Rare-variant pathogenicity triage and inclusion of synonymous variants improves analysis of disease associations of orphan G protein–coupled receptors

Ridge Dershem, Raghu P. R. Metpally, Kirk Jeffreys, Sarathbbaru Krishnamurthy, Diane T. Smelser, Michal Hershfinkel, Regeneron Genetics Center, David J. Carey, Janet D. Robishaw, and Gerda E. Breitwieser

From the Department of Molecular and Functional Genomics, Geisinger, Weis Center for Research, Danville, Pennsylvania 17822, the Faculty of Health Sciences, Ben-Gurion University of the Negev, Beer-Sheva, 8410501 Israel, the Regeneron Pharmaceuticals, Inc., Tarrytown, New York 10591, and the Schmidt College of Medicine, Florida Atlantic University, Boca Raton, Florida 33431

The pace of deorphanization of G protein–coupled receptors (GPCRs) has slowed, and new approaches are required. Small molecule targeting of orphan GPCRs can potentially be of clinical benefit even if the endogenous receptor ligand has not been identified. Many GPCRs lack common variants that lead to reproducible genome-wide disease associations, and rare-variant approaches have emerged as a viable alternative to identify disease associations for such genes. Therefore, our goal was to prioritize orphan GPCRs by determining their associations with human diseases in a large clinical population. We used sequence kernel association tests to assess the disease associations of 85 orphan or understudied GPCRs in an unselected cohort of 51,289 individuals. Using rare loss-of-function variants, missense variants predicted to be pathogenic or likely pathogenic, and a subset of rare synonymous variants that cause large changes in local codon bias as independent data sets, we found strong, phenotype-wide disease associations shared by two or more variant categories for 39% of the GPCRs. To validate the bioinformatics and sequence kernel association test analyses, we functionally characterized rare missense and synonymous variants of GPR39, a family A GPCR, revealing altered expression or Zn2+–mediated signaling for members of both variant classes. These results support the utility of rare variant analyses for identifying disease associations for GPCRs that lack impactful common variants. We highlight the importance of rare synonymous variants in human physiology and argue for their routine inclusion in any comprehensive analysis of genomic variants as potential causes of disease.

The superfamily of G protein–coupled receptors (GPCRs) translates extracellular signals from light, metabolites, and hormones into cellular changes, which makes them the targets of a significant fraction of drugs currently on the market. Genome-wide association studies (GWASs) on common variants in GPCRs have begun to identify their contributions to various disease processes. However, many GPCRs lack functionally important common variants, and alternate strategies are needed to understand their roles in human physiology. Recently, sequence kernel association tests (SKATs) have been applied to rare variants to assess disease associations. Initial approaches binned all rare variants in a genomic region or gene, but current methods group the rare variants most likely to contribute to disease associations or aggregate variants based on domain or family structures. Large-scale studies of orphan or understudied GPCRs to characterize natural genetic variation in the human population can provide insights into the biological function and/or potential causal contributions to disease processes. The DiscoveHR collaboration represents a tremendous resource because whole exome sequences can be linked to the electronic health record (EHR) in a deidentified manner. Here we used whole exome sequences and clinical information from 51,289 individuals in the DiscoveHR cohort. We identified sequence variants (common variants, mean allele frequency (MAF) ≥ 1%; and rare variants, MAF < 1%) for 85 orphan or understudied GPCRs not amenable to GWAS studies. Rare variants were binned into functional classes: loss-of-function (LOF, truncation and frameshift) variants; missense variants with predicted pathogenicity; or rare synonymous variants with altered codon bias. We performed independent SKAT analyses on the various functional classes to determine their disease associations. Remarkably, for those GPCRs with sufficient numbers of individuals for statistical analyses, we found that the top disease associations were common to all available variant classes in the GPCR.

For initial in vitro assessment of the validity of the disease associations, we focused on GPR39. Among its particular advantages, GPR39 had sufficient numbers of variants in all three functional classes to be amenable to rare variant approaches. Second, GPR39

This work was supported by Geisinger. The authors declare that they have no conflicts of interest with the contents of this article.

1 Present address: Dept. of Biomedical Science, Charles E. Schmidt College of Medicine, Florida Atlantic University, Boca Raton, FL 33431.
2 To whom correspondence should be addressed. Present address: Dept. of Biomedical Science, Charles E. Schmidt College of Medicine, Florida Atlantic University, 777 Glades Rd., BC71, Boca Raton, FL 33431. Tel.: 561-297-0060; E-mail: gbweitwieser@health.fau.edu.
3 The abbreviations used are: GPCR, G protein–coupled receptor; GWAS, genome-wide association study; SKAT, sequence kernel association test; LOF, loss of function; MISS, missense; EHR, electronic health record; MAF, mean allele frequency; pLP, predicted likely pathogenic; pP, predicted pathogenic; PheWAS, phenome-wide association studies; iP, inositol trisphosphate; IP, inositol monophosphate; CI, confidence interval; pB, predicted to be benign; WES, whole exome sequences; ASCVD, atherosclerotic cardiovascular disease.

© 2019 Dershem et al. Published under exclusive license by The American Society for Biochemistry and Molecular Biology, Inc.
**Missense and synonymous variants alter GPCR function**

is a small family A GPCR, allowing rapid generation of mutants. Third, GPR39 is activated by extracellular \( \text{Zn}^{2+} \) and is coupled to inositol phosphate production (10, 11), permitting straightforward functional analyses. Finally, we have only a rudimentary understanding of the role(s) of GPR39 in human physiology, despite its broad expression and multiple cellular functions (12, 13); thus, results may provide insights into the contribution(s) of GPR39 variation to disease. Our results demonstrate the validity of combined computational and functional approaches to provide important insights into the clinical contributions of orphan and understudied GPCRs, and focus attention on the importance of rare synonymous variants for identifying disease associations. The overall strategy can be readily applied to other classes of genes without clinically important common variants.

**Results**

**Rare variant-based disease associations of orphan or understudied GPCRs**

We identified rare variants in whole exome sequences from 51,289 individuals. To predict their functional impact, variants were annotated and sorted into three classes in order of predicted severity, i.e. LOF (premature stop codon, loss of a start or stop codon, or disruption of a canonical splice site), missense variants (predicted to be likely pathogenic (pLP) or pathogenic (pP) by RMPath scoring, derived from application of multiple pathogenicity algorithms (see “Experimental procedures”)), and synonymous variants with significantly different codon frequency relative to the reference codon (termed SYN\_\Delta CB, i.e. synonymous variants with altered codon bias).

