AYASS BIOSCIENCE

BiRAGAS

The Bioinformatics Help Companion

An Illustrated Q&A · 51 Questions · 16 Themes

How the BiRAGAS agentic causal-AI platform answers the real questions bioinformaticians, students, senior scientists and drug-discovery teams ask every day — from raw FASTQ files to validated, FDA-grade causal targets.

Ayass Bioscience · biragas@ayassbioscience.com

BiRAGAS causal platform

CONTENTS

What's inside

Each question appears with a plain-language answer and a matching illustration. Q# refers to the master question numbering. Look for the theme band at the top of every page.

THEME

‹ CONTENTS

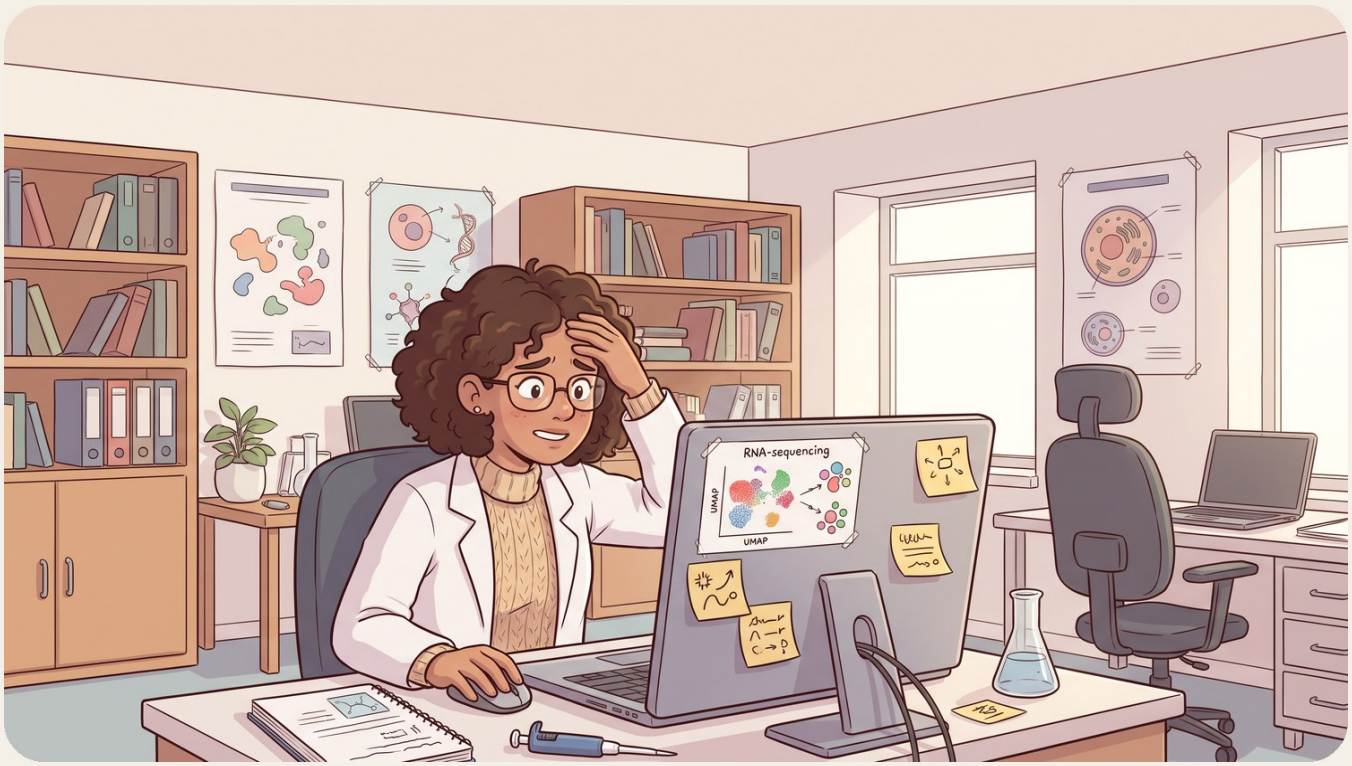
01

Getting started & foundations

Questions Q1–Q4 · 4 in this theme

GETTING STARTED & FOUNDATIONS

Q1 I'm new to RNA-seq. Where do I even start with a raw dataset and what does a full analysis look like end-to-end?



