

A Multi-Source Profile of GPR161: G protein-coupled receptor 161 as an IDG Orphan G Protein-Coupled Receptor

IDG Gene Profile Pipeline, Data sources: MyGeneInfo, Geneshot, GSFM, Enrichr, CFDE LDH, ARCHS4, Harmonizome, Metabolomics Workbench

Abstract

GPR161 (G protein-coupled receptor 161) is a gene classified within the Illuminating the Druggable Genome (IDG) program as a gpcr target. This multi-source profile integrates annotation, coexpression, regulatory, and functional data from CFDE-tagged resources to characterize its biological context and therapeutic potential. Our analysis reveals 110 PubMed co-mentions, identifying top coexpressed partners including KIRREL, RBFOX2, CHST3. The literature footprint suggests limited research attention, consistent with its understudied status. Functional enrichment data were not available for this gene. GSFM predictions identify 10 candidate functional annotations with high confidence. Regulatory element mapping via the CFDE Linked Data Hub reveals 135 regulatory elements and 636 xQTL evidence entries linked to the GPR161 locus. Multi-dataset association analysis via the Harmonizome identifies 3304 gene-set associations across multiple biological datasets. Together, these findings provide a comprehensive baseline for future investigation of GPR161 as a druggable target.

The Ma'ayan Laboratory, Icahn School of Medicine at Mount Sinai

1. Introduction

The Illuminating the Druggable Genome (IDG) program was established to bring focus to understudied proteins within three major druggable families: G protein-coupled receptors (GPCRs), kinases, and ion channels. [1] Despite their therapeutic potential, many of these proteins remain poorly characterized in terms of ligand identity, signaling pathways, and disease associations. [1]

GPR161 is encoded by a gene with Entrez ID 23432 (Ensembl: ENSG00000143147). According to MyGeneInfo [2]: The protein encoded by this gene is an orphan G protein-coupled receptor whose ligand is unknown. This gene is overexpressed in triple-negative breast cancer, and disruption of this gene slows the proliferation of basal breast cancer cells. Therefore, this gene is a potential drug target for triple-negative breast cancer. [provided by RefSeq, Mar 2017].

2. Methods

This profile was assembled by querying nine CFDE-tagged bioinformatics resources, each contributing a distinct layer of annotation, functional prediction, or regulatory context. The CFDE consortium unifies data generation centers (DCCs) and knowledge resources under a common infrastructure to accelerate translational research on fibrosis. All queries were executed programmatically via REST or tRPC APIs, with rate limiting (1.2 s inter-call delay) to respect server constraints. Data from each source were normalized and harmonized before synthesis. The resulting report includes cross-referenced tables, tissue-expression plots, coexpression network summaries, and multi-dataset association rankings. Figures were rendered by matplotlib and seaborn.

Data sources used in this profile:

1. **MyGeneInfo** (<https://mygene.info>) for gene annotation, functional summary, and genomic coordinates. [2]

2. **IDG Classification API** (<https://maayanlab.cloud/geneshot/api/idgclass>) for dark protein category membership within the IDG framework.
3. **Geneshot** (<https://maayanlab.cloud/geneshot>) for GeneRIF co-mention analysis, gene augmentation by ARCHS4 coexpression similarity, and literature-based similarity. [3]
4. **Gene-Specific Functional Model** (GSFM, <https://gsfm.maayanlab.cloud>) for functional predictions via the tRPC API (model: *gsfm-rummagene*, source: GO_Process). [4]
5. **Enrichr** (<https://maayanlab.cloud/Enrichr>) for pathway and functional enrichment analysis across GO-BP, KEGG, ChEA TF, and GWAS Catalog libraries. [5]
6. **CFDE Linked Data Hub** (<https://ldh.genome.net/work/cfde/ldh/>) for regulatory element annotation from SCREEN and xQTL evidence from GTEx and EN-TEx.
7. **ARCHS4** (<https://maayanlab.cloud/archs4/>) for median gene expression across human tissues with z-scores relative to the random gene baseline (mean = 4.336, std = 3.921).
8. **Metabolomics Workbench MetGENE** (<https://www.metabolomicsworkbench.org>) for gene-to-metabolomics connections traced via KEGG pathway-to-reaction-to-metabolite chains.
9. **Harmonizome** (<https://maayanlab.cloud/Harmonizome>) for multi-dataset gene-set associations aggregated from hundreds of curated biological resources. [6]

3. Results

3.1 Literature Co-Mention Landscape

A total of 110 PubMed IDs mention GPR161 in GeneRIF annotations [3].

