], 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*) (Fig. 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 haploblock. Therefore, we hypothesize that the functional effect of the rs204989 *GPSM3* SNP *in vivo* is most likely more subtle, yet it likely plays 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' datasets report the *GPSM3* mRNA as being expressed at an appreciable level in all human leukocyte subsets examined (GEO accession GDS62408)[40], 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 rheumatoid arthritis[37]. 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. Additionally, we will seek to establish whether one or both of the candidate transcription factors (*i.e.*, 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 rheumatoid arthritis patients (ICD-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 rheumatoid arthritis individuals included (1) at least 18 years of age, (2) positive ICD-9 diagnosis and (3) rheumatoid factor positive. Inclusion criteria for the 50 demographically matched controls included (1) at least 18 years of age, (2) not a blood relative of a volunteer enrolled in the study, and (3) no family history of autoimmune disease. For buccal swab samples, volunteers were instructed to vigorously swab the inside of both cheeks for 15 seconds on each side, and resultant samples were stored in a sterile nuclease-free container at -80°C. Blood samples were collected by a trained phlebotomist sequentially into a K_2EDTA tube (Becton Dickinson) and a PAXgene Blood RNA tube (BD); as per manufacturer's suggested protocol, the PAXgene Blood RNA tube was incubated at room temperature for 120 minutes prior to archiving at -80 °C.

Isolation and Quantification of gDNA

Isolation of genomic DNA (gDNA) from buccal swab samples was performed with the QIAamp DNA Blood Mini Kit (Qiagen) on a QIAcube automated nucleic acid extraction/purification instrument (Qiagen; Germantown, MD). Briefly, buccal swabs were immersed in 400 μl of Buffer ATL (Qiagen) and treated with proteinase K by incubation at 56 °C for 10 minutes. Following incubation, buccal swabs were removed, placed into QIAamp spin columns, and centrifuged at 13,000 RPM for 1 minute on a tabletop microfuge. Buccal swabs were then discarded. Cell lysates were then loaded into separate QIAamp spin

columns as per 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 eluate gDNA content was performed using the Investigator Quantiplex Kit on a RotorGene Q-400 thermocycler (Qiagen) with an initial denaturation phase at 95°C for 60 seconds, followed by 40 cycles of 95°C for 1 second and 60°C for 10 seconds. DNA quantity was recorded as a mean of technical replicates.

SNP Genotyping of Subjects

SNP genotypes were determined using the Type-it Fast Probe PCR Kit (Qiagen), TaqMan SNP genotyping probes, and primers for rs204989 (*GPSM3*), rs204991 (*GPSM3*), rs2812378 (*CCL21*), and rs6457620 (*HLA* gene region) (Life Technologies; cat. #C_3293828_20, C_2412452_20, C_16113556_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 multiplex qPCR reactions performed in technical replicates on a RotorGene Q-400 thermocycler. 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 hours, centrifuged at 5,000 × g for 10 minutes at 4°C, and the resulting cell pellet was resuspended in 4 ml of nuclease-free water. The solution was then centrifuged at 5,000 × g for 10 minutes at 4 °C, and the cell pellet was resuspended in 350 μl of the manufacturer's resuspension buffer (BR1). 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) 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 (cDNA) with Thermo Verso Reverse Transcriptase Kit (Thermo Scientific). cDNA was then diluted 1:5 and used immediately for qRT-PCR analyses.

Quantitative Reverse Transcriptase PCR (qRT-PCR)

A sample of cDNA, 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 VeriQuest Fast SYBR Green qPCR Master Mix (Affymetrix, Cleveland, OH) and custom-designed, intron-spanning PCR primers (Supplemental Table S1 and Figure S1). Cycling conditions were modified from the manufacturer's suggested protocol to establish

linearity of transcript detection (Supplemental Figure S2A,B). Relative quantification of gene expression was performed using the $\Delta\Delta C_T$ method[41], and the specificity of each amplicon was verified by gel electrophoresis and melt curve tests (*e.g.*, Supplemental Fig. S2C).