Determining disease associations for genes having only low-frequency variants requires binning of variants to increase statistical power. Binning methods have focused on gene or regulatory regions and/or pathways by incorporating biological information from curated knowledge databases (14). For this study, we were specifically looking for clinically impactful rare coding variants that could be validated by functional studies of the relevant GPCRs. We binned variants according to the functional classes described above and used SKAT, focusing on the disease associations that were common to more than one class of rare variant. **Fig. 1** illustrates the analysis pipeline for the entire set of 85 GPCRs. Of the 72 GPCRs with sufficient variants to allow at least 2 variant classes to be analyzed by SKAT, 28 yielded significant disease associations. Because this was an exploratory analysis, we used a significance cutoff of \( p < 0.05 \), i.e. \( p \) values \( < 6.2E-07 \) after Bonferonni correction for multiple testing. **Table 1** shows the top disease associations for the 13 GPCRs that had sufficient variants in the LOF class. Supporting validity, associations between GPR37L1 and epilepsy (15) and LGR5 and alopecia (16, 17) have been previously identified by other means. Attesting to the discovery potential, other disease associations were novel, including GPR161 with hyperhidrosis, LGR4 with Sicca syndrome, GPR153 with hyperpotassemia, GPR84 with anemia, and GPR39 with peripheral nerve damage, among others. Notably, the phenome-wide disease associations for the predicted LOF class showed congruence across the other variant classes for the subset of GPCRs having sufficient missense or SYN\_\Delta CB variants for independent analyses. Although truncation and frameshift variants are easily identified and likely to have the greatest functional impact, missense variants classed as pLP + pP can also have a significant impact on GPCR function. For the subset of GPCRs without sufficient LOF variants, we found significant disease associations for the missense and SYN\_\Delta CB variants (**Table 1**). Some of the associations validate previous studies, including GPR183 and disorders of liver (18) and GPR85 with acute myocardial infarction (19). Other associations were novel and represent potential targets for future study, including GPR132 and disturbances in sulfur-bearing amino acid metabolism, GPR176 with asthma, GPR12 with epilepsy, and GPR3SD with renal osteodystrophy. Of note, the SYN\_\Delta CB variant class had significant disease associations in common with LOF and/or missense variants and most consistently had sufficient variants to be included in the analysis. Parallel SKAT analyses of the three distinct classes of rare variants produced two important results. First, independent analysis of distinct classes of variants yielded concordant disease associations, increasing confidence in the results. Second, the most consistent source of phenome-wide disease associations was SYN\_\Delta CB. Altogether, these results provide a strong rationale for including all three functional variant classes in disease association analyses.

---

**Figure 1. Flow diagram for orphan/understudied GPCR SKAT analysis.** Rare variants in 85 GPCRs were identified in the 51,289 DiscovEHR WES data set and sorted into LOF, missense, and synonymous categories. Missense variants were further triaged to identify those predicted to be pLP or pP by RMPath scoring and the subset of synonymous variants with large changes in codon usage bias (SYN\_\Delta CB; see “Experimental procedures”). GPCRs with at least 2 categories of variant having \( \geq 7 \) distinct members were further analyzed (72 of 85 GPCRs). SKAT analysis was run independently on 123 sets of variants against 656 disease codes having \( \geq 200 \) individuals with \( \geq 3 \) independent encounters in the EHR.
and classified as described in Table 1. The common variants were

By definition, all LOF, missense, and synonymous variants were analyzed

used for Phenome-wide Association Studies (PheWAS)

51,289 individuals in the DiscovEHR cohort (9) (Fig. 2). None of

agonist and dominant signaling pathways are known (12, 13).

ized in cellular and knockout mouse models and for which both

physiology. The first step in defining a causal role for these rare

regarding the potential contributions of GPCRs to human

impact GPR39 function (Fig. 4). The three pP variants (S78L,

GPR39 common and rare variants in the 51,289 WES

DiscovEHR cohort

Disease associations provide a rich source of hypotheses regarding the potential contributions of GPCRs to human physiology. The first step in defining a causal role for these rare variants, however, is functional validation of the SKAT results. Here we focused in particular on rare synonymous variants with altered codon frequency (SYN_ΔCB), whose functional consequences are not well-understood. Many of the GPCRs have no known agonist, and their dominant signaling pathways have not been characterized (Table 1). We therefore chose to focus on GPR39, the Zn²⁺ receptor, which has been characterized in cellular and knockout mouse models and for which both agonist and dominant signaling pathways are known (12, 13).

All common and rare variants in GPR39 were identified in 51,289 individuals in the DiscovEHR cohort (9) (Fig. 2). None of the GPR39 variants were found in the ClinVar database. Accordingly, all LOF, missense, and synonymous variants were analyzed and classified as described in Table 1. The common variants were used for Phenome-wide Association Studies (PheWAS) and included a missense variant (2:133174764, pA50V) predicted to be benign, a synonymous variant (2:133174999; pT128T), which changed a common to the rarest codon (acA (28%) to acG (12%)), and a noncoding variant (2:133402607). Notably, no phenome-wide disease associations were detected (Fig. 3A).

The rare variants used for SKAT, i.e. LOF, missense (pLP + pP), and SYN_ΔCB variants, were distributed throughout the coding region of GPR39 (Fig. 4). Examination of the 28 missense variants that were categorized as pLP and pP confirmed that the RMPath pipeline identified variants which are likely to impact GPR39 function (Fig. 4). The three pP variants (S78L, C108R, and R133G) included a cysteine that participates in a disulfide bond at the extracellular face of GPR39 (20). The pLP variants included one that contributes to the Zn²⁺-binding site (H19R) (21), two variants in a disulfide bond cysteine (C191G and C191W) (20), a residue in the highly conserved NPXY motif common to family A GPCRs (N340S) (22), and a variant distal to a putative palmitoylation site in the proximal C terminus (R352Q). Other residues in the pLP category introduce charged or proline residues into transmembrane helices or extracellular loops (S49N, M87I, T122M, A167S, P179T, R193C, S222Y, T291I, A293P, R302Q, A307V, P310A, Y314R,