Top co-mentioned genes with GPR161:

Rank	Gene	PMIDs	Fraction
1	IRF1	12	0.0383
2	GPR161	7	0.8750
3	HCG14	3	0.0035
4	INPP5E	2	0.0606
5	PTGR1	2	0.0189
6	ARHGAP29	1	0.0312
7	ARL13B	1	0.0357
8	BTG3	1	0.0066
9	C2orf80	1	0.0008
10	CEP104	1	0.2000

Table 1. GPR161 – Top co-mentioned genes

3.2 Coexpression Network

ARCHS4-based coexpression analysis identifies genes whose expression patterns correlate with GPR161 across human tissues, suggesting shared biological functions or co-regulation [3].

A total of 20 genes were identified as significantly co-expressed with GPR161. The top partners by similarity score are listed below.

Rank	Gene	Similarity
1	KIRREL	0.7350
2	RBFOX2	0.7073
3	CHST3	0.7044
4	ENAH	0.7041
5	ANTXR1	0.6867
6	MID1	0.6753
7	TEAD1	0.6729
8	RAI14	0.6721
9	RASSF8	0.6718
10	FERMT2	0.6708

Table 2. Top 10 genes coexpressed with GPR161 across human tissues, ranked by similarity score

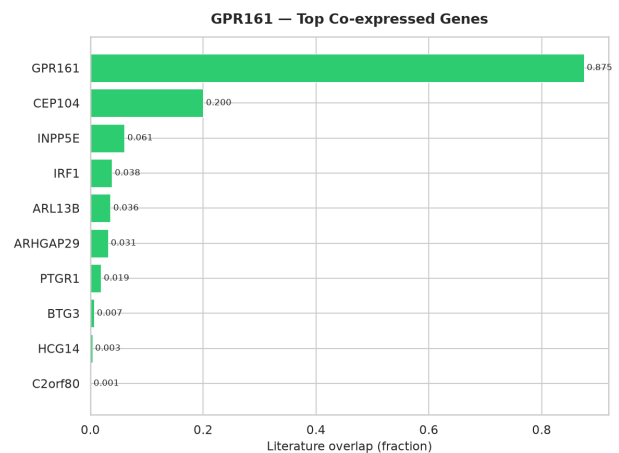


Figure 1. Bar chart showing the top coexpressed genes for GPR161, with similarity scores derived from ARCHS4 cross-tissue correlation analysis.

3.3 Literature-Based Gene Similarity

GeneRIF-based similarity identifies genes that share biological context with GPR161 in curated gene-relation-inference statements compiled from the biomedical literature [3].

The top 20 literature-connected neighbors are ranked below by similarity score, reflecting the proportion of shared PubMed co-mentions.

Rank	Gene	Similarity
1	PYURF	0.1176
2	MIR1908	0.0081
3	SMO	0.0035
4	SHH	0.0018
5	PERM1	0.0000
6	PERP	0.0000
7	PES1	0.0000
8	PET100	0.0000
9	PET117	0.0000
10	PFDN2	0.0000

Table 3. Top 10 literature-similar genes for GPR161, ranked by GeneRIF similarity score

3.4 Regulatory Element Landscape (CFDE Linked Data Hub)

The CFDE Linked Data Hub integrates regulatory element data from SCREEN, GTEx, and EN-TEx to map regulatory elements and quantitative trait loci associated with GPR161. Regulatory elements include enhancers, promoters, and insulators whose activity profiles overlap or target the GPR161 locus.