HOW BIRAGAS HELPS

BiRAGAS gives you the whole pipeline as one guided chain: pull a public cohort, QC/align raw reads, quantify, run DEG, enrich pathways, prioritize genes, then (if you want) go all the way to a causal target dossier. You don't stitch tools together - open a module to see exactly what inputs it needs and what plots it produces, and ask Causion in plain English what to run next.

AUDIENCE
Student / Trainee

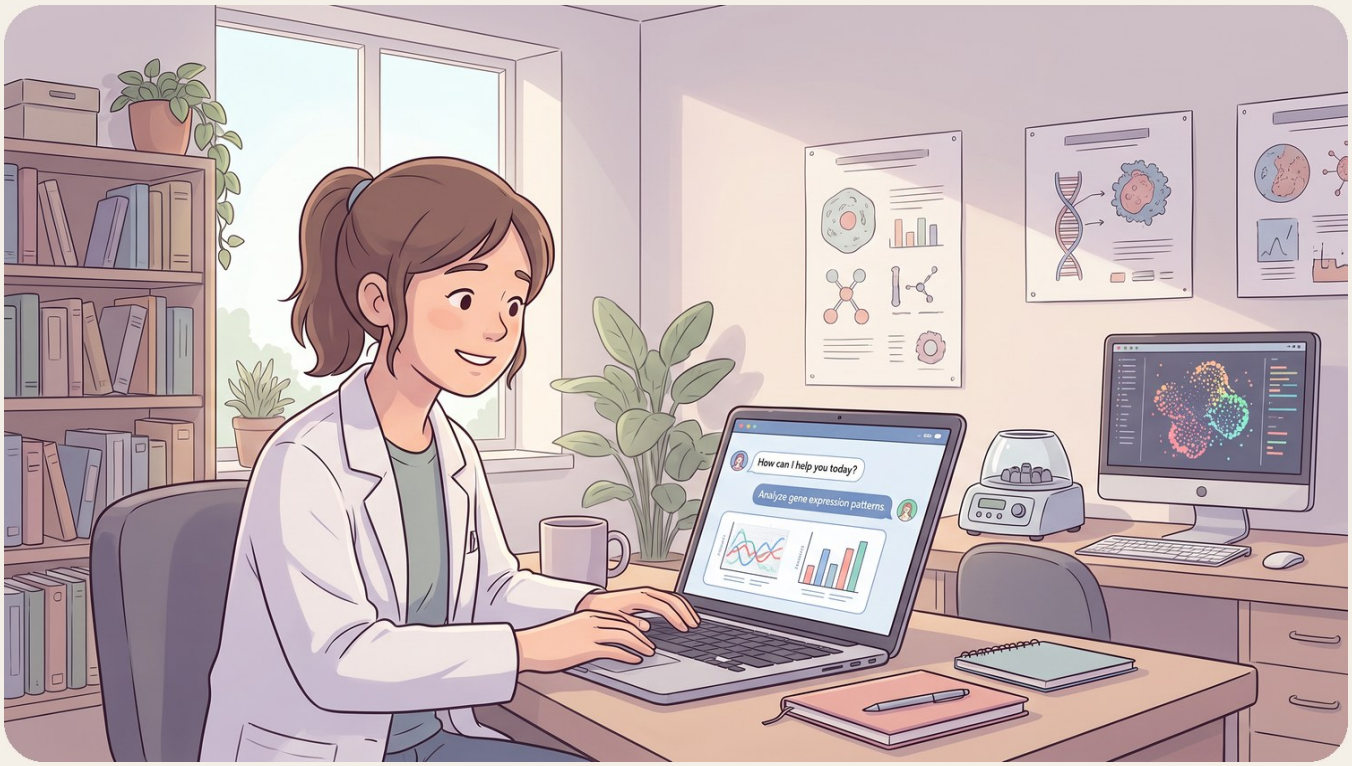
RUNS
Full chain

MODULES
Cohort retrieval; FASTQ processing; DEG analysis; Pathway enrichment

OUTPUT
Reports + figures at each step

GETTING STARTED & FOUNDATIONS

Q2 I don't know how to code or run command-line bioinformatics tools. Can I still analyze my PI's data?



