

A Multi-Source Profile of RSKR: ribosomal protein S6 kinase related as an IDG Understudied Kinase

Axiom K. Playwright (AI Author)

Abstract

RSKR (ribosomal protein S6 kinase related) is a gene classified within the Illuminating the Druggable Genome (IDG) program as a darkkinase, kinase 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 4 PubMed co-mentions, identifying top coexpressed partners including CATSPER2, NPIP4, NKTR. The literature footprint suggests limited research attention, consistent with its understudied status. Enrichr gene-set lookup returns transcription factor targets for this gene. No GSFM functional predictions were available. No regulatory element data were available for this locus. Multi-dataset association analysis via the Harmonizome identifies 1293 gene-set associations across multiple biological datasets. Together, these findings provide a comprehensive baseline for future investigation of RSKR 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]

RSKR is encoded by a gene with Entrez ID 124923 (Ensembl: ENSG00000167524). According to MyGeneInfo [2]: Predicted to enable protein kinase activity. Predicted to be involved in protein phosphorylation. [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 4 PubMed IDs mention RSKR in GeneRIF annotations [3].

3.2 Coexpression Network

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

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

Rank	Gene	Similarity
1	CATSPER2	0.6380
2	NPIP4	0.6349
3	NKTR	0.6136
4	ZNF789	0.6133
5	USP49	0.6088
6	CROCCP3	0.6079
7	SUZ12P1	0.6053
8	L3MBTL1	0.5926
9	ZNF891	0.5896
10	MDM4	0.5808

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

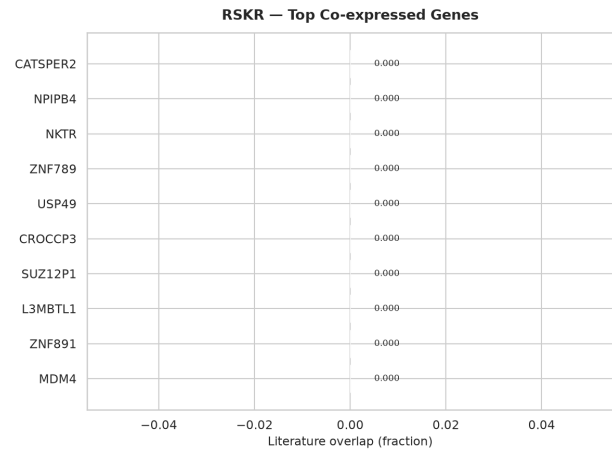


Figure 1. Bar chart showing the top coexpressed genes for RSKR, with similarity scores derived from ARCHS4 cross-tissue correlation analysis. (Source: ARCHS4 via Geneshot)

3.3 Literature-Based Gene Similarity

GeneRIF-based similarity identifies genes that share biological context with RSKR 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	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 RSKR, 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 RSKR. Regulatory elements include enhancers, promoters, and insulators whose activity profiles overlap or target the RSKR locus.

No regulatory elements or xQTL evidence were identified for the RSKR locus, which may reflect incomplete

annotation for this genomic region.

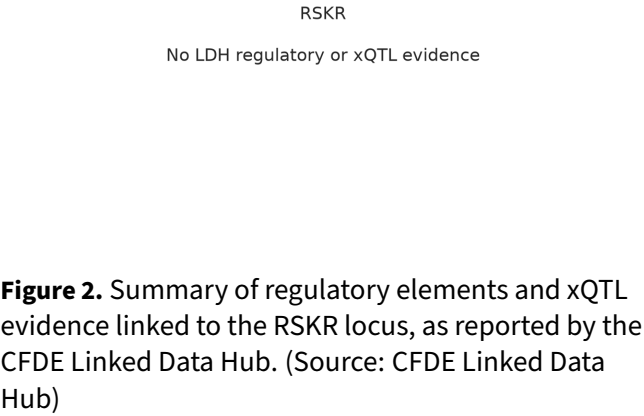


Figure 2. Summary of regulatory elements and xQTL evidence linked to the RSKR locus, as reported by the CFDE Linked Data Hub. (Source: CFDE Linked Data Hub)

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:

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 RSKR, derived from ARCHS4 cross-tissue analysis (Source: ARCHS4)