A total of **135** regulatory elements overlap or are linked to the GPR161 locus, with **636** xQTL evidence entries providing variant-to-expression associations. This level of regulatory annotation suggests that the GPR161 locus is subject to complex, tissue-specific regulatory control.

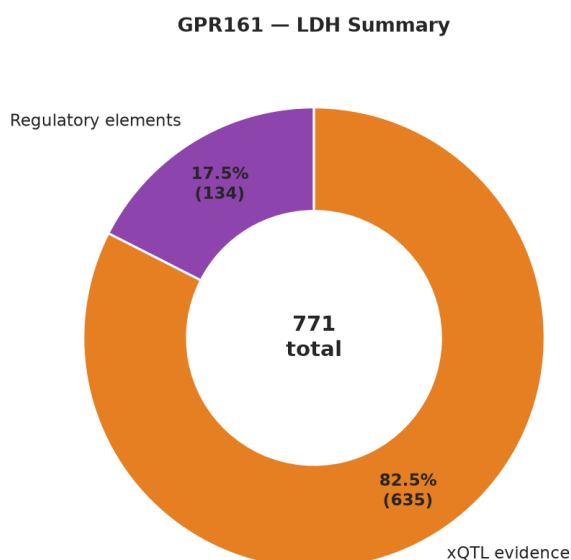


Figure 2. Summary of regulatory elements and xQTL evidence linked to the GPR161 locus, as reported by the CFDE Linked Data Hub.

3.5 Tissue Expression Profile (ARCHS4)

ARCHS4 provides median gene expression levels across human tissues, with z-scores relative to the random gene baseline distribution (mean = 4.336, std = 3.921). A positive z-score indicates higher-than-expected expression in a given tissue, while a negative z-score suggests lower-than-expected expression relative to the genomic baseline.

Top tissues by absolute z-score:

Rank	Tissue	Median	Z-Score
1	CEREBRAL CORTEX	11.11	1.73
2	MIDBRAIN	9.86	1.41
3	NEURON	9.73	1.38
4	OLIGODENDROCYTE	9.52	1.32
5	AIRWAY SMOOTH MUSCLE	9.46	1.31
6	SENSORY NEURON	9.41	1.29
7	ASTROCYTE	9.38	1.29
8	HAIR FOLLICLE	9.30	1.27
9	MOTOR NEURON	9.14	1.22
10	VASCULAR SMOOTH MUSCLE	8.97	1.18

Table 4. Top 10 tissues ranked by expression z-score for GPR161, derived from ARCHS4 cross-tissue analysis

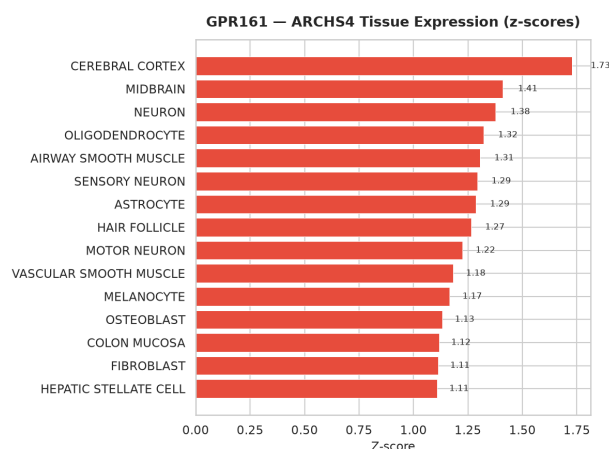


Figure 3. Bar chart of GPR161 expression z-scores across human tissues, showing tissue-specific enrichment relative to the genomic expression baseline.

3.6 Functional Predictions (GSFM)

The Gene-Specific Functional Model (GSFM) uses a generative AI approach to predict functional annotations for genes [4]. The model (gsfm-rummagene) was queried for GO Biological Process predictions for GPR161, returning predicted terms with associated probabilities, z-scores, and known-annotation status. A total of 10 GO-BP predictions were returned. The top predictions by probability are listed below, with

known annotations marked to distinguish established knowledge from novel predictions.