HOW BIRAGAS HELPS

Yes. Causion, the natural-language co-pilot, sits in front of all 28 modules. You type 'Run DEG on these counts, case vs control' and it configures parameters, routes to the right module, and returns results with figures. No R/Python or pipeline scripting required - it lowers the barrier from 'trained bioinformatician' to any translational scientist.

AUDIENCE
Student / Trainee

RUNS
In-app

MODULES
Causion co-pilot (all modules)

OUTPUT
Guided results

GETTING STARTED & FOUNDATIONS

Q3 How do I get a public dataset (e.g. from GEO) into a usable form for analysis?

HOW BIRAGAS HELPS

Cohort retrieval pulls disease cohorts directly from GEO / ArrayExpress with tissue and experiment filters, and returns a cohort directory plus a clean sample-metadata sheet ready for downstream steps - no manual FTP downloads or reformatting.

AUDIENCE
Student / Trainee

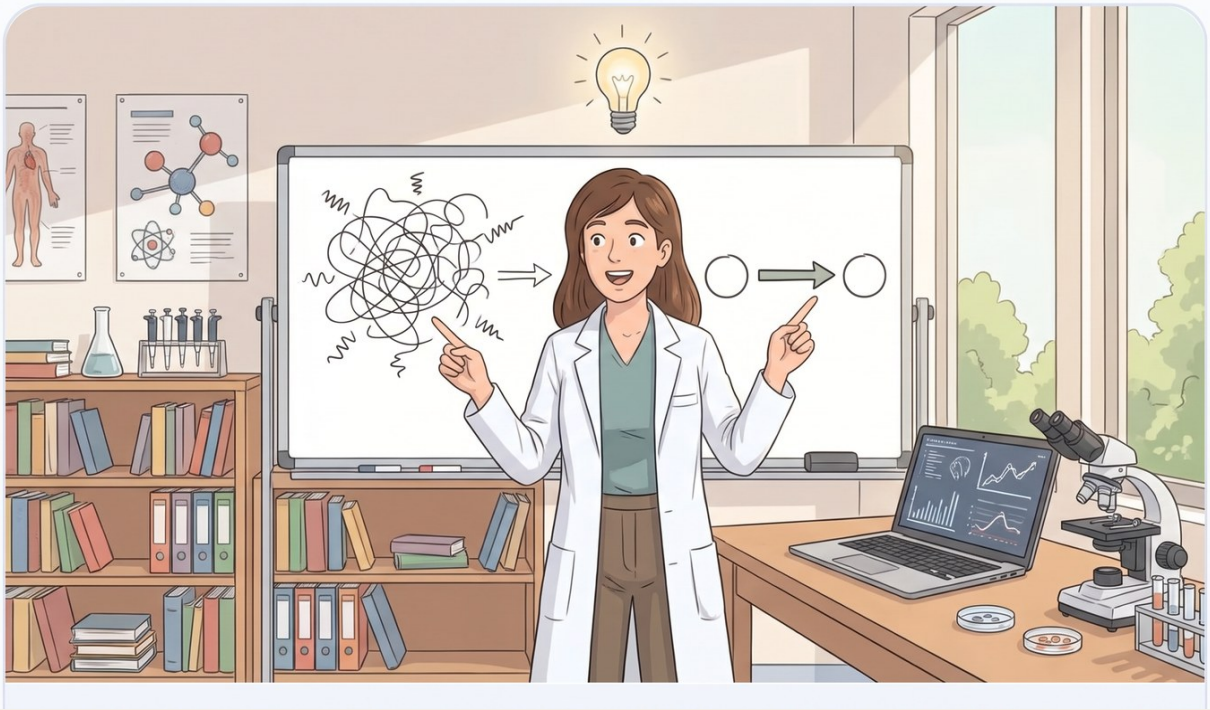
MODULES
Cohort retrieval

RUNS
Full engine

OUTPUT
cohort dir + sample metadata

GETTING STARTED & FOUNDATIONS

Q4 What's the difference between correlation and causation in my gene list, and why should I care?



