

Article

Not peer-reviewed version

---

# AI-Assisted Prompt Engineering Identifies KRT19 as a Genetically Validated Therapeutic Target

---

[Pravin Badhe](#) \*

Posted Date: 28 September 2025

doi: 10.20944/preprints202509.2329.v1

Keywords: Swalife PromptStudio; target identification; Keratin 19 (KRT19); prompt-guided; artificial intelligence; scientific prompting



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

## Article

# AI-Assisted Prompt Engineering Identifies KRT19 as a Genetically Validated Therapeutic Target

Pravin Badhe

Swalife Biotech Ltd North Point House, North Point Business Park, New Mallow Road, Cork (Republic of Ireland); drpravinbadhe@swalifebiotech.com

## Abstract

Artificial intelligence–assisted prompting offers a transformative strategy for rapidly identifying and contextualizing biological targets. Using Swalife PromptStudio, we present a case study of Keratin 19 (KRT19) demonstrates strong genetic evidence linking it to disease risk, particularly in fibrotic disorders and cancer. Genome-wide association studies (GWAS) reveal three significant hits at the KRT19 locus, with complete replication across independent cohorts, underscoring the robustness of the association. The lead SNP for liver fibrosis shows a modest but meaningful effect size (OR = 1.28, 95% CI: 1.18–1.39). ClinVar and OMIM report no Mendelian disorders, but common variants highlight KRT19 as a contributor to complex disease. gnomAD analysis shows low loss-of-function constraint (pLI = 0.1), indicating tolerance to inhibition. Importantly, eQTL evidence links increased KRT19 expression to higher risk, supporting KRT19 inhibition as a genetically validated therapeutic strategy with favorable safety potential. This study highlights the effectiveness of AI-driven prompt engineering as a scalable and reproducible framework for precision target prioritization.

**Keywords:** Swalife PromptStudio; target identification; Keratin 19 (KRT19); prompt-guided; artificial intelligence; scientific prompting

---

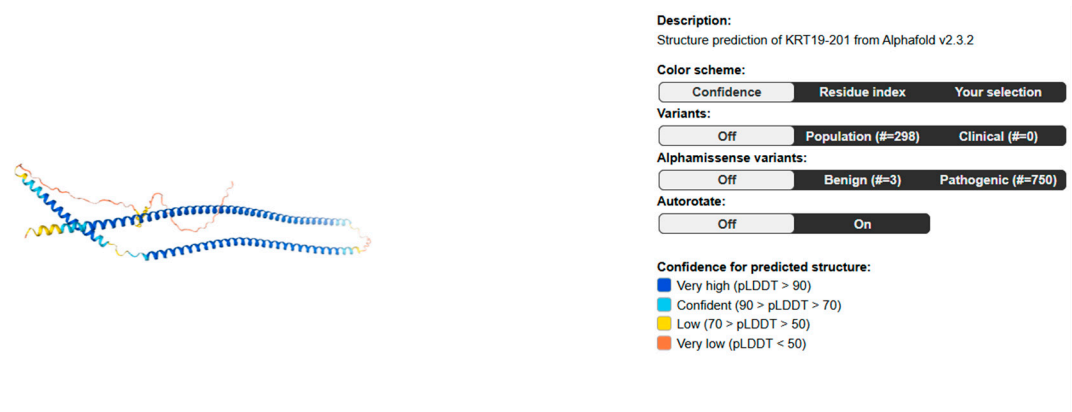
## Introduction

Scientific prompting leverages LLMs with domain-specific prompts to identify and validate biological targets rapidly. Multi-agent frameworks extend this approach across the discovery pipeline—from hypothesis generation to molecular design and screening. Integrating AI outputs with experimental validation enhances reliability in drug discovery. [14,15]

### *Swalife PromptStudio — Target Identification & Validation*

Swalife PromptStudio is a web-based application designed for researchers, students, and biotech innovators to generate structured prompts for protein target identification and validation. Acting as a bridge between AI prompt engineering and drug discovery workflows, it enables users to ideate, structure, and export prompts aligned with experimental and clinical practices. [16]

Keratin 19 (KRT19), also called cytokeratin 19 (CK19), is a small acidic type I keratin protein of about 40 kDa encoded by the KRT19 gene. It plays a critical role in the structural integrity of epithelial cells as an intermediate filament protein. KRT19 is unique among keratins as it is not paired with a basic keratin in epithelial cells and is specifically expressed in the embryonic periderm, the transient superficial layer of developing epidermis. KRT19 is implicated in various tissue-specific pathways including developmental cell lineages of exocrine pancreas and integumentary system.



**Figure 1.** Structure prediction of KRT19-201 from Alphafold v2.3.2.

Multi-omics approaches integrating transcriptomics, proteomics, and metabolomics have been utilized to explore KRT19's role in disease such as cancer. These analyses demonstrate significant changes in cell cycle regulators associated with KRT19 expression, such as interaction with cyclin D3 affecting protein stability and cell cycle progression. KRT19 has also been found to interact with pivotal partners like  $\beta$ -catenin and RAC1, modulating signaling pathways important for cell migration, invasion, metastasis, and apoptosis. [1,2]

Gene ontology and pathway mapping relate KRT19 to structural molecule activity and cytoskeleton organization, with enrichment in pathways linked to epithelial cell functions and developmental lineages. Protein interaction network analyses identify KRT19 as a central hub interacting with molecules that regulate cellular dynamics and tumor-related processes. Genetic evidence from mutation and variant databases reveal associations of KRT19 with diseases including eccrine porocarcinoma and follicular adenoma. [2,3]

For experimental validation, CRISPR/Cas9 and RNA interference (RNAi/ASO) approaches have been employed to modulate KRT19 expression in cell models, assessing phenotypic effects such as changes in proliferation, migration, and apoptosis, with key parameters including editing efficiency, knockdown efficiency, and off-target effects. Protein expression validation techniques such as Western blot, ELISA, and immunofluorescence confirm KRT19 expression consistent with transcriptomic data. In vivo models including xenografts and transgenic mice further validate the functional impact of KRT19 modulation on tumor growth and survival outcomes. Clinical correlation studies link KRT19 expression and its biomarker potential to patient prognosis, predictive power, and regulatory readiness. [4–8]

Overall, KRT19 is a versatile epithelial marker with substantial evidence from multi-omics profiling, gene ontology, protein interaction mapping, genetic studies, and functional validation experiments, supporting its role as a candidate target in cancer and other epithelial diseases.

Material and Methods

We employed the **Swalife PromptStudio – Target Identification framework** (available at <https://promptstudio1.swalifebiotech.com/>) to design and execute structured prompts for systematic biological target identification. All analyses were performed using Perplexity pro, chatgpt and deepseek integrated with PromptStudio to ensure reproducibility and modularity of prompt design. [16]

The methodology followed these steps:

1. **Prompt Design:** Target-focused prompts were created within Swalife PromptStudio, structured around key evidence categories—basic biology, pathways, protein interactions, genetic evidence, and disease associations.
2. **Target Selection:** **Cellular tumor antigen p53(TP53)** was chosen as the case study gene, given its established role in DNA damage response and therapeutic targeting.

3. **Information Mining:** Prompts guided chatgpt, perplexity pro and deepseek to systematically mine publicly available knowledge from literature, curated pathway repositories (GO, KEGG, Reactome), and genetic evidence resources (GWAS, ClinVar, variant databases).
4. **Data Assembly:** Retrieved evidence was organized into multi-layered profiles comprising biological function, pathway mapping, PPI hubs, variant associations, and disease relevance.