Cell lines and culture conditions

The human embryonic kidney 293T cell line (HEK293T) was obtained from the American Type Tissue Collection (ATCC). Cell lines were maintained at 37 °C in a humidified atmosphere containing 5% CO_2. The HEK293T cell line was maintained in DMEM supplemented with 10% fetal bovine serum, 100 IU penicillin, and 100 μg/ml streptomycin. Transient transfections of HEK293T cells were performed as previously described [42]. Briefly, transfection of monolayers grown to 60-70% confluence was performed by incubating cells 15-18 hours with $Ca_3(PO_4)_2$ 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 streptomycin, and 2 μg/ml 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: TRCN0000036819 and TRCN0000036820). An additional two THP-1 monoclonal cell lines were established stably-expressing scrambled control (a gift from Dr. David Sabatini, ref. [43]; Addgene plasmid # 1864). Briefly, transduction was performed at a multiplicity of infection of 10 using the spinoculation protocol (described in ref. [44]), transduced cells were selected with 2 μg/ml puromycin, and monoclonal populations were established by limiting dilution. All THP-1 monoclonal cell lines and HEK 293T cells were confirmed free of *Mycoplasma spp.* and *Acholeplasma spp.* contamination (Supplemental Figure S3) using Lookout® Mycoplasma PCR Detection kit following manufacturer's protocols (Sigma-Aldrich, St. Louis, MO).

Cloning and mutagenesis of plasmids

DNA from plasmids pGL3-basic-luc2P (Promega, Madison, WI) and pcDNA3.1(+) were digested at the *Hin*dIII and *Xho*I sites using FastDigest enzymes and buffers (Thermo Scientific, Pittsburgh, PA), and restricted DNA ends were phosphatase treated with calf intestinal alkaline phosphatase (Promega, Madison, WI). 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 c32195523-32190766). A 5010-bp region of the *GPSM3* promoter was amplified from gDNA ("*GPSM3* amplification primers"; Supplemental Table S2) from one individual with the major *GPSM3* alleles ("pGL3-M") and another individual with all minor polymorphic alleles ("pGL3-m"); nested primers (Supplemental Table S2) were used to amplify a 3512-bp span of this promoter region (NC_000006.12 32192523 - 32196035) and thereby introduce flanking *Hin*dIII and *Xho*I cut sites ("*GPSM3* cloning primers"). The resultant amplicons were digested with *Hin*dIII and *Xho*I, and ligated into both pGL3-luc2P and pcDNA3.1(+) vectors (Rapid DNA ligation kit; Roche, Basel, Switzerland). Mutagenesis was performed using the QuikChange Lightning Mutagenesis kit (Agilent Technologies, Santa Clara, CA) and manufacturer's

suggested protocols with primers (Supplemental 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[42,45], with 5 μg of pGL3 firefly luciferase-based reporter vector and 1 μg of pRL-TK *Renilla* luciferase-based reporter control vector (Promega, Madison, WI). 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, Madison, WI). Briefly, cells were harvested, washed in PBS pH 7.4, and lysed using the manufacturer's “passive lysis” buffer. Using automatic injectors within a Flexstation-3 (Molecular Devices, Sunnyvale, CA), firefly luciferase substrate was added to the cellular lysate and luminescence was recorded in a white, clear-bottom 96-well plate for 10 seconds. Firefly luminescence was then quenched and *Renilla* luciferase substrate was added to the lysate and luminescence was recorded for 10 seconds. Data were normalized to luminescence induced by *Renilla* luciferase activity and reported as fold-change from pGL3-M transfected HEK 293T 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 HBSS + 1% BSA at 37°C for one hour and labeled with calcein-AM cell permeant fluorescent dye (494 nm/517 nm). Cells were washed and suspended in HBSS + 1% BSA + 10 mM HEPES, and 5×10^7 cells/ml were loaded into the upper well of a Falcon™ HTS FluoroBlok 96-Multiwell Insert System chemotaxis chambers (Corning, NY). Cells were allowed to migrate toward vehicle or 100 ng/mL MCP-1/CCL2 (R&D Systems, Minneapolis, MN) in the lower chamber. Fluorescence was measured at 2 minute intervals over the course of 60 minutes at 37°C using a Flexstation 3 microplate reader (Molecular Devices, Sunnyvale, CA). All data are representative of biological triplicates with technical replicates (i.e., n = 6).