| GPCR | ICD9 | Description | LOF | MISS | SYN_ΔCB |
|------|------|-------------|-----|------|---------|
| LGR4 | 710.2 | Sicca syndrome | 1.58E-31 | 1.28E-13 | 2.31E-14 |
| GPR150 | 372.14 | Other chronic allergic conjunctivitis | 2.59E-31 | 1.45E-32 | 5.48E-07 |
| GPR153 | 276.7 | Hyperpotassemia | 2.32E-30 | 8.73E-16 | 1.14E-22 |
| GPR149 | 374.3 | Phytos of eyelid | 5.89E-30 | 4.62E-08 | 1.53E-07 |
| GPR161 | 780.8 | Generalized hyperhidrosis | 3.78E-26 | × | 9.03E-27 |
| GPR37L1 | 345.1 | Generalized convulsive epilepsy,without intractable epilepsy | 2.04E-18 | 1.92E-19 | 4.95E-20 |
| GPR84 | 280.8 | Other specified iron deficiency anemias | 2.31E-18 | 2.99E-08 | 2.07E-15 |
| GPR39 | 354.2 | Lesion of ulnar nerve | 8.04E-15 | 1.62E-09 | 8.69E-15 |
| GPR6 | 583.81 | Nephritis and nephropathy, not specified as acute or chronic | 4.52E-11 | × | 1.31E-19 |
| GPR5 | 704 | Alopecia NOS | 5.73E-13 | × | 1.36E-12 |
| GPR1 | 287.5 | Thrombocytopenia NOS | 1.5E-09 | 1.26E-07 | 1.5E-09 |
| GPR18 | 583.81 | Nephritis and nephropathy, not specified as acute or chronic | 2.61E-08 | 1.01E-08 | 3.77E-09 |
| GPR179 | 270.4 | Disturbances of sulfur-bearing amino acid metabolism | 1.38E-07 | × | 1.05E-07 |
| GPR176 | 493.91 | Asthma, unspecified type, with status asthmaticus | × | 1.35E-33 | 6.03E-26 |
| GPR132 | 216.3 | Disturbances of sulfur-bearing amino acid metabolism | × | 4.13E-30 | 9.61E-24 |
| GPR85 | 410.1 | Acute myocardial infarction of other anterior wall | × | 2.06E-26 | 1.33E-35 |
| GPR183 | 573.8 | Other specified disorders of liver | × | 1.34E-19 | 1.5E-19 |
| GPR12 | 345.1 | Generalized convulsive epilepsy, without intractable epilepsy | × | 5.64E-14 | 5.48E-14 |
| GPR135 | 350.1 | Trigeminal neuralgia | × | 1.91E-12 | 1.99E-12 |
| GPR26 | 372.14 | Other chronic allergic conjunctivitis | × | 4.56E-12 | 4.55E-12 |
| MRGPR1 | 401.1 | Benign essential hypertension | × | 2.17E-10 | 8.61E-08 |
| GPR15 | 455.6 | Hemorrhoids NOS | × | 1.61E-10 | 1.97E-10 |
| GPR83 | 250.8 | Diabetes mellitus type II (non-insulin-dependent) | × | 1.29E-09 | 5.71E-12 |
| GPR5D | 588 | Renal osteodystrophy | × | 7.5E-09 | 2.0E-20 |
| GPCR8 | 381.8 | Dysfunction of Eustachian tube | × | 3.17E-09 | 2.76E-10 |
| GPR141 | 786.07 | Wheezing | × | 1.82E-07 | 2.51E-09 |
| GPR6 | 214.1 | Lipoma of other skin and subcutaneous tissue | × | 2.04E-07 | 2.04E-07 |
| GPR31 | 41.02 | Group B Streptococcus infection | × | 2.04E-07 | 2.04E-07 |

Figure 2. Pathogenicity triage of 173 coding variants in GPR39 identified in 51,289 WES. Common variants with MAF 1% were categorized as likely benign (MAF between 1 and 5%) and benign (MAF 5%). All variants having MAF of <1% were sorted to synonymous, LOF (frameshift or premature stop codon), and missense. Missense variants were thus variants of unknown significance (VUS) and subjected to bioinformatics pathogenicity triage using RMPath scoring ("Experimental procedures") and classified as pB, pLB, pLP, or pP and the subset of SYN variants having a significant change in codon usage bias (SYN_ΔCB).
R320Y, S336R, R352Q, R362C, H367R, R386C, and R390C). We also identified the subset of 21 variants that introduced at least a 2-fold change in site-specific codon frequency (SYN\_H9004\_CB; denoted by asterisks in Fig. 4).

Disease associations of GPR39 rare variant classes

The initial SKAT screen for GPR39 disease associations compared top associations for the three independent classes of rare variants. To more fully explore the results for GPR39 and increase power to identify all significant disease associations, we combined all individuals heterozygous for rare LOF, pLP, and pP variants (termed LOF/MISS). Analysis was restricted to only those ICD9 codes with \( \geq 200 \) individuals independent of genotype (23), and the results are plotted against ICD9 codes in Fig. 3B. SKAT analysis was also run on rare SYN\_\Delta CB variants, plotted in Fig. 3C. Both the LOF/MISS and SYN\_\Delta CB analyses had a significant number of phenome-wide associations in common. All results for both variant classes that were above the significance cutoff for multiple testing (dotted lines in Fig. 3, B and C) are listed in Table 2. Top associations for GPR39 included lesion of the ulnar nerve and benign prostate hyperplasia with lower urinary tract obstruction. Note that individuals who had both rare LOF/MISS and SYN\_\Delta CB variants were excluded from analyses, because the sequence data are not phased, and we therefore cannot distinguish between 1 allele having both variants versus a variant on each allele. Also of note in Table 2, some significant disease associations were specific to either the LOF/MISS or SYN\_\Delta CB variant classes. To assess the importance of rare variant pathogenicity triage for binning prior to SKAT analysis, we combined all rare (MAF \( \leq 1\% \)) variants (LOF + missense + synonymous = 168 variants) and ran SKAT. The only significant disease association was upper quadrant abdominal pain (ICD9 789.02; \( p \) value 8.98E\(-06\); significance
Figure 4. GPR39 topology indicating rare variants identified in 51,289 WES. The GPR39 sequence was color-coded to indicate reference codon usage frequency. For all amino acids having multiple potential codons, red indicates the most common codon, with progression from common to rare colored as gradations to light pink (rare). Asterisks indicate the locations of rare synonymous variants exhibiting the largest changes in codon bias, either from common to rare or rare to common. indicates rare LOF variants (truncations or frameshifts), and @ indicates locations of rare missense variants combined for SKAT analysis.
**Validation of the functional impact of rare GPR39 variants**

Point mutations can cause changes in GPR39 function by altering protein expression and/or by altering essential features of receptor activation. Therefore, to explore the functional basis for the observed disease associations, we generated the rare missense and SYN_{ACB} variants in the reference human GPR39 background. The cDNA construct included an N-terminal 3X FLAG epitope followed by a minimal bungarotoxin-binding site, for ease of expression analysis and subcellular localization, respectively. Expression in HEK293 cells of equivalent amounts of cDNA and subsequent Western blotting of lysates showed significant differences in net expression levels of the rare missense and SYN_{ACB} variants. Fourteen missense variants had reduced and one had significantly increased expression relative to WT GPR39 (Fig. 5, A and C). Rare missense variants with altered expression were distributed throughout GPR39 structural domains as color-coded in the quantitation of multiple blots (Fig. 5C). Likewise, SYN_{ACB} variant expression was significantly greater than WT GPR39 for three variants, whereas expression of fourteen variants was significantly less than WT (Fig. 5B), as quantified in Fig. 5D. Surprisingly, despite the change from rare to common codon frequency for the majority of SYN_{ACB} variants, only four showed increased expression relative to WT. Four variants went from common to rare codons, denoted by asterisks on the x axis in Fig. 5D. Contrary to expectations, Gln-265 was expressed at significantly higher levels than WT (~150%), despite the large decrease in codon frequency. Overall, results argue that factors in addition to site-specific codon usage bias determine net expression levels for rare synonymous variants.