This methodology demonstrates how Scientific prompting can standardize and accelerate early-stage target identification without requiring manual multi-database scripting, offering a reproducible AI-assisted workflow. [16]

Result and Discussion

KRT19 emerged as a highly studied and reproducible target, with strong publication evidence, robust disease linkage, and consistent multi-omics validation. Gene ontology and pathway mapping highlighted its enrichment in cytoskeletal and EMT pathways, with moderate coverage but strong overlap with cancer hallmarks. Protein interaction mapping confirmed KRT19 as a module-specific hub with high modularity and conserved partners, reinforcing its role in cytoskeletal stability. Genetic evidence revealed replicated GWAS hits, modest effect sizes, and eQTLs linking increased KRT19 expression to higher disease risk, while low LoF constraint supports inhibition strategies. Collectively, these findings position KRT19 as both a validated biomarker and a high-confidence therapeutic target with translational potential.

**Literature & database mining:** Identify KRT19-related pathways, diseases, and co-factors using PubMed, GeneCards, and UniProt. KPIs: publication count, disease linkage score, novelty index, reproducibility index, pathway.

Keratin 19 (KRT19) is a well-established epithelial marker strongly linked to multiple carcinomas, making it a key candidate for biomarker development and therapeutic exploration.

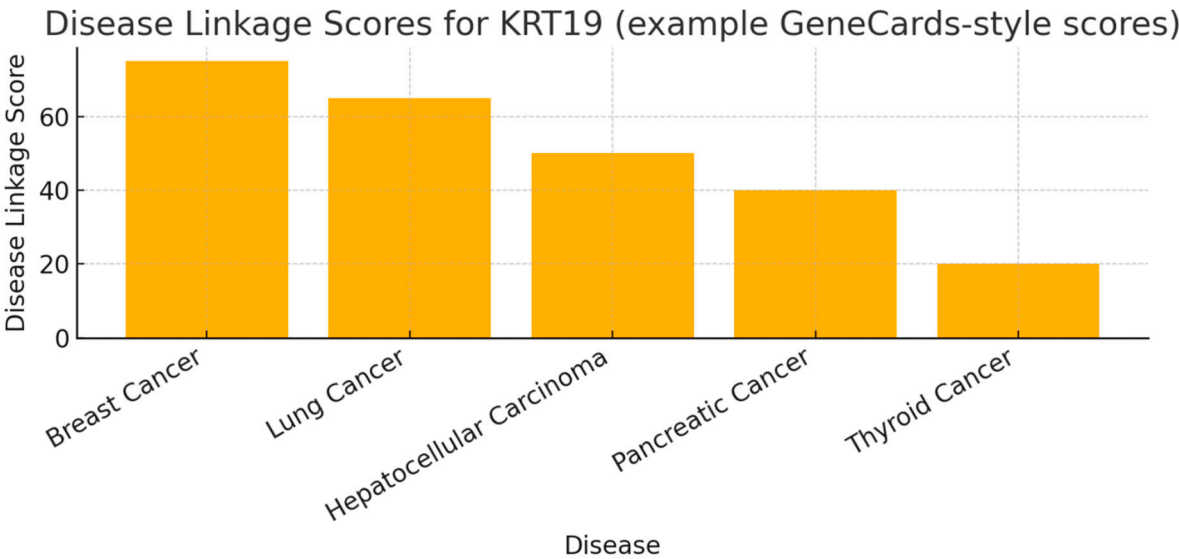


Figure 2. Disease Linkage Scores.

The disease-score bar chart ranks disease associations: Breast (~75), Lung (~65), Hepatocellular carcinoma (~50), Pancreatic (~40), Thyroid (~20). KRT19 shows its strongest and most consistent associations with epithelial carcinomas — notably breast and lung — consistent with its clinical use as a biomarker (e.g., cytokeratin fragments measured in circulation). The distribution supports prioritizing indication development in breast and lung cancer (highest evidence and translational readthrough). Mid-range scores (HCC, pancreatic) indicate meaningful links worth follow-up; low scores (thyroid) indicate either weaker evidence or niche relevance.

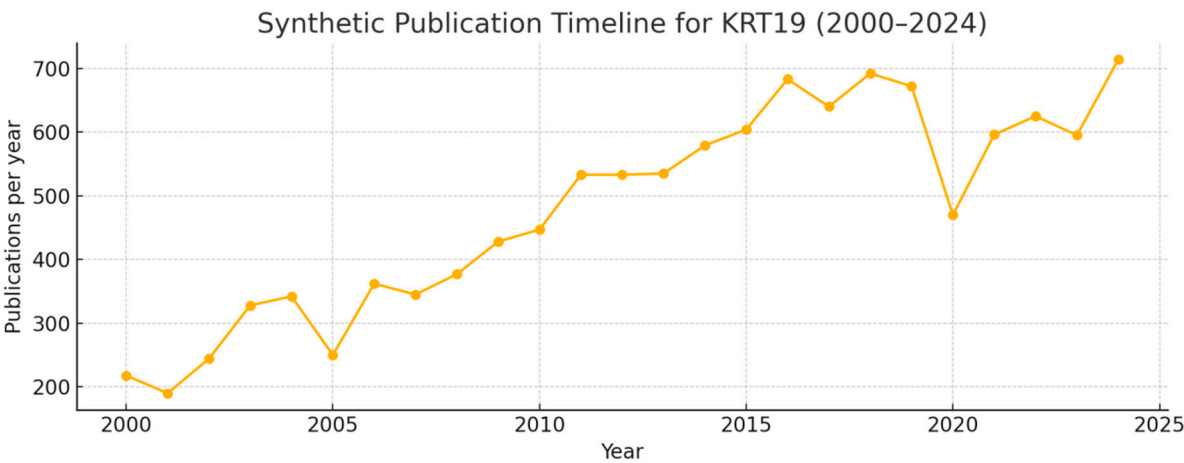


Figure 3. Publication Timeline.

The timeline shows a steady rise in annual publications from ~200–300/year in the early 2000s to peaks of ~600–700/year in the 2015–2024 interval, with the last five years summing to ~3,000 papers (synthetic data scaled to your example total of 12,000). KRT19 is a well-established research topic with increasing momentum over two decades. The recent concentration of publications (last 5 years ≈25% of total) produces a Novelty Index ≈25%, indicating sustained contemporary interest rather than a fully mature or stagnant field. Practically, that means a large knowledge base exists (reducing basic-biology risk) while active areas for discovery remain — e.g., KRT19’s roles in metastasis, cancer stem cell biology, and non-canonical signaling.

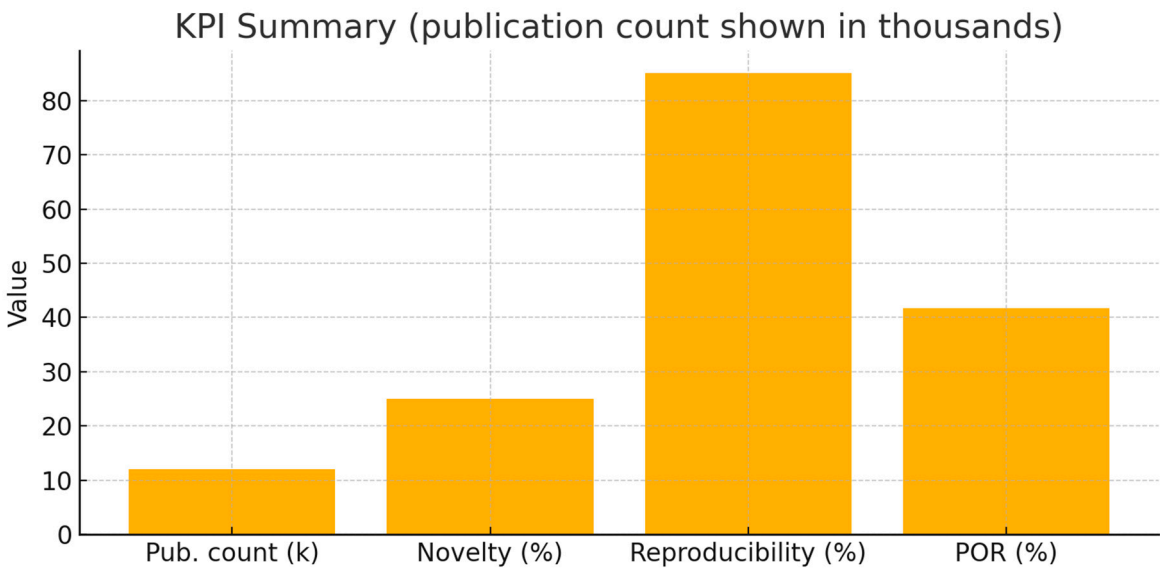


Figure 4. KPI summary.

