TargetSage: Identifying Therapeutic Target Genes with Interpretable and Robust LLM Reasoning

Ziheng Duan* UC Irvine zihend1@uci.edu

Tong Wu* UC Irvine twu0955@gmail.com

Xi Li UC Irvine xil43@uci.edu

Simon Dongbo Sun UC Irvine dongbos@uci.edu

Chaoyang Wang UC Irvine chaoyaw2@uci.edu

Junhao Liu UC Irvine junhaoliu17@gmail.com

Martin Renqiang Min NEC Laboratories America renqiang@gmail.com

Jing Zhang† UC Irvine zhang.jing@uci.edu

Abstract

Therapeutic target identification—determining which genes or proteins to modulate for a desired clinical effect—is a foundational step in drug discovery. However, existing computational methods that rank candidates at genome scale fail to capture mechanistic knowledge locked in the biomedical literature, lack interpretability, and suffer from the false-negative problem by treating unlabeled but potentially druggable genes as definite negatives. We propose TargetSage, an LLM reasoning agent for therapeutic target identification built on positive-unlabeled (PU) learning. Our approach consists of three modules and provides an evidence-grounded, continuously updatable foundation for interpretable target prioritization at genome scale. Specifically, first, AGENTIC PROFILING enriches each gene with LLM-gathered biomedical evidence; second, GUIDED REASONING produces gene-specific rationales and explicit reasoning attributes through a GRPO-trained policy rewarded by downstream Adjusted F1; third, PU-AWARE SCORING produces genome-wide druggability scores with an LLM-informed class-prior. Evaluated on 15 benchmark tasks spanning 19,032 human protein-coding genes, TargetSage consistently outperforms ten baselines, surpassing the strongest by 37% in macro-average Adjusted F1 and generalizing to independent held-out benchmarks. Using only the training data collected in 2021, TargetSage outperforms all ten baselines at ranking genes that gain clinical approval by 2025. On an independent validation set of Phase II/III clinical candidates, it further achieves a 51% higher enrichment odds ratio at the top-25 shortlist than the runner-up baseline.

1 Introduction

Therapeutic target identification—the task of determining which genes or proteins to modulate for a desired clinical effect—is a foundational bottleneck in drug discovery, because it shapes key downstream decisions, including mechanism selection, assay development, lead optimization, biomarker strategy, and patient stratification in clinical trials [1, 2]. However, despite decades of progress, fewer than 1,000 of the roughly 20,000 human protein-coding genes have been validated as therapeutic targets [3, 4], leaving most of the genome unexplored from a translational perspective. At the same time, the evidence relevant to target discovery has expanded rapidly across molecular databases, functional studies, and biomedical literature, making this gap increasingly difficult to close through manual curation alone. Reliable computational prioritization is therefore essential both to broaden the tractable target landscape and to reduce the costly attrition that continues to constrain modern drug development [5, 6].

*Equal contribution.

†Corresponding author.

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

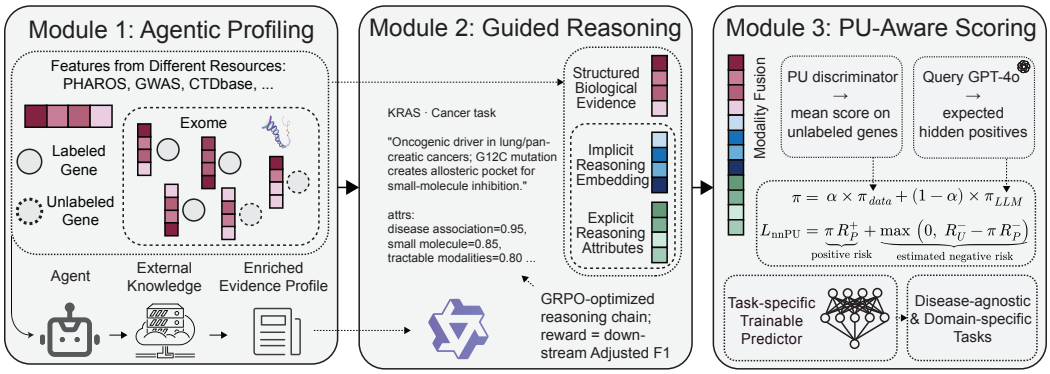

Figure 1: **TargetSage overview.** TargetSage prioritizes each gene through three sequential modules. **(i) Agentic Profiling** semanticizes structured biological features and retrieves gene-specific biomedical evidence to build an enriched gene profile. **(ii) Guided Reasoning** applies a GRPO-trained policy, rewarded by downstream Adjusted F1, to produce gene-specific reasoning chains and explicit reasoning attributes. **(iii) PU-Aware Scoring** combines structured evidence, explicit reasoning attributes, and implicit reasoning embeddings via gated fusion with a PU prior to produce druggability scores.

Machine learning has progressively broadened the evidence base for therapeutic target prioritization, moving from single-modality descriptors with classical classifiers [7, 8], to multi-modal integration of genomic, network, and expression-level signals [9, 10], and more recently to deep models over these integrated features [11–13]. However, despite these advances, the core formulation remains largely unchanged: genes are still encoded as fixed-dimensional structured feature vectors, and prediction is learned from supervised labels defined by current target annotations. This formulation leaves three structural gaps. First, it depends on a frozen feature space that strips away biological context and makes incorporating new evidence costly. Second, it provides only population-level interpretability, offering feature importance across the training set but little rationale for why any specific gene is prioritized. Third, it introduces false negatives into the training signal by treating unlabeled genes as negatives, even though many are positives awaiting validation. As a result, current methods remain difficult to trust, update, and translate into actionable target-discovery decisions.

Here, we introduce TargetSage (**Fig. 1**), a reasoning-augmented framework that addresses these three gaps through three dedicated modules. First, AGENTIC PROFILING replaces frozen feature vectors by enriching each gene with retrieved biomedical evidence, building an evidence profile that captures biological context invisible to static numerical representations. Second, GUIDED REASONING moves beyond population-level interpretability by training a local reasoning policy with Group Relative Policy Optimization (GRPO) and a downstream classification reward, producing gene-specific rationales, explicit reasoning attributes, and implicit reasoning embeddings. Third, PU-AWARE SCORING integrates all three modalities via gene-adaptive multimodal gating and corrects incomplete supervision by treating unlabeled genes as potential positives under a hybrid class prior combining empirical label statistics with LLM-derived knowledge. Together, these three modules transform target prioritization from static supervised scoring into an evidence-grounded, updateable, and interpretable framework for therapeutic discovery.

To summarize, we make four contributions to therapeutic target identification at genome scale:

1. **TargetSage, a reasoning-augmented framework for target identification.** A unified design of AGENTIC PROFILING, GUIDED REASONING, and PU-AWARE SCORING that transforms genome-wide target prioritization into an evidence-grounded, auditable pipeline.

2. **A standardized positive-unlabeled benchmark for target discovery.** Fifteen benchmark tasks across four clinically motivated categories spanning 19,032 human protein-coding genes, evaluated with Adjusted F1 [14], a PU-aware metric robust to unlabeled positives.

3. **Strong and generalizable performance against diverse baselines.** TargetSage achieves a macro-average Adjusted F1 of 11.1 across all 15 tasks, outperforming the strongest of ten baselines by 37% relative and generalizing to five independent held-out druggability benchmarks.

4. **Temporal and independent validation beyond training labels.** Trained solely on 2021 annotations, TargetSage outperforms all ten baselines at prospectively ranking genes that gain clinical approval by 2025. On an independent validation set of Phase II/III clinical candidates, it further achieves a 51% higher enrichment odds ratio at the top-25 shortlist than the runner-up baseline.

2 Related Work

Computational target prioritization. Computational target prioritization has evolved from single-modality descriptors with classical classifiers [7, 8] to multi-modal integration of genomic, network, and expression-level evidence [9, 10]. More recently, deep supervised frameworks such as TargetDB [11], DrugnomeAI [12], and Hybrid-DNN [13] have raised the ceiling on this paradigm. These methods have improved the scale and consistency of target scoring, but they still largely follow the same paradigm: genes are encoded as fixed structured feature vectors and trained against current target annotations. More recently, sequence- and expression-level foundation models such as ESM-2 [15] and Geneformer [16] provide pre-trained gene and protein representations, yet they still require task-specific fine-tuning with binary labels and do not address the PU supervision or interpretability gaps. TargetSage is the first framework to unify all three corrections in a single genome-scale pipeline.

Positive-unlabeled learning. Positive-unlabeled (PU) learning treats unlabeled examples as a mixture of hidden positives and true negatives, making it well suited to incomplete-label settings [17, 18]. Its key challenge is class-prior estimation, i.e., estimating how many positives are hidden in the unlabeled set. Existing estimators such as KM2 [19] and AlphaMax [20] rely mainly on score-distribution statistics, which can underestimate the positive fraction when true positives are rare or heterogeneous. TargetSage instead uses a hybrid prior that combines empirical label statistics with LLM-derived biomedical knowledge, grounding the estimate in both data and domain context.

LLM reasoning over biomedical evidence. LLMs have recently been used for biomedical evidence retrieval, annotation, and reasoning [21–24]. However, most systems are not designed for calibrated target prioritization: they generate labels, summaries, or explanations, but do not integrate reasoning into a PU-aware predictive model. Reinforcement-learning methods such as GRPO [25] show that reward optimization can elicit step-by-step reasoning, but mainly in domains with verifiable rewards such as mathematics and code. TargetSage adapts reward-guided reasoning to target discovery, using a PU-aware reward to produce gene-specific rationales for calibrated prioritization.

3 Methods

As shown in **Fig. 1**, TargetSage addresses the three gaps above through three modules: **Agentic Profiling** enriches each gene with LLM-gathered biomedical evidence beyond frozen feature spaces; **Guided Reasoning** produces auditable gene-level rationales through a GRPO-trained policy; and **PU-Aware Scoring** corrects incomplete supervision via gated multi-modal fusion and LLM-informed PU learning. Each module is described in the following subsections.

3.1 Problem Formulation

Formally, let $\mathcal{G}$ denote the $n = 19{,}032$ human protein-coding genes and $t \in \{1, \ldots, 15\}$ the target-identification tasks. Each task provides observed positives $P_t \subset \mathcal{G}$, with all other genes treated as unlabeled, $U_t = \mathcal{G} \setminus P_t$. Because U_t may contain undiscovered targets, TargetSage learns a PU scoring function $f_t : \mathcal{G} \to [0, 1]$ that prioritizes likely positives without assuming unlabeled genes are negatives. See Appendix A for task details.

3.2 Module 1: Agentic Profiling

To compute the per-task scoring function f_t, conventional approaches encode each gene as a fixed numerical feature vector, a representation that strips biological meaning and makes absorbing new evidence costly. We instead build an *enriched evidence profile* that preserves biological context and supports task-specific augmentation. For each gene $g_i \in \mathcal{G}$, we start from structured features $\mathbf{x}_i^{\text{bio}} \in$

$\mathbb{R}^{482}$, aggregated from 16 public databases as compiled by the DrugnomeAI feature pipeline [12, 26] (**Table 8**; Appendix C). The module proceeds in three steps: converting informative numerical features into value-paired evidence statements, using an LLM agent to retrieve missing task-relevant evidence from external sources, and filtering label leakage before assembling the unified enriched profile. This profile decouples reasoning from source-specific schemas, allowing new data sources without architectural changes (Appendix C).

3.3 Module 2: Guided Reasoning

Given the enriched evidence profile from Module 1 and the current task context, Guided Reasoning generates task-specific representations through a **reasoning policy** (trained via GRPO; see below). For each gene-task pair, the policy produces (i) a step-by-step rationale that provides an auditable gene-level explanation; (ii) an explicit attribute vector $\mathbf{x}_i^{\text{attr}} \in \mathbb{R}^{d_a}$ scoring d_a=23 therapeutically relevant attributes (vocabulary derived by running the agent across all 19,032 genes and retaining attributes with $\geq 5\%$ coverage after deduplication; Appendix D); and (iii) an implicit reasoning embedding $\mathbf{z}_i^{\text{emb}} \in \mathbb{R}^{256}$ obtained by encoding the rationale with `text-embedding-3-large` [27] and PCA-compressing to 256 dimensions (Appendix B; training details in Appendix F). The explicit attributes $\mathbf{x}_i^{\text{attr}}$ and implicit reasoning embedding $\mathbf{z}_i^{\text{emb}}$, together with $\mathbf{x}_i^{\text{bio}}$ from Module 1, form the three modalities passed to PU-Aware Scoring; the rationale provides gene-level interpretability.

GRPO optimization. We train the Guided Reasoning policy π_θ with **GRPO** [25], instantiated as Qwen2.5-1.5B-Instruct [28] with LoRA adapters [29] (implementation details in Appendix F). For each gene g_i, the policy samples G rollouts $y_1, \ldots, y_G$, where each rollout contains a reasoning chain τ_k and an explicit attribute vector $\mathbf{a}_k \in [0,1]^{d_a}$. Rollout k's attribute vector $\mathbf{a}_k$ replaces gene i's entry in the attribute matrix; a set of frozen proxy Module 3 discriminators (pre-trained once on pre-GRPO features before any RL update) evaluates the modified matrix via a forward pass to produce $\hat{\mathbf{p}}^{(t,k)} \in [0,1]^n$ per task, and the reward is macro-avg Adjusted F1:

$$r(y_k) = \frac{1}{T}\sum_{t=1}^{T} \mathrm{F1}_{\mathrm{adj}}\big(\hat{\mathbf{p}}^{(t,k)}\big). \tag{1}$$

GRPO converts these rewards into centered advantages $\hat{A}_k$ and updates the policy with the advantage-weighted policy-gradient objective (full derivation in Appendix F).

3.4 Module 3: PU-Aware Scoring

PU-Aware Scoring converts outputs from Agentic Profiling and Guided Reasoning into the final PU-based target score through three steps: gene-adaptive multi-modal fusion, hybrid class-prior estimation, and non-negative PU training.

Gated multi-modal fusion. For each gene g_i, we fuse three representations: structured biological evidence $\mathbf{x}_i^{\text{bio}}$, explicit reasoning attributes $\mathbf{x}_i^{\text{attr}}$, and the implicit reasoning embedding $\mathbf{z}_i^{\text{emb}}$. Each modality is projected into a shared latent space and then combined through learned modality gates:

$$\mathbf{h}_i^m = \sigma(\mathbf{W}_m \mathbf{u}_i^m + \mathbf{b}_m), \quad \boldsymbol{\alpha}_i = \mathrm{softmax}\big(\mathbf{W}_g[\mathbf{h}_i^{\text{bio}}; \mathbf{h}_i^{\text{attr}}; \mathbf{h}_i^{\text{emb}}]\big), \quad \mathbf{h}_i^{\text{fuse}} = \sum_{m\in\mathcal{M}} \alpha_i^m \mathbf{h}_i^m, \tag{2}$$

where $\mathcal{M} = \{\text{bio}, \text{attr}, \text{emb}\}$ and $(\mathbf{u}_i^{\text{bio}}, \mathbf{u}_i^{\text{attr}}, \mathbf{u}_i^{\text{emb}}) = (\mathbf{x}_i^{\text{bio}}, \mathbf{x}_i^{\text{attr}}, \mathbf{z}_i^{\text{emb}})$. Input-dependent gates reweight each modality per gene—structured biological evidence, explicit reasoning attributes, and implicit reasoning embeddings—dynamically suppressing noise in high-dimensional semantic vectors while preserving sparse biological signals. The fused representation $\mathbf{h}_i^{\text{fuse}}$ is passed to a task-specific discriminator $D_t(\mathbf{h}_i^{\text{fuse}}) \in [0,1]$, which defines the scoring function $f_t(g_i)$.