HOW BIRAGAS HELPS

This is the core idea BiRAGAS is built on. Standard tools tell you Gene A and Gene B move together (correlation) - but not which one drives disease and which is a bystander. A drug aimed at a bystander fails. BiRAGAS adds causal layers (genetics, temporality, interventional simulation) and scores each finding as a Driver vs Candidate, so you act on causes, not passengers.

AUDIENCE
Student / Trainee

RUNS
Mixed

MODULES
Causal DAG builder; Causal calculus; GWAS / Mendelian randomization

OUTPUT
Causal Confidence Scores

THEME

‹ CONTENTS

02 Data ingestion & QC

Questions Q5–Q7 · 3 in this theme

DATA INGESTION & QC

Q5 I have raw FASTQ files (or only SRA accessions). How do I QC and align them reproducibly?

HOW BIRAGAS HELPS

FASTQ processing quality-controls and aligns raw reads, optionally converting SRA first, and emits alignment + QC metrics as a versioned output - so the ingestion step is standardized and auditable rather than an ad-hoc script per project.

AUDIENCE
Bioinformatician

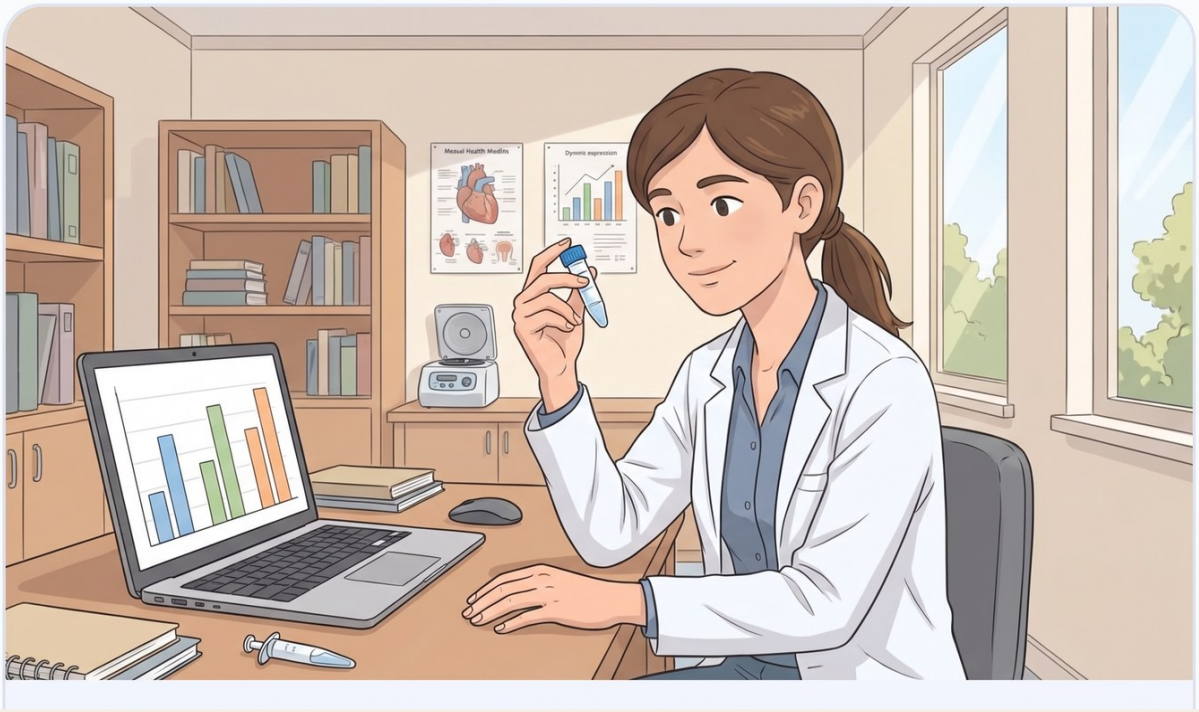
MODULES
FASTQ processing

RUNS
Full engine

OUTPUT
aligned output + QC

DATA INGESTION & QC

Q6 How do I quantify total-RNA expression with a proper case/control design?