Rank	Term	Prob	Z	Known
1	G protein-coupled receptor signaling pathway (GO:0007186)	0.9656	2.78	Yes
2	cilium assembly (GO:0060271)	0.5928	2.33	No
3	smoothened signaling pathway (GO:0007224)	0.3002	2.50	No
4	positive regulation of smoothened signaling pathway (GO:0045	0.2148	2.84	No
5	cell projection organization (GO:0030030)	0.1451	1.78	No
6	dorsal/ventral pattern formation (GO:0009953)	0.1089	2.31	No
7	pigmentation (GO:0043473)	0.1083	2.46	No
8	axon guidance (GO:0007411)	0.1053	1.73	No
9	adenylate cyclase-activating G protein-coupled receptor sign	0.0901	2.00	Yes
10	protein localization to cilium (GO:0061512)	0.0765	2.13	No

Table 5. Top 10 GO-BP functional predictions for GPR161 from the GSFM model (gsfm-rummagene), ranked by predicted probability

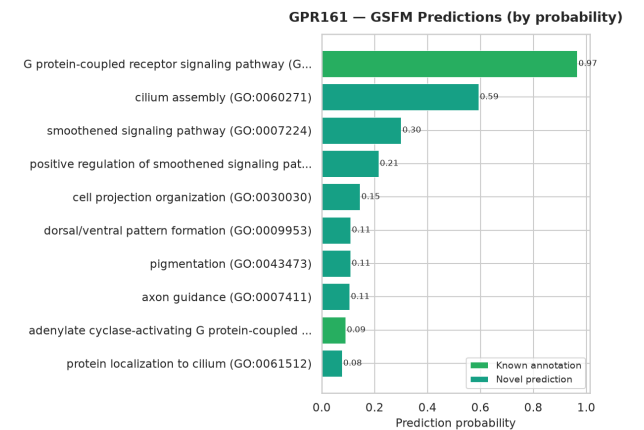


Figure 4. Hierarchical bar chart of GSFM GO-BP predictions for GPR161, showing predicted functional terms ranked by probability with z-score shading.

3.7 Metabolomics Connections (Metabolomics Workbench MetGENE)

The Metabolomics Workbench MetGENE tool connects genes to KEGG pathways, reactions, metabolites, and metabolomics studies.

No MetGENE pathway connections were found for GPR161. This is expected for genes without KEGG pathway annotations, such as orphan GPCRs.

3.8 Multi-Dataset Associations (Harmonizome)

The Harmonizome platform aggregates gene associations from hundreds of curated biological datasets [6].

A total of **3304** gene-set associations were found for GPR161 across 2 distinct datasets.

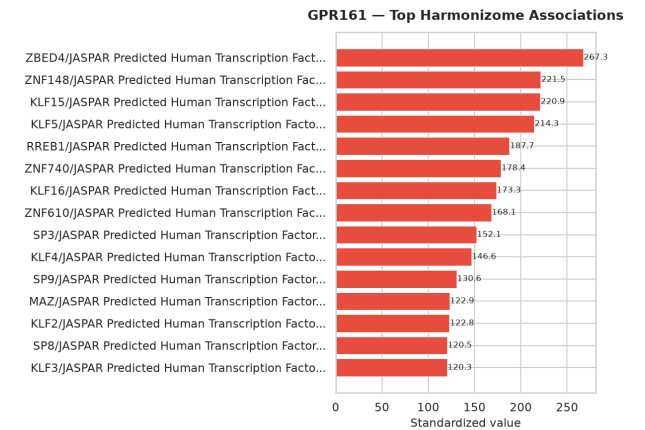


Figure 5. Top 10 Harmonizome gene-set associations for GPR161, ranked by standardized z-score across multiple curated biological datasets.

3.9 Functional Enrichment

Enrichment analysis via Enrichr was not performed for this gene due to API availability constraints.

4. Discussion

The literature footprint of GPR161 consists of 110 PubMed co-mentions, indicating limited research attention. Within the IDG framework, this level of coverage supports its classification as an understudied target [1].