Hybrid class-prior estimation. PU learning requires a task-specific class prior $\pi_t = \Pr_t(y{=}1)$, the expected fraction of true positives among all genes for task t; since observed positives P_t are only a lower bound on the true positive set, TargetSage estimates π_t by combining discriminator evidence on unlabeled genes with LLM-derived domain knowledge:

$$\pi_t = \alpha \pi_{t,\text{data}} + (1-\alpha)\pi_{t,\text{LLM}}. \tag{3}$$

Here, $\pi_{t,\text{data}} = \frac{1}{|U_t|}\sum_{g_j \in U_t} D_t(\mathbf{h}_j^{\text{fuse}})$ is the mean discriminator score over the unlabeled set (following [17]), and $\pi_{t,\text{LLM}} = m^*_{t,\text{med}}\,|P_t|/n$ anchors an LLM-derived multiplier to the observed positive rate. The multiplier $m^*_{t,\text{med}}$ is the median of multiple offline LLM queries with task metadata (Appendix E). The mixing coefficient $\alpha \in [0,1]$ is fixed at α=0.6 in all experiments; sensitivity analysis in Appendix M confirms performance is stable across $\alpha \in [0.2, 0.8]$.

Non-negative PU loss. Given π_t, we train the discriminator with nnPU learning rather than treating genes in U_t as true negatives. For each $g_j \in U_t$, we define a soft confidence weight using discriminator confidence and semantic similarity to labeled positives:

$$w_{j,t} = \beta D_t(\mathbf{h}_j^{\text{fuse}}) + (1-\beta)\cos(\mathbf{z}_j^{\text{emb}}, \bar{\mathbf{z}}_{P_t}^{\text{emb}}), \tag{4}$$

where $\mathbf{z}_j^{\text{emb}}$ and $\bar{\mathbf{z}}_{P_t}^{\text{emb}}$ are the PCA-compressed embeddings ($d_p = 256$) of gene g_j and the positive-set centroid for task t, and $\beta \in [0,1]$ is a fixed semantic guidance weight set to β=0.6 (Appendix M). We optimize the nnPU objective [30]:

$$\mathcal{L}^{(t)}_{\text{nnPU}} = \pi_t R^+_{P_t} + \max\big(0,\ R^-_{U_t}(w_{j,t}) - \pi_t R^-_{P_t}\big)\,, \tag{5}$$

where $R^+_{P_t}$, $R^-_{P_t}$ are positive-set risks under positive/negative loss, and $R^-_{U_t}(w_{j,t})$ is the weighted negative risk over U_t. Backpropagation through $\mathcal{L}^{(t)}_{\text{nnPU}}$ trains the discriminator, projection heads, and gates, yielding scores for target ranking and the GRPO reward (§3.3).

3.5 Evaluation Metric

We evaluate with Adjusted F1 [14], a PU-aware metric that avoids treating unlabeled genes as true negatives:

$$F^{(t)}_{\text{1-adj}} = \frac{\big(R^{(t)}_{\text{soft}}\big)^2}{\bar{p}^{(t)}}, \qquad R^{(t)}_{\text{soft}} = \frac{1}{|P_t|}\sum_{i \in P_t}\hat{p}^{(t)}_i, \qquad \bar{p}^{(t)} = \frac{1}{n}\sum_{i=1}^{n}\hat{p}^{(t)}_i. \tag{6}$$

Here $R^{(t)}_{\text{soft}}$ is the mean predicted score over labeled positives and $\bar{p}^{(t)}$ is the genome-wide average; the formula rewards concentration of probability mass on true positives without penalizing high scores on unlabeled genes that may be undiscovered targets, and also serves as the GRPO reward (§3.3). Adjusted F1 is the soft form of $r^2/\Pr(\hat{y}{=}1)$, the standard criterion for PU data [31, 14], and is computed only on the held-out test fold, whose positives are excluded from training, so a high score requires ranking unseen positives highly.

4 Experiments and Results

We benchmark TargetSage against ten baselines across 15 tasks and five independent reference sets, and probe prospective validity against clinical targets not seen during training (**Tables 1–4**); ablations (**Fig. 2**) confirm that each architectural component contributes independently, and case studies show that Guided Reasoning surfaces evidence-grounded distinctions—non-enzymatic signaling, antibody tractability—invisible to conventional features. Subsections 4.1–4.5 present experimental setup, main results, prospective validation, ablations, and case studies in order.

4.1 Experimental Setup

Benchmark tasks. We evaluate TargetSage on 15 target-identification tasks covering 19,032 human protein-coding genes, constructed from public PHAROS/TCRD [32] and Open Targets [33] snapshots (full definitions in Appendix A). The tasks span two axes—*scope* (disease-agnostic vs. domain-specific) and *evidence type* (clinically validated vs. tractability-predicted)—and form four categories: PHAROS disease-agnostic targets (2 tasks), druggability triage (2), cancer druggability (6), and drug modality (5). This design reduces dependence on any single annotation scheme and covers positive rates from 0.4% to 13.5%, spanning rare and common class-prior regimes.

Table 1: **Adjusted F1 on 15 benchmark tasks.** Adjusted F1 $= R^2_{\text{soft}}/\bar{p}$ (PU-aware; §3.5). Mean $\pm$ std over 5 seeds with stratified 80/20 splits. **Bold** = best per row.

Task	Classical ML Baselines							Deep Tabular Baselines			TargetSage
	LR	SVM	KNN	NB	RF	GB	MLP	ResNet	TabNet	FT-Trans	
PHAROS (Disease-agnostic)											
Clinical Targets	$7.6_{\pm 0.7}$	$6.1_{\pm 0.9}$	$5.6_{\pm 0.9}$	$1.3_{\pm 0.0}$	$4.0_{\pm 0.3}$	$10.2_{\pm 0.5}$	$9.7_{\pm 1.2}$	$9.4_{\pm 0.7}$	$7.3_{\pm 2.6}$	$9.6_{\pm 0.6}$	$\mathbf{11.2_{\pm 1.0}}$
Clinical & Chemical	$2.5_{\pm 0.1}$	$2.1_{\pm 0.1}$	$1.6_{\pm 0.1}$	$2.0_{\pm 0.1}$	$1.9_{\pm 0.0}$	$2.5_{\pm 0.0}$	$2.9_{\pm 0.3}$	$3.1_{\pm 0.2}$	$2.4_{\pm 0.3}$	$2.5_{\pm 0.2}$	$\mathbf{3.2_{\pm 0.3}}$
Druggability Triage (Disease-agnostic)											
Top-Tier Targets	$4.0_{\pm 0.2}$	$3.3_{\pm 0.2}$	$2.9_{\pm 0.2}$	$2.9_{\pm 0.3}$	$2.6_{\pm 0.1}$	$4.2_{\pm 0.1}$	$4.2_{\pm 0.4}$	$5.0_{\pm 0.4}$	$3.9_{\pm 0.7}$	$4.3_{\pm 0.8}$	$\mathbf{6.5_{\pm 0.2}}$
High-Confidence	$3.2_{\pm 0.1}$	$2.8_{\pm 0.1}$	$2.1_{\pm 0.1}$	$2.6_{\pm 0.2}$	$2.3_{\pm 0.0}$	$3.3_{\pm 0.1}$	$3.5_{\pm 0.1}$	$3.9_{\pm 0.1}$	$3.3_{\pm 0.1}$	$3.3_{\pm 0.2}$	$\mathbf{4.2_{\pm 0.2}}$
Cancer Druggability (Domain-specific)											
Cancer-Relevant	$4.9_{\pm 1.3}$	$2.7_{\pm 0.4}$	$2.3_{\pm 0.3}$	$1.0_{\pm 0.0}$	$2.5_{\pm 0.4}$	$6.3_{\pm 0.7}$	$5.3_{\pm 1.9}$	$7.5_{\pm 1.9}$	$4.6_{\pm 2.1}$	$4.7_{\pm 0.7}$	$\mathbf{14.0_{\pm 2.5}}$
Type-Specific	$1.4_{\pm 0.3}$	$0.3_{\pm 0.2}$	$0.7_{\pm 0.2}$	$1.0_{\pm 0.1}$	$0.7_{\pm 0.2}$	$1.9_{\pm 0.2}$	$1.3_{\pm 0.5}$	$2.8_{\pm 1.2}$	$1.9_{\pm 0.4}$	$2.6_{\pm 0.4}$	$\mathbf{5.4_{\pm 0.5}}$
Pan-Cancer	$1.8_{\pm 0.7}$	$0.2_{\pm 0.1}$	$0.7_{\pm 0.5}$	$0.9_{\pm 0.2}$	$0.7_{\pm 0.2}$	$2.2_{\pm 0.6}$	$0.8_{\pm 0.5}$	$3.0_{\pm 1.3}$	$2.1_{\pm 0.3}$	$2.6_{\pm 0.7}$	$\mathbf{5.5_{\pm 1.1}}$
T1 Cancer	$12.0_{\pm 1.7}$	$4.1_{\pm 2.9}$	$5.1_{\pm 2.3}$	$1.1_{\pm 0.2}$	$4.1_{\pm 0.9}$	$13.9_{\pm 2.4}$	$7.0_{\pm 1.8}$	$16.9_{\pm 5.1}$	$7.0_{\pm 2.6}$	$11.0_{\pm 2.4}$	$\mathbf{24.2_{\pm 7.4}}$
T1–T2 Cancer	$4.5_{\pm 0.4}$	$3.2_{\pm 0.4}$	$2.6_{\pm 0.3}$	$1.2_{\pm 0.1}$	$2.2_{\pm 0.1}$	$5.5_{\pm 0.3}$	$4.7_{\pm 1.3}$	$5.7_{\pm 0.5}$	$4.7_{\pm 1.2}$	$4.3_{\pm 0.7}$	$\mathbf{7.8_{\pm 0.8}}$
T1–T3 Cancer	$3.4_{\pm 0.2}$	$2.5_{\pm 0.1}$	$1.8_{\pm 0.2}$	$2.1_{\pm 0.3}$	$2.3_{\pm 0.1}$	$4.9_{\pm 0.3}$	$4.0_{\pm 0.8}$	$4.0_{\pm 0.1}$	$3.7_{\pm 0.9}$	$4.6_{\pm 0.8}$	$\mathbf{5.4_{\pm 0.3}}$
Drug Modality (Domain-specific)											
SM (Appr.)	$8.8_{\pm 0.5}$	$7.9_{\pm 0.7}$	$7.9_{\pm 0.9}$	$1.3_{\pm 0.0}$	$5.3_{\pm 0.3}$	$10.6_{\pm 0.2}$	$9.6_{\pm 1.6}$	$11.4_{\pm 1.1}$	$8.6_{\pm 3.6}$	$9.4_{\pm 1.7}$	$\mathbf{13.2_{\pm 1.1}}$
SM (Clin+)	$6.0_{\pm 0.3}$	$5.2_{\pm 0.4}$	$4.8_{\pm 0.4}$	$3.2_{\pm 0.3}$	$4.1_{\pm 0.2}$	$7.7_{\pm 0.1}$	$7.5_{\pm 0.8}$	$7.6_{\pm 0.4}$	$7.3_{\pm 0.4}$	$7.6_{\pm 0.7}$	$\mathbf{9.1_{\pm 0.3}}$
Ab (Appr.)	$11.3_{\pm 6.1}$	$1.4_{\pm 1.3}$	$6.2_{\pm 3.4}$	$1.2_{\pm 0.4}$	$3.7_{\pm 1.4}$	$13.8_{\pm 8.3}$	$12.5_{\pm 9.9}$	$16.2_{\pm 9.1}$	$8.1_{\pm 5.1}$	$13.8_{\pm 7.0}$	$\mathbf{22.6_{\pm 13.4}}$
Ab (Clin+)	$9.0_{\pm 2.3}$	$4.1_{\pm 1.8}$	$6.9_{\pm 0.6}$	$1.6_{\pm 0.1}$	$5.3_{\pm 0.6}$	$12.9_{\pm 1.5}$	$8.5_{\pm 2.0}$	$11.0_{\pm 1.1}$	$5.3_{\pm 1.6}$	$7.9_{\pm 1.4}$	$\mathbf{17.5_{\pm 1.2}}$
PROTAC	$8.6_{\pm 0.7}$	$3.8_{\pm 0.8}$	$3.6_{\pm 0.5}$	$1.2_{\pm 0.1}$	$4.3_{\pm 0.5}$	$13.5_{\pm 1.5}$	$8.0_{\pm 2.7}$	$13.6_{\pm 1.8}$	$6.6_{\pm 4.1}$	$12.8_{\pm 3.0}$	$\mathbf{17.2_{\pm 3.3}}$
Macro-Avg	6.0	3.3	3.7	1.6	3.0	7.6	6.0	8.1	5.1	6.7	**11.1**

Baselines, Metrics and Protocol. We compare TargetSage with ten baselines: seven classical models commonly used in target-identification studies [12, 11, 13]—logistic regression (LR), support vector machine (SVM), k-nearest neighbors (KNN), naïve Bayes (NB), random forest (RF), gradient boosting (GB), and multilayer perceptron (MLP)—and three modern deep tabular models: ResNet [34], TabNet [35], and FT-Transformer [36]. All methods are evaluated with Adjusted F1 (Sec. 3.5) [14], a PU-aware metric that avoids treating high-scoring unlabeled genes as false positives. Ablations in Sec. 4.4 further evaluate architectural variants and simpler baselines—across feature modalities, fusion operators, learning objectives, and attribute extraction strategies—to isolate each component's contribution. We report mean $\pm$ std over 5 seeds using stratified 80/20 gene-level train/test splits shared by all methods; hyperparameters are tuned on a 10% held-out set from the training fold. Pairwise comparisons use one-sided Wilcoxon signed-rank tests over the 15 tasks with Bonferroni correction across the ten head-to-head comparisons. Full training details and per-task ablations are provided in Appendices B and H.

4.2 Main benchmark results: TargetSage outperforms all baselines on 15 tasks and generalizes to independent drug-target benchmarks

We evaluate TargetSage against ten baselines on the 15 benchmark tasks (**Table 1**) and further test generalization to five independent druggability reference sets. TargetSage ranks first on all 15 tasks and in all four categories, achieving a macro-average Adjusted F1 of 11.1—3.0 points (37% relative) above the strongest baseline overall. The macro-average gap over the strongest classical baseline (GB: 7.6) and the strongest deep tabular model (ResNet: 8.1) is significant at $p<0.001$ across all ten pairwise comparisons (one-sided Wilcoxon signed-rank, Bonferroni correction); per-task margins range from 0.1 to 7.3 (median 2.1).

The largest gains appear in settings where literature-grounded evidence synthesis is especially important. For cancer druggability, TargetSage nearly doubles the strongest baseline across all three subtasks—cancer-relevant genes, type-specific cancer targets, and pan-cancer genes (14.0 vs. 7.5, 5.4 vs. 2.8, and 5.5 vs. 3.0, respectively). Drug-modality tasks follow the same pattern: TargetSage leads on antibody-approved, antibody clinical-stage, and PROTAC-degradable targets (22.6 vs. 16.2, 17.5 vs. 12.9, and 17.2 vs. 13.6, respectively). These improvements are consistent with the reasoning module's role: by retrieving disease-context evidence for cancer tasks and modality-specific tractability signals for drug-modality tasks, it surfaces information that fixed numerical features can only encode indirectly.

Table 2: **Robustness of the comparison** (15 tasks, 5 seeds). *Top*: macro Adjusted F1 when every method receives the structured features only, or additionally the explicit reasoning attributes and implicit embedding. *Bottom*: ranking metrics that are not the GRPO reward.