GPR39 signaling is complex, with multiple signaling outputs. However, in both endogenous and heterologous expression systems, GPR39 activates inositol trisphosphate (IP_{3})–dependent release of intracellular Ca^{2+} (10, 11, 24). The first step in the signaling pathway is G_{q}-mediated activation of phospholipase C, which produces diacylglycerol and IP_{3}. Here we exposed HEK293 cells transiently expressing WT or variant GPR39 to ZnCl_{2} and used a FRET assay to quantify accumulation of inositol monophosphate (IP_{1}) in the presence of LiCl. Fig. 6A illustrates normalized maximal levels of IP_{1} for WT and representative missense variants having decreased (R386C) or increased (W314R) EC_{50} for activation by ZnCl_{2}. Fig. 6C plots the EC_{50} values and 95% confidence intervals (CIs) for WT and all other missense variants, color-coded for their locations in GPR39 domains. Three variants, S78L, T122M, and R133G, had no measurable activity, likely accounted for by low levels of expression (recall Fig. 5, A and C), whereas M87L was well...

---

**Table 2**

SKAT results for rare GPR39 variants in 51,289 DiscovEHR cohort

| Code     | Description                                      | LOF + pLP + pP | SYN_{ACB} |
|----------|--------------------------------------------------|----------------|-----------|
| 354.2    | Lesion of ulnar nerve                            | 6.2E-08        | 8.69E-15  |
| 995.3    | Allergy NOS                                       | 6.67E-06       | 7.22E-14  |
| 600.21   | Benign prostate hyperplasia                       | 1.24E-07       | 2.57E-09  |
| 362.56   | Macular puckering of retina                       | ×              | 7.65E-07  |
| 786.6    | Swelling, mass, or lump in chest                  | 3.83E-05       | 2.4E-07   |
| 704.8    | Other specified diseases of hair and hair follicles | ×              | 6.72E-07  |
| 696      | Psoriatic arthropathy                             | 3.24E-13       | 1.03E-06  |
| 727.05   | Other tenosynovitis of hand or wrist              | ×              | 1.78E-06  |
| 372.3    | Conjunctivitis NOS                                | 1.68E-05       | 2.75E-06  |
| 998.59   | Other postoperative infection                     | ×              | 5.0E-06   |
| 401.1    | Benign essential hypertension                     | ×              | 9.55E-06  |
| 70.54    | Chronic hepatitis C without mention of hepatic coma | ×              | 1.09E-05  |
| 655.83   | Other known or suspected fetal abnormality        | 9.71E-06       | 1.41E-05  |
| 591      | Hydrophoreosis                                    | ×              | 2.06E-05  |
| 110.4    | Dermatophytosis of foot                           | 1.62E-06       | ×         |
| 726.2    | Other affections of shoulder region NOS           | 4.53E-06       | ×         |
| 429.2    | Cardiovascular disease                           | 2.26E-05       | ×         |
| 707.19   | Ulcer of lower limb                              | 2.46E-05       | ×         |

---

**Table 3**

SKAT analysis of all rare synonymous GPR39 variants in 51,289 DiscovEHR cohort

| Code     | Description                                      | p value |
|----------|--------------------------------------------------|---------|
| 696      | Psoriatic arthropathy                             | 3.87E-11|
| 354.2    | Lesion of ulnar nerve                            | 1.08E-07|
| 110.4    | Dermatophytosis of foot                           | 8.34E-07|
| 726.2    | Other affections of shoulder, NOS                 | 2.24E-06|
| 995.3    | Allergy, NOS                                      | 3.51E-06|
| 655.83   | Other known or suspected fetal abnormality        | 4.38E-06|
| 707.19   | Ulcer of lower limb                              | 1.06E-05|
| 786.6    | Swelling, mass, or lump in chest                  | 2.06E-05|
| 429.2    | Cardiovascular disease (ASCVD)                   | 3.03E-05|
| 649.13   | Obesity complicating pregnancy, childbirth, antepartum | 3.58E-05|
| 600.21   | Benign prostate hyperplasia with lower urinary tract symptoms | 3.79E-05|
| 440.9    | Atherosclerosis NOS                               | 6.31E-05|
| 649      | Tobacco use disorder complicating pregnancy, childbirth | 6.8E-05|
| 368.2    | Diplopia                                          | 7.46E-05|
expressed but inactive. One variant, R193C, had a dose-response curve that was significantly right-shifted, precluding determination of EC$_{50}$ or $V_{\text{max}}$, despite expression levels not significantly different from WT. Likewise, Fig. 6B illustrates the normalized dose-response relations for WT GPR39 and representative SYN$_{\text{ACB}}$ variants. Expression levels varied and to ensure quantitation relative to WT, blots were segregated for low (upper) and high (lower) expressing variants. The blots were cut (dotted line) to allow simultaneous probing for GPR39 and GAPDH. C, quantitation of expression levels of GPR39 rare missense variants normalized to WT (dotted line). The data are presented as percentages of WT $\pm$ S.D. ($n$ = 2–4 independent transfections), with all points plotted. D, quantitation of expression levels of GPR39 rare SYN$_{\text{ACB}}$ variants normalized to WT (dotted line). The details are the same as in C. Statistical analysis for both figures was two-tailed t test relative to WT. *, $p < 0.05$;**, $p < 0.01$;***, $p < 0.001$; ****, $p < 0.0001$.

Figure 5. Protein expression of rare missense and synonymous GPR39 variants. A, representative Western blots of rare missense variants identified in the 51,289 WES cohort. Western blots of HEK293 cells lysates expressing WT or rare variant missense cDNAs were run as described under “Experimental procedures.” Each blot had HEK and WT for background subtraction and normalization, respectively. The blots were probed for GAPDH to confirm equal loading. B, representative blots for HEK293 cells expressing WT and rare SYN$_{\text{ACB}}$ variants. Expression levels varied and to ensure quantitation relative to WT, blots were segregated for low (upper) and high (lower) expressing variants. The blots were cut (dotted line) to allow simultaneous probing for GPR39 and GAPDH. C, quantitation of expression levels of GPR39 rare missense variants normalized to WT (dotted line). The data are presented as percentages of WT $\pm$ S.D. ($n$ = 2–4 independent transfections), with all points plotted. D, quantitation of expression levels of GPR39 rare SYN$_{\text{ACB}}$ variants normalized to WT (dotted line). The details are the same as in C. Statistical analysis for both figures was two-tailed t test relative to WT. *, $p < 0.05$;**, $p < 0.01$;***, $p < 0.001$; ****, $p < 0.0001$.

Missense and synonymous variants alter GPCR function

J. Biol. Chem. (2019) 294(48) 18109 –18121 18115

Alexa594 binding on nonpermeabilized cells (Fig. 6, E and H), and generally correlated with the extrapolated $V_{\text{max}}$ values for the IP$_{3}$ assays (Fig. 6, D and G). The data were corrected for the very low levels of unspecific BgTx–Alexa 594 binding to untransfected HEK293 cells (6.03 ± 1.8% of WT ($n$ = 6 independent experiments)). Total (Western blotting) and plasma membrane localization (BgTx binding) were comparable for many of the missense variants (H19R, A167S, S222Y, A293P, R302Q, P310A, R193C, and S336R), but total expression was greater than plasma membrane localization for five variants (M871, C191W, R193C, A307V, and S336R), suggesting potential alterations in trafficking, whereas the remaining 12 variants had reduced total expression relative to plasma.
membrane localization, suggesting a potential increase in quality control-mediated degradation at the endoplasmic reticulum.

