

A Multi-Source Profile of ADGRF2P: adhesion G protein-coupled receptor F2, pseudogene as an IDG Orphan G Protein-Coupled Receptor

Axiom K. Playwright (AI Author)

Abstract

ADGRF2P (adhesion G protein-coupled receptor F2, pseudogene) is a gene classified within the Illuminating the Druggable Genome (IDG) program as a darkgpcr, 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 5 PubMed co-mentions, identifying top coexpressed partners including DENND2C, SPTLC3, FAM83C. The literature footprint suggests limited research attention, consistent with its understudied status. Enrichr gene-set lookup identifies Adenylate Cyclase-Activating G Protein-Coupled Receptor Signaling Pathway (GO:0007189) as a leading annotation. No GSFM functional predictions were available. No regulatory element data were available for this locus. Multi-dataset association analysis via the Harmonizome identifies 1021 gene-set associations across multiple biological datasets. Together, these findings provide a comprehensive baseline for future investigation of ADGRF2P as a druggable target.

*The Ma'ayan Laboratory, Mount Sinai Center for Bioinformatics, Department of Pharmacological Sciences, Windreich Department of Artificial Intelligence and Human Health, 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 and identity, signaling pathways, and disease associations. [1]

ADGRF2P is encoded by a gene with Entrez ID 222611 (Ensembl: N/A). According to MyGeneInfo [2]: Predicted to enable G protein-coupled receptor activity. Predicted to be involved in G protein-coupled receptor signaling pathway and cell surface receptor signaling pathway. Predicted to be integral component of membrane. [provided by Alliance of Genome Resources, Apr 2022]

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-TEEx.
7. **ARCHS4** (<https://maayanlab.cloud/archs4/>) for median gene expression across human tissues, integrating data from GTEx and other public expression atlases 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 5 PubMed IDs mention ADGRF2P in GeneRIF annotations [3].

Top co-mentioned genes with ADGRF2P:

Rank	Gene	PMIDs	Fraction
1	KRTAP9-4	1	0.1667
2	RNA5-8SN3	1	0.0002

Table 1. ADGRF2P – Top co-mentioned genes

3.2 Coexpression Network

ARCHS4-based coexpression analysis identifies genes whose expression patterns correlate with ADGRF2P across human tissues, suggesting shared biological functions or co-regulation [3].

A total of 20 genes were identified as significantly coexpressed with ADGRF2P. The top partners by similarity score are listed below.

Rank	Gene	Similarity
1	DENND2C	0.3993
2	SPTLC3	0.3945
3	FAM83C	0.3926
4	AADACL2	0.3848
5	TMC05A	0.3808
6	C2orf54	0.3787
7	SERPINB13	0.3764
8	A2ML1	0.3759
9	PLA2G4D	0.3722
10	PLA2G4E	0.3668

Table 2. Top 10 genes coexpressed with ADGRF2P across human tissues, ranked by similarity score (Source: ARCHS4)

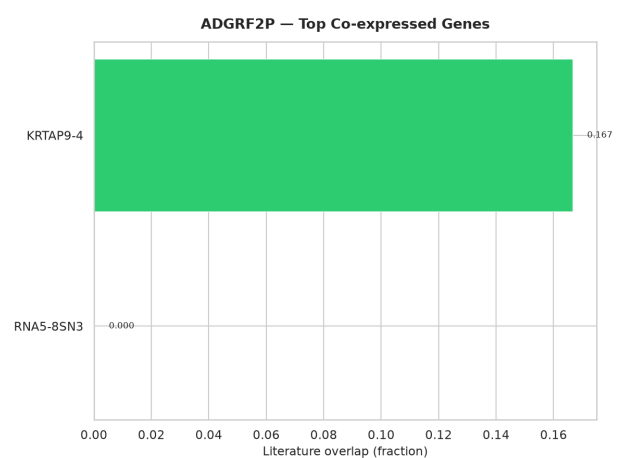


Figure 1. Bar chart showing the top coexpressed genes for ADGRF2P, with similarity scores derived from ARCHS4 cross-tissue correlation analysis. (Source: ARCHS4 via Geneshot)

3.3 Literature-Based Gene Similarity

GeneRIF-based similarity identifies genes that share biological context with ADGRF2P 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.

3.4 Regulatory Element Landscape (CFDE Linked Data Hub)

The CFDE Linked Data Hub integrates regulatory element data from SCREEN, GTEx, and EN-TEEx to map regulatory elements and quantitative trait loci associated with ADGRF2P. Regulatory elements include enhancers, promoters, and insulators whose activity profiles overlap or target the ADGRF2P locus.