	LR	SVM	KNN	NB	RF	GB	MLP	ResNet	TabNet	FT-Tr.	**TargetSage**
Structured only	6.0	3.3	3.7	1.6	3.0	7.6	6.0	8.1	5.1	6.7	**10.7**
+ attr. & emb.	7.6	5.0	4.4	1.7	2.6	7.9	7.1	8.9	6.4	7.8	**11.1**
AUROC	0.869	0.915	0.805	0.688	**0.946**	0.939	0.877	0.877	0.896	0.944	**0.946**
AUPRC	0.446	0.470	0.348	0.116	0.564	0.504	0.488	0.428	0.423	0.499	**0.590**
P@25	0.640	0.659	0.620	0.131	0.732	0.624	0.701	0.619	0.585	0.647	**0.748**

Table 3: **Cross-dataset validation: TargetSage generalizes to independent benchmarks.** Top-5% overlap (951 genes) across five external reference sets. **Bold**: best. *Appr.*: approved targets [38]; *MAB*/*SM*: Open Targets tractability. B1 = approved; B1–3 = +clinical.

Method	Appr.	MAB-B1	MAB-B1-3	SM-B1	SM-B1-3
TargetSage (*Tclin*)	**378**	**65**	129	**557**	**636**
TargetSage (Tier1)	323	45	**140**	503	628
DrugnomeAI (*Tclin*)	331	53	122	388	522
DrugnomeAI (Tier1)	304	42	130	349	480
TargetDB	178	27	91	171	245
Hybrid-DNN	205	25	89	219	254
MetaDT	28	0	10	46	83

Table 4: **Prospective validation: TargetSage ranks future targets early.** *Left*: rank percentile of *Tclin* (n=90) or *Tclin*/*Tchem* (n=456) upgrades, 2021–2025 (lower = better). *Right*: OR at top-25/50 *Tbio* genes vs. ChEMBL Phase 2/3 targets [39] (higher = better). **Bold** = best.

Method	*Tclin*	*Tclin*/*Tchem*	OR@25	OR@50
TargetSage	**7.1**	**22.1**	**43.2**	**26.4**
FT-Trans	7.4	22.4	16.9	20.2
GB	7.7	22.8	12.0	7.7
RF	8.1	23.6	28.5	23.2
LR	10.5	24.0	22.4	14.8
SVM	11.2	24.5	0.0	5.6
MLP	21.8	30.5	0.0	1.8
TabNet	28.1	37.1	0.0	3.6
ResNet	36.6	29.8	12.0	7.7
KNN	41.3	29.7	0.0	3.6
NB	50.5	25.3	0.0	3.6

Robustness of the comparison. The lead holds when the baselines receive the same information and under other metrics (**Table 2**). Giving all ten baselines the identical enriched profile—structured features plus the explicit reasoning attributes and implicit embedding—raises almost every baseline, so the retrieved evidence carries real signal, yet TargetSage stays ahead (11.1 vs. 8.9); a direct GPT-4o-mini [37] ranker reaches only 1.2, and biomedical text encoders over the same profiles without generated reasoning reach 10.8 (Appendix J). TargetSage is also best or tied-best on AUROC, AUPRC and precision@25, none of which is the GRPO reward, and against each baseline individually it wins 15/15 tasks (sign test p=3.05×10^{-5}; Appendix L). On the smallest tasks ($\sim$14 test positives) seed variance is intrinsic: TargetSage's coefficient of variation ($\approx$0.6) matches the baselines' (0.56–0.79).

We also compare against four purpose-built druggability tools—DrugnomeAI [12], TargetDB [11], Hybrid-DNN [13], and MetaDT [8]—which publish fixed genome-wide rankings and are evaluated by top-5% overlap with five independent reference sets (**Table 3**). TargetSage (*Tclin*) achieves the highest overlap in four out of five reference sets, recovering 378 approved drug targets and 557 small-molecule-approved genes—exceeding the next-best method (DrugnomeAI-*Tclin*) by 14% and 44%, respectively. TargetSage (Tier1) leads on antibody clinical-stage targets (MAB-B1–3: 140 vs. 130 for DrugnomeAI-Tier1), consistent with Tier1's broader training label, which also includes the targets of biotherapeutics and clinical-phase drug candidates. Across the five reference sets, TargetSage's top 5% is enriched 9.3–18.9$\times$ over the genome base rate (hypergeometric $p < 10^{-80}$). Together, these results confirm that TargetSage's rankings generalize to independently curated druggability criteria.

4.3 Prospective Target Discovery and Independent Validation: TargetSage Recovers Future Targets and Shows External Clinical Enrichment

We assess TargetSage's prospective validity through temporal holdout and independent enrichment experiments. For the temporal experiment, models are trained exclusively on PHAROS 2021 annotations, with every input pinned to its $\leq$2021 release, and used to rank all 19,032 genes genome-wide, defining two prospective evaluation sets: Upgrade to *Tclin* (n=90) and Upgrade to *Tclin*/*Tchem* (n=456). Between the two releases, *Tclin* grew by 14.8% (610$\rightarrow$700 genes), *Tchem* by 18.4% (1,572$\rightarrow$1,862), and *Tdark* decreased by 20.7% (5,774$\rightarrow$4,579). TargetSage achieves the lowest

Table 5: **Per-source $\leq$ 2021 constraint in the temporal experiment.** The ten baselines read only the February-2021 structured features; TargetSage's text sources are each constrained to $\leq$ 2021 and embedded with BioBERT, and the generated attributes are dropped.

Source	$\leq$ 2021 constraint	Retained coverage
NCBI Gene summary	post-2021 summaries blanked	12,987 summaries (6,045 blanked)
PubMed	published $\leq$ 2021	19.8% of genes with a $\leq$ 2021 drug-target abstract
UniProt	release 2021_01	18,801 genes with an entry
Open Targets	release 21.02	11,521 genes (tractability)

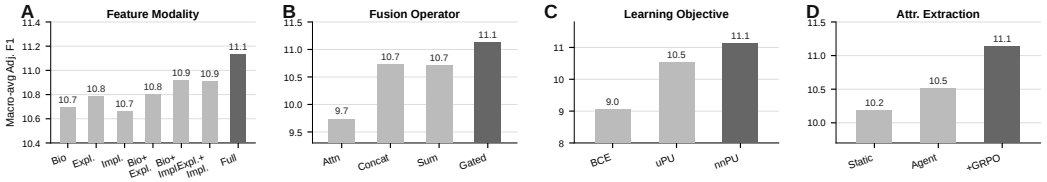

Figure 2: **Component ablations.** Macro-avg Adj. F1 over 15 tasks; each panel varies one design choice. Full per-task tables in Appendix H.

median rank percentile on both tasks (**Table 4**): **7.1%** for Upgrade to *Tclin* and **22.1%** for Upgrade to *Tclin*/*Tchem*, both best among all baselines (next-best: FT-Trans at 7.4% and 22.4%, respectively). Because models never observe 2025 labels, this confirms prospective predictive power rather than label interpolation; see Table 5 and Appendix G. Hidden positives also score highly: the 90 upgraded genes receive $25\times$ the background mean probability, and with all drug and clinical evidence blinded, established targets still rank in the top 3.3% (Appendix K).

For independent validation, all models are trained on 2025 PHAROS *Tclin* labels and evaluated against ChEMBL Phase 2/3 targets [39]—an external signal neither used in training nor derived from PHAROS annotations. We focus on the 11,891 understudied *Tbio* proteins, of which 136 are Phase 2/3 targets at a background rate of 1.1%; enrichment is measured by odds ratio via Fisher's exact test at the top 25 and top 50 shortlists. TargetSage achieves OR=43.2 at the top 25, outperforming all baselines and 51% above the runner-up RF at OR=28.5 (p=2.2×10^{-10}); at the top 50 it achieves OR=26.4 (p=7.4×10^{-12}). TargetSage also leads every method on top-25 and top-50 precision (0.32 and 0.22; Appendix I), whereas directly prompting GPT-4o-mini to rank genes reaches only OR@25 $\approx$ 3.6, so the enrichment is not LLM label recall (Appendix K). Together, the two experiments confirm that TargetSage's rankings generalize beyond its training labels.

4.4 Component Ablations: Each Module Contributes Distinct, Complementary Gains

To evaluate each design choice from Section 3, we ablate components one at a time (results in **Fig. 2**).

Feature modality and fusion. All three modalities are necessary: removing any single one degrades the macro-average (Bio alone: 10.7; Expl./Impl. alone: 10.7–10.8), and no two-modality pairing reaches the full combination, which wins on 9/15 tasks (11.1). Among fusion operators, gated fusion (11.1) outperforms concatenation (10.7), summation (10.7), and attention (9.7), with input-dependent gates effectively suppressing noisy dimensions on low-prevalence tasks.

Learning objective and attribute extraction. The choice of PU objective matters: class-balanced BCE (9.0) already surpasses all baselines (best: ResNet 8.1), confirming that the multi-modal architecture contributes independently of the PU objective; nnPU then adds 2.1 points over BCE and surpasses uPU (10.5) by clipping negative risk estimates on rare-positive tasks. GRPO-guided attribute extraction adds a further 0.6 points over agent-discovered vocabularies without GRPO ($+1.4$ points on T1 Cancer), confirming that reward-shaped reasoning yields task-discriminative features beyond fixed schemas. Per-task breakdowns for all ablation conditions are in Appendix H.

4.5 Interpretability: TargetSage Generates Gene-Specific Rationales for Expert Review

We present two concrete cases where structured features are insufficient to distinguish proteins that require fundamentally different drug strategies, and show that the reasoning module fills this gap by synthesizing evidence from biomedical literature into auditable, gene-specific rationales; in both cases, the model's evidence-grounded conclusions are consistent with established or investigational therapeutic strategies (eight additional examples in Appendix N).

Case 1: Feature-space collision (ERBB2 vs. ERBB3). In the 482-dimensional feature vector, ERBB2 and ERBB3 carry the same kinase annotations: both have the InterPro protein-kinase-domain flag (`IPR_d_Protein_kinase_domain` $= 1$) and the tyrosine-kinase catalytic-domain flag set, so the structured features describe ERBB3 as a kinase. TargetSage identifies that ERBB3 binds neuregulins but cannot transmit signals because its kinase domain is *catalytically inactive*—absent from any structured feature flag, recoverable only from biomedical literature:

> *"ERBB3 is unique in that it lacks a functional kinase domain, which means it cannot directly initiate signaling through phosphorylation. Instead, it acts primarily as a co-receptor that forms heterodimers with other EGFR family members that possess kinase activity, such as ERBB2."* — TargetSage reasoning trace

Small-molecule kinase inhibitors require an active enzyme to block, which ERBB3 lacks, so antibody-based therapies are pursued instead. This matches therapeutic practice: ERBB2 is targeted by approved small-molecule kinase inhibitors such as lapatinib [40] and neratinib [41], whereas HER3-directed therapeutic development includes antibody–drug conjugates such as patritumab deruxtecan [42].

Case 2: Missing context (CD19). CD19 is expressed almost exclusively on B-cells, including the malignant cells responsible for certain leukemias; structured features record its antibody counts and expression specificity but do not explain *why* this protein is druggable. TargetSage identifies that CD19 has no intrinsic enzymatic activity and functions as a *lineage marker*—a surface tag identifying a specific cell population:

> *"CD19 is a B-lymphocyte co-receptor that modulates B-cell receptor signaling and promotes B-cell activation and proliferation; its selective expression on B-lineage cells makes it a prime target for antibody-based and CAR-T immunotherapy."* — TargetSage reasoning trace

This directs therapeutic strategy toward CAR-T cells and antibodies that destroy CD19-bearing cells rather than small molecules. FDA-approved CD19-directed CAR-T products include tisagenlecleucel [43], axicabtagene [44], lisocabtagene [45], and brexucabtagene [46], consistent with this therapeutic profile.

These cases demonstrate that TargetSage's reasoning captures information structurally absent from the feature space. This is consistent with the modality ablation (§4.4): adding LLM reasoning representations to structured features raises Adj. F1 on small-molecule classification from 12.7 to 13.2 and on antibody classification from 21.7 to 22.6. More broadly, these examples suggest that the gap between feature-based and reasoning-augmented methods is largest precisely where the biology is most nuanced—the settings where accurate target identification matters most.

Beyond these cases, every statement in a profile is tagged to its source (per-gene coverage: NCBI 100%, UniProt 93.4%, Open Targets 86.5%, PubMed 29%), and the fact-checker flags cross-source inconsistencies in 385 genes (2.0%) and strips label-revealing sentences (Appendix C). The attribute scores follow the retrieved evidence rather than the model's prior: deleting the single evidence sentence behind a kinase-, ion-channel- or GPCR-family score (fixed-vocabulary attributes, Appendix H) lowers that attribute by 0.28 on average versus 0.04 for the gene's other attributes ($\sim 7\times$; paired t-test $p < 10^{-11}$; Appendix C).

5 Discussion and Conclusion

In this study, we introduced TargetSage to address three persistent gaps in computational therapeutic target identification: static feature design that discards mechanistic knowledge, limited gene-level interpretability, and the false-negative problem from incomplete supervision. The three modules tackle these gaps in sequence: AGENTIC PROFILING distills 2,261 candidate attributes into a 23-

dimensional reasoning vocabulary that encodes evidence-grounded signals invisible to numerical features alone; GUIDED REASONING couples interpretability with predictive utility via a GRPO policy rewarded by downstream Adjusted F1, producing auditable gene-specific rationales; and PU-AWARE SCORING integrates all three representation types under a hybrid class-prior that corrects for score underestimation caused by unlabeled druggable genes.

Evaluated on 15 benchmark tasks spanning 19,032 human protein-coding genes, TargetSage consistently outperformed ten baselines across all four task categories, achieving a macro-average Adjusted F1 of 11.1—37% above the strongest baseline—with cross-dataset evaluation further confirming generalization to independent druggability reference sets. Temporal validation demonstrated prospective predictive power by outperforming all baselines at ranking subsequently validated *Tclin* targets, while independent validation confirmed significant enrichment of Phase 2/3 clinical candidates among *Tbio* proteins, achieving OR=43.2 at the top-25 shortlist—51% above the runner-up baseline.

The case studies further illustrate a structural limitation of feature-based methods that statistics alone cannot reveal: two genes can share the same structured annotations yet differ fundamentally in their biological mechanism and therapeutic tractability. ERBB3's catalytic inactivity and CD19's lineage-marker role are not encoded in any structured database column, yet they determine the entire therapeutic strategy for these targets. This suggests the gains from Guided Reasoning reflect access to a qualitatively different class of evidence than any structured feature can encode.

Despite its advantages, TargetSage has certain limitations. LLM-derived profiles depend on literature coverage, which may disadvantage understudied *Tdark* proteins—the very targets of greatest scientific interest; future work could address this by incorporating protein language model embeddings or AlphaFold structural features as evidence sources independent of publication density. This is a data-level limitation shared by every method: a gene with little measured evidence and sparse literature is hard to separate from most of the genome. Within what the data support, the $\leq$2021 temporal experiment shows TargetSage ranking the 2021$\rightarrow$2025 upgrades, including genes that were understudied in 2021, ahead of every baseline, and the agentic profiling picks up new evidence as it accumulates. Furthermore, TargetSage scores each gene independently, without modeling PPI topology, co-expression modules, or pathway co-regulation; incorporating network structure and patient genomics data are natural extensions. The 15-task benchmark and evaluation protocol introduced here are intended as a community resource for standardized comparison of future target identification methods. TargetSage's rankings are hypotheses for prioritizing experiments, not substitutes for experimental validation, and should not be acted on without it. We hope TargetSage serves as an evidence-grounded, updatable foundation for genome-scale interpretable target prioritization.

Data and Code Availability

Structured biological features are derived from publicly available databases; gene labels are derived from PHAROS TDL classifications. Code and benchmarks are available at `https://github.com/aicb-ZhangLabs/TargetSage`.

