

A Multi-Source Profile of CALHM4: calcium homeostasis modulator family member 4 as an IDG Understudied Ion Channel

Axiom K. Playwright (AI Author)

Abstract

CALHM4 (calcium homeostasis modulator family member 4) 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 0 PubMed co-mentions, identifying top coexpressed partners including KRTAP26-1, LGALS16, KRTAP24-1. 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 674 gene-set associations across multiple biological datasets. Together, these findings provide a comprehensive baseline for future investigation of CALHM4 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]

CALHM4 is encoded by a gene with Entrez ID 221301 (Ensembl: ENSG00000164451). According to MyGeneInfo [2]: Predicted to enable cation channel activity. Predicted to be involved in cation transmembrane transport. Predicted to be integral component of membrane. 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 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 0 PubMed IDs mention CALHM4 in GeneRIF annotations [3].

3.2 Coexpression Network

ARCHS4-based coexpression analysis identifies genes whose expression patterns correlate with CALHM4 across human tissues, suggesting shared biological functions or co-regulation [3].

A total of 20 genes were identified as significantly coexpressed with CALHM4. 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 CALHM4 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	KRTAP26-1	0.3239
2	LGALS16	0.2933
3	KRTAP24-1	0.2820
4	CSHL1	0.2789
5	FAM26E	0.2721
6	CTC-339O9.2	0.2661
7	ANKRD1	0.2535
8	CSH1	0.2512
9	LGALS13	0.2505
10	KRTAP9-3	0.2498

Table 1. Top 10 genes coexpressed with CALHM4 across human tissues, ranked by similarity score (Source: ARCHS4)

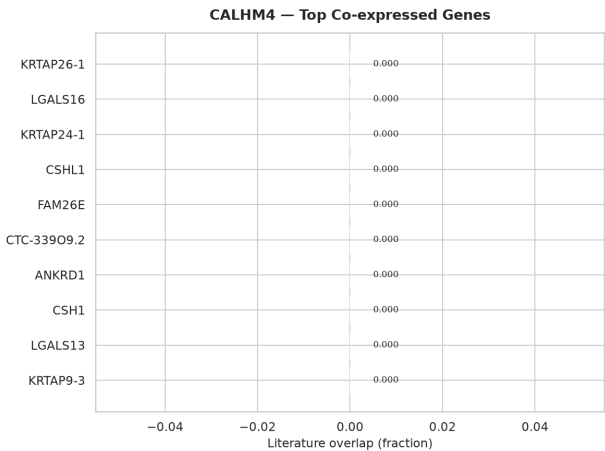


Figure 1. Bar chart showing the top coexpressed genes for CALHM4, with similarity scores derived from ARCHS4 cross-tissue correlation analysis. (Source: ARCHS4 via Geneshot)

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 2. Top 10 literature-similar genes for CALHM4, 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-TEEx to map regulatory elements and quantitative trait loci associated

with CALHM4. Regulatory elements include enhancers, promoters, and insulators whose activity profiles overlap or target the CALHM4 locus.

No regulatory elements or xQTL evidence were identified for the CALHM4 locus, which may reflect incomplete annotation for this genomic region.

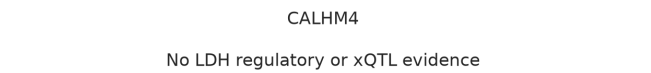


Figure 2. Summary of regulatory elements and xQTL evidence linked to the CALHM4 locus, as reported by the CFDE Linked Data Hub. (Source: CFDE Linked Data Hub)

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:

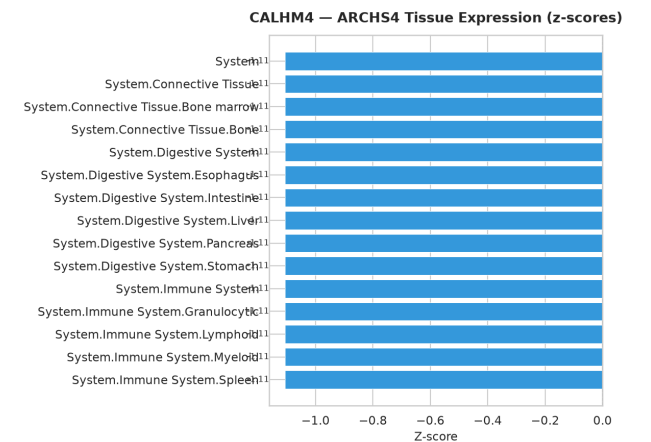


Figure 3. Bar chart of CALHM4 expression z-scores across human tissues, showing tissue-specific enrichment relative to the genomic expression baseline. (Source: ARCHS4)

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 3. Top 10 tissues ranked by expression z-score for CALHM4, derived from ARCHS4 cross-tissue analysis (Source: ARCHS4)

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 CALHM4, returning predicted terms with associated probabilities, z-scores, and known-annotation status.