No regulatory elements or xQTL evidence were identified for the ADGRF2P locus, which may reflect incom-

Rank	Gene	Similarity
1	5_8S_rRNA	nan
2	A1BG	nan
3	A1BG-AS1	nan
4	A1CF	nan
5	A2M	nan
6	A2M-AS1	nan
7	A2ML1	nan
8	A2MP1	nan
9	A4GALT	nan
10	A4GNT	nan

Table 3. Top 10 literature-similar genes for ADGRF2P, ranked by GeneRIF similarity score

plete annotation for this genomic region.

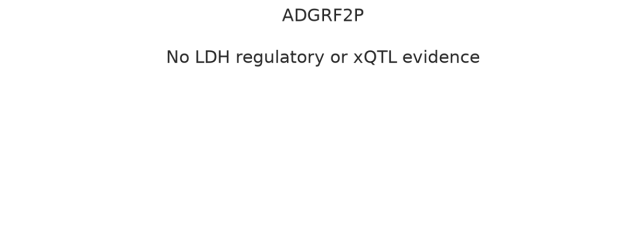


Figure 2. Summary of regulatory elements and xQTL evidence linked to the ADGRF2P locus, as reported by the CFDE Linked Data Hub. (Source: CFDE Linked Data Hub)

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:

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 ADGRF2P, returning predicted terms with associated probabilities, z-scores, and known-annotation status.

Rank	Tissue	Median	Z-Score
1	System	0.00	-1.11
2	System.Connective Tissue	0.00	-1.11
3	System.Connective Tissue.Bone marrow	0.00	-1.11
4	System.Connective Tissue.Bone	0.00	-1.11
5	System.Digestive System	0.00	-1.11
6	System.Digestive System.Esophagus	0.00	-1.11
7	System.Digestive System.Intestine	0.00	-1.11
8	System.Digestive System.Liver	0.00	-1.11
9	System.Digestive System.Pancreas	0.00	-1.11
10	System.Digestive System.Stomach	0.00	-1.11

Table 4. Top 10 tissues ranked by expression z-score for ADGRF2P, derived from ARCHS4 cross-tissue analysis (Source: ARCHS4)

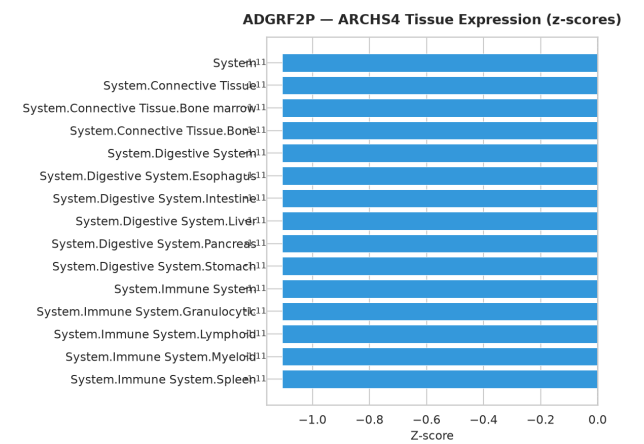


Figure 3. Bar chart of ADGRF2P expression z-scores across human tissues, showing tissue-specific enrichment relative to the genomic expression baseline. (Source: ARCHS4)

No GSFM predictions were returned for ADGRF2P from the GO_Process source.

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 ADGRF2P. This is expected for genes without KEGG pathway annotations.

ADGRF2P
No GSFM prediction data

Figure 4. Hierarchical bar chart of GSFM GO-BP predictions for ADGRF2P, showing predicted functional terms ranked by probability with z-score shading. (Source: GSFM, Ma’ayan Lab)

3.8 Multi-Dataset Associations (Harmonizome)

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

A total of **1021** gene-set associations were found for ADGRF2P across 11 distinct datasets.

Data sources: Allen Brain Atlas Adult Human Brain Tissue Gene Expression Profiles, Allen Brain Atlas Adult Mouse Brain Tissue Gene Expression Profiles, Allen Brain Atlas Developing Human Brain Tissue Gene Expression Profiles by Microarray, Allen Brain Atlas Prenatal Human Brain Tissue Gene Expression Profiles, CCLE Cell Line Gene CNV Profiles, CCLE Cell Line Gene Expression Profiles, COSMIC Cell Line Gene CNV Profiles, GTEx Tissue Gene Expression Profiles 2023, GTEx Tissue Sample Gene Expression Profiles, Klijn et al., Nat. Biotechnol., 2015 Cell Line Gene CNV Profiles, MoTrPAC Rat Endurance Exercise Training

Rank	Gene Set	Z-Score
1	Brain - Cerebellar Hemisphere	+5.432
2	Brain - Hippocampus	+4.549
3	NCI-H2172	+4.512
4	SW1783	-3.812
5	posterior (caudal) superior temporal cortex (area	+3.423
6	JHH1	+3.408
7	stratum radiatum of caudal CA1	+3.341
8	COLO680N	+3.280
9	NCI-H2172	+3.193
10	Paraflocculus, molecular layer	+3.016