Acknowledgments and Disclosure of Funding

This work was supported by the National Institutes of Health grants R01HG012572, R01NS128523, and R01DA063316. The authors declare no competing interests.

References

[1] Ismail Kola and John Landis. Can the pharmaceutical industry reduce attrition rates? *Nature reviews Drug discovery*, 3(8):711–716, 2004.

[2] Steven M Paul, Daniel S Mytelka, Christopher T Dunwiddie, Charles C Persinger, Bernard H Munos, Stacy R Lindborg, and Aaron L Schacht. How to improve r&d productivity: the pharmaceutical industry's grand challenge. *Nature reviews Drug discovery*, 9(3):203–214, 2010.

[3] Andrew L Hopkins and Colin R Groom. The druggable genome. *Nature reviews Drug discovery*, 1(9):727–730, 2002.

[4] Chris J Radoux, Francesca Vianello, Jake McGreig, Nikita Desai, and Anthony R Bradley. The druggable genome: Twenty years later. *Frontiers in Bioinformatics*, 2:958378, 2022.

[5] Jessica Vamathevan, Dominic Clark, Paul Czodrowski, Ian Dunham, Edgardo Ferran, George Lee, Bin Li, Anant Madabhushi, Parantu Shah, Michaela Spitzer, et al. Applications of machine learning in drug discovery and development. *Nature reviews Drug discovery*, 18(6):463–477, 2019.

[6] Jake P. Taylor-King, Michael Bronstein, and David Roblin. The future of machine learning within target identification: Causality, reversibility, and druggability. *Clinical Pharmacology & Therapeutics*, 115(4):655–657, 2024.

[7] Tala M Bakheet and Andrew J Doig. Properties and identification of human protein drug targets. *Bioinformatics*, 25(4):451–457, 2009.

[8] Pedro R Costa, Marcio L Acencio, and Ney Lemke. A machine learning approach for genome-wide prediction of morbid and druggable human genes based on systems-level data. *BMC genomics*, 11(Suppl 5):S9, 2010.

[9] Dorothea Emig, Alexander Ivliev, Olga Pustovalova, Lee Lancashire, Svetlana Bureeva, Yuri Nikolsky, and Marina Bessarabova. Drug target prediction and repositioning using an integrated network-based approach. *PloS one*, 8(4):e60618, 2013.

[10] Zerrin Isik, Christoph Baldow, Carlo Vittorio Cannistraci, and Michael Schroeder. Drug target prioritization by perturbed gene expression and network information. *Scientific reports*, 5(1): 17417, 2015.

[11] Stephane De Cesco, John B Davis, and Paul E Brennan. Targetdb: a target information aggregation tool and tractability predictor. *PloS one*, 15(9):e0232644, 2020.

[12] Arwa Raies, Ewa Tulodziecka, James Stainer, Lawrence Middleton, Ryan S Dhindsa, Pamela Hill, Ola Engkvist, Andrew R Harper, Slavé Petrovski, and Dimitrios Vitsios. Drugnomeai is an ensemble machine-learning framework for predicting druggability of candidate drug targets. *Communications Biology*, 5(1):1291, 2022.

[13] Lezheng Yu, Li Xue, Fengjuan Liu, Yizhou Li, Runyu Jing, and Jiesi Luo. The applications of deep learning algorithms on in silico druggable proteins identification. *Journal of Advanced Research*, 41:219–231, 2022.

[14] Jessa Bekker and Jesse Davis. Learning from positive and unlabeled data: A survey. *Machine learning*, 109(4):719–760, 2020.

[15] Zeming Lin, Halil Akin, Roshan Rao, Brian Hie, Zhongkai Zhu, Wenting Lu, Nikita Smetanin, Robert Verkuil, Ori Kabeli, Yaniv Shmueli, et al. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science*, 379(6637):1123–1130, 2023.

[16] Christina V Theodoris, Ling Xiao, Anant Chopra, Mark D Chaffin, Zeina R Al Sayed, Matthew C Hill, Helene Mantineo, Elizabeth M Brydon, Zexian Zeng, X Shirley Liu, et al. Transfer learning enables predictions in network biology. *Nature*, 618(7965):616–624, 2023.

[17] Charles Elkan and Keith Noto. Learning classifiers from only positive and unlabeled data. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 213–220, 2008.

[18] Marthinus C Du Plessis, Gang Niu, and Masashi Sugiyama. Analysis of learning from positive and unlabeled data. *Advances in neural information processing systems*, 27, 2014.

[19] Harish Ramaswamy, Clayton Scott, and Ambuj Tewari. Mixture proportion estimation via kernel embeddings of distributions. In *International conference on machine learning*, pages 2052–2060. PMLR, 2016.

[20] Shantanu Jain, Martha White, and Predrag Radivojac. Estimating the class prior and posterior from noisy positives and unlabeled data. *Advances in neural information processing systems*, 29, 2016.

[21] Qiao Jin, Yifan Yang, Qingyu Chen, and Zhiyong Lu. Genegpt: augmenting large language models with domain tools for improved access to biomedical information. *Bioinformatics*, 40 (2):btae075, 2024.

[22] Eric Wang, Samuel Schmidgall, Paul F Jaeger, Fan Zhang, Rory Pilgrim, Yossi Matias, Joelle Barral, David Fleet, and Shekoofeh Azizi. Txgemma: Efficient and agentic llms for therapeutics. *arXiv preprint arXiv:2504.06196*, 2025.

[23] Adibvafa Fallahpour, Andrew Magnuson, Purav Gupta, Shihao Ma, Jack Naimer, Arnav Shah, Haonan Duan, Omar Ibrahim, Hani Goodarzi, Chris Maddison, et al. Bioreason: Incentivizing multimodal biological reasoning within a dna-llm model. *Advances in Neural Information Processing Systems*, 38:111644–111671, 2026.

[24] Bernardo P de Almeida, Guillaume Richard, Hugo Dalla-Torre, Christopher Blum, Lorenz Hexemer, Priyanka Pandey, Stefan Laurent, Chandana Rajesh, Marie Lopez, Alexandre Laterre, et al. A multimodal conversational agent for dna, rna and protein tasks. *Nature Machine Intelligence*, 7(6):928–941, 2025.

[25] Zhihong Shao, Peiyi Wang, Qihao Zhu, Runxin Xu, Junxiao Song, Xiao Bi, Haowei Zhang, Mingchuan Zhang, YK Li, Yang Wu, et al. Deepseekmath: Pushing the limits of mathematical reasoning in open language models. *arXiv preprint arXiv:2402.03300*, 2024.

[26] Dimitrios Vitsios and Slavé Petrovski. Mantis-ml: disease-agnostic gene prioritization from high-throughput genomic screens by stochastic semi-supervised learning. *The American Journal of Human Genetics*, 106(5):659–678, 2020.

[27] OpenAI. New embedding models and api updates, January 2024. URL `https://openai.com/index/new-embedding-models-and-api-updates/`.

[28] An Yang, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chengyuan Li, Dayiheng Liu, Fei Huang, Haoran Wei, et al. Qwen2.5 technical report. *arXiv preprint arXiv:2412.15115*, 2024.

[29] Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. LoRA: Low-rank adaptation of large language models. In *International Conference on Learning Representations*, 2022.

[30] Ryuichi Kiryo, Gang Niu, Marthinus C. du Plessis, and Masashi Sugiyama. Positive-unlabeled learning with non-negative risk estimator. In *Advances in Neural Information Processing Systems*, volume 30, 2017.

[31] Wee Sun Lee and Bing Liu. Learning with positive and unlabeled examples using weighted logistic regression. In *Proceedings of the Twentieth International Conference on International Conference on Machine Learning*, ICML'03, page 448–455. AAAI Press, 2003.

[32] Keith J Kelleher, Timothy K Sheils, Stephen L Mathias, Jeremy J Yang, Vincent T Metzger, Vishal B Siramshetty, Dac-Trung Nguyen, Lars Juhl Jensen, Dušica Vidović, Stephan C Schürer, et al. Pharos 2023: an integrated resource for the understudied human proteome. *Nucleic Acids Research*, 51(D1):D1405–D1416, 2023.

[33] David Ochoa, Andrew Hercules, Miguel Carmona, Daniel Suveges, Jarrod Baker, Cinzia Malangone, Irene Lopez, Alfredo Miranda, Carlos Cruz-Castillo, Luca Fumis, et al. The next-generation Open Targets platform: reimagined, redesigned, rebuilt. *Nucleic Acids Research*, 51 (D1):D1353–D1359, 2023.

[34] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778. IEEE, 2016.

[35] Sercan Ö. Arik and Tomas Pfister. TabNet: Attentive interpretable tabular learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(8):6679–6687, 2021.

[36] Yury Gorishniy, Ivan Rubachev, Valentin Khrulkov, and Artem Babenko. Revisiting deep learning models for tabular data. In *Advances in Neural Information Processing Systems*, volume 34, pages 18932–18943, 2021.

[37] OpenAI. GPT-4o mini: advancing cost-efficient intelligence, 2024. URL `https://openai.com/index/gpt-4o-mini-advancing-cost-efficient-intelligence/`.

[38] Emily A. King, J. Wade Davis, and Jacob F. Degner. Are drug targets with genetic support twice as likely to be approved? Revised estimates of the impact of genetic support for drug mechanisms on the probability of drug approval. *PLOS Genetics*, 15(12):e1008489, 2019.

[39] David Mendez, Anna Gaulton, A Patrícia Bento, Jon Chambers, Marleen De Veij, Eloy Félix, María Paula Magariños, Juan F Mosquera, Prudence Mutowo, Michał Nowotka, et al. ChEMBL: towards direct deposition of bioassay data. *Nucleic Acids Research*, 47(D1):D930–D940, 2019.

[40] Qin Ryan, Amna Ibrahim, Martin H. Cohen, John Johnson, Chia-wen Ko, Rajeshwari Sridhara, Robert Justice, and Richard Pazdur. FDA drug approval summary: Lapatinib in combination with capecitabine for previously treated metastatic breast cancer that overexpresses HER-2. *The Oncologist*, 13(10):1114–1119, 2008.

[41] Harpreet Singh, Amanda J. Walker, Laleh Amiri-Kordestani, Joyce Cheng, Shenghui Tang, Pamela Balcazar, Kimberly Barnett-Ringgold, Todd R. Palmby, Xianhua Cao, Nan Zheng, et al. U.S. Food and Drug Administration approval: Neratinib for the extended adjuvant treatment of early-stage HER2-positive breast cancer. *Clinical Cancer Research*, 24(15):3486–3491, 2018.

[42] Pasi A. Jänne, Christina Baik, Wu-Chou Su, Melissa L. Johnson, Hidetoshi Hayashi, Makoto Nishio, Dong-Wan Kim, Marianna Koczywas, Kathryn A. Gold, Conor E. Steuer, et al. Efficacy and safety of patritumab deruxtecan (HER3-DXd) in EGFR inhibitor-resistant, EGFR-mutated non-small cell lung cancer. *Cancer Discovery*, 12(1):74–89, 2022.

[43] Shannon L Maude, Theodore W Laetsch, Jochen Buechner, Susana Rives, Michael Boyer, Henrique Bittencourt, Peter Bader, Michael R Verneris, Heather E Stefanski, Gary D Myers, et al. Tisagenlecleucel in children and young adults with b-cell lymphoblastic leukemia. *New England Journal of Medicine*, 378(5):439–448, 2018.

[44] Sattva S. Neelapu, Frederick L. Locke, Nancy L. Bartlett, Lazaros J. Lekakis, David B. Miklos, Caron A. Jacobson, Ira Braunschweig, Olalekan O. Oluwole, Tanya Siddiqi, Yi Lin, et al. Axicabtagene ciloleucel CAR T-cell therapy in refractory large B-cell lymphoma. *New England Journal of Medicine*, 377(26):2531–2544, 2017.

[45] Jeremy S Abramson, M Lia Palomba, Leo I Gordon, Matthew A Lunning, Michael Wang, Jon Arnason, Amitkumar Mehta, Enkhtsetseg Purev, David G Maloney, Charalambos Andreadis, et al. Lisocabtagene maraleucel for patients with relapsed or refractory large B-cell lymphomas (TRANSCEND NHL 001): a multicentre seamless design study. *The Lancet*, 396(10254): 839–852, 2020.

[46] Michael Wang, Javier Munoz, Andre Goy, Frederick L. Locke, Caron A. Jacobson, Brian T. Hill, John M. Timmerman, Houston Holmes, Samantha Jaglowski, Ian W. Flinn, et al. KTE-X19 CAR T-cell therapy in relapsed or refractory mantle-cell lymphoma. *New England Journal of Medicine*, 382(14):1331–1342, 2020.

[47] Chris Finan, Anna Gaulton, Felix A. Kruger, R. Thomas Lumbers, Tina Shah, Jorgen Engmann, Luana Galver, Ryan Kelley, Anneli Karlsson, Rita Santos, et al. The druggable genome and support for target identification and validation in drug development. *Science Translational Medicine*, 9(383):eaag1166, 2017.

[48] Matthew G Rees, Brinton Seashore-Ludlow, Jaime H Cheah, Drew J Adams, Edmund V Price, Shubhroz Gill, Sarah Javaid, Matthew E Coletti, Victor L Jones, Nicole E Bodycombe, et al. Correlating chemical sensitivity and basal gene expression reveals mechanism of action. *Nature Chemical Biology*, 12(2):109–116, 2016.

[49] Clare Pacini, Emma Duncan, Emanuel Gonçalves, James Gilbert, Shriram Bhosle, Stuart Horswell, Emre Karakoc, Howard Lightfoot, Ed Curry, Francesc Muyas, et al. A comprehensive clinically informed map of dependencies in cancer cells and framework for target prioritization. *Cancer Cell*, 42(2):301–316.e9, 2024.

[50] Sara R. Savage, Xinpei Yi, Jonathan T. Lei, Bo Wen, Hongwei Zhao, Yuxing Liao, Eric J. Jaehnig, Lauren K. Somes, Paul W. Shafer, Tobie D. Lee, et al. Pan-cancer proteogenomics expands the landscape of therapeutic targets. *Cell*, 187(16):4389–4407.e15, 2024.

[51] Melanie Schneider, Chris J. Radoux, Andrew Hercules, David Ochoa, Ian Dunham, Lykourgos-Panagiotis Zalmas, Gerhard Hessler, Sven Ruf, Veerabahu Shanmugasundaram, Michael M. Hann, et al. The PROTACtable genome. *Nature Reviews Drug Discovery*, 20(10):789–797, 2021.

[52] Allan Peter Davis, Cynthia J Grondin, Robin J Johnson, Daniela Sciaky, Jolene Wiegers, Thomas C Wiegers, and Carolyn J Mattingly. Comparative Toxicogenomics Database (CTD): update 2021. *Nucleic Acids Research*, 49(D1):D1138–D1143, 2021.

[53] Matthias Blum, Hsin-Yu Chang, Sara Chuguransky, Tiago Grego, Swaathi Kandasaamy, Alex Mitchell, Gift Nuka, Typhaine Paysan-Lafosse, Matloob Qureshi, Shriya Raj, et al. The InterPro protein families and domains database: 20 years on. *Nucleic Acids Research*, 49(D1): D344–D354, 2021.

[54] Damian Szklarczyk, Annika L Gable, David Lyon, Alexander Junge, Stefan Wyder, Jaime Huerta-Cepas, Milan Simonovic, Nadezhda T Doncheva, John H Morris, Peer Bork, et al. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research*, 47(D1): D607–D613, 2019.

[55] Taibo Li, Rasmus Wernersson, Rasmus B Hansen, Heiko Horn, Johnathan Mercer, Greg Slodkowicz, Christopher T Workman, Olga Rigina, Kristoffer Rapacki, Hans H Stærfeldt, et al. A scored human protein–protein interaction network to catalyze genomic interpretation. *Nature Methods*, 14(1):61–64, 2017.

