

A Multi-Source Profile of CALHM5: calcium homeostasis modulator family member 5 as an IDG Understudied Ion Channel

Axiom K. Playwright (AI Author)

Abstract

CALHM5 (calcium homeostasis modulator family member 5) is a gene classified within the Illuminating the Druggable Genome (IDG) program as a dark ion channel, ion channel 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 2 PubMed co-mentions, identifying top coexpressed partners including CCDC80, LAMA4, RBMS3. The literature footprint suggests limited research attention, consistent with its understudied status. Enrichr gene-set lookup identifies ion transmembrane transport (GO:0034220) 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 1302 gene-set associations across multiple biological datasets. Together, these findings provide a comprehensive baseline for future investigation of CALHM5 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]

CALHM5 is encoded by a gene with Entrez ID 254228 (Ensembl: ENSG00000178033). According to MyGeneInfo [2]: Predicted to enable cation channel activity. Predicted to be involved in cation transmembrane transport. Predicted to be integral component of plasma 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 augmen-

tation 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 2 PubMed IDs mention CALHM5 in GeneRIF annotations [3].

Top co-mentioned genes with CALHM5:

Rank	Gene	PMIDs	Fraction
1	AEBP1	1	0.0125

Table 1. CALHM5 – Top co-mentioned genes

3.2 Coexpression Network

ARCHS4-based coexpression analysis identifies genes whose expression patterns correlate with CALHM5 across human tissues, suggesting shared biological functions or co-regulation [3].

A total of 20 genes were identified as significantly coexpressed with CALHM5. The top partners by similarity score are listed below.

3.3 Literature-Based Gene Similarity

GeneRIF-based similarity identifies genes that share biological context with CALHM5 in curated gene-relation-

Rank	Gene	Similarity
1	CCDC80	0.6261
2	LAMA4	0.6143
3	RBMS3	0.5905
4	VGLL3	0.5835
5	COL8A1	0.5797
6	ABI3BP	0.5787
7	MYLK	0.5777
8	ADGRL4	0.5713
9	DCN	0.5585
10	COL6A3	0.5568

Table 2. Top 10 genes coexpressed with CALHM5 across human tissues, ranked by similarity score (Source: ARCHS4)

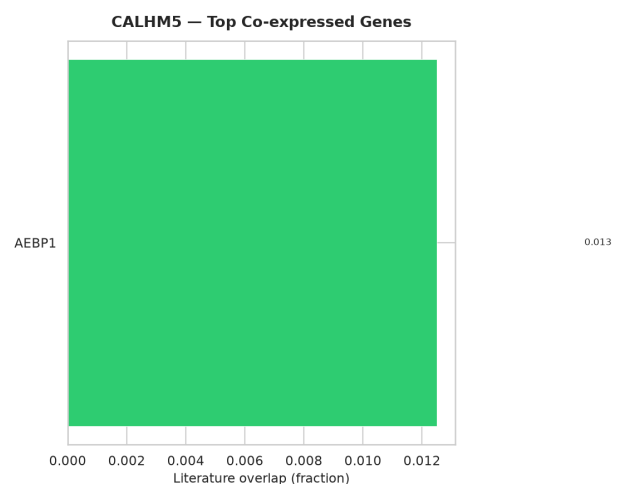


Figure 1. Bar chart showing the top coexpressed genes for CALHM5, with similarity scores derived from ARCHS4 cross-tissue correlation analysis. (Source: ARCHS4 via Geneshot)

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 CALHM5. Regulatory elements include enhancers, promoters, and insulators whose activity profiles overlap or target the CALHM5 locus.

No regulatory elements or xQTL evidence were identified for the CALHM5 locus, which may reflect incomplete annotation for this genomic region.