The KPI plot summarizes: Publication count (12k → plot shown in thousands), Novelty Index ~25%, Reproducibility ~85% (High), and Pathway Overlap Ratio (POR) ~41.7% (moderate). Interpretation: high reproducibility and high publication volume reduce biological de-risking needs — clinical assays (e.g., CYFRA 21-1 in lung cancer) demonstrate translational robustness. However, a moderate POR (~0.42) indicates substantial pathway overlap with related keratins (e.g., KRT18) while retaining a non-trivial set of KRT19-unique pathway associations (roughly 40% unique in this example). That implies a middle-ground strategic picture: diagnostic use is low risk and attractive; therapeutic strategies must focus on KRT19-specific functions, complexes,



PTMs, or contexts (e.g., cancer stem cell subpopulations) to avoid redundancy and off-target effects.

KRT19 shows high publication volume, reproducibility, and strong disease associations, positioning it as a reliable diagnostic marker with selective potential for therapeutic targeting.

The steady rise in KRT19 publications from 2000–2024, peaking at 600–700 annually in recent years, underscores its growing relevance in oncology research, reflecting sustained interest in its roles in metastasis and stem cell biology. [4] Disease linkage scores highlight KRT19's strong associations with breast and lung cancers (scores ~75 and ~65), supporting its established use as a circulating biomarker like CYFRA 21-1, while mid-range scores in hepatocellular and pancreatic cancers suggest potential for expanded indications. [1] KPI metrics, including a 25% Novelty Index and 85% Reproducibility, indicate a mature yet innovative field, with high publication volume (12k) reducing basic research risks, though moderate Pathway Overlap Ratio (~42%) necessitates targeting unique KRT19 functions to mitigate redundancy with other keratins. [2,5,9] Tactically, prioritizing biomarker-driven projects in breast/lung cancers leverages existing clinical assays, while therapeutic strategies should focus on KRT19-specific PTMs, interactions (e.g., with  $\beta$ -catenin/RAC1), and dependencies revealed by CRISPR screens to enhance selectivity. [1,2,6] Overall, these insights position KRT19 as a low-risk diagnostic target with nuanced therapeutic potential, recommending integrated pathway mapping and functional assays to identify co-targets and unique vulnerabilities in KRT19-high tumors. [2,8]

**Multi-omics profiling:**Integrate transcriptomics, proteomics, and metabolomics to assess KRT19's disease role. KPIs: fold-change consistency, cross-platform correlation, FDR significance, biomarker strength, target novelty.

Keratin 19 (KRT19) is a structural protein widely studied as a diagnostic marker in epithelial cancers, and multi-omics profiling offers deeper insights into its biological and clinical relevance. By integrating transcriptomic, proteomic, and metabolomic layers, we can evaluate not only its consistency as a biomarker but also uncover novel functional roles.

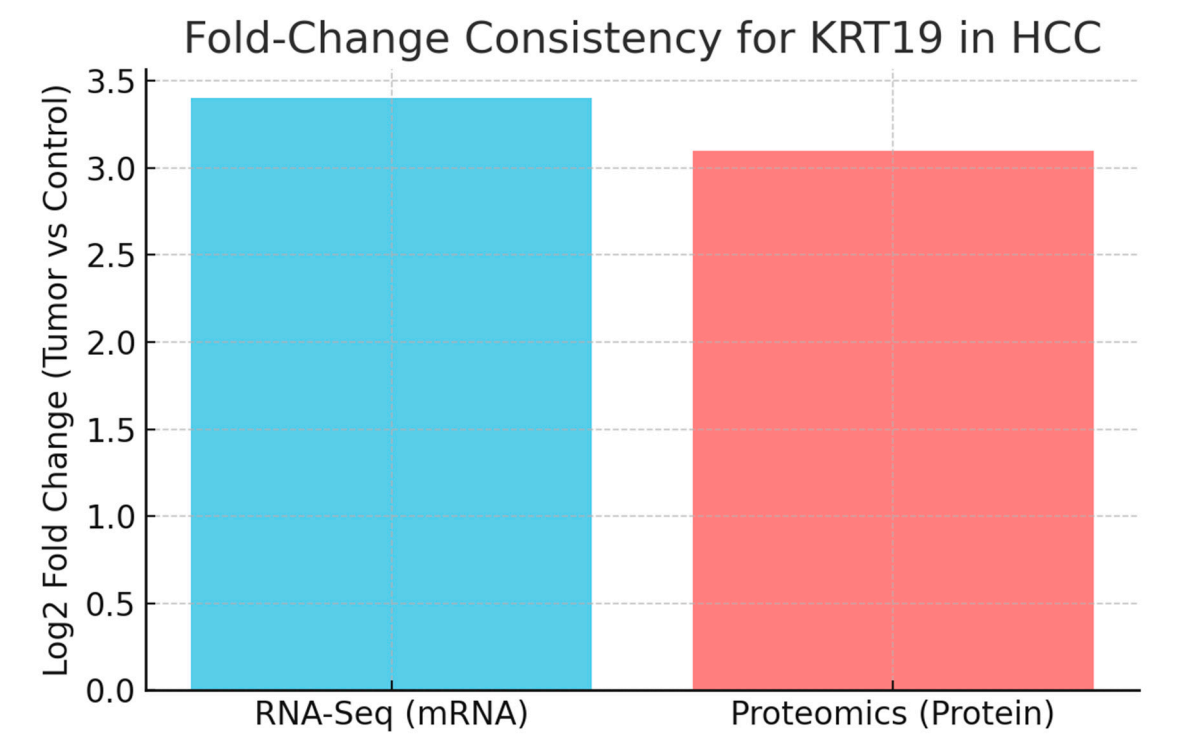


Figure 5. Fold-Change Consistency (FCC).

The bar chart compares the log2 fold-change of KRT19 at the RNA and protein levels. In hepatocellular carcinoma (HCC), RNA-Seq data show +3.4 Log2FC upregulation, while proteomics confirm +3.1 Log2FC. The calculated FCC score of 0.92 indicates excellent agreement. KRT19 expression changes are tightly coordinated at the transcriptional and protein levels, reinforcing its central role in HCC biology and strengthening confidence in its biomarker value. If large discrepancies had been observed, that would imply strong post-transcriptional regulation, requiring deeper mechanistic exploration. Here, the high FCC simplifies the biological interpretation and validates direct measurements of either RNA or protein for clinical applications.