[56] Bijay Jassal, Lisa Matthews, Guilherme Viteri, Chuqiao Gong, Pascual Lorente, Antonio Fabregat, Konstantinos Sidiropoulos, Justin Cook, Marc Gillespie, Robin Haw, et al. The reactome pathway knowledgebase. *Nucleic Acids Research*, 48(D1):D498–D503, 2020.

[57] Sharon L Freshour, Susanna Kiwala, Kelsy C Cotto, Adam C Coffman, Joshua F McMichael, Jonathan J Song, Malachi Griffith, Obi L Griffith, and Alex H Wagner. Integration of the Drug–Gene Interaction Database (DGIdb 4.0) with open crowdsource efforts. *Nucleic Acids Research*, 49(D1):D1144–D1151, 2021.

[58] Monkol Lek, Konrad J. Karczewski, Eric V. Minikel, Kaitlin E. Samocha, Eric Banks, Timothy Fennell, Anne H. O'Donnell-Luria, James S. Ware, Andrew J. Hill, Beryl B. Cummings, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*, 536(7616):285–291, 2016.

[59] Konrad J. Karczewski, Laurent C. Francioli, Grace Tiao, Beryl B. Cummings, Jessica Alföldi, Qingbo Wang, Ryan L. Collins, Kristen M. Laricchia, Andrea Ganna, Daniel P. Birnbaum, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*, 581 (7809):434–443, 2020.

[60] Slavé Petrovski, Quanli Wang, Erin L. Heinzen, Andrew S. Allen, and David B. Goldstein. Genic intolerance to functional variation and the interpretation of personal genomes. *PLoS Genetics*, 9(8):e1003709, 2013.

[61] Joshua Traynelis, Michael Silk, Quanli Wang, Samuel F. Berkovic, Liping Liu, David B. Ascher, David J. Balding, and Slavé Petrovski. Optimizing genomic medicine in epilepsy through a gene-customized approach to missense variant interpretation. *Genome Research*, 27(10): 1715–1729, 2017.

[62] Carol J Bult, Judith A Blake, Cynthia L Smith, James A Kadin, Joel E Richardson, and the Mouse Genome Database Group. Mouse Genome Database (MGD) 2019. *Nucleic Acids Research*, 47(D1):D801–D806, 2019.

[63] Joanna S Amberger, Carol A Bocchini, Alan F Scott, and Ada Hamosh. OMIM.org: leveraging knowledge across phenotype–gene relationships. *Nucleic Acids Research*, 47(D1):D1038–D1043, 2019.

[64] The UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Research*, 49(D1):D480–D489, 2021.

[65] John Lonsdale, Jeffrey Thomas, Mike Salvatore, Rebecca Phillips, Edmund Lo, Saboor Shad, Richard Hasz, Gary Walters, Fernando Garcia, Nancy Young, et al. The genotype-tissue expression (GTEx) project. *Nature genetics*, 45(6):580–585, 2013.

[66] Garth R. Brown, Vichet Hem, Kenneth S. Katz, Michael Ovetsky, Craig Wallin, Olga Ermolaeva, Igor Tolstoy, Tatiana Tatusova, Kim D. Pruitt, Donna R. Maglott, and Terence D. Murphy. Gene: a gene-centered information resource at NCBI. *Nucleic Acids Research*, 43(D1):D36–D42, 2015.

[67] Aaron Hurst, Adam Lerer, Adam P Goucher, Adam Perelman, Aditya Ramesh, Aidan Clark, AJ Ostrow, Akila Welihinda, Alan Hayes, Alec Radford, et al. GPT-4o system card. *arXiv preprint arXiv:2410.21276*, 2024.

[68] Jinhyuk Lee, Wonjin Yoon, Sungdong Kim, Donghyeon Kim, Sunkyu Kim, Chan Ho So, and Jaewoo Kang. BioBERT: a pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics*, 36(4):1234–1240, 2020.

[69] Fangyu Liu, Ehsan Shareghi, Zaiqiao Meng, Marco Basaldella, and Nigel Collier. Self-alignment pretraining for biomedical entity representations. In *Proceedings of the 2021 conference of the North American chapter of the association for computational linguistics: human language technologies*, pages 4228–4238, 2021.

[70] Iz Beltagy, Kyle Lo, and Arman Cohan. SciBERT: A pretrained language model for scientific text. In *Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (EMNLP-IJCNLP)*, pages 3615–3620, 2019.

A 15-Task Benchmark Details

To enable systematic evaluation of computational therapeutic target identification, we curated a standardized benchmark of 15 binary classification tasks over 19,032 human protein-coding genes. **Table 6** summarizes task definitions, positive counts $|P|$, and source resources.

Category definitions and distinctions. The four categories differ in both their evidence type and clinical scope. **PHAROS tasks** use *clinical validation* as the criterion: a gene is positive if it has at least one approved drug of any modality (small molecule, antibody, biologic), as recorded in PHAROS/TCRD [32]. **Druggability triage tasks** follow the druggable-genome tiers of Finan et al. [47], used as published without further filtering: Tier 1 comprises the efficacy targets of approved small-molecule and biotherapeutic drugs and of clinical-phase drug candidates, and Tier 2 adds proteins with known bioactive drug-like small-molecule binding partners or high sequence similarity to approved drug targets. Tier 1 (1,399 genes) is about twice as large as *Tclin* (700 genes); 75% of *Tclin* genes fall in Tier 1, and the two positive sets have a Jaccard similarity of 0.33. **Cancer druggability tasks** use *oncology-specific annotation* across three stringency tiers (T1 $\subset$ T1–T2 $\subset$ T1–T3) and three scope definitions (type-specific, pan-cancer, cancer-relevant) [48–50]. **Drug modality tasks** further stratify tractability by therapeutic modality (small molecule, antibody, PROTAC) and evidence level (approved vs. clinical-stage) [33, 51]. The overlap structure across all 15 tasks is visualized in **Figure 3**.

Positive-unlabeled framing and benchmark design. All 15 tasks are formulated as positive-unlabeled (PU) learning problems: known druggable genes form the positive set P, while the remaining protein-coding genes constitute the unlabeled set U—a mixture of true negatives and undiscovered targets whose absence of annotation reflects limited evidence rather than confirmed non-druggability. The positive rate $|P|/n$ ranges from 0.4% (Antibody Approved, $|P|$=69) to 13.5% (Clinical & Chemical, $|P|$=2,562), spanning a broad diversity of annotation density across modalities and stringency levels. This multi-task structure evaluates methods across complementary axes—evidence type, disease scope, therapeutic modality, and annotation stringency—with macro-averaged Adjusted F1 as the summary metric robust to task-level variation in class balance.

Table 6: **Task formulations and detailed statistics for target identification benchmarks.**

Task	Definition	$\|P\|$	$\|U\|$	Refs.
PHAROS (Disease-agnostic)				
Clinical Targets	*Tclin*-labeled genes in PHAROS/TCRD.	700	18,332	[32]
Clinical & Chemical	*Tclin*- or *Tchem*-labeled genes in PHAROS/TCRD.	2,562	16,470	[32]
Druggability Triage (Disease-agnostic)				
Top-Tier Targets	Genes mapped to Tier 1 druggable targets.	1,399	17,633	[47]
High-Confidence	Genes mapped to Tier 1 or Tier 2 druggable targets.	2,048	16,984	[47]
Cancer Druggability (Domain-specific)				
Cancer-Relevant	Genes in the oncology-relevant target panel.	319	18,713	[48–50]
Type-Specific	Cancer-type-specific targets in the derived panel.	297	18,735	[49]
Pan-Cancer	Pan-cancer targets in the derived panel.	195	18,837	[49]
T1 Cancer	Genes matched to approved cancer-drug targets.	156	18,876	[50]
T1–T2 Cancer	Approved or other-indication cancer-drug targets.	620	18,412	[50]
T1–T3 Cancer	Approved, investigational, or experimental cancer-drug targets.	1,066	17,966	[50]
Drug Modality (Domain-specific)				
Small-Molecule (Approved)	Approved small-molecule targets in Open Targets.	777	18,255	[33]
Small-Molecule (Clinical+)	Approved or clinical-stage small-molecule targets.	1,103	17,929	[33]
Antibody (Approved)	Approved antibody targets in Open Targets.	69	18,963	[33]
Antibody (Clinical+)	Approved or clinical-stage antibody targets.	279	18,753	[33]
PROTAC	Genes in the PROTACtable target set.	269	18,763	[51]

B Implementation Details

B.1 Data sources, releases and access dates

Table 7 records the exact release or access date of every source. The 482 structured features $\mathbf{x}^{\text{bio}}$ are built from February-2021 releases throughout. For the main benchmark the four LLM-retrieved sources use the releases current at retrieval time; for the strict temporal validation (§4.3) every source is pinned to its $\leq$ 2021 release and the generative-LLM attribute channel is dropped entirely, so no post-2021 information can enter through any modality.

B.2 Training configurations

All baselines treat unlabeled genes as negatives. The classical baselines use their standard scikit-learn objectives: LR and SVM (RBF kernel) with class-balanced weights, KNN (k=10), Gaussian NB, RF (200 trees), and GB (200 trees, depth 5) and MLP (256–128) with log-loss. ResNet and FT-Transformer are trained with class-weighted binary cross-entropy. TabNet uses the official `pytorch-tabnet` library with AUC-based early stopping (max 100 epochs, patience 20).

The TargetSage architecture uses three independent MLP transformation heads ($d_{\text{latent}} = 256$, inner hidden size 512) with gated fusion. LLM embeddings are PCA-compressed to 256 dimensions before fusion. Training proceeds in two stages: (1) 10-epoch warm-up with weighted BCE on observed labels (class-balanced), and (2) 25-epoch nnPU optimization with semantic reweighting. We use AdamW with learning rate 2×10^{-4}, batch size 512, dropout 0.2, and report results over 5 random

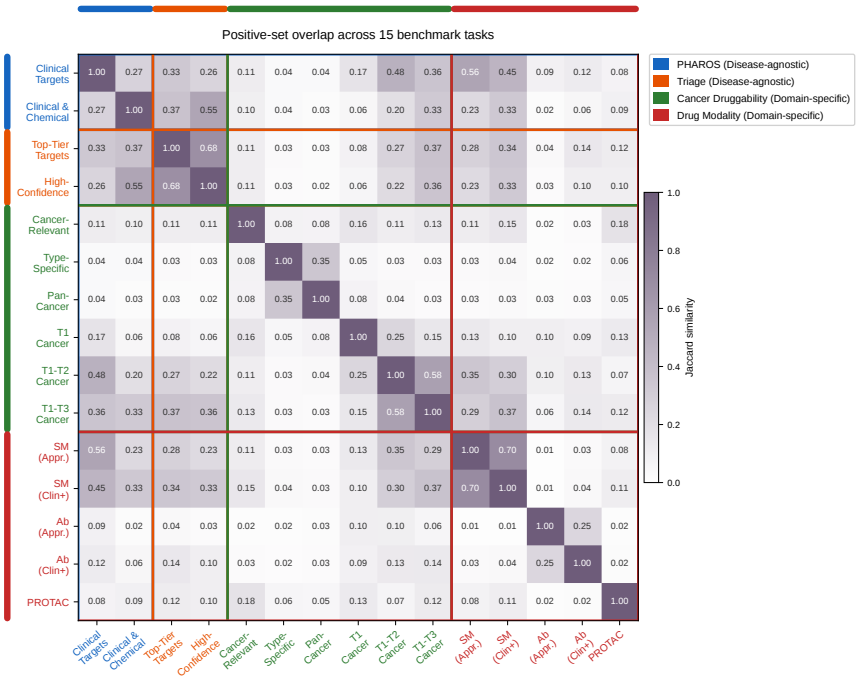

Figure 3: **Positive-set overlap across 15 benchmark tasks (Jaccard similarity).** Within-category overlap (diagonal blocks) reflects intentional stringency-level nesting, e.g., T1 $\subset$ T1–T2 $\subset$ T1–T3 and SM (Appr.) $\subset$ SM (Clin+). Cross-category overlap varies: the mean Jaccard of the six category pairs ranges from 0.11 (Cancer $\times$ Drug Modality) to 0.38 (PHAROS $\times$ Druggability Triage), with the greatest overlap between the two disease-agnostic categories. The highest cross-category value (Clinical Targets $\times$ SM (Appr.) $= 0.56$) is expected: FDA-approved small-molecule targets are largely a subset of *Tclin* by definition.

Table 7: Source releases and access dates. Structured features are February-2021 releases in both settings; the retrieved-evidence channels differ between the main benchmark and the temporal validation.

Channel / source	**Main benchmark**	**Temporal validation ($\leq$2021)**
Structured features $\mathbf{x}^{\text{bio}}$		
Open Targets	21.02 (2021-02)	21.02 (2021-02)
Pharos / TCRD	v6.7	v6.7
gnomAD	v2.1.1	v2.1.1
STRING	v11.0	v11.0
CTDbase	2021-02 snapshot	2021-02 snapshot
InterPro	84.0 (2021-02)	84.0 (2021-02)
DGIdb	v4.0	v4.0
Retrieved evidence (Module 1) and *implicit embedding* $\mathbf{z}^{\text{emb}}$		
NCBI Gene	accessed 2026-02-21	$\leq$2021 (post-2021 summaries blanked)
PubMed	accessed 2026-04-07	abstracts $\leq$2021
UniProt	release 2025_04 (retrieved 2026-04-07)	release 2021_01
Open Targets	release 26.03 (retrieved 2026-04-07)	release 21.02
Text encoder	`text-embedding-3-large`	BioBERT `biobert-base-cased-v1.2`
Explicit reasoning attributes $\mathbf{x}^{\text{attr}}$		
Generative LLM	Module 2 policy (§3)	dropped (no generative LLM)

Table 8: **Summary of structured biological features ($N = 482$) aggregated for the human exome.** Features are taken from the DrugnomeAI feature set [12], which builds on the mantis-ml compilation [26]; expression-specificity and antibody counts are obtained via Pharos.

Category	Source	Dim.	Description
Function & Networks	CTDbase [52]	239	Pathway / function annotations
	InterPro [53]	97	Protein families and domains
	STRING [54]	10	L1/L2 seed-gene overlap‡
	InWeb [55]	4	L1/L2 seed-gene overlap‡
	Reactome [56]	2	L1/L2 seed-gene overlap‡
	Pharos [32]	1	Protein interaction count
Chemical & Antibody	CTDbase [52]	68	Chemical–gene interaction annotations
	Pharos [32]	2	Antibody and monoclonal counts
	DGIdb [57]	1	Gene–drug interaction types
Genomic Constraint	ExAC [58]	19	Constraint / CNV (18), gene size (1)
	GnomAD [59]	15	Mutation constraint metrics
	Genic Intol. [60, 61]	6	RVIS, MTR, LoF-FDR, gene coverage
	Mouse Genes	4	Essentiality, overexpression, GC%
Phenotype & Disease	GWAS	6	Hits, P values, odds ratios, trait flag
	MGI [62]	1	Mouse–human phenotypes
	OMIM [63]	1	Disease association count
Expression & Sequence	Pharos [32]	4	HPM/HPA tissue specificity
	UniProt [64]	1	Sequence length
	GTEx [65]	1	Tissue expression profile

‡Fraction of a gene's direct (L1) or two-hop (L2) network neighbours that are known targets.

seeds with stratified 80/20 splits. For temporal validation, models were trained on 2021 PHAROS labels and evaluated against 2025 annotations.

C Agentic Tool Pipeline Details

The agentic pipeline uses six tools to semanticize structured biological evidence, retrieve complementary external knowledge, and assemble a unified profile for each gene.