Statistical analyses

Descriptive data from the case-control cohort were analyzed using SAS University Edition 3.1 (SAS Institute Inc., Cary, NC). Differences in 2 × 2 discrete variables were determined using Fisher's exact test, while 2 × 3 discrete variables were compared using the Fisher-Freeman-Halton exact test[46]. Regression analyses in Transwell migration assays were performed using 4 parameter logistic nonlinear regression analyses over the period of

migration with post-hoc comparison by one-way ANOVA 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-Wilk normality test ($p \geq 0.05$). All other analyses were performed in Prism 6.0 using a two-tailed independent t-test or one-way ANOVA with Bonnferroni or Dunnett's post-hoc tests where indicated.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

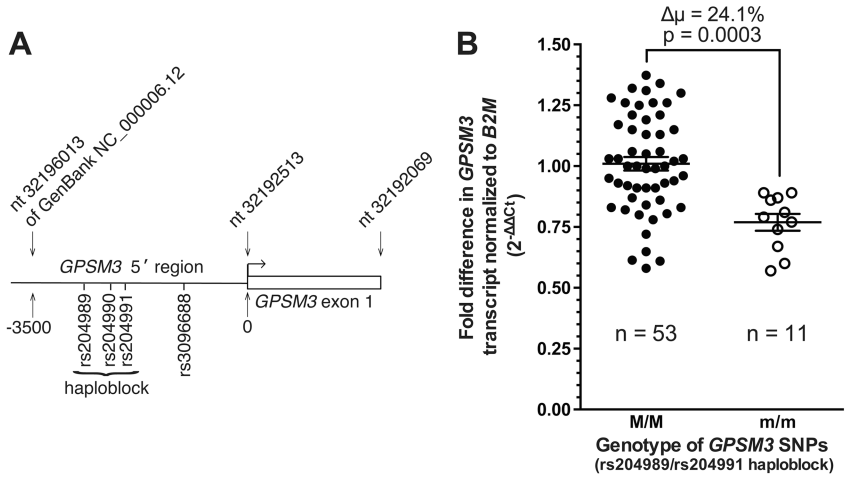
This work was supported in part by the National Institutes of Health under Award Number U54GM104942. B.J.G. and D.P.S. thank Dr. Teresa Tarrant (UNC-Chapel Hill TARC) for helpful discussions on this topic.

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. [PubMed: 23300131]
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. [PubMed: 24655300]
3. Kehrl JH. Chemoattractant receptor signaling and the control of lymphocyte migration. Immunol Res. 2006; 34:211–227. [PubMed: 16891672]
4. Cho, H.; Kehrl, JH. Regulation of immune function by G protein-coupled receptors, trimeric G proteins, and RGS proteins. 1st. Elsevier Inc.; 2009.
5. Weninger W, Biro M, Jain R. Leukocyte migration in the interstitial space of non-lymphoid organs. Nat Rev Immunol. 2014; 14:232–46. [PubMed: 24603165]
6. Willard FS, Kimple RJ, Siderovski DP. Return of the GDI: the GoLoco motif in cell division. Annu Rev Biochem. 2004; 73:925–951. [PubMed: 15189163]
7. Lambert NA, Johnston CA, Cappell SD, Kuravi S, Kimple 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 U S A. 2010; 107:7066–7071. [PubMed: 20351284]
8. Takesono A, Cismowski 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. [PubMed: 10559191]
9. Siderovski DP, Diverse-Pierluissi MA, De Vries L. The GoLoco motif: a Gα-i/o binding motif and potential guanine-nucleotide exchange factor. Trends Biochem Sci. 1999; 24:340–341. [PubMed: 10470031]
10. Kimple RJ, Kimple ME, Betts L, Sondek J, Siderovski DP. Structural determinants for GoLoco-induced inhibition of nucleotide release by Galpha subunits. Nature. 2002; 416:878–881. [PubMed: 11976690]
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; 289:17791–801. [PubMed: 24808183]
12. Lehmann DM, Seneviratne AMPB, Smrcka AV, York N. Small Molecule Disruption of G Protein βγ Subunit Signaling Inhibits Neutrophil Chemotaxis and Inflammation. Mol Pharmacol. 2008; 73:410–418. [PubMed: 18006643]
13. Kinoshita-Kawada M, Oberdick J, Xi Zhu M. A Purkinje cell specific GoLoco 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. [PubMed: 15548431]