SYN/H9004 CB plasma membrane localization varied more widely relative to WT, with a subset of variants having significantly elevated plasma membrane localization (Fig. 6H), which was not a reflection of increased expression assessed by Western blotting (Fig. 5D). Six variants had proportional total and plasma membrane localization (Ser-78, Ala-321, Ser-248, Leu-363, Pro-391, and Ala-394), whereas three had higher total expression versus plasma membrane localization (Ser-147, Gln-265, and Gly-4370), and the remainder had reduced total expression but elevated plasma membrane targeting. The most striking SYN/H9004 CB variant, Gln-265, had an increased EC_{50} value relative to WT (Fig. 6F), increased V_{max}, and surface localization relative to WT (Fig. 6G and H), as well as increased total expression (Fig. 5D). Comparing total and plasma membrane expression of rare missense and SYN/H9004 CB variants thus identified a subset of high-value variants that likely have significantly altered biosynthesis and/or trafficking, which could be confirmed by mechanistic studies.

Overall, the results demonstrate a striking parallel in functional impact on expression and IP_{1} signaling for a subset of both rare missense and SYN/H9004 CB variants of GPR39, rationalizing the disease associations identified for both classes of variants. The congruence of top disease associations between rare LOF, pathogenic missense, and SYN/H9004 CB variants for a significant subset of GPCRs (recall Table 1) argues strongly for the utility of these analytical approaches for identification of functionally impactful rare variants and the

Figure 6. Functional analysis of rare missense and synonymous GPR39 variants. A, IP_{1} responses to various concentration of ZnCl_{2} for all variants that expressed ≥50% of WT. Representative curves and fits for WT, W314R, and R386C are illustrated. Plotted are the means ± S.D. of three independent experiments, and lines show fits to Michaelis–Menten equation. The dotted line indicates 95% CI of WT. B, IP_{1} responses of representative SYN\_ΔCB variants showing altered EC_{50}, compared with WT, Ile-29, and Val-354. The details are as in A. C, EC_{50} values for dose-response relations of all functional rare missense variants (as in A), plotted as means ± 95% CI for n = 3 independent experiments. The dotted lines indicate the 95% CI range for WT. D, extrapolated V_{max} of Michaelis–Menten fits to IP_{1} dose-response relations of rare missense variants, plotted as mean of three independent experiments plus 95% CI. The dotted lines indicate the 95% CI range for WT. E, bungarotoxin binding was used to assess surface localization of rare missense variants, plotted as mean plus all points, normalized to WT (dotted line), and corrected for unspecific binding in untransfected HEK293 cells. F, EC_{50} values of all rare SYN\_ΔCB variants with measurable activity, analyzed and plotted as in C. G, extrapolated V_{max} of fits to IP_{1} dose-response relations for rare SYN\_ΔCB variants, analyzed, and plotted as in D. H, bungarotoxin binding to assess surface localization of rare SYN\_ΔCB variants, analyzed, and plotted as in E. For all plots, *, p < 0.05; **, p < 0.01; ***, p < 0.001; ****, p < 0.0001, by two-tailed t test relative to WT, assuming unequal variances.
imperative to include rare synonymous variants in genomic and functional analyses.

Discussion

The related GWAS and PheWAS approaches, which utilize common variants, have produced strong, validated targets for drug development (2). However, many genes do not have functionally important common variants, and additional approaches are required. The advantages of rare coding variant approaches are 2-fold. First, rare variants can have stronger effects on protein function, and second, assessment of the impact of rare coding variants on protein expression and function is straightforward. However, because of their low frequency, it is necessary to combine rare coding variants to increase the statistical power for detecting disease associations.

In this study, we used pathogenicity triage as a basis for rare variant binning, based on the hypothesis that aggregating variants with similar directions of effect on protein function would amplify disease associations. We considered, in order of the strength of their potential effects, predicted LOF variants, predicted likely pathogenic or pathogenic missense variants, and synonymous variants with altered codon frequency bias. As a test of the importance of rare variant triage on disease associations, we compared agnostic binning of all rare variants in HEK293 cells, these variants had the expected effects on EC50 values and maximal activities. Of the 47 synonymous variants, the codon usage table identified 21 variants having significantly altered codon usage frequency, and many had significant effects on expression, surface localization, and Zn2+–induced IP3 production. Surprisingly, the distributions of rare missense and SYN_ΔCB variants in the GPR39 structural domains were distinct. The majority of missense variants were localized to the transmembrane domain (77%), half in the extracellular loops (10 of 26), one in a transmembrane helix, and the remainder in the intracellular loops (9 of 26). In contrast, ~50% of SYN_ΔCB variants were in the C terminus; the remainder were distributed among transmembrane helices (3 of 20), intracellular (2 of 20), and extracellular (4 of 20) loops. The C-terminal variants in both classes had significant effects on expression and signaling, primarily reducing Vmax. In relation to many family A GPCRs, the C terminus of GPR39 is relatively large (108 amino acids) and has 22 putative phosphorylation sites. GPR39 signaling outputs are regulated by agonist-toggled binding of protein kinase inhibitor β (38). In addition, phosphorylation by Rho kinase mediates desensitization of GPR39 (39). Thus, C-terminal missense or synonymous variants may alter secondary structure and/or phosphorylation to impact signaling. The results focus attention on the C terminus as a critical domain mediating distinct aspects of GPR39 signaling. It should be noted that GPR39 signaling is complex, including Gq–mediated cAMP generation, Gi–mediated IP3 production, and serum response factor–serum response element (SRF–SRE)–mediated transcriptional regulation via G12/13 (39). The current study examined only Gq–mediated IP3 generation, leaving open the possible impact of missense and/or synonymous variants on signaling bias. Further, our results suggest that structural studies on a subset of rare missense and SYN_ΔCB variants having strong effects on signaling may reveal important structural determinants for optimal GPR39 function.

Based on ClinVar and related databases, GPR39 has not been causally linked to disease. The present study focused on the general patient population, and all individuals in the present
study were heterozygous for GPR39 rare variants. Nevertheless, both LOF/MISS and SYN_ΔCB variants yielded disease associations by SKAT, which suggests a role for GPR39 in nerve function, benign hyperplasia of the prostate, psoriatic arthropathy, diseases of hair and hair follicles, and benign essential hypertension. These associations are in agreement with cellular and animal studies suggesting that GPR39 regulates neuronal transport of potassium and chloride, attenuating postsynaptic excitability (12, 40–42). Animal models demonstrate GPR39-mediated up-regulation of the neuronal K⁺/Cl⁻ cotransporter attenuates seizure activity (42), and knockout of GPR39 promotes depression (43). GPR39 is up-regulated by some antidepressants (44) in mouse models. One study in humans showed that GPR39 levels were significantly reduced in the hippocampus and cortex of suicide victims (45). The highest levels of Zn²⁺ in the body are in the prostate and seminal fluid (46), and GPR39 is differentially expressed in normal versus malignant prostate cells, where it regulates cell growth and proliferation (10, 12, 13), potentially contributing to the association with benign prostate hyperplasia observed in the present work. Finally, the observation that individuals heterozygous for rare variants in GPR39 have a significant association with diseases of hair and hair follicles should be considered in light of a recent study that found that GPR39 marks a class of stem cells in the sebaceous gland and contributes to wound healing in mice (47). Overall, the identification of a subset of clinical phenotypes that echo cell culture and animal model studies supports disparate roles for GPR39 in human physiology and validates the SKAT approaches used herein to identify disease associations. Additional phenotypes, not previously associated with GPR39 including psoriatic arthropathy and macular puckerings of the retina, deserve further attention.