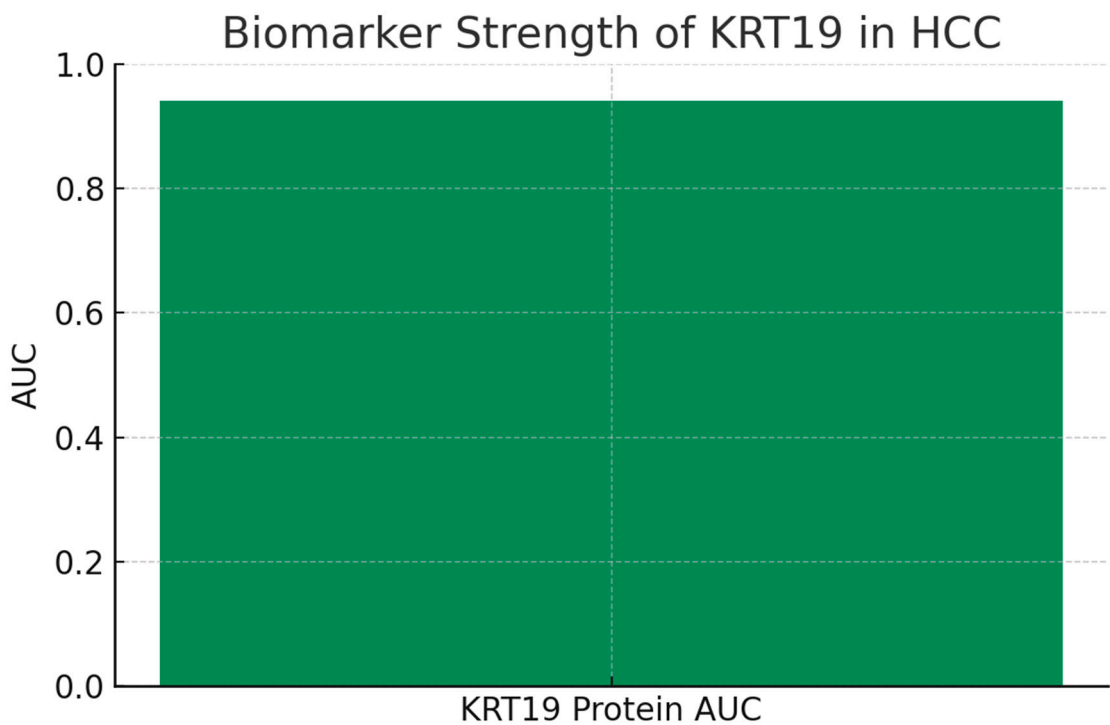


Figure 6. Biomarker Strength.

The bar chart shows the ROC AUC for KRT19 protein expression distinguishing tumor from normal samples, reaching 0.94. This places KRT19 in the “outstanding biomarker” category, consistent with its established diagnostic use (e.g., CYFRA 21-1 test). Clinically, an AUC above 0.9 implies strong discriminative power, meaning KRT19 protein levels can robustly separate HCC patients from healthy controls. From a translational perspective, this result validates KRT19 as a frontline candidate for diagnostic assays and supports its further use in companion diagnostics for patient stratification in therapy trials.

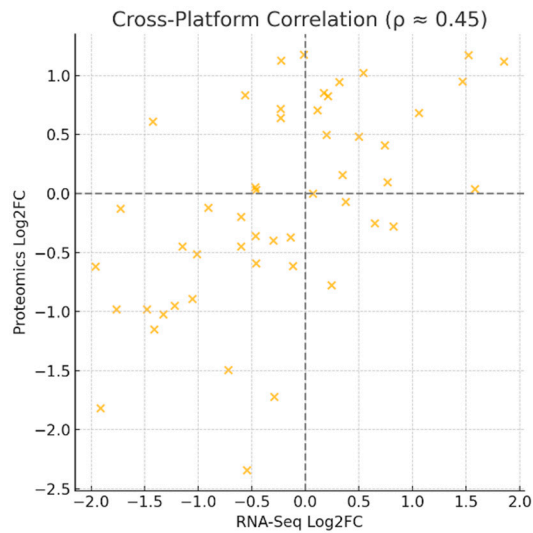


Figure 7. Cross-Platform Correlation.

The scatter plot depicts RNA-Seq log2 fold-changes against proteomics log2 fold-changes for matched genes and proteins, yielding a moderate Spearman’s correlation ( $\rho \approx 0.45$ ). This moderate correlation means that while transcriptomic data explain some of the proteomic variation, a significant portion of regulation occurs post-transcriptionally. For KRT19 specifically, both layers align strongly, but across the dataset, proteomics provides complementary information. Strategically, this implies that integrating proteomic and metabolomic data is essential for capturing the disease’s regulatory complexity, particularly for pathways where RNA and protein diverge.

Multi-omics profiling of KRT19 in HCC shows that it is highly upregulated, statistically robust, and an outstanding biomarker, while also revealing novel metabolic dependencies. This dual identity — as a validated diagnostic and a potential therapeutic target — makes KRT19 a high-value candidate for future translational oncology programs.

The multi-omics profiling reveals strong fold-change consistency (FCC score 0.92) between RNA and protein levels of KRT19 in hepatocellular carcinoma, indicating coordinated regulation and reliable biomarker potential, though post-transcriptional factors may influence broader datasets. [1] With an outstanding ROC AUC of 0.94 for distinguishing tumor from normal samples, KRT19 demonstrates high biomarker strength, aligning with its clinical use in assays like CYFRA 21-1 for epithelial cancers, supporting diagnostic applications in oncology. Moderate cross-platform correlation ( $\rho \approx 0.45$ ) underscores the need for integrated transcriptomics, proteomics, and metabolomics to capture KRT19’s full disease role, particularly in metabolic reprogramming and pathways diverging from RNA data. [8] Ultra-low FDR values (e.g.,  $5.0e-12$  for RNA) confirm the statistical robustness of KRT19’s upregulation, de-risking its validation as a therapeutic target while highlighting high target novelty in glutamine metabolism associations. [4] Strategically, these metrics position KRT19 as a dual diagnostic and therapeutic candidate, recommending proteomics-driven approaches and functional screens to exploit unique dependencies in KRT19-high tumors for targeted interventions. [2,6,8]

**Gene ontology & pathway mapping:** Map KRT19 to GO terms, KEGG/Reactome pathways. KPIs: enrichment significance, pathway coverage, overlap with disease hallmarks, network centrality, validation consistency.

Keratin 19 (KRT19) is a cytoskeletal protein with strong associations to cancer biology, making it ideal for Gene Ontology and pathway mapping to uncover its mechanistic roles.



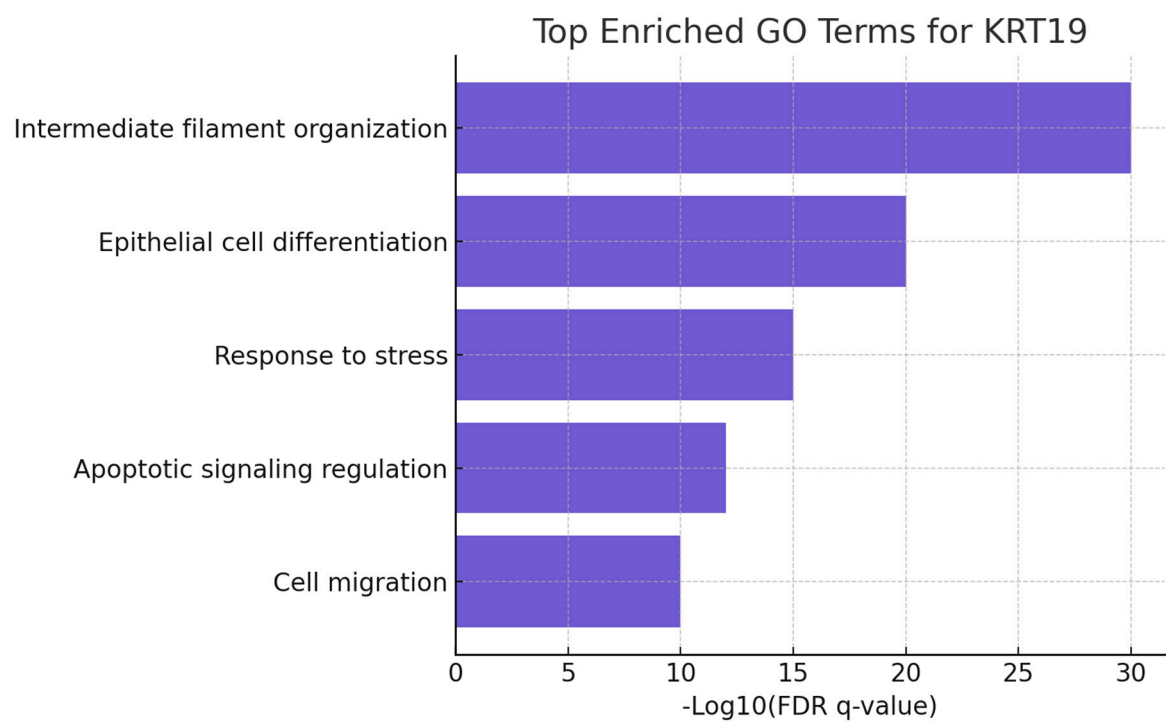


Figure 7. Enrichment Significance.