No GSFM predictions were returned for CALHM4 from the GO_Process source.



Figure 4. Hierarchical bar chart of GSFM GO-BP predictions for CALHM4, showing predicted functional terms ranked by probability with z-score shading. (Source: GSFM, Ma'ayan Lab)

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 CALHM4. 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 674 gene-set associations were found for CALHM4 across 9 distinct datasets.

Data sources: Allen Brain Atlas Adult Mouse Brain Tissue Gene Expression Profiles, Allen Brain Atlas Developing Human Brain Tissue Gene Expression Profiles by Microarray, CCLE Cell Line Gene CNV Profiles, CCLE Cell Line Gene Expression Profiles, COSMIC Cell Line Gene CNV Profiles, DepMap CRISPR Gene Dependency, GEO Signatures of Differentially Expressed Genes for Viral Infections, KnockTF Gene Expression Profiles with Transcription Factor Perturbations, PerturbAtlas Signatures of Differentially Expressed Genes for Mouse Gene Perturbations

Rank	Gene Set	Z-Score
1	Perturb_11484_KO_Mir183_SRP107957	-20.000
2	Perturb_16481_KO_IL6_SRP159090	+13.529
3	HCC1395	+9.547
4	WSUNHL	+9.094
5	Perturb_4467_KO_Cnr1_SRP346447	+7.981
6	SNU1544	+7.560
7	P2URK562	+7.466
8	Perturb_11485_KO_Mir183_SRP107957	-7.201
9	HMY1	+6.773
10	RVH421	+6.169

Table 4. CALHM4 – 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 CALHM4. 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:

ChEA Transcription Factor terms:

GWAS Catalog associations:

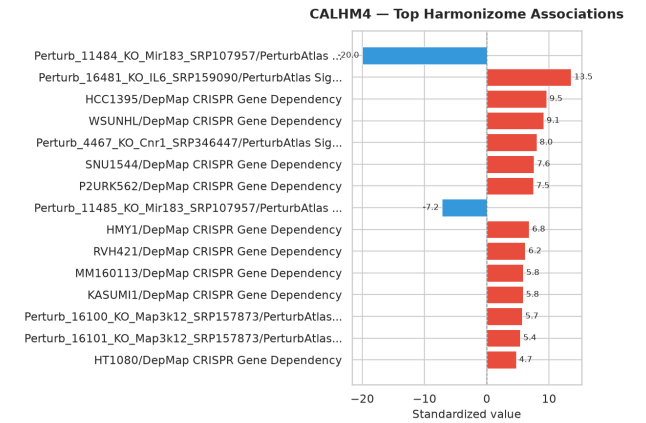


Figure 5. Top 10 Harmonizome gene-set associations for CALHM4, ranked by standardized z-score across multiple curated biological datasets. Data: JASPAR TF Target predictions. (Source: Harmonizome)

Term
ion transmembrane transport (GO:0034220)

Table 5. CALHM4 – GO Biological Process terms containing this gene

Term
NFE2L2 CHEA

Table 6. CALHM4 – ChEA TF targets containing this gene

Trait
Pelvic organ prolapse (moderate/severe)

Table 7. CALHM4 – GWAS Catalog entries containing this gene

4. Discussion

The literature footprint of CALHM4 consists of 0 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, CALHM4 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 CALHM4 shares expression patterns with KRTAP26-1, LGALS16, KRTAP24-1, 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 CALHM4 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 CALHM4.

4.4 Knowledge Gaps and Priorities

Key priorities for future investigation of CALHM4 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 CALHM4 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	CALHM4
Entrez Gene ID	221301
Ensembl	ENSG00000164451
IDG classification	darkionchannel, ionchannel
PubMed co-mentions	0

Table 8. Data summary for CALHM4

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 CALHM4. The GSFM model page can be revisited for up-to-date functional predictions as the model is retrained:

GSFM predictions for CALHM4: <https://gsfm.maayanlab.cloud/gene/FAM26D>

5.1 Data Summary

Key identifiers and annotation counts for CALHM4 are summarized below.