14. Dennis, EA.; Bradshaw, RA. Transduction Mechanisms in Cellular Signaling. Academic Press; Waltham, MA: 2011. Signal Transduction by G Proteins: Basic Principles, Molecular Diversity, and Structural Basis of Their Actions; p. 499-466.
15. Wu C, Orozco C, Boyer J, Leglise M, Goodale J, Batalov S, et al. BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. Genome Biol. 2009; 10:R130. [PubMed: 19919682]
16. Giguère P, Billard M, Laroche G, Buckley B, Timoshchenko 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. [PubMed: 23280397]
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. [PubMed: 25143870]
18. Sirota M, Schaub Ma, Batzoglou S, Robinson WH, Butte AJ. Autoimmune disease classification by inverse association with SNP alleles. PLoS Genet. 2009; 5:e1000792. [PubMed: 20041220]
19. Corona E, Dudley JT, Butte AJ. Extreme evolutionary disparities seen in positive selection across seven complex diseases. PLoS One. 2010; 5:e12236. [PubMed: 20808933]
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. N Engl J Med. 2009; 357:1199–1209. [PubMed: 17804836]
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; 30:1205–13. [PubMed: 2446635]
22. Holoshitz J. The rheumatoid arthritis HLA-DRB1 shared epitope. Curr Opin Rheumatol. 2010; 22:293–8. [PubMed: 20061955]
23. Wellcome T, Case T, Consortium C. Genomewide association study of 14,000 cases of seven common diseases and 3,000 shared controls Supplementary Information. 2007:1–56.
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. [PubMed: 18794853]
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. [PubMed: 11125122]
26. Hobl E, Mader R, Erlacher L, Duhm B, Mustak M, Broll H, et al. The influence of methotrexate on the gene expression of the pro-inflammatory cytokine IL-12A in the therapy of rheumatoid arthritis. Clin Exp Rheumatol. 2011; 29:963–969. [PubMed: 22133036]
27. Goldminz AM, Suárez-Fariñas M, Wang AC, Dumont N, Krueger JG, Gottlieb AB. *CCL20* and *IL22* Messenger RNA Expression After Adalimumab vs Methotrexate Treatment of Psoriasis. JAMA Dermatology. 2015; 02111:1–10.
28. The 1000 Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012; 491:56–65. [PubMed: 23128226]
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; 42:1–6. [PubMed: 24376271]
30. Ward LD, Kellis M. Interpreting noncoding genetic variation in complex traits and human disease. Nat Biotechnol. 2012; 30:1095–106. [PubMed: 23138309]
31. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007; 447:661–78. [PubMed: 17554300]
32. Han TU, Bang SY, Kang C, Bae SC. TRAF1 polymorphisms associated with rheumatoid arthritis susceptibility in Asians and in Caucasians. Arthritis Rheum. 2009; 60:2577–84. [PubMed: 19714643]
33. Kochi Y, Suzuki A, Yamamoto K. Genetic basis of rheumatoid arthritis: a current review. Biochem Biophys Res Commun. 2014; 452:254–62. [PubMed: 25078624]

34. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, et al. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res. 2012; 22:1790–7. [PubMed: 22955989]
35. Dausset J, Cann H, Cohen D, Lathrop M, Lalouel JM, White R. Centre d'etude du polymorphisme humain (CEPH): collaborative genetic mapping of the human genome. Genomics. 1990; 6:575–7. [PubMed: 2184120]
36. Brisbin, RA.; Dilger, RJ.; Hammock, AS.; Plein, C. West Virginia Politics and Government. 2nd. University of Nebraska Press; Lincoln, NE: 2009.
37. Tarrant TK, Patel DD. Chemokines and leukocyte trafficking in rheumatoid arthritis. Pathophysiology. 2006; 13:1–14. [PubMed: 16380240]
38. Brühl H, Cihak 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. [PubMed: 14707060]
39. 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; 52:273–8. [PubMed: 22750227]
40. Hrdlickova B, Kumar V, Kanduri K, Zhernakova 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. [PubMed: 25419237]
41. 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; 25:402–8. [PubMed: 11846609]
42. Kingston, RE.; Chen, CA.; Rose, JK. Current Protocols in Molecular Biology. John Wiley & Sons, Inc.; 2001. Calcium Phosphate Transfection; p. 10.13.1-10.13.9.
43. Sarbassov DD, Guertin DA, Ali SM, Sabatini DM. Phosphorylation and regulation of Akt/PKB by the rictor-mTOR complex. Science. 2005; 307:1098–101. [PubMed: 15718470]
44. O'Doherty U, Swiggard WJ, Malim MH. Human immunodeficiency virus type 1 spinoculation enhances infection through virus binding. J Virol. 2000; 74:10074–80. [PubMed: 11024136]
45. Kingston R, Chen C, Okayama H. Transfection of DNA into eukaryotic cells. Curr Protoc Immunol. 2003:10.13.1–10.13.9.
46. Freeman GH, Halton JH. Note on an exact treatment of contingency, goodness of fit and other problems of significance. Biometrika. 1951; 38:141–9. [PubMed: 14848119]
47. 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; 287:4863–74. [PubMed: 22167191]