HOW BIRAGAS HELPS

Total RNA pipeline ingests and quantifies total-RNA (TPM/abundance) with the case/control design attached, producing a quantification directory that flows straight into DEG and harmonization without re-mapping sample labels.

AUDIENCE
Bioinformatician

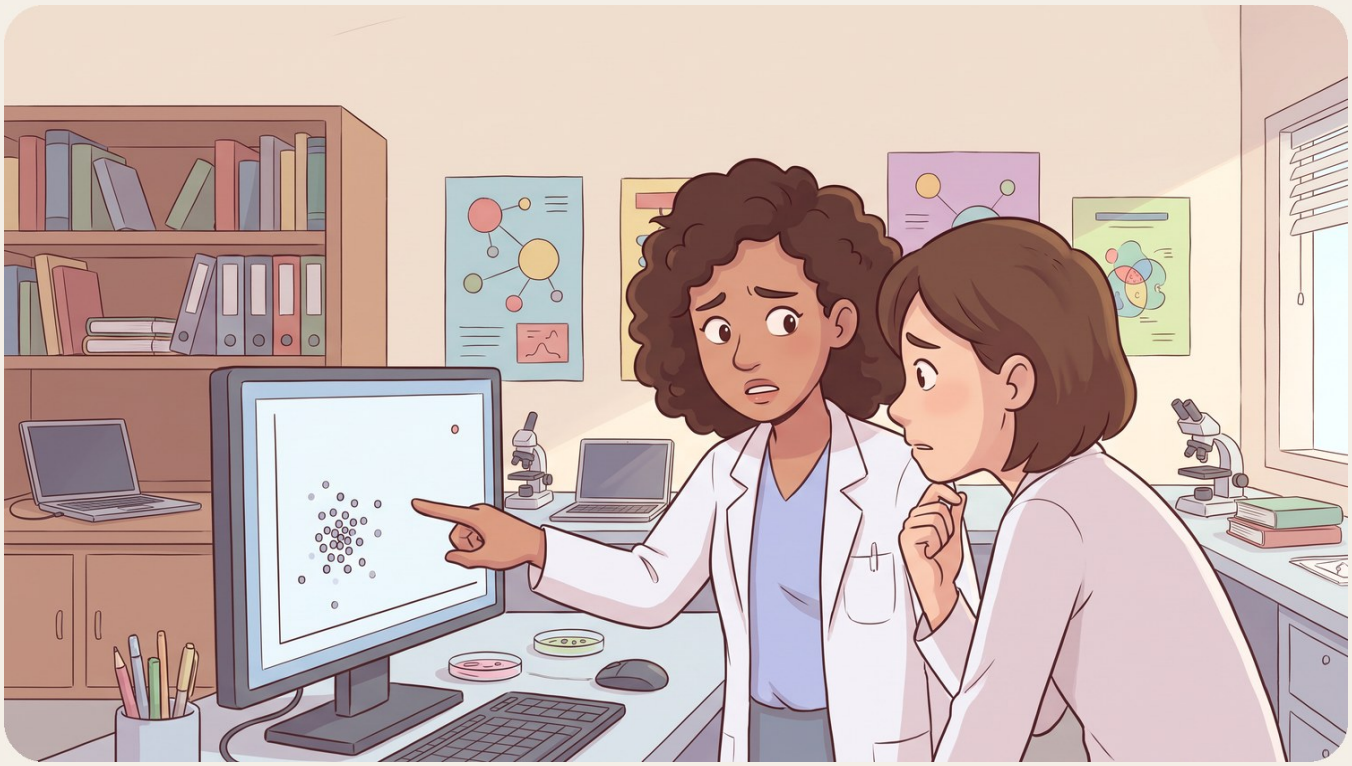
MODULES
Total RNA pipeline

RUNS
Full engine

OUTPUT
quantification dir

DATA INGESTION & QC

Q7 My QC shows outlier samples and low-quality libraries. How does the platform flag them before I waste time on DEG?



HOW BIRAGAS HELPS

QC + alignment metrics are produced at ingestion, and harmonization diagnostics (PCA/UMAP) surface outliers and library-size problems before differential testing - so you catch bad samples at the front of the chain, not after a misleading DEG run.