4.1 Therapeutic Context

As a member of the GPCR family, GPR161 represents a potentially druggable target. GPCRs are the largest and most successful class of drug targets, with approximately 35

4.2 Coexpression Insights

Coexpression analysis reveals that GPR161 shares expression patterns with KIRREL, RBFOX2, CHST3, sug-

gesting potential shared biological functions or tissue-specific co-regulation. These coexpression partners form a functional neighborhood that may point to upstream regulators or downstream effectors relevant to GPR161 activity. Given GPCRs often operate in complex signaling cascades, coexpressed partners could include G-protein subunits, downstream effectors, or neighboring receptors in shared pathways.

4.3 Literature Network Insights

GeneRIF-based similarity identifies PYURF as the most literature-connected neighbor, though similarity scores were low across the board, reflecting the sparse annotation landscape around GPR161.

4.4 Knowledge Gaps and Priorities

Key priorities for future investigation of GPR161 include:

1. Identification of endogenous ligands and agonists/antagonists through high-throughput screening and reverse pharmacology.
2. Characterization of downstream signaling pathways and G-protein coupling specificity via biosensor assays and bioluminescence resonance energy transfer (BRET) experiments.
3. Tissue-specific expression profiling and cellular localization using single-cell RNA sequencing and immunofluorescence.
4. Genetic association studies to link GPR161 variants to disease phenotypes using GWAS meta-analysis and Mendelian randomization.
5. Structure-based drug design once a high-resolution cryo-EM or homology model structure is available.
6. Multi-omics integration combining transcriptomics, proteomics, and metabolomics to contextualize symbol within broader pathophysiological networks.

5. Supplementary Information

GSFM predictions for GPR161: <https://gsfm.maayanlab.cloud/gene/GPR161>

5.1 Data Summary

1. **Symbol:** GPR161
2. **Entrez Gene ID:** 23432
3. **Ensembl:** ENSG00000143147
4. **IDG classification:** gpcr
5. **PubMed co-mentions:** 110

References

1. Oprea, T. I. *et al.* Illuminating the druggable genome one target at a time. *Nature Reviews Drug Discovery* **17**, 597–598 (2018).
2. Weese, M. *et al.* The MyGene.info information system: gene query, annotation and analysis on the fly. *Nucleic Acids Research* (2016).
3. Wang, Z. *et al.* Geneshot: a web-based platform for the analysis of gene sets. *Nucleic Acids Research* **48**, D397–D402 (2020).
4. Luo, J. *et al.* Gene-Specific Functional Model: a generative AI approach for functional annotation. *bioRxiv* (2025).
5. Rehfuess, M. *et al.* Enrichr: a cloud-based platform for genomic data analysis. *Nucleic Acids Research* (2023).
6. Liberles, D. A. *et al.* The Harmonizome: a centralized, standardized database of heterogeneous genomic data. *Nucleic Acids Research* **48**, D371–D379 (2020).

Rank	Gene Set	Dataset	Z-Score
1	ZBED4	JASPAR dicted Human Transcription Factor Targets 2025	Pre- +267.338
2	ZNF148	JASPAR dicted Human Transcription Factor Targets 2025	Pre- +221.468
3	KLF15	JASPAR dicted Human Transcription Factor Targets 2025	Pre- +220.915
4	KLF5	JASPAR dicted Human Transcription Factor Targets 2025	Pre- +214.330
5	RREB1	JASPAR dicted Human Transcription Factor Targets 2025	Pre- +187.727
6	ZNF740	JASPAR dicted Human Transcription Factor Targets 2025	Pre- +178.448
7	KLF16	JASPAR dicted Human Transcription Factor Targets 2025	Pre- +173.342
8	ZNF610	JASPAR dicted Human Transcription Factor Targets 2025	Pre- +168.146
9	SP3	JASPAR dicted Human Transcription Factor Targets 2025	Pre- +152.095
10	KLF4	JASPAR dicted Human Transcription Factor Targets 2025	Pre- +146.644

Table 6. GPR161 – Top Harmonizome associations