C

	1000 Genomes Project (n = 2504)	Control (n=50)			Rheumatoid Arthritis (n=50)			Significance levels of observed differences between n = 50 groups	Preliminary Cohort (n=100)		
GPSM3 Haploblock Genotype		M/M	M/m	m/m	M/M	M/m	m/m		M/M	M/m	m/m
Genotype Occurrence		33	16	1	30	17	3		52	38	10
Genotype Frequency		66%	32%	2%	60%	34%	6%		52%	38%	10%
	Minor Allele Frequency	Minor Allele Frequency			Minor Allele Frequency			Fisher's exact p-value	Minor Allele Frequency		
GPSM3 rs204989 / rs204991	23.3%	18.0%			23.0%			p = 0.4839	29.0%		
CCL21 rs2812378 [C]	30.1%	25.0%			30.0%			p = 0.5267	31.5%		
HLA region rs6457620 [G]	46.5%	45.0%			35.0%			p = 0.1938	56.0%		
Mean Age in Years (S.D.)		43 (15.3)			57.8 (11.1)			t=5.543 df=89.66 p<0.0001	28 (11.5)		
Sex (% Female)		58%			76%			p=0.0103	75%		

Figure 1.

GPSM3 transcript abundance in whole blood from individuals homozygous for the minor allele of the rs204989/rs204991 haploblock (the “m/m” genotype) is significantly lower than in whole blood from individuals homozygous for the major allele of the haploblock (the “M/M” genotype). (**A**) Schematic representation of identified polymorphisms with respect to the predicted transcription start site of the human *GPSM3* gene. (**B**) Scatter plot representation of *GPSM3* mRNA abundance, measured in whole blood by qRT-PCR, as normalized to the average level measured in M/M individuals. Individual results for homozygous major allele samples (M/M; n = 53) are depicted by solid circles, and homozygous minor allele samples (m/m; n=11) are depicted by open circles. An average decrease (Δμ) in *GPSM3* transcript abundance of 24.1% (t = 3.80, df = 62, p = 0.0003) was observed in homozygous minor allele samples. Error bars represent S.E.M. (**C**) Descriptive statistics of the three cohorts of volunteers genotyped in this study: Minor allele frequencies were 23% for rs204989 and rs204991 in the rheumatoid arthritis samples versus 18% in the disease-free controls (Fisher's exact test; p = 0.4839). Risk/minor allele frequency for rs2812378 was 30.0% in the rheumatoid arthris samples versus 25.0% in the disease-free controls (p = 0.5267). Risk/minor allele frequency for the *HLA* gene region biallelic SNP

rs6457620 was 35.0% in the rheumatoid arthritis samples and 45.0% in the disease-free matched controls (p = 0.1938).