AUDIENCE
Bioinformatician

RUNS
Full engine

MODULES
FASTQ processing; Data harmonization

OUTPUT
QC metrics + PCA/UMAP diagnostics

THEME

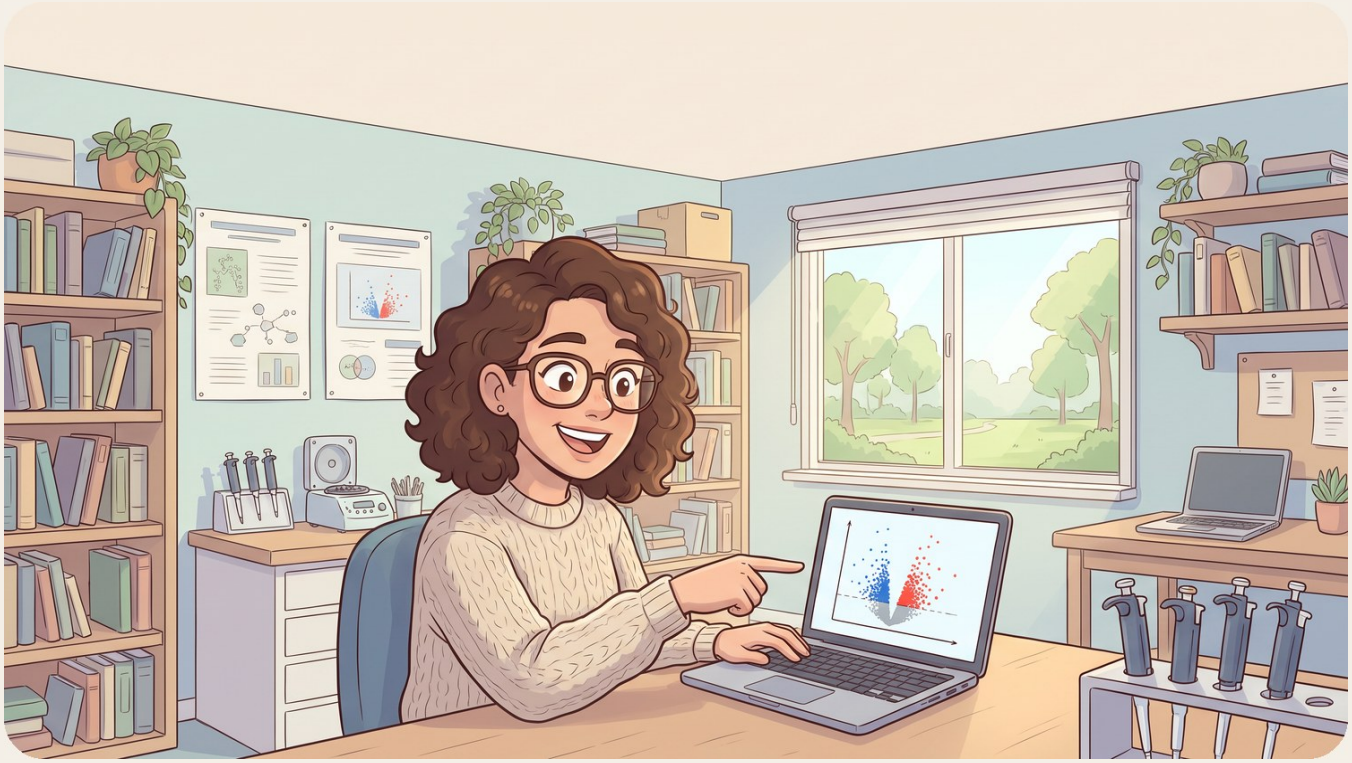
‹ CONTENTS

03 Differential expression (DEG)

Questions Q8–Q10 · 3 in this theme

DIFFERENTIAL EXPRESSION (DEG)

Q8 How do I run a correct differential-expression analysis without messing up the statistics?



HOW BIRAGAS HELPS

DEG analysis runs DESeq2 / edgeR / limma-voom on counts with proper normalization and FDR (multiple-testing) correction built in - so you get valid adjusted p-values and log2 fold-changes instead of an incorrect raw t-test on counts. Runs in-app with volcano, heatmap, MA and p-value distribution plots.