Tool 1: Structured Feature Reader. The agent first reads the 482 numerical features aggregated from 16 databases (**Table 8**) and converts informative entries into human-readable evidence statements, so that downstream reasoning operates on contextualized biological evidence rather than isolated numerical values. For example, selected entries for ABL1 can be rendered as "*highly loss-of-function intolerant (pLI=1.00); observed/expected LoF ratio: 0.07; essential in mouse; protein length 1130 aa; 43 DGIdb interaction types; protein kinase, SH2, and SH3 domains*." This semanticization provides a more informative input for downstream reasoning.

Tools 2–5: External Knowledge Collection. The agent then queries four external sources to collect information not fully captured by the structured feature matrix:

- **NCBI Gene Summary** [66] — gene summaries providing broad functional and disease context.
- **UniProt Function Annotation** [64] — expert-curated molecular function descriptions, subcellular localization, and tissue specificity.
- **PubMed Literature** — recent abstracts providing mechanistic or therapeutic context for the gene.
- **Open Targets Platform** — tractability assessments (small molecule, antibody, PROTAC) and known safety liabilities.

Together, these sources provide functional context, localization, tractability, safety annotations, and literature evidence, which are merged into a unified multi-source profile for downstream reasoning. Across the 19,032-gene corpus, structured features and NCBI summaries are available for all genes, UniProt annotations for 93.4% and Open Targets tractability for 86.5% of genes; PubMed abstracts are retrieved selectively when additional literature evidence is needed and are present for 29% of genes.

Tool 6: Fact-Checking Agent. Collecting information from multiple sources introduces two risks: (1) *label leakage*, where external sources contain statements directly revealing the prediction target (e.g., "FDA-approved drug target"); and (2) *inconsistent statements across sources*. Such inconsistencies are expected in multi-source biomedical evidence, because annotations may reflect different experimental settings, biological contexts, or database update cycles. The fact-checker scans collected text for leakage patterns, removes sentences that directly reveal the target label from the assembled profile, and flags inconsistent statements for downstream interpretation. For example, a PubMed abstract for CA4 contained the phrase "FDA approved," which would have directly disclosed clinical validation status and was therefore removed. In our corpus, we detected direct label leakage in 1 of 19,032 genes and inconsistent cross-source statements in 385 genes (2.0%), where different sources provided non-identical descriptions of the same gene, such as differing localization or functional annotations. **Counterfactual faithfulness.** To test whether attribute scores follow the retrieved evidence rather than the model's prior knowledge, we take three attributes that are each supported by a single evidence sentence in the gene summary—the kinase-, ion-channel- and GPCR-family scores of the fixed attribute vocabulary (Appendix H)—remove that sentence, and rescore 50 genes per attribute. The corresponding attribute drops by 0.28 on average, against 0.04 for the gene's other attributes ($\sim$7$\times$ larger; paired t-test $p < 10^{-11}$).

D Reasoning Attribute Vocabulary

The agentic pipeline surfaces 2,261 candidate attribute dimensions across all 19,032 genes; we retain those with coverage $\geq 5\%$, yielding 28 attributes; after deduplication of naming variants (five RVIS-related names collapsed to *rvis score*; *tissue specificity* merged with *expression tissue specificity*), $d_a = 23$ unique attributes remain. The 23 deduplicated attributes are: *disease association*, *essential gene*, *loss of function tolerance*, *loss of function constraint*, *observed expected lof ratio*, *rvis score*, *gwas associations*, *protein protein interactions*, *chemical gene interactions*, *drug interaction types*, *known antibodies*, *antibody*, *antibody availability*, *small molecule*, *tractable modalities*, *overall therapeutic potential*, *functional characterization*, *protein domains*, *protein length*, *expression tissue specificity*, *expression specificity*, *expression broadly expressed*, and *alternative splicing*.

E LLM Prior Estimation Details

We estimate the task-specific LLM prior $\pi_{t,\text{LLM}}$ using GPT-4o [67] (Azure OpenAI endpoint) at temperature 0.15. For each of the 15 tasks, we issue **10 independent queries**; each query supplies the task name, the total gene pool size $n = 19{,}032$, the observed positive count $|P_t|$, and the empirical lower bound $\ell_t = |P_t|/n$.

Multiplier formulation. Rather than requesting a raw class-prior probability—which is susceptible to hallucination and tends to produce degenerate round-number outputs—the model is instructed to return a triplet of *multipliers* $(m_{\min}, m_{\text{med}}, m_{\max})$ representing how much larger the true positive rate is likely to be relative to the observable lower bound ℓ_t. All three values are constrained to $[1.0,\ 3.5]$, and the system prompt explicitly guides the model to use non-round values (e.g., 1.18, 1.43, 2.05) to avoid degenerate responses such as 1.0, 2.0, or 3.0. Anchoring the output to the observable lower bound keeps the estimate in a biologically plausible range and reduces sensitivity to the model's internal calibration.

Aggregation. From the k-th run we retain the median multiplier $m_{\text{med}}^{(k)}$. The final multiplier is the median across all ten runs:

$$m_{t,\text{med}}^{*} = \text{median}_{k=1}^{10}\ m_{\text{med}}^{(k)}, \tag{7}$$

giving $\pi_{t,\text{LLM}} = m^*_{t,\text{med}} \cdot \ell_t$. The repeated-query strategy reduces variance from any single stochastic draw. The resulting $\pi_{t,\text{LLM}}$ values are computed once before training and held fixed throughout all experiments.

F GRPO Full Derivation

The task-specific reasoning policy receives the enriched multi-source profile for gene g_i (assembled by the 6-tool agentic pipeline in Appendix C) together with the current task context. The policy is parameterized by Qwen2.5-1.5B-Instruct with LoRA adapters (r=32, α=64, dropout 0.05) and optimized directly from the instruct checkpoint via GRPO [25], without supervised fine-tuning on teacher traces. GRPO updates the LoRA parameters while the base Qwen weights remain fixed.

Rollout generation. For each gene state s_i, we sample G rollouts from the old policy:

$$y_{i,1}, \ldots, y_{i,G} \sim \pi_{\theta_{\text{old}}}(\cdot \mid s_i), \qquad y_{i,k} = (\tau_{i,k}, \mathbf{a}_{i,k}). \tag{8}$$

Here s_i denotes the task-conditioned textual input state for gene g_i, $\tau_{i,k} = (\tau_{i,k,1}, \ldots, \tau_{i,k,L_{i,k}})$ is the k-th sampled reasoning trace with variable length $L_{i,k}$, and $\mathbf{a}_{i,k} \in [0,1]^{d_a}$ is the corresponding explicit reasoning-attribute vector over the fixed d_a-dimensional vocabulary.

Rollout-derived multimodal representation. Each sampled trace $\tau_{i,k}$ is encoded into a dynamic implicit embedding $\mathbf{z}^{\text{emb}}_{i,k} \in \mathbb{R}^{d_{\text{proj}}}$ by mean-pooling the last-layer hidden states of the reasoning policy over the generated tokens and projecting via a fixed random matrix $\mathbf{W}_{\text{proj}} \in \mathbb{R}^{d_{\text{llm}} \times d_{\text{proj}}}$ (initialized once, never updated). This *self-supervised* embedding is computed for free from the same forward pass used for policy-gradient computation and requires no external API calls; it is used only to score rollouts during GRPO, and the final $\mathbf{z}^{\text{emb}}$ passed to Module 3 is computed with `text-embedding-3-large` (§3.3). The frozen proxy M3 discriminators are pre-trained on these same hidden-state embeddings (extracted from the base model before any RL update), so their embedding input distribution matches what they receive during GRPO reward evaluation. The downstream predictor therefore receives, for each sampled rollout, the same three modalities as in the main method:

$$\mathbf{x}^{\text{bio}}_i, \qquad \mathbf{a}_{i,k}, \qquad \mathbf{z}^{\text{emb}}_{i,k}.$$

Across the gene set, the rollout-specific explicit-attribute and implicit-embedding matrices are

$$\mathbf{A}^{(k)} = [\mathbf{a}_{1,k}, \ldots, \mathbf{a}_{n,k}]^\top \in \mathbb{R}^{n \times d_a}, \qquad \mathbf{Z}^{(k)} = [\mathbf{z}^{\text{emb}}_{1,k}, \ldots, \mathbf{z}^{\text{emb}}_{n,k}]^\top \in \mathbb{R}^{n \times d_{\text{proj}}}. \tag{9}$$

For each gene g_i, Eq. (2) maps the rollout-specific input triple

$$(\mathbf{x}^{\text{bio}}_i, \mathbf{a}_{i,k}, \mathbf{z}^{\text{emb}}_{i,k})$$

to a fused latent representation $\mathbf{h}^{(k)}_i \in \mathbb{R}^{d_{\text{latent}}}$. Stacking these vectors gives

$$\mathbf{H}^{(k)} = [\mathbf{h}^{(k)}_1, \ldots, \mathbf{h}^{(k)}_n]^\top \in \mathbb{R}^{n \times d_{\text{latent}}}.$$

Downstream reward. For each task t, applying the PU discriminator to $\mathbf{H}^{(k)}$ yields gene-level prediction scores

$$\hat{p}^{(t,k)}_i = f_t\big(\mathbf{h}^{(k)}_i\big), \qquad i = 1, \ldots, n. \tag{10}$$

Thus $\hat{\mathbf{p}}^{(t,k)} = (\hat{p}^{(t,k)}_1, \ldots, \hat{p}^{(t,k)}_n)^\top \in [0,1]^n$ is the resulting task-t score vector. The scalar reward assigned to rollout $y_{i,k}$ is the macro-averaged Adjusted F1 over all T tasks:

$$r(s_i, y_{i,k}) = \frac{1}{T} \sum_{t=1}^{T} F^{(t)}_{\text{1-adj}}\big(\hat{\mathbf{p}}^{(t,k)}\big). \tag{11}$$

Thus, reward is assigned to the complete rollout $y_{i,k} = (\tau_{i,k}, \mathbf{a}_{i,k})$, so that optimization depends jointly on the sampled reasoning trace and the explicit reasoning attributes that it induces.

Advantage estimation. Following GRPO, rewards are normalized within each gene's rollout group:

$$\hat{A}_{i,k} = \frac{r(s_i, y_{i,k}) - \bar{r}_i}{\text{std}\big(\{r(s_i, y_{i,j})\}_{j=1}^{G}\big) + \varepsilon_{\text{std}}}, \qquad \bar{r}_i = \frac{1}{G} \sum_{j=1}^{G} r(s_i, y_{i,j}). \tag{12}$$

Policy update. Let $\tau_{i,k,m}$ denote the m-th token in the sampled reasoning trace $\tau_{i,k}$. GRPO then weights the autoregressive log-probability of each sampled trace by its group-centered advantage:

$$\mathcal{L}_{\text{PG}}(\theta) = \mathbb{E}_{s_i \sim \mathcal{S},\, \{y_{i,k}\}_{k=1}^{G} \sim \pi_{\theta_{\text{old}}}} \left[\frac{1}{G} \sum_{k=1}^{G} \big(-\hat{A}_{i,k}\big) \sum_{m=1}^{|\tau_{i,k}|} \log \pi_\theta(\tau_{i,k,m} \mid s_i, \tau_{i,k,<m}) \right]. \tag{13}$$

In this way, sampled traces and attribute outputs that improve downstream three-modal PU-based prediction receive positive policy pressure, whereas weaker rollouts are suppressed. We use this on-policy form without ratio clipping or a KL penalty, and clip gradient norms at 1.0.

Training details. After each RL step, the rollout set is refreshed and the policy is updated on sampled genes from the labeled subset. We use $G{=}4$ rollouts per gene, optimize with AdamW at learning rate 5×10^{-5}, use temperature 0.8, and train on a single NVIDIA RTX 3090.

G Temporal Validation: Pinning Every Input to $\leq$ 2021

The temporal experiment in Section 4.3 restricts every input to data available by February 2021 (Table 7). The ten baselines consume only the 482 structured features, which are February-2021 releases in every setting; TargetSage additionally reads text, so each text source is constrained to $\leq$ 2021 as listed in Table 5. NCBI Gene summaries have no dated archive, so the 6,045 summaries whose provenance tag is dated 2022 or later are blanked rather than back-dated; the remaining 12,987 carry a $\leq$ 2021 tag (11,602) or no tag (1,385).

No generative LLM existed at the 2021 cutoff, so the explicit reasoning attributes are dropped and the text channel is re-embedded with BioBERT (`dmis-lab/biobert-base-cased-v1.2`). Under these inputs TargetSage ranks the 2021$\rightarrow$2025 upgrades best of all methods: median rank percentile 7.1 for Upgrade to *Tclin* ($n{=}90$) and 22.1 for Upgrade to *Tclin/Tchem* ($n{=}456$), against 7.4–50.5 and 22.4–37.1 across the baselines.

H Ablation Study

We ablate each TargetSage component to isolate its contribution (**Tables 9–10**).

Feature modalities. **Table 9** ablates all combinations of the three input modalities—structured biological features (Bio), explicit reasoning attributes (Expl.), and implicit reasoning embeddings (Impl.). All modality configurations achieve competitive macro-average performance (Bio: 10.7, Expl.: 10.8, Impl.: 10.7), indicating that structured biological features and LLM-derived representations each carry meaningful signal. The relative ranking of individual modalities varies by task: Impl. falls slightly below Bio on Top-Tier Targets (6.1 vs. 6.2) and Cancer-Relevant (13.3 vs. 13.5), while Bio+Impl. outperforms the full combination on Clinical & Chemical (3.3 vs. 3.2) and Ab (Clin+) (17.6 vs. 17.5), and Expl.+Impl. leads on text-rich tasks such as Top-Tier Targets, Pan-Cancer, and Ab (Appr.). Two-modality combinations (Bio+Expl.: 10.8, Bio+Impl.: 10.9, Expl.+Impl.: 10.9) match or exceed every single modality, confirming complementarity across feature types. The full combination Bio+Expl.+Impl. achieves the best macro-average of 11.1 and wins on 9 of 15 tasks; the six tasks where a two-modality config prevails suggest that adding the third modality does not always yield further improvement, possibly due to redundancy between explicit and implicit reasoning representations on those tasks.

Fusion operators. All four variants share the same per-modality MLP projections $\mathbf{h}_i^m = \sigma(\mathbf{W}_m \mathbf{u}_i^m + \mathbf{b}_m)$ and differ only in how the projected representations are aggregated into $\mathbf{h}_i^{\text{fuse}}$: *Summation*: $\mathbf{h}_i^{\text{fuse}} = \frac{1}{|\mathcal{M}|} \sum_m \mathbf{h}_i^m$; *Concatenation*: $\mathbf{h}_i^{\text{fuse}} = \mathbf{W}_c[\mathbf{h}_i^{\text{bio}}; \mathbf{h}_i^{\text{attr}}; \mathbf{h}_i^{\text{emb}}]$ with $\mathbf{W}_c \in \mathbb{R}^{d_{\text{latent}} \times 3d_{\text{latent}}}$; *Attention*: $\mathbf{h}_i^{\text{fuse}} = \sum_m a_m \mathbf{h}_i^m$ where $a_m = \text{softmax}_m(\mathbf{q}^\top \mathbf{h}_i^m / \sqrt{d_{\text{latent}}})$ with a learnable query $\mathbf{q}$; *Gated* (default): input-dependent gates $\boldsymbol{\alpha} = \text{softmax}(\mathbf{W}_g[\mathbf{h}_i^{\text{bio}}; \mathbf{h}_i^{\text{attr}}; \mathbf{h}_i^{\text{emb}}])$ modulate each modality contribution (Eq. 2). Gated fusion achieves the best macro-average Adjusted F1 (11.1) against concatenation (10.7), summation (10.7), and vanilla attention (9.7) (**Table 10**). The ranking is not uniform across tasks: concatenation outperforms gated on Clinical Targets, Clinical & Chemical, Cancer-Relevant, SM (Appr.), Ab (Appr.), and Ab (Clin+), and summation leads on Pan-Cancer, T1 Cancer, and PROTAC; gated fusion's macro-average advantage comes from its

Table 9: **Modality ablation.** Adjusted F1 for all combinations of structured biological features (Bio), explicit reasoning attributes (Expl.), and implicit reasoning embeddings (Impl.). Mean ± std over 5 seeds. **Bold** = best per row.