A	Homozygous Ancestral *GPSM3* (M/M; n = 115)			Heterozygous *GPSM3* (M/m; n = 71)			Homozygous SNP *GPSM3* (m/m; n = 14)			Significance levels of observed differences
***HLA* SNP (rs6457620) Genotype**	**C/C**	**C/G**	**G/G**	**C/C**	**C/G**	**G/G**	**C/C**	**C/G**	**G/G**	
Genotype Occurrence	41	53	21	10	42	19	3	5	6	
Genotype Frequency	35.6%	46.1%	18.3%	14.1%	59.1%	26.8%	21.4%	35.7%	42.9%	
	Minor Allele Frequency			Minor Allele Frequency			Minor Allele Frequency			Fisher-Freeman-Halton Exact Test
***HLA* SNP (rs6457620) [G]**	41.3%			56.3%			60.7%			p = 0.0066

B	Rheumatoid Arthritis (n=50)			
***HLA* SNP (rs6457620) Genotype**	**C/C**	**C/G**	**G/G**	**Significance levels of observed differences**
M/M	14	15	1	
M/m	7	7	3	
m/m	0	1	2	
	GPSM3 Minor Allele Frequency			Fisher-Freeman-Halton Exact Test
rs204989 / rs204991	16.7%	19.6%	58.3%	p = 0.0123

C	Control (n=50)			
***HLA* SNP (rs6457620) Genotype**	**C/C**	**C/G**	**G/G**	**Significance levels of observed differences**
M/M	11	16	6	
M/m	2	13	1	
m/m	0	0	1	
	GPSM3 Minor Allele Frequency			Fisher-Freeman-Halton Exact Test
rs204989 / rs204991	7.7%	22.4%	18.8%	p = 0.2739

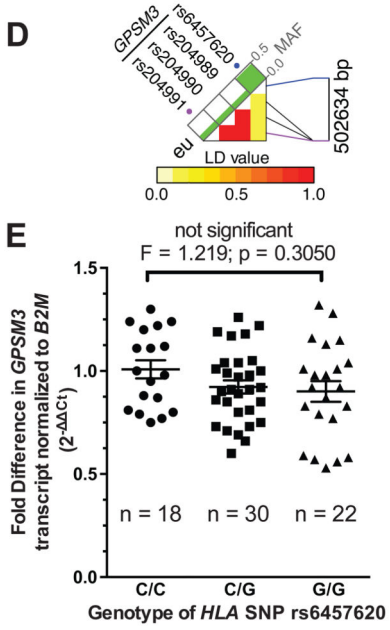


Figure 2.

Linkage of the *GPSM3* SNP haploblock 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 the current study. These data are stratified by *GPSM3* SNP haplotype, showing that the rs6457620 [G] risk allele frequency amongst homozygous ancestral *GPSM3* haploblock individuals (M/M) is 41.3%, heterozygous individuals (M/m) is 56.3%, and homozygous minor *GPSM3* haploblock (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 haploblock 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 minor allele frequencies (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 haploblock (rs204989, rs204990, rs204991) and the *HLA* gene region SNP rs6547620, as obtained using the NIEHS SNPinfo webserver (http://snpinfo.niehs.nih.gov/) from NCBI dbSNP data (ref. [25]). (**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 ANOVA with Bonferroni *post-hoc*.