In conclusion, we have identified strong disease associations for GPR39 using either LOF/missense or rare synonymous variants that confirm and extend cellular and animal studies. Rare variants from both the missense and synonymous classes altered GPR39 expression and function, providing a rationale for the observed disease associations. The approach applied to a larger set of orphan or understudied GPCRs also yielded robust associations with phenome-wide significance, providing potential novel targets for further investigation. The methods described in this report are easily applied to any set of genes of potential clinical importance with available genomic and integrated longitudinal electronic health records. Finally, this study highlights the importance of rare synonymous variants in human physiology and argues for their routine inclusion in any comprehensive analysis of genomic variants as potential causes of disease.

**Experimental procedures**

The MyCode® Community Health Initiative recruits Geisinger patients from a broad range of inpatient and outpatient clinics to integrate genetic information with their clinical EHR data to foster discoveries in health and disease (23). This study used only deidentified data and does not therefore involve “human subjects” as defined in 45CFR46.102(f). Although this study was not subject to oversight by the Geisinger Institutional Review Board, the specific uses of the genetic and associated clinical data described herein were approved by the Geisinger Institutional Review Board. We used WES from 51,289 participants (9), which includes patients enrolled through primary care and specialty outpatient clinics. Participants were 59% female, with a median age of 61 years, and predominantly Caucasian (98%) and non-Hispanic/Latino (95%). Patient demographics are in Table 4. Whole exome sequencing was performed using 75-bp paired-end sequencing on an Illumina v4 HiSeq 2500 to a coverage depth sufficient to provide greater than 20× haploid read depth of over 85% of targeted bases for at least 96% of samples. Sample preparation, sequencing, sequence alignment, variant identification, genotype assignment, and quality control steps were carried out as previously described (23).

Rare variants were separated into two groups: 1) nonsense and frameshift variants and 2) missense variants. Nonsense and frameshift variants were scored by SnpEff fields LOF[*] or PERC (percent) ≥ 0.5 or (nonsense mediated decay NMD)[*]. Number of transcripts (NUMTR) values >3 are considered potential LOF variants. To score missense variants, we summed numerical values derived from pathogenicity prediction tools to generate a cumulative score as follows (RMPath, in silico radical mutations pathogenicity prediction; maximum score = 24): 1) SIFT scoring: D (damaging) = 2; all others = 0; 2) PolyPhen2 scoring: (i) HDIV: probably damaging (D) = 2; possibly damaging (P) = 1; all other results = 0. (ii) HVAR: probably damaging (D) = 2; possibly damaging (P) = 1; all other results = 0; 3) MutationTaster scoring: Disease causing (D) = 2; all other results = 0; 4) MutationAssessor scoring: high (H) = 2; medium (M) = 1; all other results = 0; 5) FATHMM scoring: D (damaging) = 2; all other results = 0; 6) LRT: D (deleterious) → score 2; All others → score 0; 7) PROVEAN scoring: D (damaging) = 2; all other results = 0; 8) GRANTHAM scoring: ≥140 = 4; ≥120 = 3; ≥90 = 2; ≥70 = 1; all other results = 0; and 9) Genomic Evolutionary Rate Profiling GERP++ rejected substitutions (RS) scoring: ≥6 = 4; ≥5 = 3; ≥3.5 = 2; ≥2 = 1; all other results = 0. Variants with RMPath scores ≥8 were pB, those with scores in the range from 6 to 11 were predicted to be Likely Benign (pLB); those with scores between 11 and 17 were predicted to be likely pathogenic (pLP); and those with scores ≥17 were predicted to be pathogenic (pP). For clinical purposes, all variants are considered variants of unknown significance until family cosegregation and/or functional analyses definitively establish a variant as pathogenic. Our goal in this study was to identify variants likely to impact GPR39 expression or function.

For all rare synonymous codons, we determined the ratio of the genome-wide codon fraction of the GPR39 reference codon over the synonymous codon variant at that position using

---

**Table 4**

**Demographics of the DiscoveEHR cohort used in the present study**

| Basic demographics | DiscoveEHR sequenced patients |
|--------------------|-----------------------------|
| Total patients, N  | 51,289                      |
| Female, N (%)      | 30,290 (59%)                |
| Median age, years  | 61 (range 48–72)            |
| Median body mass index, kg/m² | 30 (range 26–36) |
| Median years of electronic health record data | 14 (range 11–17) |
| Median medication orders/patient | 129 (range 37–221) |

...
human-specific codon usage tables (48, 49), e.g. for S147S (tcG/tcA), 0.15/0.06 = 2.5, i.e. a 2.5-fold change. All synonymous variants with at least a 2-fold change in site-specific codon usage were used in the SKAT analyses.

Deidentified EHR data were obtained from an approved data broker. All unique ICD9 codes with ≥200 patients (regardless of genotype) with at least three independent incidences of the ICD9 code were considered as controls. All non-European individuals hav- ing one or two calls were excluded, and those having no calls of a particular ICD9 code were considered as controls. All non-
Europeans were excluded from the analysis, as was one sample from pairs of closely related individuals up to first cousins. All models were adjusted for sex, age, age², and the first four principal components. The PhenoWAS R package was used for association analyses, which were performed for three GPR39 common variants (2:133174764, pA50V; 2:133174999, pT128T; and 2:133402607, noncoding), and plotting of results was done using graphpad Prism (version 6). We used the sequence kernel association test with default weights (SKAT R package) to compare the burden of rare variants in cases and controls. Cases and controls were defined as for PhenoWAS analysis (50).

All non-European samples were excluded. We used principal components calculated from high quality common SNPs to estimate the ancestry of our samples. Using a kernel density estimator, the likelihood that an individual belonged to each HapMap class (African, admixed American, East Asian, European, or south Asian) was estimated; samples that fell into an unknown group were excluded (51).

Statistical significance was determined using the Bonferroni correction, \( \alpha/m \), where \( \alpha \) is 0.05, and \( m \) is the number of variant sets and disease codes, as indicated for each analysis (52).