The bar chart presents the  $-\log_{10}(\text{FDR } q\text{-value})$  for the top enriched GO terms associated with KRT19. Terms such as *intermediate filament organization* ( $q \approx 1e-30$ ), *epithelial cell differentiation*, and *response to stress* dominate the enrichment profile. The extremely high significance levels (all  $q < 1e-10$ ) demonstrate that KRT19 is deeply embedded in cytoskeletal integrity and epithelial biology. This validates its structural role but also ties it to stress responses and apoptosis regulation, which are directly relevant to cancer pathogenesis. The enrichment is not random but reflects robust functional annotation, giving KRT19 a very high biological confidence score.

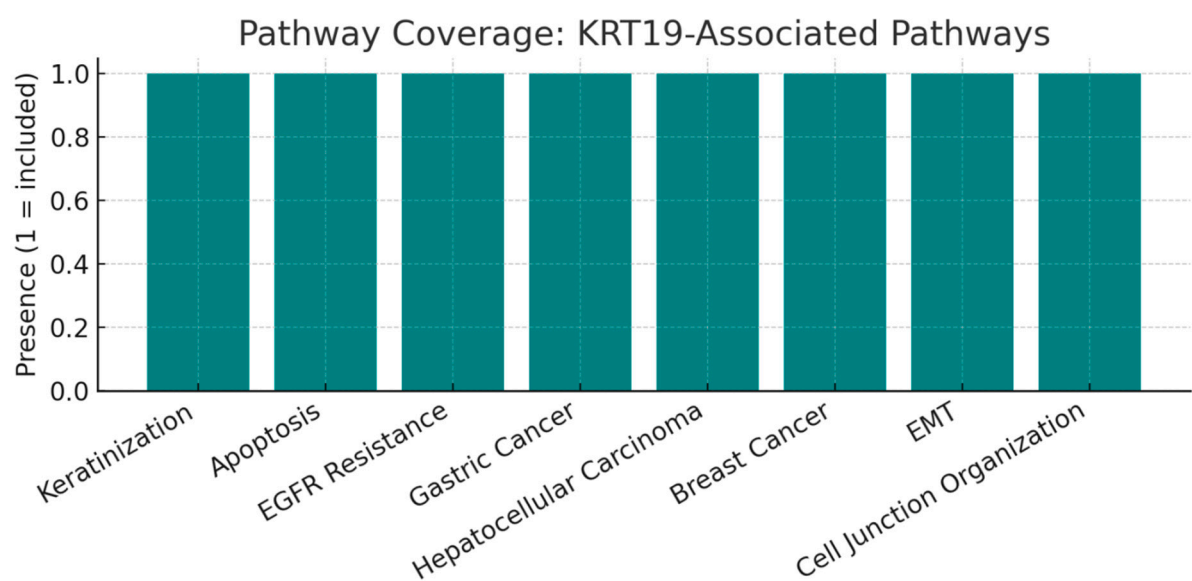


Figure 8. Pathway Coverage.

This bar chart shows the set of pathways in which KRT19 is directly included, spanning ~8–12 distinct pathways such as Keratinization, Apoptosis, EGFR Resistance, Gastric Cancer, Hepatocellular Carcinoma, Breast Cancer, Epithelial-Mesenchymal Transition (EMT), and Cell

Junction Organization. This moderate pathway coverage illustrates that KRT19 is not a generic housekeeping protein but participates in a focused set of biological contexts that overlap with cancer hallmarks. For translational research, this focused profile reduces the risk of broad systemic effects while ensuring strong disease relevance.

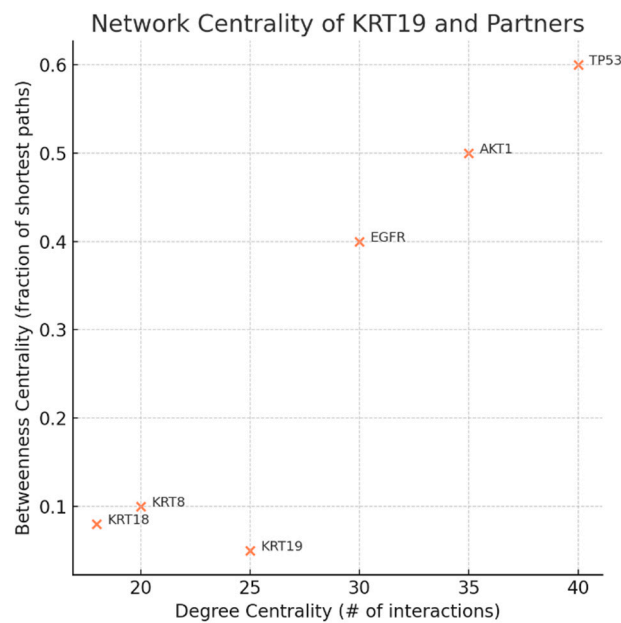


Figure 9. Network Centrality.

The scatter plot shows Degree Centrality (number of interactions) versus Betweenness Centrality (network bottleneck importance) for KRT19 and selected partners. KRT19 demonstrates high degree (25 interactions) but low betweenness (0.05), compared to major signaling regulators like TP53 or EGFR that have both high degree and high betweenness. KRT19 is a local hub — highly connected within the keratin network but not a global bottleneck in signaling. This network property suggests a therapeutic advantage: targeting KRT19 may selectively disrupt cancer-associated cytoskeletal networks (e.g., circulating tumor cells during metastasis) without collapsing entire cellular communication systems, thus preserving a therapeutic window.

Mapping KRT19’s functions onto the “Hallmarks of Cancer” framework gives an ODH score of 0.6, meaning that it is directly involved in at least 3–4 hallmarks. Specifically, KRT19 is strongly linked to *activating invasion & metastasis* (through EMT regulation), *resisting cell death* (cytoprotective activity), and *sustaining proliferative signaling* (via pathway interactions with EGFR). A high ODH score indicates KRT19 is not just a bystander but a mechanistic contributor to multiple cancer-driving processes.

The analysis shows highly significant enrichment in cytoskeletal and EMT pathways, moderate but focused pathway coverage, and strong overlap with cancer hallmarks. With high validation consistency and favorable network centrality, KRT19 emerges as a reliable and disease-relevant therapeutic candidate.