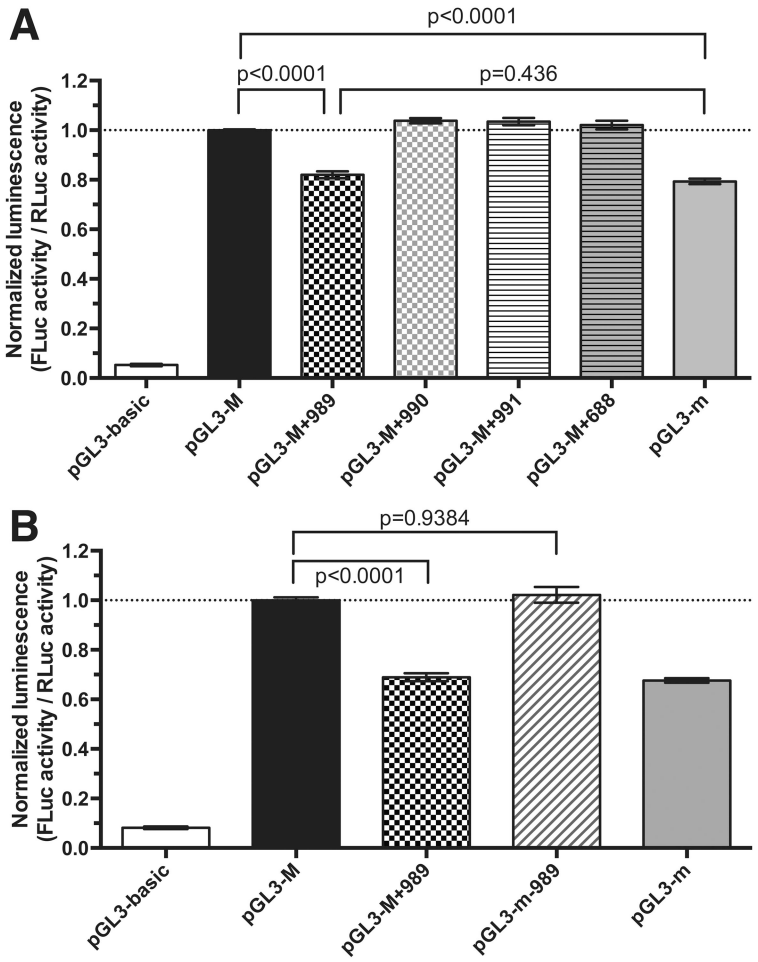
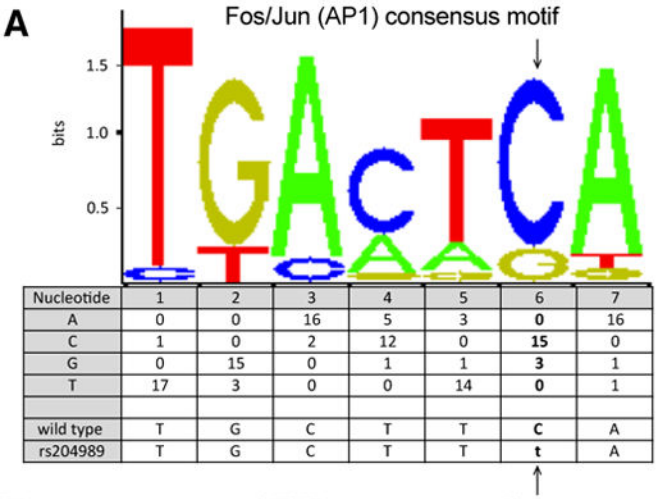
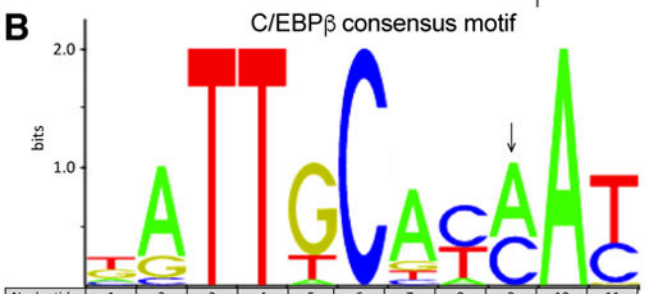


Figure 3.

Functional consequences of SNPs rs204989, rs204990, rs204991, and rs3096688 on *GPSM3* promoter activity. (**A**) Functional results upon subcloning a 3.5 kb-region immediately 5′ to the *GPSM3* transcription start site from M/M and m/m genotype volunteers, leading to wild type (pGL3-M) and "minor"/polymorphic (pGL3-m) promoter-driven expression of firefly luciferase in HEK293T cells. Luciferase activity is reported as a ratio of *GPSM3* promoter-dependent firefly luciferase activity (FLuc) to control *Renilla* luciferase activity (pRL-TK control vector) and normalized to the average level measured from the wild type pGL3-M vector (set to 1.00; dotted line). Introduction of independent minor alleles (denoted "+###") 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 (*i.e.*, "pGL-m-989") restores wild type *GPSM3* promoter activity. All data are compiled from three independent experiments. Error measure is S.E.M. Significance was determined using one-way ANOVA after controlling for multiple comparisons with Dunnett's test using pGL3-M as the defined control.



Nucleotide	1	2	3	4	5	6	7
A	0	0	16	5	3	**0**	16
C	1	0	2	12	0	**15**	0
G	0	15	0	1	1	**3**	1
T	17	3	0	0	14	**0**	1
wild type	T	G	C	T	T	**C**	A
rs204989	T	G	C	T	T	**t**	A



Nucleotide	1	2	3	4	5	6	7	8	9	10	11
A	13 006	75 198	0	0	4 556	0	74 715	8 654	**60 151**	99 494	0
C	10 026	5 868	0	0	0	99 494	5 478	51 954	**39 343**	0	36 038
G	33 617	18 428	0	0	75 531	0	10 015	0	**0**	0	2 043
T	42 845	0	99 494	99 494	19 407	0	9 286	38 886	**0**	0	61 413
wild type	A	A	G	T	G	C	T	T	**C**	A	T
rs204989	A	A	G	T	G	C	T	T	**t**	A	T

Figure 4.