Task	Bio	Expl.	Impl.	Bio+Expl.	Bio+Impl.	Expl.+Impl.	Bio+Expl.+Impl.
PHAROS (Disease-agnostic)							
Clinical Targets	10.7±0.9	10.9±0.9	10.8±0.9	11.0±1.0	11.1±1.0	10.9±1.0	**11.2**±1.0
Clinical & Chemical	3.1±0.3	3.1±0.3	3.1±0.3	3.1±0.3	**3.3**±0.3	3.1±0.3	3.2±0.3
Druggability Triage (Disease-agnostic)							
Top-Tier Targets	6.2±0.2	6.3±0.2	6.1±0.2	6.3±0.2	6.3±0.2	**6.6**±0.2	6.5±0.2
High-Confidence	4.0±0.2	4.1±0.2	4.0±0.2	4.1±0.2	4.1±0.2	4.1±0.2	**4.2**±0.2
Cancer Druggability (Domain-specific)							
Cancer-Relevant	13.5±2.1	13.6±2.1	13.3±2.2	13.6±2.2	13.6±2.3	13.5±2.4	**14.0**±2.5
Type-Specific	5.2±0.5	5.2±0.5	5.1±0.5	5.2±0.5	5.3±0.5	5.2±0.5	**5.4**±0.5
Pan-Cancer	5.3±0.9	5.3±0.9	5.3±0.9	5.3±1.0	5.3±1.0	**5.6**±1.0	5.5±1.1
T1 Cancer	23.3±6.5	23.6±6.6	23.2±6.8	23.6±6.9	23.6±7.1	23.6±7.2	**24.2**±7.4
T1–T2 Cancer	7.5±0.6	7.5±0.6	7.5±0.7	7.5±0.7	7.6±0.7	7.6±0.7	**7.8**±0.8
T1–T3 Cancer	5.2±0.3	5.2±0.3	5.2±0.3	5.2±0.3	**5.5**±0.3	5.2±0.3	5.4±0.3
Drug Modality (Domain-specific)							
SM (Appr.)	12.7±0.9	12.8±0.9	12.7±0.9	12.9±1.0	12.9±1.0	12.9±1.0	**13.2**±1.1
SM (Clin+)	8.7±0.3	8.8±0.3	8.7±0.3	8.8±0.3	8.9±0.3	8.8±0.3	**9.1**±0.3
Ab (Appr.)	21.7±11.0	21.9±11.2	21.7±11.5	21.9±11.7	21.9±12.0	**22.8**±13.1	22.6±13.4
Ab (Clin+)	16.8±1.0	16.9±1.0	16.8±1.0	16.9±1.1	**17.6**±1.1	17.1±1.1	17.5±1.2
PROTAC	16.5±2.7	16.6±2.8	16.4±2.8	16.6±2.9	16.8±2.9	16.6±3.0	**17.2**±3.3
Macro-Avg	10.7±1.9	10.8±1.9	10.7±2.0	10.8±2.0	10.9±2.1	10.9±2.1	**11.1**±2.1

consistently strong performance on tail tasks with sparse positive labels, where the learned gate suppresses noisy dimensions more reliably than fixed aggregation schemes.

Learning objectives. The three objectives share the same risk notation as Eq. 5 and differ only in prior correction and non-negativity enforcement: *BCE* uses class-balanced binary cross-entropy with no prior correction (pos_weight $= n_{\text{unl}}/n_{\text{pos}}$), equivalent to running only the warm-up stage; *uPU* [18] applies the prior-corrected risk without the non-negativity clip ($\mathcal{L}^{(t)}_{\text{uPU}} = \pi_t R^+_{P_t} + R^-_{U_t}(w_{j,t}) - \pi_t R^-_{P_t}$); *nnPU* (default) adds $\max(0, \cdot)$ to prevent negative risk estimates (Eq. 5). nnPU achieves the best macro-average Adjusted F1 (11.1), followed by uPU (10.5) and class-balanced BCE (9.0) (**Table 10**). Class-balanced BCE (9.0) already surpasses every baseline in Table 1 (best: ResNet 8.1), confirming that the TargetSage architecture—multi-modal fusion and reasoning embeddings—contributes independently of the PU objective; nnPU adds a further 2.1 points by eliminating residual false-negative bias that persists even under class rebalancing, with the largest relative gains on T1 Cancer (24.2 vs. 16.2) and Pan-Cancer (5.5 vs. 3.7) where rare positives are most vulnerable. uPU (10.5) matches or outperforms nnPU on tasks with moderate positive rates—Cancer-Relevant (14.1 vs. 14.0), Type-Specific, SM (Clin+), and Ab (Clin+)—but its risk estimator turns negative under very low positive prevalence, degrading Pan-Cancer and T1–T3 Cancer below nnPU; the non-negativity correction in nnPU stabilises these tail tasks and yields the highest macro-average overall.

Attribute extraction strategy. We compared three attribute configurations: a fixed-vocabulary baseline (Static, 12-dim; a manually curated set covering established drug discovery assessment axes—druggability, target class membership (kinase, GPCR, ion channel), disease association, clinical evidence, pathway centrality, and safety profile—scored by Qwen2.5-1.5B-Instruct via direct prompting without reasoning chain or vocabulary discovery), open-ended agent extraction (Agent, 23-dim; vocabulary discovered by the evidence-gathering agent across all genes, after deduplication), and Agent+GRPO (**Table 10**). Agent (10.5) over Static (10.2) confirms that open-ended attribute discovery surfaces therapeutically relevant dimensions beyond the fixed vocabulary, with the largest gains on cancer-specific tasks where mechanistic diversity matters most. GRPO optimization achieves the highest macro-avg of 11.1, with the largest absolute gains on cancer-specific tasks (T1 Cancer: 24.2 vs 22.8 for Agent) where reward-guided refinement best aligns attribute extraction with the downstream prediction objective.

Table 10: Component ablations: **(a)** fusion operator (nnPU loss, full modalities), **(b)** learning objective (gated fusion, full modalities), **(c)** attribute extraction strategy (gated fusion, nnPU loss). Adjusted F1. **Bold** = best per row.

(a) Fusion operator

Task	Attn	Concat	Sum	Gated
Clinical Targets	9.8	**11.3**	10.7	11.2
Clin. & Chem.	3.0	**3.3**	3.1	3.2
Top-Tier	5.7	6.1	6.0	**6.5**
High-Conf.	3.6	3.9	3.8	**4.2**
Cancer-Relev.	11.8	**14.1**	13.5	14.0
Type-Specific	4.5	4.9	4.9	**5.4**
Pan-Cancer	4.8	5.1	**5.6**	5.5
T1 Cancer	21.1	22.6	**24.4**	24.2
T1–T2 Cancer	6.8	7.2	7.3	**7.8**
T1–T3 Cancer	4.7	5.0	5.0	**5.4**
SM (Appr.)	11.4	**13.3**	12.5	13.2
SM (Clin+)	7.8	8.5	8.4	**9.1**
Ab (Appr.)	19.8	**22.7**	21.6	22.6
Ab (Clin+)	15.2	**17.6**	16.5	17.5
PROTAC	16.1	15.4	**17.3**	17.2
Macro	9.7	10.7	10.7	**11.1**

(b) Learning objective

Task	BCE	uPU	nnPU
Clinical Targets	10.3	9.5	**11.2**
Clin. & Chem.	3.2	2.9	**3.2**
Top-Tier	5.9	5.2	**6.5**
High-Conf.	**4.4**	4.0	4.2
Cancer-Relev.	10.1	**14.1**	14.0
Type-Specific	3.7	**5.5**	5.4
Pan-Cancer	3.7	4.2	**5.5**
T1 Cancer	16.2	23.6	**24.2**
T1–T2 Cancer	6.5	6.8	**7.8**
T1–T3 Cancer	4.9	4.9	**5.4**
SM (Appr.)	12.4	12.7	**13.2**
SM (Clin+)	8.5	**9.2**	9.1
Ab (Appr.)	17.5	22.0	**22.6**
Ab (Clin+)	13.5	**18.0**	17.5
PROTAC	14.9	15.5	**17.2**
Macro	9.0	10.5	**11.1**

(c) Attribute extraction

Task	Static	Agent	+GRPO
Clinical Targets	10.3	10.6	**11.2**
Clin. & Chem.	2.9	3.0	**3.2**
Top-Tier	6.0	6.1	**6.5**
High-Conf.	3.8	4.0	**4.2**
Cancer-Relev.	12.8	13.2	**14.0**
Type-Specific	4.9	5.1	**5.4**
Pan-Cancer	5.0	5.2	**5.5**
T1 Cancer	22.2	22.8	**24.2**
T1–T2 Cancer	7.1	7.4	**7.8**
T1–T3 Cancer	4.9	5.1	**5.4**
SM (Appr.)	12.1	12.5	**13.2**
SM (Clin+)	8.3	8.6	**9.1**
Ab (Appr.)	20.7	21.3	**22.6**
Ab (Clin+)	16.0	16.5	**17.5**
PROTAC	15.8	16.2	**17.2**
Macro	10.2	10.5	**11.1**

I Metric Independence: AUROC, AUPRC and Early Retrieval

Adjusted F1 is both the evaluation metric and the GRPO reward, so we verify that the ranking of methods does not depend on it. Table 11 reports AUROC, AUPRC and precision@25 for all eleven methods, averaged over the 15 tasks (mean $\pm$ std over 5 seeds). TargetSage is best or tied-best on every metric: it ties the strongest baseline (RF) on AUROC and leads on both AUPRC and precision@25. Because it leads on three metrics that are *not* the training reward, its advantage is not an artifact of optimizing Adjusted F1.

We nonetheless keep Adjusted F1 as the primary metric because AUROC and AUPRC treat every unlabeled gene as a confirmed negative, which is precisely the assumption the PU setting rejects; Adjusted F1 corrects for latent positives in the unlabeled pool. It is the soft form of $r^2/\Pr(\hat{y}{=}1)$, the standard criterion for PU data [31, 14], and is computed only on the held-out test fold, whose positives are excluded from training, so a high score requires ranking unseen test positives highly.

Table 11: Metric sensitivity, averaged over the 15 tasks. Mean $\pm$ std over 5 seeds; bold = best per column.

Method	AUROC	AUPRC	P@25
LR	0.869 ± 0.011	0.446 ± 0.014	0.640 ± 0.021
SVM	0.915 ± 0.006	0.470 ± 0.011	0.659 ± 0.010
KNN	0.805 ± 0.003	0.348 ± 0.009	0.620 ± 0.013
NB	0.688 ± 0.005	0.116 ± 0.004	0.131 ± 0.014
RF	**0.946** ± 0.005	0.564 ± 0.004	0.732 ± 0.009
GB	0.939 ± 0.010	0.504 ± 0.010	0.624 ± 0.016
MLP	0.877 ± 0.005	0.488 ± 0.010	0.701 ± 0.017
ResNet	0.877 ± 0.007	0.428 ± 0.004	0.619 ± 0.008
TabNet	0.896 ± 0.011	0.423 ± 0.020	0.585 ± 0.024
FT-Trans	0.944 ± 0.004	0.499 ± 0.017	0.647 ± 0.017
TargetSage	**0.946** ± 0.005	**0.590** ± 0.011	**0.748** ± 0.011

External positives. On external positives never seen in training, TargetSage's top 5% is enriched 9.3–$18.9\times$ over the genome base rate for all five reference sets of Table 3 (hypergeometric $p < 10^{-80}$), recovering 94% of antibody-approved (65/69) and 72% of small-molecule-approved (557/777) targets. On the ChEMBL Phase II/III set TargetSage also leads every method on top-25 and top-50 precision (Table 12).

Calibration. As a prior-independent reliability check we ask whether higher scores correspond to a higher density of true positives, using hidden positives: the 90 genes upgraded to *Tclin* during 2021–2025, which sat in the unlabeled set the 2021 model was trained to score low. Their mean

Table 12: ChEMBL Phase II/III enrichment among understudied *Tbio* genes: odds ratio and precision at the top-25 and top-50 shortlists. NB assigns every gene the same score and KNN's top scores are heavily tied, so their shortlists depend on tie-breaking.

	LR	SVM	KNN	NB	RF	GB	MLP	ResNet	TabNet	FT-Tr.	**TargetSage**
OR@25	22.4	0.0	0.0	0.0	28.5	12.0	0.0	12.0	0.0	16.9	**43.2**
OR@50	14.8	5.6	3.6	3.6	23.2	7.7	1.8	7.7	3.6	20.2	**26.4**
P@25	0.20	0.00	0.00	0.00	0.24	0.12	0.00	0.12	0.00	0.16	**0.32**
P@50	0.14	0.06	0.04	0.04	0.20	0.08	0.02	0.08	0.04	0.18	**0.22**

predicted probability is about $25\times$ that of background genes, and the top 5% of the 2021-unlabeled pool captures 38% of them, a $7.6\times$ enrichment.

J Separating "More Information" from "Better Model"

TargetSage both retrieves more evidence than the baselines and processes it differently, so we isolate the two factors with four controls. The conclusion is that both contribute and neither alone explains the result.

Control 1: give every baseline the same evidence. Supplying all ten baselines with the identical enriched profile—the 482 structured features, the explicit reasoning attributes and the implicit reasoning embedding—raises almost every baseline (Table 13), confirming that the retrieved evidence carries real signal. TargetSage still leads at both levels: 10.7 versus 8.1 on structured features alone, and 11.1 versus 8.9 on the enriched profile.

Table 13: Evidence-matched comparison, macro-average Adjusted F1 over the 15 tasks. Every method sees the same features within a row.

Features	LR	SVM	KNN	NB	RF	GB	MLP	ResNet	TabNet	FT-Tr.	**TargetSage**
Structured (482)	6.0	3.3	3.7	1.6	3.0	7.6	6.0	8.1	5.1	6.7	**10.7**
+ reasoning attr. & emb.	7.6	5.0	4.4	1.7	2.6	7.9	7.1	8.9	6.4	7.8	**11.1**

Control 2: a direct LLM ranker. Prompting GPT-4o-mini to score genes directly reaches only 1.2 macro Adjusted F1, so the result does not follow from an LLM's prior knowledge of which genes are drug targets.

Control 3: match the data access, remove the reasoning. Encoding the same retrieved profiles with biomedical text encoders—BioBERT [68], SapBERT [69] and SciBERT [70]—with no reasoning trace and no generated attributes, then ranking with the same PU model, gives 10.78 ± 1.5 (SciBERT), 10.82 ± 1.1 (BioBERT) and 10.83 ± 1.7 (SapBERT), all below the full model's 11.1. With data access held fixed, the generated reasoning is what supplies the remainder of the gap.

Control 4: generic foundation-model representations. Used alone, pre-trained representations reach a macro Adjusted F1 of 4.8 for ESM-2 [15] protein embeddings and 0.7 for Geneformer [16] single-cell expression embeddings. These encoders describe a protein's biology rather than its target status, which depends on binding-site druggability, disease genetics and clinical precedent.

K Recovering Blinded Targets Without Drug Evidence

A prediction that merely echoes existing drug annotations would be useless for *undiscovered* targets. We test this directly by blinding well-established targets and asking whether intrinsic biology alone still ranks them highly.