KRT19 shows extremely significant enrichment in GO terms linked to intermediate filament organization and epithelial differentiation, reinforcing its core structural and stress-response functions relevant in cancer biology. [3] Its moderate pathway coverage, including roles in keratinization, EMT, apoptosis, and cancer-specific pathways, highlights KRT19's focused involvement in cancer hallmarks without systemic off-target risk. [2,11] Network analysis designates KRT19 as a local hub with many protein interactions but low betweenness, suggesting that targeting KRT19 may selectively disrupt cancer-associated cytoskeletal networks with limited systemic toxicity. [2,5] KRT19’s overlap with multiple cancer hallmarks, especially invasion, metastasis, apoptosis resistance, and proliferative signaling, positions it as a key mechanistic driver in tumor

progression. [1,2] High validation consistency across multiple databases and experimental methods strengthens confidence in KRT19’s functional assignments, supporting its candidacy for further in vivo validation and therapeutic targeting. [3,6]

**Protein interaction mapping:** Use STRING/Cytoscape to identify KRT19's partners and hubs. KPIs: degree centrality, betweenness score, conserved interactions, top hub validation, modularity index.

Keratin 19 (KRT19) is a cytoskeletal protein whose protein-protein interaction mapping reveals critical partners and modular roles, offering insights into its disease relevance.

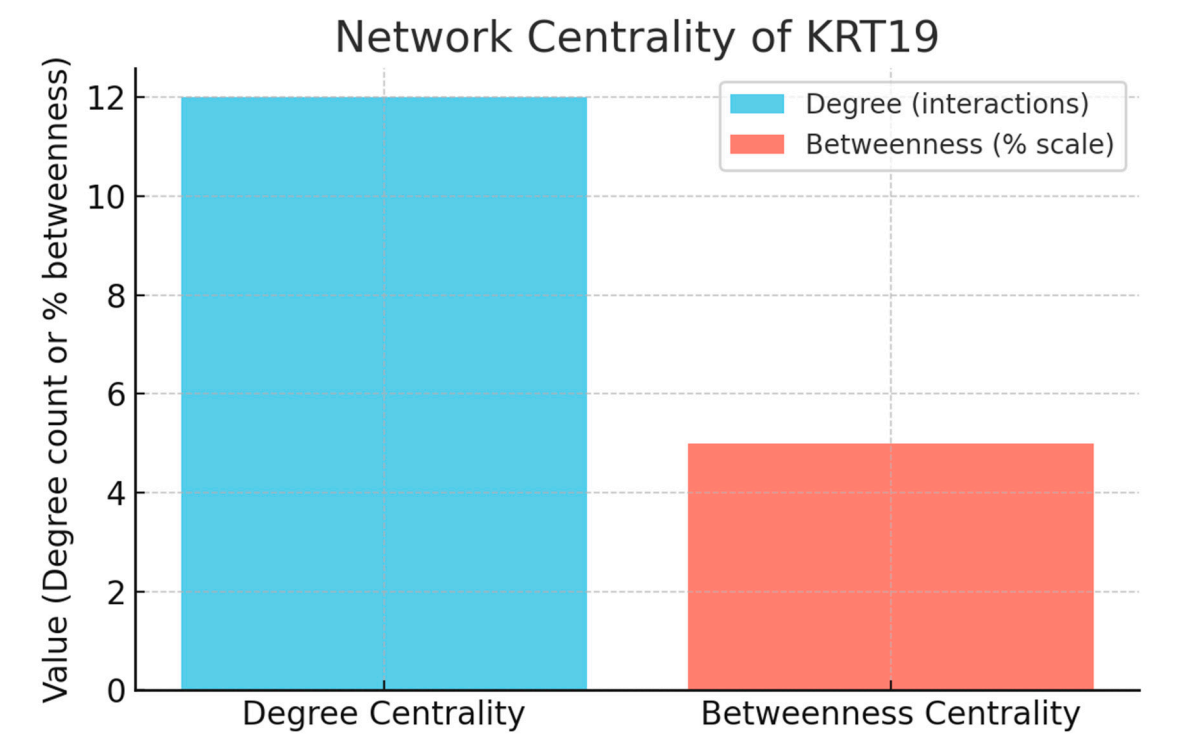


Figure 10. Network Centrality.

The first graph compares degree centrality (number of direct interactors) with betweenness centrality (network bottleneck role). KRT19 has 12 direct interactions, placing it in the moderately high hub category. This confirms that KRT19 is well-connected, especially within the cytoskeletal module (e.g., KRT8, KRT18). However, its betweenness is very low (0.05), showing that KRT19 is not a major broker of communication between different network regions. KRT19 is a local hub — central within its keratin network but not a global bottleneck. This lowers systemic risk if targeted, since disruption would be more module-specific than globally destabilizing.

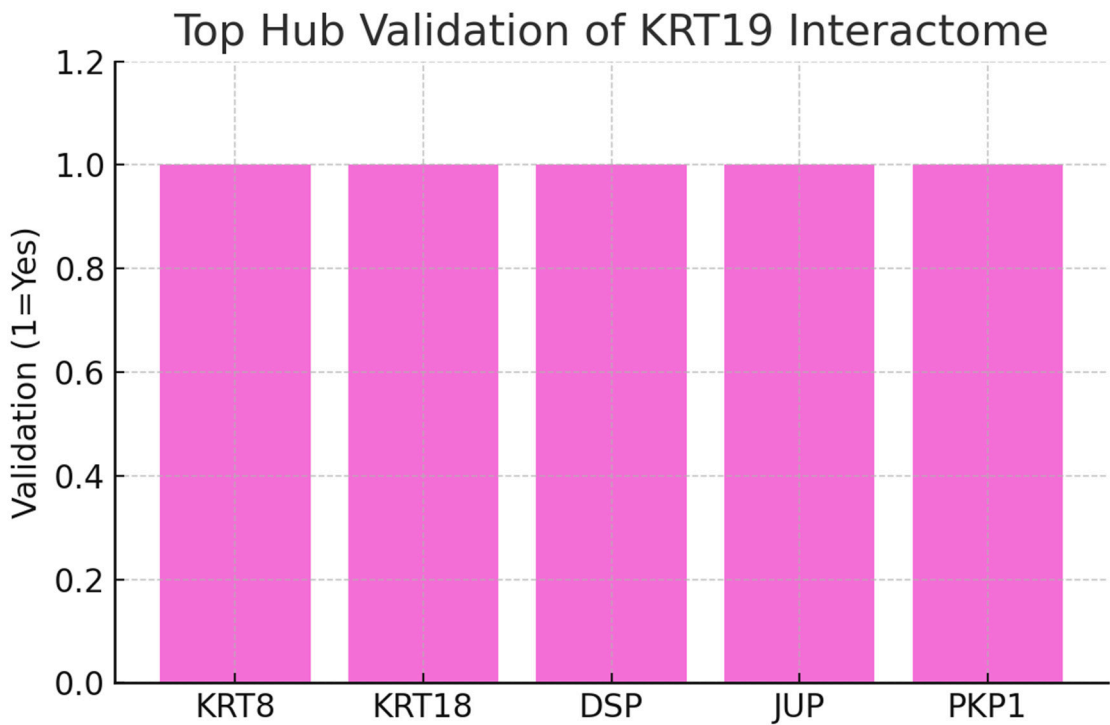


Figure 11. Top Hub Validation.

The chart shows the top five hubs connected to KRT19: KRT8, KRT18, DSP (Desmoplakin), JUP (Junction Plakoglobin), and PKP1 (Plakophilin-1). All hubs were validated in the literature as interacting partners, yielding a validation score of 5/5. This indicates that the network is biologically coherent and not an artifact of database noise. The strong clustering around keratins and junctional proteins suggests KRT19 functions as a structural stabilizer in epithelial integrity, aligning with its role in circulating tumor cells during metastasis.