Putative transcription factor binding sites within the human *GPSM3* promoter potentially disrupted by the minor allele of SNP rs204989 (arrows), as predicted by the JASPAR database: (**A**) Fos/Jun (AP1) heterodimer; (B) C/EBP-β.

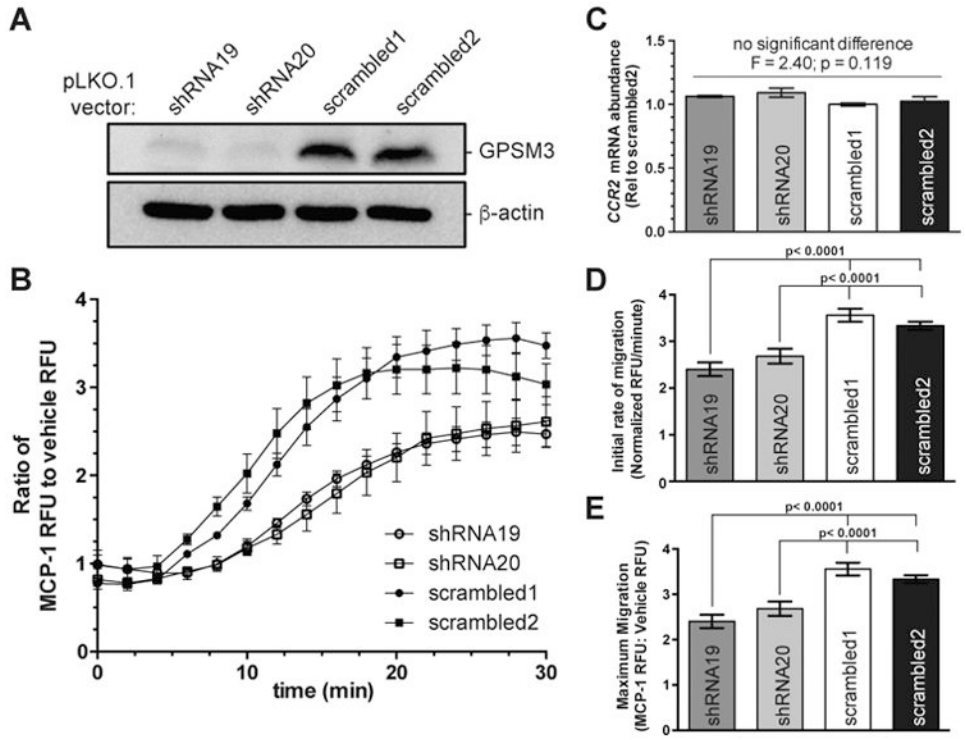


Figure 5.

Lentiviral-mediated shRNA knockdown of endogenous GPSM3 expression in the human monocytic THP-1 cell line disrupts migration toward MCP-1 (**A**) Western blot showing whole lysate (40 μg) immunoreactivity with mouse monoclonal anti-GPSM3 antibody 35.5.1 (ref. [47]) and with anti-β-actin as a loading control. (**B**) Real-time Transwell migration analysis of indicated calcein-AM stained THP-1 cell lines as shown by change in the ratio of MCP-1-induced to vehicle-induced relative fluorescence units (RFU) over a 30 minute timeframe. (**C**) qRT-PCR data from indicated stable knockdown or scrambled control THP-1 cell line total RNA preparations. There was no significant difference in *CCR2* mRNA abundance in the THP-1 cell lines as determined by one-way ANOVA (F = 2.40; p = 0.119). (**D**) Regression analysis-derived rate of THP-1 cell line transmigration over initial migratory period; differences between each cell line was calculated by one-way ANOVA. (**E**) Regression analysis-derived maximum THP-1 cell transmigration over 30 minute timecourese; differences between each cell line were calculated by one-way ANOVA. All data are reported as means with error bars representing S.E.M.