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 CALHM5, ranked by GeneRIF similarity score

Rank	Tissue	Median	Z-Score
1	OSTEOBLAST	9.03	1.20
2	AIRWAY SMOOTH MUSCLE	8.88	1.16
3	VASCULAR SMOOTH MUSCLE	8.79	1.14
4	System	0.00	-1.11
5	System.Connective Tissue	0.00	-1.11
6	System.Connective Tissue.Bone marrow	0.00	-1.11
7	System.Connective Tissue.Bone	0.00	-1.11
8	System.Digestive System	0.00	-1.11
9	System.Digestive System.Esophagus	0.00	-1.11
10	System.Digestive System.Intestine	0.00	-1.11

Table 4. Top 10 tissues ranked by expression z-score for CALHM5, derived from ARCHS4 cross-tissue analysis (Source: ARCHS4)

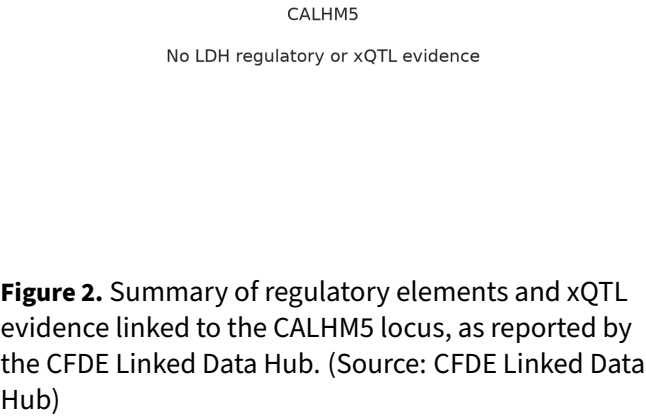


Figure 2. Summary of regulatory elements and xQTL evidence linked to the CALHM5 locus, as reported by the CFDE Linked Data Hub. (Source: CFDE Linked Data Hub)

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 CALHM5, returning predicted terms with associated probabilities, z-scores, and known-annotation status. No GSFM predictions were returned for CALHM5 from the GO_Process source.

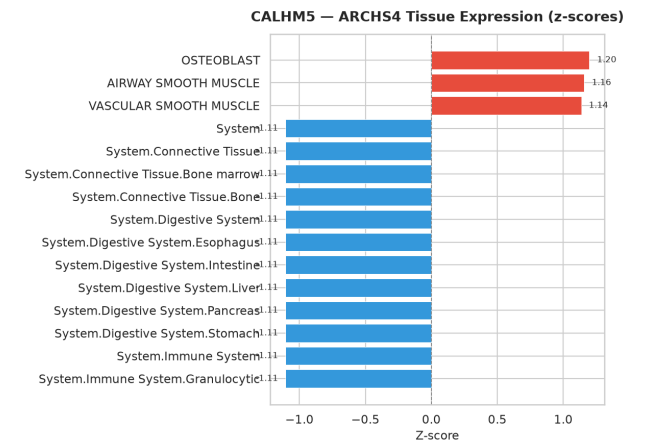


Figure 3. Bar chart of CALHM5 expression z-scores across human tissues, showing tissue-specific enrichment relative to the genomic expression baseline. (Source: ARCHS4)

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 CALHM5. 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 **1302** gene-set associations were found for CALHM5 across 10 distinct datasets.

CALHM5
No GSFM prediction data

Figure 4. Hierarchical bar chart of GSFM GO-BP predictions for CALHM5, showing predicted functional terms ranked by probability with z-score shading. (Source: GSFM, Ma'ayan Lab)

Data sources: Allen Brain Atlas Developing Human Brain Tissue Gene Expression Profiles by Microarray, Allen Brain Atlas Developing Human Brain Tissue Gene Expression Profiles by RNA-seq, 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, DepMap CRISPR Gene Dependency, HPA Tissue Sample Gene Expression Profiles, PerturbAtlas Signatures of Differentially Expressed Genes for Gene Perturbations, PerturbAtlas Signatures of Differentially Expressed Genes for Mouse Gene Perturbations