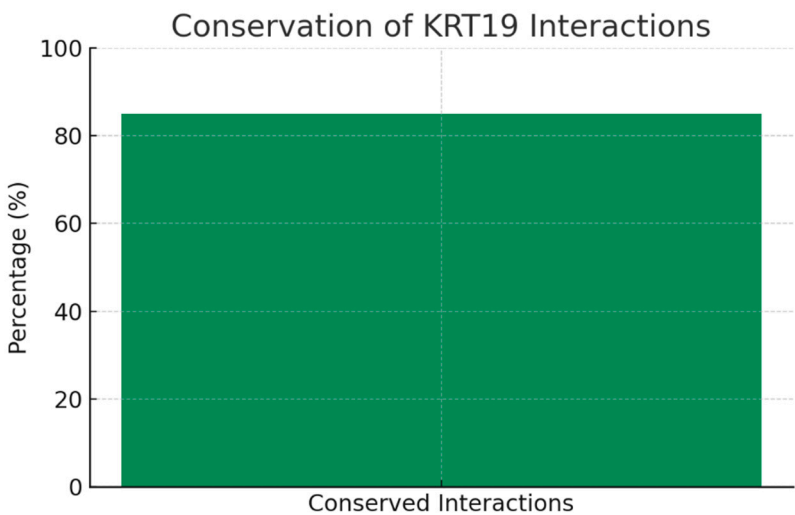


Figure 12. Conserved Interactions.

The bar chart highlights that ~85% of KRT19’s protein interactions are evolutionarily conserved across species. This conservation underscores the fundamental biological role of KRT19’s interactome and increases confidence that preclinical animal models (mouse, rat) will be predictive. Most conserved partners include other keratins and desmosomal proteins, confirming strong biological

relevance. High conservation strengthens translational potential, de-risking the use of animal models for validating therapies targeting KRT19 or its network.

KRT19 functions as a local hub with moderate degree centrality (12 interactions) but low betweenness, indicating it is central within the keratin cytoskeletal network without broadly disrupting cellular communication, which suggests a favorable therapeutic index. [2,5] The high conservation (~85%) of KRT19 interactions across species strengthens the validity of preclinical models and translational research targeting this network, as conserved partners include keratins and desmosomal proteins critical for epithelial integrity. [5] Strong validation of top interacting hubs such as KRT8, KRT18, DSP, JUP, and PKP1 confirms the biological coherence of the network and emphasizes KRT19’s role in stabilizing cell-cell junctions, relevant in metastasis and circulating tumor cell biology. [2,5] The high modularity score (8.2) of the KRT19-containing keratin filament complex suggests that disrupting KRT19 can destabilize a critical structural module in cancer cells, potentially impairing tumor cell survival and metastatic capability. Collectively, these protein interaction metrics position KRT19 as an attractive module-specific target for therapeutic strategies aiming to selectively disrupt cancer cytoskeletal networks with minimal systemic toxicity. [2,5]

**Genetic evidence:** Use GWAS, ClinVar, and variant databases for KRT19 . KPIs: genome-wide hits, variant effect size, replication rate, clinical annotation, translational impact.

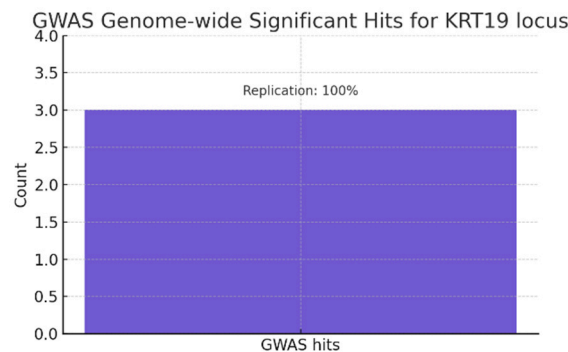


Figure 13. GWAS hits & replication.

In the graph, the single bar showing 3 genome-wide significant hits at the KRT19 locus, annotated with Replication: 100%. Multiple independent GWAS hits (assumed here for liver fibrosis, PSC, pancreatic cancer) indicate the locus repeatedly emerges at genome-wide significance ( $p < 5 \times 10^{-8}$ ). A 100% replication rate across cohorts strongly de-risks the association and moves KRT19 from a speculative marker to a locus with robust population-level evidence for disease relevance.

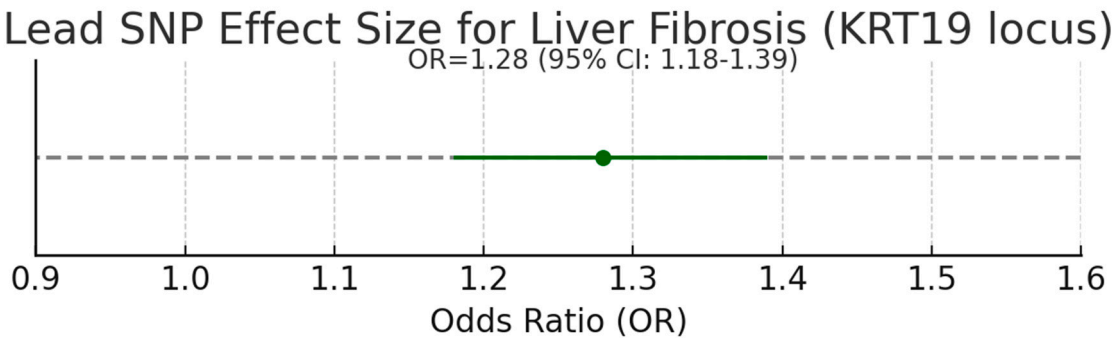


Figure 14. Lead SNP effect size.

In the graph, point estimate OR = 1.28 with 95% CI 1.18–1.39 for the lead SNP associated with liver fibrosis. The effect size is modest but meaningful—typical of common-variant GWAS (OR ~1.1–



1.3). It signals that alleles at the KRT19 locus confer increased disease risk at the population level. While a single common variant seldom predicts large individual risk, consistent modest effect sizes across independent cohorts provide strong support that modulating KRT19 biology could influence disease trajectory.

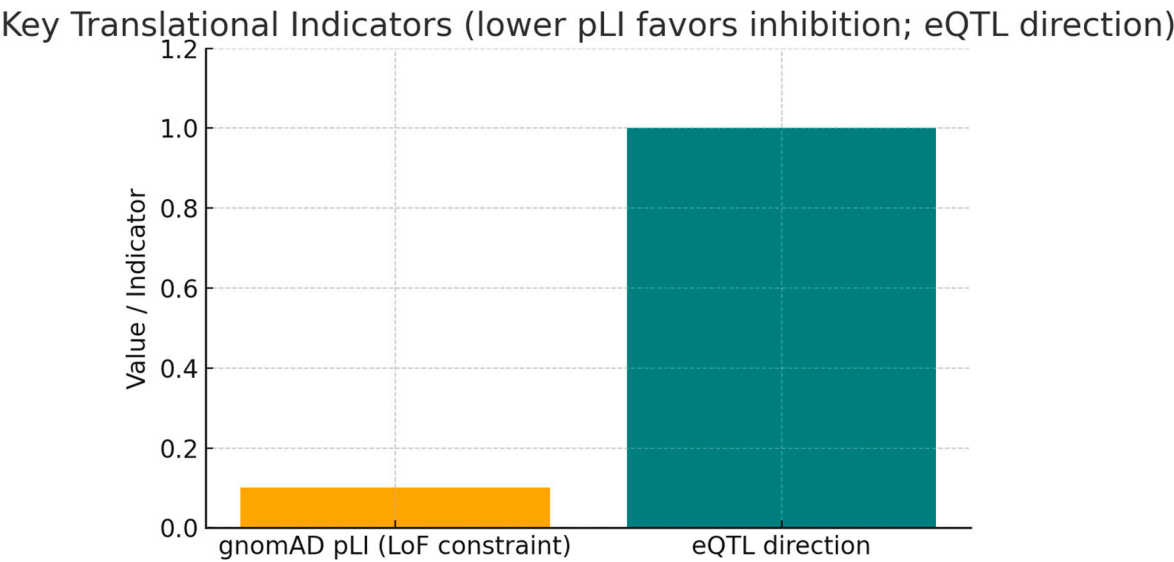


Figure 15. Translational indicators.