Setup. For each blinded gene we remove every feature and textual mention that reveals established-target status—drug and clinical annotations (antibody/monoclonal counts, DGIdb, CTD), drug-related reasoning attributes, and the text embedding—replacing them with the genome-wide median or mean. We retain only the intrinsic biology that an undiscovered target would also possess: protein domains, network properties, population-genetic constraint and expression.

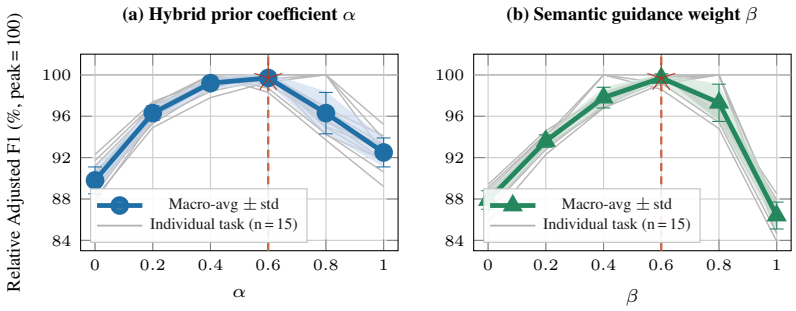

Figure 4: **Parameter sensitivity analysis.** Relative Adjusted F1 (%) for all 15 tasks (thin gray) and macro-average (colored, ±1 std across 5 seeds), normalized so that each task's peak equals 100%. (a) Hybrid prior coefficient α; (b) semantic guidance weight β. The dashed line and ★ mark the selected value (α=0.6, β=0.6). The macro-average peaks at 0.6 for both parameters; individual tasks peak between 0.4 and 0.8.

Result. Re-ranking all 19,032 genes by predicted *Tclin* score (5-seed ensemble), the blinded targets remain in the top 1.5–3.3%: ERBB2 25 → 325, EGFR 13 → 320, ESR1 5 → 463, AR 21 → 629, ERBB3 201 → 358, ABL1 137 → 288. The ranks do fall, so the drug and text evidence genuinely contributes; but intrinsic biology alone still places these genes in the top few percent, which is the property required to surface targets that carry no drug annotation yet.

Can the LLM reconstruct the external labels? The five reference sets of Table 3 and the ChEMBL Phase II/III set are never used in training or as the GRPO reward. To rule out indirect or parametric knowledge, we rank every gene by the therapeutic-potential score obtained by directly prompting GPT-4o-mini. Its top-5% overlap with the reference sets is on par with a random ranking and an order of magnitude below TargetSage (Table 14); on ChEMBL Phase II/III it reaches OR@25 ≈ 3.6 and OR@50 ≈ 1.8, against 43.2 and 26.4 for TargetSage.

Table 14: Top-5% overlap with the five reference sets of Table 3.

	Appr.	MAB-B1	MAB-B1-3	SM-B1	SM-B1-3
LLM only (direct prompt)	25	4	15	47	67
Random (expected)	23	3	14	39	55
TargetSage	**378**	**65**	**129**	**557**	**636**

L Statistical Significance of the Benchmark Lead

On the smallest tasks (∼14 test positives) the seed-to-seed spread can exceed the gap between methods, so we establish the lead at the aggregate level rather than task by task.

The variance is intrinsic to the label count, not specific to TargetSage. On Antibody-Approved the strongest methods (GB, MLP, ResNet, FT-Transformer and TargetSage) have stds of 7–13 points. Measured relative to the mean, TargetSage's spread is typical of the field: its coefficient of variation (std/mean ≈ 0.6) is comparable to ResNet (0.56), GB (0.60) and MLP (0.79).

The aggregate lead is significant. Against each of the ten baselines individually, TargetSage wins all 15 of 15 tasks (exact sign test $p = 3.05 \times 10^{-5}$; Wilcoxon signed-rank $p \leq 3.3 \times 10^{-4}$).

M Parameter Sensitivity

The hybrid prior coefficient α weights the data-driven prior $\pi_{t,\mathrm{data}}$ against the LLM-informed prior $\pi_{t,\mathrm{LLM}}$ (Eq. 3). As shown in **Fig. 4(a)**, performance is broadly stable across $\alpha \in [0.2, 0.8]$ and peaks at $\alpha \approx 0.6$ (macro-avg 11.1). All 15 tasks follow the same inverted-U trend, with relative Adjusted F1 degrading at both extremes: by about 10% at $\alpha = 0$ (purely LLM-informed) and by about 7%

at $\alpha = 1$ (purely data-driven). This confirms that neither source alone is sufficient — the hybrid blending is essential.

The semantic guidance weight β modulates how strongly the nnPU loss is regularized by LLM-derived class membership scores. **Fig. 4(b)** shows a clearer inverted-U shape peaking at $\beta = 0.6$ (macro-avg 11.1). The curve is asymmetric: performance falls more sharply toward $\beta = 1.0$ (-13.3% relative) than toward $\beta = 0$ (-11.8% relative). The steeper right-side drop reflects the self-confirmation trap in which the discriminator reinforces its own high-confidence errors on the unlabeled set. Tasks with rare positives (Ab (Approved), T1 Cancer) are most affected by this pathology, but 14 of the 15 tasks share the same asymmetric shape.

N Example Reasoning Chains from the RL-Optimized Module

We show verbatim reasoning chains produced by the GRPO-optimized reasoning module on eight therapeutically diverse genes at inference temperature 0.5. Each output contains an evidence-grounded reasoning chain followed by a structured attribute JSON; we show the reasoning and the top-ranked attributes below. All outputs are generated zero-shot at inference time on the trained RL policy.

TP53 (tumor suppressor).

> *"TP53 regulates cell cycle arrest, apoptosis, senescence, DNA repair, or changes in metabolism under stress conditions."*
> *Top attrs: disease association=0.95, gwas associations=0.93, protein protein interactions=0.91, loss of function tolerance=0.08, essential gene=0.88, overall therapeutic potential=0.62*

KRAS (oncogene, small molecule target).

> *"KRAS is a critical driver of lung cancer progression, linking Ras signaling to downstream pathways involved in cell proliferation and survival."*
> *Top attrs: disease association=0.96, small molecule=0.88, chemical gene interactions=0.87, gwas associations=0.85, overall therapeutic potential=0.90, drug interaction types=0.82*

HMGCR (metabolic enzyme, statin target).

> *"HMGCR is an important enzyme linking cholesterol biosynthesis to cell signaling pathways, but its role can be altered by environmental factors such as diet."*
> *Top attrs: small molecule=0.97, drug interaction types=0.95, chemical gene interactions=0.93, disease association=0.90, overall therapeutic potential=0.92, gwas associations=0.74*

CD19 (B-cell antigen, CAR-T target).

> *"CD19 is a B-lymphocyte co-receptor that modulates B-cell receptor signaling and promotes B-cell activation and proliferation; its selective expression on B-lineage cells, including malignant B cells in leukemia and lymphoma, makes it a prime target for antibody-based and CAR-T immunotherapy."*
> *Top attrs: antibody=0.97, antibody availability=0.95, tractable modalities=0.93, disease association=0.92, expression tissue specificity=0.88*

EGFR (receptor tyrosine kinase).

> *"EGFR is a receptor tyrosine kinase involved in signaling pathways critical for cellular proliferation and survival."*
> *Top attrs: small molecule=0.96, chemical gene interactions=0.94, drug interaction types=0.93, disease association=0.95, overall therapeutic potential=0.94, gwas associations=0.88*

ERBB2/HER2 (kinase, dual modality target).

> *"ERBB2 is a receptor tyrosine kinase that can form heterodimers with other EGFR family members, potentially affecting downstream signaling pathways."*
> *Top attrs: small molecule=0.96, antibody=0.94, tractable modalities=0.97, drug interaction types=0.92, overall therapeutic potential=0.95, disease association=0.90*

MYC (oncogene, transcription factor).

> *"MYC is a key oncogene involved in cell proliferation, apoptosis, and transformation. It interacts with other proteins and regulates downstream pathways."*
> *Top attrs: disease association=0.96, protein protein interactions=0.93, gwas associations=0.88, small molecule=0.22, overall therapeutic potential=0.65, essential gene=0.85*

IL3RA/CD123 (cytokine receptor, AML target).

> *"IL3RA is involved in hematopoiesis, but its role may vary between tissues and cell types."*
> *Top attrs: antibody=0.91, antibody availability=0.88, disease association=0.90, tractable modalities=0.85, expression tissue specificity=0.82*

Observations. These chains demonstrate two properties of the GRPO-optimized reasoning module. First, the reasoning is *gene-specific and biologically grounded*: TP53 receives high *gwas associations* and *protein protein interactions* consistent with its role as a central stress-response hub, while HMGCR receives the highest *small molecule* and *drug interaction types* scores consistent with statins being among the most widely prescribed drugs—none of these distinctions are derivable from the gene name alone. Second, the attribute profiles discriminate therapeutic modality: EGFR and ERBB2 are assigned high *small molecule* scores reflecting their established small-molecule inhibitor landscape, CD19 and IL3RA receive high *antibody*/*antibody availability* scores consistent with antibody-based therapies, and the notoriously undruggable transcription factor MYC is correctly assigned a low *small molecule* score (0.22) despite its high *disease association*. The reasoning module provides interpretable justifications that domain experts can inspect and challenge—fulfilling the auditability claim of contribution 1.

O Case Studies

ERBB3/HER3 (rank 201 of 19,032 on Clinical Targets, 5-seed ensemble; top 1.1%) is prioritized by the model as a therapeutically relevant receptor with strong antibody-oriented tractability. Specifically, TargetSage assigned high scores for *antibody* (0.94), *antibody availability* (0.91), *tractable modalities* (0.90), *disease association* (0.88), and *protein protein interactions* (0.85), while correctly assigning low *small molecule* (0.12) consistent with its weak intrinsic kinase activity and limited suitability for direct catalytic inhibition. The LLM reasoning trace shows:

> *"ERBB3 is unique in that it lacks a functional kinase domain, which means it cannot directly initiate signaling through phosphorylation. Instead, it acts primarily as a co-receptor that forms heterodimers with other EGFR family members that possess kinase activity, such as ERBB2. This heterodimerization is crucial for the activation of downstream signaling pathways, including the PI3K/AKT and MAPK pathways [. . .] Overexpression or mutations in ERBB3 have been linked to resistance to therapies targeting other members of the EGFR family, particularly in HER2-positive breast cancer." — TargetSage reasoning trace (Module 2)*

The model correctly identified that ERBB3's therapeutic value derives from its role as a heterodimerization partner—a distinction purely structure-based predictors would miss—demonstrating both the transparency of auditable LLM reasoning traces and the ability of gated multi-modal fusion to integrate complementary signals from structured features and reasoning representations.

P Computational Cost

Table 15 summarizes computational cost per component. M1 issues one GPT-4o-mini [37] API call per gene per tool (6 tools $\times$ 19,032 genes $=$ 114,192 calls total), batched at up to 50 concurrent requests; the $\approx$\$15 total cost makes routine re-runs feasible as databases are updated. M2 fine-tunes Qwen2.5-1.5B with LoRA (r=32, α=64, dropout 0.05) via GRPO for 100 gradient steps with G=4 rollouts per gene; peak GPU memory is 18 GB on a single RTX 3090, and inference over all 19,032 genes adds approximately 1 h. M3 trains the gated fusion model (10 warm-up and 25 nnPU epochs; Appendix B) across 15 tasks and 5 random seeds, completing in $\sim$3 h; classical ML baselines (LR, SVM, RF, GB, etc.) run on CPU in under 1 h per task. The bottleneck for a full TargetSage run is M1 throughput, which is bounded by API rate limits rather than local compute.

Table 15: **Computational cost per component.** GPU: NVIDIA RTX 3090 (24 GB). Wall-clock is for one full run over all 15 tasks. API cost uses Azure OpenAI GPT-4o-mini pricing.

Component	Hardware	GPU Mem.	Wall-clock	Notes
Baselines				
LR/SVM/KNN/NB/RF/GB/MLP	CPU	—	$<$1 h	per task
ResNet/TabNet/FT-Transformer	RTX 3090	$<$4 GB	1–3 h	per task, 5 seeds
TargetSage				
M1: Agentic Profiling	API	—	$\sim$12 h	$\approx$\$15 total
M2: Guided Reasoning	RTX 3090	18 GB	$\sim$8 h	Qwen2.5-1.5B+LoRA
M3: PU-Aware Scoring	RTX 3090	$<$6 GB	$\sim$3 h	15 tasks $\times$ 5 seeds
Total (TargetSage)	RTX 3090	18 GB	$\sim$23 h	excl. API concurrency

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract and introduction clearly state the four contributions (the TargetSage framework, a standardized positive-unlabeled benchmark, performance against diverse baselines, and temporal and independent validation) and the scope (therapeutic target identification across 15 benchmark tasks). All claims are supported by experiments in Section 4.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: Limitations are discussed in Section 5 (Discussion and Conclusion), including the dependence of LLM-derived profiles on literature coverage, which may disadvantage understudied *Tdark* proteins, and the independent scoring of genes without modeling network topology, co-expression, or pathway co-regulation.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [N/A]

   Justification: The paper does not include theoretical proofs. The nnPU loss formulation follows Kiryo et al. [30] and is cited appropriately.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: Full implementation details (hyperparameters, data splits, training procedure, hardware) are provided in Appendix B. GRPO hyperparameters are in Appendix F. Code is available at `https://github.com/aicb-ZhangLabs/TargetSage`.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [Yes]

   Justification: Code and benchmark data are available at `https://github.com/aicb-ZhangLabs/TargetSage`. All biological feature data are derived from publicly available databases (PHAROS, Open Targets, ExAC, gnomAD, STRING, CTDbase, InterPro).

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [Yes]

   Justification: Training details (AdamW optimizer, learning rate, batch size, dropout, 5-seed stratified 80/20 splits) are reported in Appendix B. Baseline configurations are described in Section 4.1 with a pointer to Appendix B.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes]

   Justification: Table 1 reports mean $\pm$ std over 5 random seeds with stratified splits. One-sided Wilcoxon signed-rank tests over the 15 tasks confirm statistical significance of TargetSage over each baseline individually (all $p < 0.001$).

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: Computational costs for all components (baseline models, TargetSage fusion, GRPO training) are reported in Appendix P, including GPU type (NVIDIA RTX 3090), memory, and wall-clock time.

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: This work uses only publicly available genomic and pharmacological databases. No human subjects are involved. The research aims to accelerate therapeutic target discovery, with no foreseeable dual-use risks.

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

    Justification: The work aims to help prioritize therapeutic targets for experimental study. Section 5 notes that the rankings are hypotheses rather than substitutes for experimental validation, which addresses the risk of over-reliance on computational predictions. No significant negative societal impacts are anticipated.

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [N/A]

    Justification: The work does not release pre-trained generative models or scraped datasets. All data sources are publicly available pharmacological databases with established usage policies.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

Justification: All databases used (PHAROS, Open Targets, ExAC, gnomAD, STRING, CTDbase, InterPro) are cited with their original publications. Third-party tools (TabNet, FT-Transformer) are cited. All are publicly available under open-access licenses.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [Yes]

    Justification: The 15-task benchmark suite is described in detail in Appendix A with task definitions, label sources, and positive rates. The released code at `https://github.com/aicb-ZhangLabs/TargetSage` includes the benchmark labels.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: This paper does not involve crowdsourcing or research with human subjects.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: This paper does not involve human subjects research. IRB approval is not applicable.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research?

    Answer: [Yes]

    Justification: LLM usage is a core and novel component of TargetSage. The paper describes LLM-based semanticization and agentic evidence collection in Module 1, a Qwen2.5-1.5B-Instruct reasoning policy optimized with GRPO in Module 2, re-encoding of reasoning traces into implicit embeddings, and GPT-4o for offline task-specific prior estimation. These components are described in Section 3 and Appendix C, Appendix F, and Appendix E.