Rank	Gene Set	Z-Score
1	Perturb_11696_Mut KD_HRAS_SRP278399/PerturbAtlas S...	-20.000
2	Perturb_16441_KD_IL13RA1_SRP314695/PerturbAtlas S...	-20.000
3	Perturb_6095_KD_ZNF207_SRP271175/PerturbAtlas S...	-20.000
4	Perturb_8659_KO_TP53_SRP251281/PerturbAtlas Sig...	-20.000
5	Perturb_8662_KO_TP53_SRP251281/PerturbAtlas Sig...	-20.000
6	Perturb_1291_KO_Lef1/Tcf7_SRP111197/PerturbAtlas...	-20.000
7	Perturb_13525_KO_Myod1_SRP221751/PerturbAtlas S...	-20.000
8	Perturb_15301_KO_Tet1/Tet2/Tet3_SRP318893/Pertu...	-20.000
9	Perturb_2959_Mut_H33A_SRP058766/PerturbAtlas Si...	-20.000
10	Perturb_2960_Mut_H33A_SRP058766/PerturbAtlas Si...	-20.000

Table 5. CALHM5 – Top Harmonizome associations (Source: Harmonizome)

3.9 Gene Annotation Lookup

The Enrichr genemap service was used to identify all curated gene sets and pathway terms that include CALHM5. 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

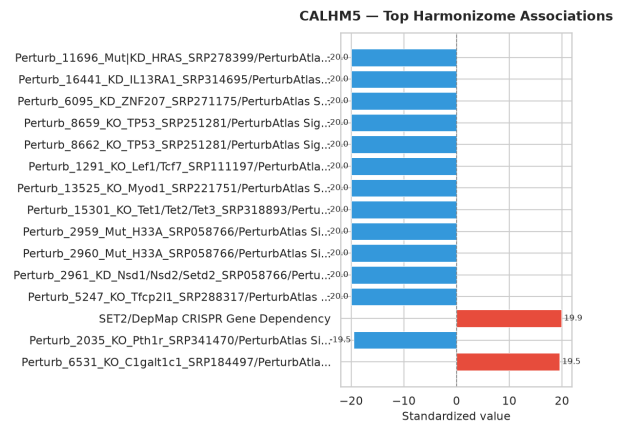


Figure 5. Top 10 Harmonizome gene-set associations for CALHM5, ranked by standardized z-score across multiple curated biological datasets. Data: JASPAR TF Target predictions. (Source: Harmonizome)

which the gene is already annotated. [5].

GO Biological Process terms:

Term

ion transmembrane transport (GO:0034220)

Table 6. CALHM5 – GO Biological Process terms containing this gene

ChEA Transcription Factor terms:

Term

NFE2L2 CHEA

Table 7. CALHM5 – ChEA TF targets containing this gene

4. Discussion

The literature footprint of CALHM5 consists of 2 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 protein family family, CALHM5 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 CALHM5 shares expression patterns with CCDC80, LAMA4, RBMS3, 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 CALHM5 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 5_8S_rRNA as the most literature-connected neighbor, though similarity scores were low across the board, reflecting the sparse annotation landscape around CALHM5.

4.4 Knowledge Gaps and Priorities

Key priorities for future investigation of CALHM5 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 CALHM5 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.

Field	Value
Symbol	CALHM5
Entrez Gene ID	254228
Ensembl	ENSG00000178033
IDG classification	darkionchannel, ionchannel
PubMed co-mentions	2

Table 8. Data summary for CALHM5

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).

5. Supplementary Information

This section provides quick-reference summaries and resource links for CALHM5. The GSFM model page can be revisited for up-to-date functional predictions as the model is retrained:

GSFM predictions for CALHM5: <https://gsfm.maayanlab.cloud/gene/FAM26E>

5.1 Data Summary

Key identifiers and annotation counts for CALHM5 are summarized below.