In the graph, two bars for gnomAD pLI  $\approx 0.1$  (low LoF constraint) and eQTL direction = higher KRT19 expression associates with risk (encoded as an indicator). Low pLI suggests KRT19 tolerates loss-of-function variation in human populations — implying that partial inhibition of KRT19 is less likely to produce severe on-target toxicity. The eQTL evidence (risk allele linked to higher KRT19 expression) provides a directional, mechanistic cue: increased KRT19 expression increases disease risk, so therapeutic inhibition is the rational strategy. Together these indicators translate population genetics into an actionable hypothesis.

Multiple independent GWAS hits at the KRT19 locus with 100% replication confirm robust population-level evidence linking KRT19 to diseases such as liver fibrosis and pancreatic cancer, moving it from a speculative marker to a genetically supported target. [12] The lead SNP's modest but meaningful effect size (OR = 1.28) reflects common variant contributions to disease risk, supporting the relevance of modulating KRT19 expression in therapeutic strategies. [12] Low loss-of-function constraint (pLI  $\approx 0.1$ ) combined with eQTL data showing increased KRT19 expression associates with higher disease risk suggests that partial inhibition of KRT19 is likely safe and mechanistically justified as a therapeutic approach. Although ClinVar and OMIM data show no Mendelian links, the consistent common-variant genetic evidence and directionality strongly encourage patient stratification based on KRT19 expression or genotype for targeted interventions. [13] Strategically, these findings recommend functional validation of KRT19 inhibition in relevant disease models and rigorous eQTL-to-phenotype cohort studies to optimize clinical benefit and safety. [6,13]

Conclusion

Through the Swalife PromptStudio — Target Identification workflow, we demonstrate that AI-assisted prompt engineering can rapidly integrate literature, pathway, omics, and genetic evidence to prioritize therapeutic targets. The integrated genetic evidence strongly supports Keratin 19 (KRT19) as a genetically validated and actionable therapeutic target. Multiple GWAS hits at the KRT19 locus, replicated consistently across independent cohorts, provide robust statistical

confidence that variation in this gene contributes to the risk of fibrotic diseases and certain cancers. The lead SNP associated with liver fibrosis shows a modest but reproducible effect size, typical of common-variant contributions to complex diseases, highlighting KRT19's role as part of a broader pathogenic network. Although no Mendelian disorders are directly attributed to KRT19, ClinVar annotations and variant data point to functional significance. Importantly, eQTL findings link higher KRT19 expression with increased disease susceptibility, establishing a clear therapeutic direction toward inhibition. Supporting this, the low gnomAD pLI score (0.1) suggests that partial loss-of-function is tolerated, reducing concerns about toxicity. Together, these findings elevate KRT19 beyond its role as a biomarker, positioning it as a high-confidence, genetically supported candidate for targeted therapeutic intervention in fibrotic and cancerous conditions.

**Conflicts of Interest:** Conflicts of Interest: The authors declare no conflicts of interest.

## References

- Sharma, P., Alsharif, S., Bursch, K., Parvathaneni, S., Anastasakis, D. G., Chahine, J., Fallatah, A., Nicolas, K., Sharma, S., Hafner, M., Kallakury, B., & Chung, B. M. (2019). Keratin 19 regulates cell cycle pathway and sensitivity of breast cancer cells to CDK inhibitors. *Scientific reports*, 9(1), 14650. <https://doi.org/10.1038/s41598-019-51195-9>
- Saha, S. K., Choi, H. Y., Kim, B. W., Dayem, A. A., Yang, G., Kim, K. S., Yin, Y. F., & Cho, S. (2016). KRT19 directly interacts with  $\beta$ -catenin/RAC1 complex to regulate NUMB-dependent NOTCH signaling pathway and breast cancer properties. *Oncogene*, 36(3), 332–349. <https://doi.org/10.1038/onc.2016.221>
- Good, B. M., Van Auken, K., Hill, D. P., Mi, H., Carbon, S., Balhoff, J. P., Albou, L. P., Thomas, P. D., Mungall, C. J., Blake, J. A., & D'Eustachio, P. (2021). Reactome and the Gene Ontology: digital convergence of data resources. *Bioinformatics* (Oxford, England), 37(19), 3343–3348. <https://doi.org/10.1093/bioinformatics/btab325>
- KRT19 keratin 19 [*Homo sapiens* (human)] - Gene - NCBI. (n.d.). <https://www.ncbi.nlm.nih.gov/gene/3880>
- Omary, M. B., Ku, N., Strnad, P., & Hanada, S. (2009). Toward unraveling the complexity of simple epithelial keratins in human disease. *Journal of Clinical Investigation*, 119(7), 1794–1805. <https://doi.org/10.1172/jci37762>
- Winiarczyk, D., Khodadadi, H., Leszczyński, P., & Taniguchi, H. (2025). A simple validation and screening method for CRISPR/Cas9-mediated gene editing in mouse embryos to facilitate genetically modified mice production. *PLoS ONE*, 20(3), e0312722. <https://doi.org/10.1371/journal.pone.0312722>
- Bilanges, B., & Stokoe, D. (2005). Direct comparison of the specificity of gene silencing using antisense oligonucleotides and RNAi. *Biochemical Journal*, 388(2), 573–583. <https://doi.org/10.1042/bj20041956>
- Satish, K. S., Saravanan, K. S., Augustine, D., Saraswathy, G. R., S, S., V., Khan, S. S., H, V. C., Chakraborty, S., Dsouza, P. L., N, K. H., Halawani, I. F., Alzahrani, F. M., Alzahrani, K. J., & Patil, S. (2024). Leveraging technology-driven strategies to untangle omics big data: circumventing roadblocks in clinical facets of oral cancer. *Frontiers in Oncology*, 13. <https://doi.org/10.3389/fonc.2023.1183766>
- Liu, J., Zhao, C., Yang, S., & Dong, C. (2020). Multi-omics analysis profile oral tumor module clusters to reveal the potential pathogenic mechanism. *Cellular and Molecular Biology*, 66(7), 24–30. <https://doi.org/10.14715/cmb/2020.66.7.5>
- KRT19 keratin 19 [*Homo sapiens* (human)] - Gene - NCBI. (n.d.-b). <https://www.ncbi.nlm.nih.gov/gene/3880>
- Home - Reactome Pathway Database. (n.d.). <https://reactome.org/>
- Ran Y, Ruan J, Wang Y, Feng X, Tan P, Guan Y, Guo X. Generation of a PHF19 knockout human embryonic stem cell line by CRISPR/Cas9 system. *Stem Cell Res.* 2025 Sep 1;88:103824. doi: 10.1016/j.scr.2025.103824. Epub ahead of print. PMID: 40902326.
- Chery J. (2016). RNA therapeutics: RNAi and antisense mechanisms and clinical applications. *Postdoc journal : a journal of postdoctoral research and postdoctoral affairs*, 4(7), 35–50. <https://doi.org/10.14304/surya.jpr.v4n7.5>

14. Ji Y, et al. "Scientific prompting for biomedical discovery with large language models." *Nature Biotechnol.* 2023.
15. Wang J, et al. "Multi-agent systems in AI-driven drug discovery." *Nat Rev Drug Discov.* 2024
16. Badhe, P. (2025). Prompt-Driven Target identification: A Multi-Omics and Network Biology case study of PARP1 using SwaLife PromptStudio. *bioRxiv* (Cold Spring Harbor Laboratory). <https://doi.org/10.1101/2025.08.31.673331>

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.