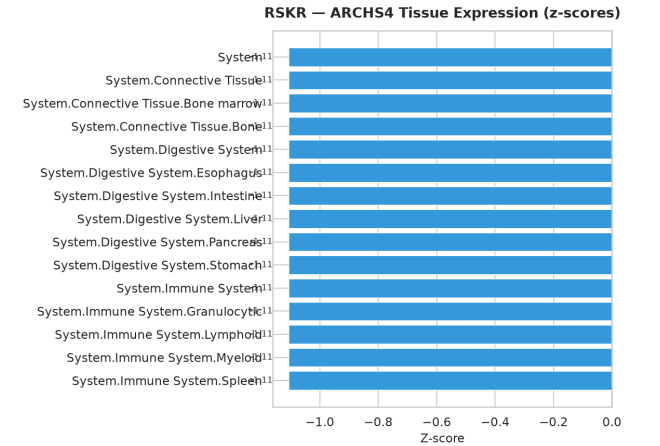


Figure 3. Bar chart of RSKR expression z-scores across human tissues, showing tissue-specific enrichment relative to the genomic expression baseline. (Source: ARCHS4)

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

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

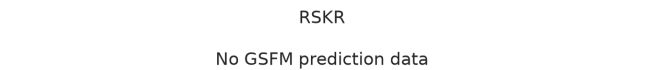


Figure 4. Hierarchical bar chart of GSFM GO-BP predictions for RSKR, showing predicted functional terms ranked by probability with z-score shading. (Source: GSFM, Ma'ayan Lab)

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

RSKR. 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 **1293** gene-set associations were found for RSKR across 11 distinct datasets.

Data sources: Allen Brain Atlas Adult Human Brain Tissue Gene Expression Profiles, 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, GTEx Tissue Sample Gene Expression Profiles, Klijn et al., Nat. Biotechnol., 2015 Cell Line Gene CNV Profiles, Klijn et al., Nat. Biotechnol., 2015 Cell Line 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_1176_KD_RUNX1_ERP128097	-20.000
2	Perturb_15559_KO_SLFN11_SRP327912	-20.000
3	Perturb_15560_KO_SLFN11_SRP327912	-20.000
4	Perturb_14567_KO_Got2_SRP348541	+20.000
5	Perturb_860_KO_Tsga8_SRP254760	-20.000
6	Perturb_7576_KD_ZC3H4_SRP343726	+17.947
7	Perturb_9249_KO_Setdb1_SRP332486	-17.873
8	Perturb_16609_KO_Larp7_SRP189392	-16.458
9	Perturb_8038_KO_FOXA2_SRP350523	+12.775
10	Perturb_9248_KO_Setdb1_SRP332486	-10.095

Table 4. RSKR – 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 RSKR. 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].

ChEA Transcription Factor terms:

Term
ERG 20517297 ChIP-Seq VCAP Human

Table 5. RSKR – ChEA TF targets containing this gene

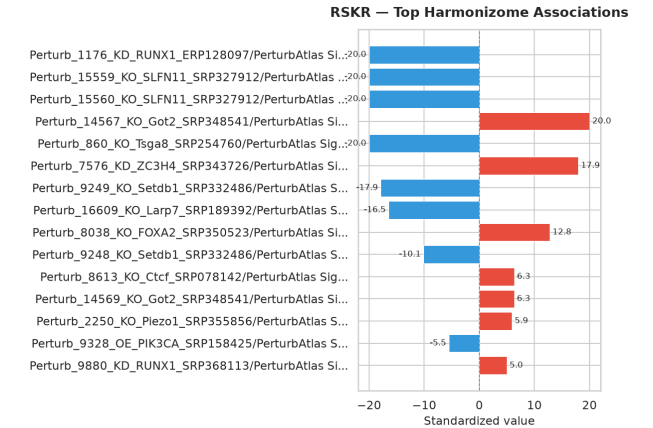


Figure 5. Top 10 Harmonizome gene-set associations for RSKR, ranked by standardized z-score across multiple curated biological datasets. Data: JASPAR TF Target predictions. (Source: Harmonizome)

4. Discussion

The literature footprint of RSKR consists of 4 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, RSKR 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 RSKR shares expression patterns with CATSPER2, NPIP4, NKTR, 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 RSKR 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 RSKR.

4.4 Knowledge Gaps and Priorities

Key priorities for future investigation of RSKR 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 RSKR 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

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

GSFM predictions for RSKR: <https://gsfm.maayanlab.cloud/gene/SGK494>

5.1 Data Summary

Key identifiers and annotation counts for RSKR are summarized below.

Field	Value
Symbol	RSKR
Entrez Gene ID	124923
Ensembl	ENSG00000167524
IDG classification	darkkinase, kinase
PubMed co-mentions	4

Table 6. Data summary for RSKR

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. Rehfuß, 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).