All missense and SYN_ΔCB variants were generated by gene synthesis (Thermo Fisher Scientific, Gene Art) based on the NP_001499.1 reference sequence (GenelID 2863, 453 amino acids) and included tandem tags at the N terminus to facilitate expression and functional studies. After the initiation methionine codon, we added a 3× FLAG epitope plus a minimal bungarotoxin-binding site as shown: GACTACAAGACACCAGCCGTGAATGAAAGATCATGATCAGTACAAAGGA-
TGACGATGACAGGTTGGAGATACGAGAGCTTCCCTGAGCCCTACCTGACGTT.

WT and variant constructs were transiently transfected into HEK293 cells. The cells were lysed after 48 h, and 25 µg of lyses were separated on 4–15% SDS-PAGE gels (Bio-Rad), blotted to nitrocellulose, blocked with 5% milk in TBS-T, and exposed for 1 h at room temperature to anti-FLAG mAb conjugated to HRP (Sigma). The blots were exposed to SuperSignal West Pico Chemiluminescence Substrate (Thermo Fisher), and the images were recorded on a FUJIFILM LAS-4000 mini chemiluminescence analyzer and processed with Image-Gauge version 3.0. GPR39 expression was quantified using Multi-Gauge software, normalized to WT expression, and corrected for HEK293 background on the same blot. Each variant was assayed in two to four independent transfections. To accurately quantify the range of expression observed for SYN_ΔCB variants, we segregated high- and low-expressing variants plus WT onto separate gels, allowing differential exposure times for chemiluminescence acquisition.

WT and variant constructs were transiently transfected into HEK293 cells and cultured for 24 h, and equivalent numbers of cells were plated in poly-D-lysine–coated 96-well plates for a further 24 h. Replicates were treated with various concentrations of ZnCl₂ for 10 min at 37 °C. Levels of IP₁ were determined as recommended by manufacturer (IP-One ELISA kit, Cisbio USA, Inc.). The plates were read on a POLARStar Omega (BMG Labtech) plate reader. We made no assumptions regarding signaling mechanism and its potential cooperativity and estimated \( V_{\text{max}} \) and EC₅₀ from Zn²⁺ versus IP₁ data combined from three independent experiments using GraphPad PRISM (V6) and a general dose-response equation with the potential for cooperativity (\( Y = Y_0 \left[ (V_{\text{max}} - Y_0)/(1 + 10^{(\log EC_{50} - [Zn^{2+}]_n)}) \right] \)), where \( Y_0 \) is baseline IP₁ in the absence of Zn²⁺, \( V_{\text{max}} \) is the maximal response, \( EC_{50} \) is the Zn²⁺ concentration at 50% of \( V_{\text{max}} \) and \( n \) is the Hill coefficient (slope). Parallel plates from three independent transfections were exposed to Alexa 594–conjugated bungarotoxin (3 µg/ml, 5 min at 37 °C, 5% CO₂) to quantify surface expression of GPR39 variants. After removal of bungarotoxin and washes with PBS, fluorescence was read on the POLARStar Omega (BMG Labtech) plate reader. WT, standards, and solution blanks were included in each assay, and unspecific binding of BgTx–Alexa 594 to HEK293 cells was determined relative to cells transfected with WT GPR39 in the same experiment.

Acknowledgments—We are grateful to Geisinger and the Regeneron Genetics Center (G. Ahecasis, A. Baras, M. Cantor, G. Coppola, A. Economides, J.D. Overton, J. G. Reid, A. Shuldiner, C. Beechert, C. Forsythe, E. D. Fuller, Z. Gu, M. Lattari, A. Lopez, T. D. Schleicher, M. S. Padilla, K. Toledo, L. Widom, S. E. Wolf, M. Pradhan, K. Mano-
chehr, R. H. Ullom, X. Bai, S. Balasubramanian, L. Barnard, A. Blu-
enfeld, Y. Chai, G. Eom, L. Habeggy, Y. Hahn, A. Hawes, S. Khalid, E. K. Maxwell, J. Penn, J. C. Staples, A. Yadav, P. M. Guzzardo, M. B. Jones, and L. J. Mitnau) for access to the DiscovEHR cohort and to the patients in the MyCode Biobank.

References

1. Hauser, A. S., Attwood, M. M., Rask-Andersen, M., Schiöth, H. B., and Gloriam, D. E. (2017) Trends in GPCR drug discovery: new agents, targets and indications. Nat. Rev. Drug Discov. 16, 829–842 CrossRef Medline

2. Roth, B. L., and Kroeze, W. K. (2015) Integrated approaches for genome-wide interrogation of the druggable non-olfactory G protein–coupled receptor family. J. Biol. Chem. 290, 19471–19477 CrossRef Medline

3. Kovacs, P., and Schönberg, T. (2016) The relevance of genomic signatures at adhesion GPCR loci in humans. Handb. Exp. Pharmacol. 234, 179–217 CrossRef Medline
Missense and synonymous variants alter GPCR function

4. Wu, M. C., Lee, S., Cai, T., Li, Y., Boehnke, M., and Lin, X. (2011) Rare-variant association testing for sequencing data with the sequence kernel association test. *Am. J. Hum. Genet.* **89**, 82–93 CrossRef Medline

5. Lee, S., Emond, M. J., Bamshad, M. J., Barnes, K. C., Rieder, M. J., Nicaise, S., Eischen, D. A., NHLBI GO Exome Sequencing Project–ESP Lung Project Team, Christiani, D. C., Wurfel, M. M., and Lin, X. (2012) Optimal unified approach for rare-variant association testing with application to small-sample case-control whole-exome sequencing studies. *Am. J. Hum. Genet.* **91**, 224–237 CrossRef Medline

6. Richardson, T. G., Shihab, H. A., Rivas, M. A., McCarthy, M. I., Campbell, C., Timpon, N. J., and Gaunt, T. R. (2016) A protein domain and family based approach to rare variant association analysis. *PLoS One* **11**, e0153803 CrossRef Medline

7. Cohen, L., Asraf, H., Sekler, I., and Hershfinkel, M. (2012) Extracellular pH regulates zinc signaling via an Asp residue of the zinc-sensing receptor (ZnR/GPR39). *J. Biol. Chem.* **287**, 33339–33350 CrossRef Medline

8. Dewey, F. E., Murray, M. F., Overton, J. D., Habegger, L., Leader, J. B., Fetterolf, S. N., O’Dushlaine, C., Van Hout, C. V., Staples, J., Gonzaga-Jauregui, C., Metally, P., Pendergrass, S. A., Giovannoni, M. A., Kirchner, H. L., Balasubramanian, S., et al. (2016) Distribution and clinical impact of functional variants in 50,726 whole-exome sequences from the DiscoEHR study. *Science* **354**, aaf6814 CrossRef Medline

9. Afzal, H., Salomon, S., Nevo, A., Sekler, I., Mayer, D., and Hershflink, M. (2014) The ZnR/GPR39 interacts with the CaSR to enhance signaling in prostate and salivary epithelia. *J. Cell. Physiology* **229**, 868–877 CrossRef Medline