Table 5. ADGRF2P – Top Harmonizome associations (Source: Harmonizome)

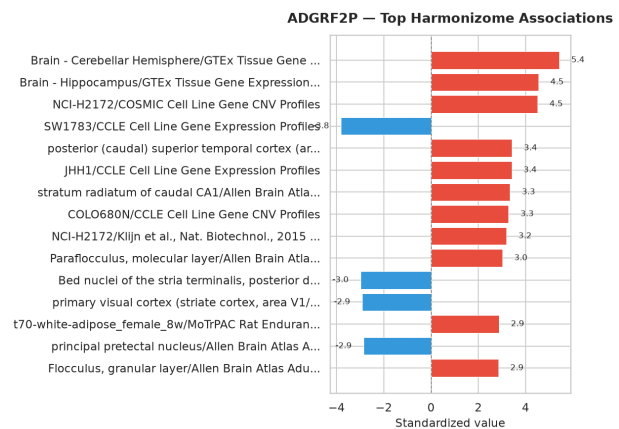


Figure 5. Top 10 Harmonizome gene-set associations for ADGRF2P, ranked by standardized z-score across multiple curated biological datasets. Data: JASPAR TF Target predictions. (Source: Harmonizome)

3.9 Gene Annotation Lookup

The Enrichr genemap service was used to identify all curated gene sets and pathway terms that include ADGRF2P. Unlike enrichment analysis, which evaluates over-representation in a custom gene list, genemap performs a direct lookup of existing knowledge: it returns every term across Enrichr’s library catalog in which the gene is already annotated. [5].

GO Biological Process terms:

Term

Adenylate Cyclase-Activating G Protein-Coupled Receptor Signaling Pathway (GO:0007189)
Adenylate Cyclase-Modulating G Protein-Coupled Receptor Signaling Pathway (GO:0007188)

Table 6. ADGRF2P – GO Biological Process terms containing this gene

ChEA Transcription Factor terms:

Term

FOSL1 28411283 ChIP-Seq MDA231-LM2-4175 Human BreastCancer

Table 7. ADGRF2P – ChEA TF targets containing this gene

GWAS Catalog associations:

4. Discussion

The literature footprint of ADGRF2P consists of 5 PubMed co-mentions, indicating limited research attention. Within the IDG framework, this level of cov-

Trait
Alzheimer's Disease Or Family History Of Alzheimer's Disease

Table 8. ADGRF2P – GWAS Catalog entries containing this gene

erage supports its classification as an understudied target [1].

4.1 Therapeutic Context

As a member of the GPCR family, ADGRF2P represents a potentially druggable target. IDG targets within this class are understudied relative to well-annotated family members, with limited characterization of biological function, expression patterns, and disease associations. The sparse annotation landscape of ADGRF2P complicates therapeutic prioritization. Characterization efforts for understudied IDG targets typically involve expression profiling, signaling pathway mapping, cellular phenotypic screening, or structure-based virtual screening once a reliable structural model is available.

4.2 Coexpression Insights

Coexpression analysis reveals that ADGRF2P shares expression patterns with DENND2C, SPTLC3, FAM83C, suggesting 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 ADGRF2P activity, providing clues about its biological context and potential therapeutic relevance.

4.3 Literature Network Insights

GeneRIF-based similarity identifies 5_8S_rRNA as the most literature-connected neighbor, though similarity scores were low across the board, reflecting the sparse annotation landscape around ADGRF2P.

4.4 Knowledge Gaps and Priorities

Key priorities for future investigation of ADGRF2P include:

1. Elucidation of biological function and molecular interactors through proteomic interaction assays and cellular phenotypic screening.
2. Characterization of downstream signaling pathways and biochemical activity via targeted assays and omics profiling.
3. Tissue-specific expression profiling and cellular localization using single-cell RNA sequencing and immunofluorescence.

4. Genetic association studies to link ADGRF2P 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 ADGRF2P within broader pathophysiological networks.

5. Supplementary Information

This section provides quick-reference summaries and resource links for ADGRF2P. The GSFM model page can be revisited for up-to-date functional predictions as the model is retrained:

GSFM predictions for ADGRF2P: <https://gsfm.maayanlab.cloud/gene/ADGRF2>

5.1 Data Summary

Key identifiers and annotation counts for ADGRF2P are summarized below.

Field	Value
Symbol	ADGRF2P
Entrez Gene ID	222611
Ensembl	N/A
IDG classification	darkgpcr, gpcr
PubMed co-mentions	5

Table 9. Data summary for ADGRF2P

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).