10. Basile, A. O., Byrns-Bishop, M., Wallace, J., Frase, A. T., and Ritchie, M. D. (2018) Extracellular pH regulates zinc signaling via an Asp residue of the zinc-sensing receptor (ZnR/GPR39). *J. Biol. Chem.* **287**, 33339–33350 CrossRef Medline

11. Sunuwar, L., Gilad, D., and Hershflink, M. (2017) The zinc sensing receptor, ZnR/GPR39, in health and disease. *Front. Biosci. (Landmark Ed)* **22**, 1469–1492 CrossRef Medline

12. Giddens, M. M., Wong, J. C., Schroeder, J. P., Farrow, E. G., Smith, B. M., O’Nion, S., Soden, E. S., Meyer, R. C., Saunders, C., LePichon, J. B., Weinshenker, D., Escayg, A., and Hall, R. A. (2017) GPR37L1 modulates seizure susceptibility: evidence from mouse studies and analyses of a human GPR37L1 variant. *Neurobiol. Dis.* **106**, 181–190 CrossRef Medline

13. Hoek, J. D., Biens, B., Kurtova, A. V., Kljavin, N. M., de Sousa, E., Melo F., Alicke, B., Koeppen, H., Modrusan, Z., Piskol, R., and de Sauvage, F. J. (2017) Stem cell plasticity enables hair regeneration following LGR5 stasis in a mouse model of early onset diet-induced obesity. *J. Neuroendocrinol.* **29**, 106, 1–11 CrossRef Medline

14. Shah, K., Cheng, Y., Hahn, B., Bridges, R., Bradbury, N. A., and Mueller, D. M. (2015) Synonymous codon usage affects the expression of wild type and F508del CFTR. *J. Mol. Biol.* **427**, 1464–1479 CrossRef Medline

15. Zhang, T., Wu, Y., Lan, Z., Shi, Q., Yang, Y., and Guo, J. (2017) Syntool: a novel region-based intolerance score to single nucleotide substitution for synonymous mutations prediction based on 12,136 individuals. *BioMed Res. Int.* **2017**, 5096208 Medline

16. Kovacs, Z., Schacht, T., Herrmann, A. K., Albrecht, P., Lefkimmiatis, K., and Mether, A. (2014) Protein kinase inhibitor β enhances the constitutive activity of G-protein-coupled zinc receptor GPR39. *Biochem. J.* **462**, 125–132 CrossRef Medline

17. Shimizu, Y., Koyama, R., and Kawamoto, T. (2017) Rho kinase-dependent novel region-based intolerance score to single nucleotide substitution for synonymous mutations prediction based on 12,136 individuals. *BioMed Res. Int.* **2017**, 5096208 Medline
40. Chorin, E., Vinograd, O., Fleidervish, I., Gilad, D., Herrmann, S., Sekler, I., Aizenman, E., and Hershfinkel, M. (2011) Upregulation of KCC2 activity by zinc-mediated neurotransmission via the mZnR/GPR39 receptor. *J. Neurosci.* **31**, 12916–12926 CrossRef Medline

41. Saadi, R. A., He, K., Hartnett, K. A., Kandler, K., Hershfinkel, M., and Aizenman, E. (2012) SNARE-dependent upregulation of potassium chloride co-transporter 2 activity after metabotropic zinc receptor activation in rat cortical neurons *in vitro*. *Neuroscience* **210**, 38–46 CrossRef Medline

42. Gilad, D., Shorer, S., Ketzef, M., Friedman, A., Sekler, I., Aizenman, E., and Hershfinkel, M. (2015) Homeostatic regulation of KCC2 activity by the zinc receptor mZnR/GPR39 during seizures. *Neurobiol. Dis.* **81**, 4–13 CrossRef Medline

43. Mlyniec, K., Budziszewska, B., Holst, B., Ostachowicz, B., and Nowak, G. (2014) GPR39 (zinc receptor) knockout mice exhibit depression-like behavior and CREB/BDNF down-regulation in the hippocampus. *Int. J. Neuropsychopharmacol.* **18**, pyu002 Medline

44. Mlyniec, K., and Nowak, G. (2015) Up-regulation of the GPR39 Zn$^2+$-sensing receptor and CREB/BDNF/TrkB pathway after chronic but not acute antidepressant treatment in the frontal cortex of zinc-deficient mice. *Pharmacol. Rep.* **67**, 1135–1140 CrossRef Medline

45. Mlyniec, K., Doboszewska, U., Szewczyk, B., Sowa-Kucma, M., Misztak, P., Piekoszewski, W., Trela, F., Ostachowicz, B., and Nowak, G. (2014) The involvement of the GPR39-Zn$^2+$-sensing receptor in the pathophysiology of depression: studies in rodent models and suicide victims. *Neuropharmacology* **79**, 290–297 CrossRef Medline

46. Costello, L. C., and Franklin, R. B. (2016) A comprehensive review of the role of zinc in normal prostate function and metabolism; and its implications in prostate cancer. *Arch. Biochem. Biophys.* **611**, 100–112 CrossRef Medline

47. Zhao, H., Qiao, J., Zhang, S., Zhang, H., Lei, H., Lei, X., Wang, X., Deng, Z., Ning, L., Cao, Y., Guo, Y., Liu, S., and Duan, E. (2015) GPR39 marks specific cells within the sebaceous gland and contributes to skin wound healing. *Sci. Rep.* **5**, 7913 CrossRef Medline

48. Nakamura, Y., Gojobori, T., and Ikemura, T. (2000) Codon usage tabulated from international DNA sequence databases: status for the year 2000. *Nucleic Acids Res.* **28**, 292 CrossRef Medline

49. Athey, J., Alexaki, A., Osipova, E., Rostovtsev, A., Santana-Quintero, L. V., Katneni, U., Simonyan, V., and Kimchi-Sarfaty, C. (2017) A new and updated resource for codon usage tables. *BMC Bioinformatics* **18**, 391 CrossRef Medline

50. Verma, A., Bradford, Y., Dudek, S., Lucas, A. M., Verma, S. S., Pendergrass, S. A., and Ritchie, M. D. (2018) A simulation study investigating power estimates in phenome-wide association studies. *BMC Bioinformatics* **19**, 120 CrossRef Medline

51. Staples, J., Maxwell, E. K., Gosalia, N., Gonzaga-Jauregui, C., Snyder, C., Hawes, A., Penn, J., Ulloa, R., Bai, X., Lopez. A. E., Van Hout, C. V., O’Dushlaine, C., Teslovich, T. M., McCarthy, S. E., Balasubramanian, S., et al. (2018) Profiling and leveraging relatedness in a Precision Medicine Cohort of 92,455 exomes. *Am. J. Hum. Genet.* **102**, 874–889 CrossRef Medline

52. Verma, A., Lucas, A., Verma, S. S., Zhang, Y., Josyula, N., Khan, A., Hartzel, D. N., Lavage, D. R., Leader, J., Ritchie, M. D., and Pendergrass, S. A. (2018) PheWAS and beyond: the landscape of associations with medical diagnoses and clinical measures across 38,662 individuals from Geisinger. *Am. J. Hum. Genet.* **102**, 592–608 CrossRef Medline