AUDIENCE
Student / Trainee

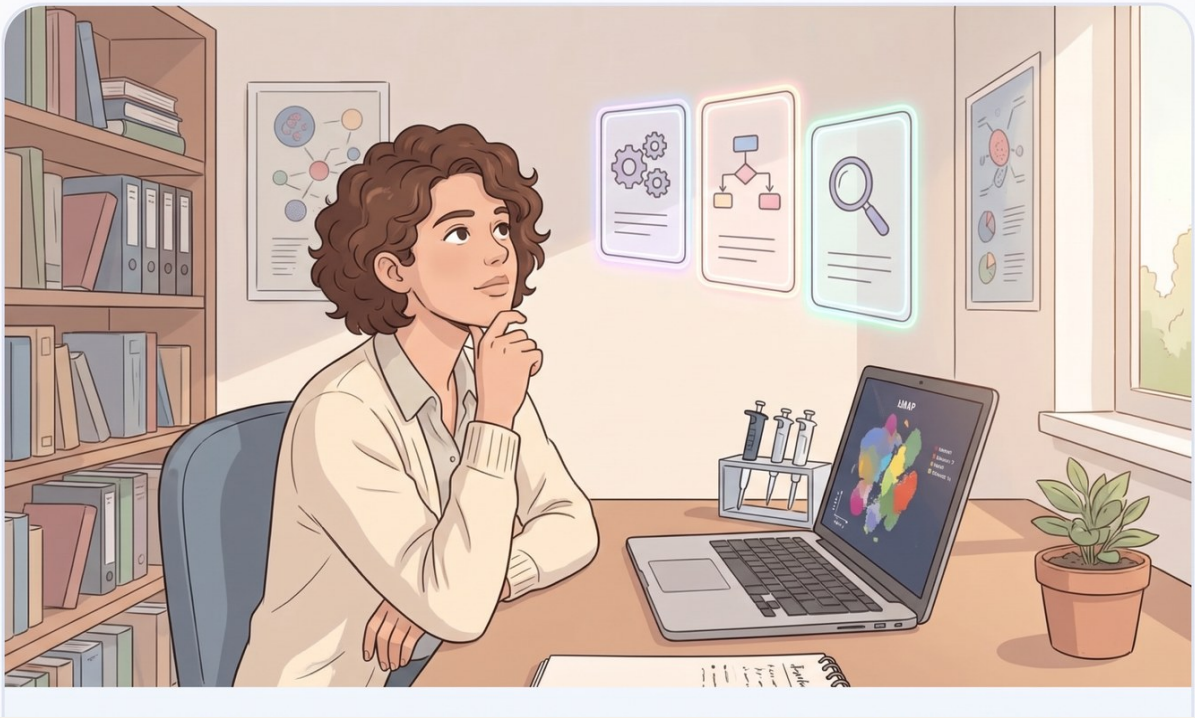
RUNS
In-app

MODULES
DEG analysis

OUTPUT
DEG base dir + volcano/heatmap/MA

DIFFERENTIAL EXPRESSION (DEG)

Q9 Which DE method should I use - DESeq2, edgeR, or limma - and can I compare them?



HOW BIRAGAS HELPS

The DEG module supports DESeq2, edgeR and limma so you can pick the method appropriate to your design (or cross-check concordance). All apply correct normalization and FDR, and report effect sizes (log2FC) alongside significance.

AUDIENCE
Bioinformatician

RUNS
In-app

MODULES
DEG analysis

OUTPUT
DEG tables

DIFFERENTIAL EXPRESSION (DEG)

Q10 My DEG list is hundreds of genes long and I can't tell which ones matter. How do I get a defensible short list?

HOW BIRAGAS HELPS

Gene prioritization composites log2FC, statistical significance and biological context into a single ranked score - collapsing a noisy DEG list into a defensible short list. That prioritized set then feeds enrichment, network and causal modules.

AUDIENCE
Senior Scientist

MODULES
DEG analysis; Gene prioritization

RUNS
In-app

OUTPUT
prioritized genes

THEME

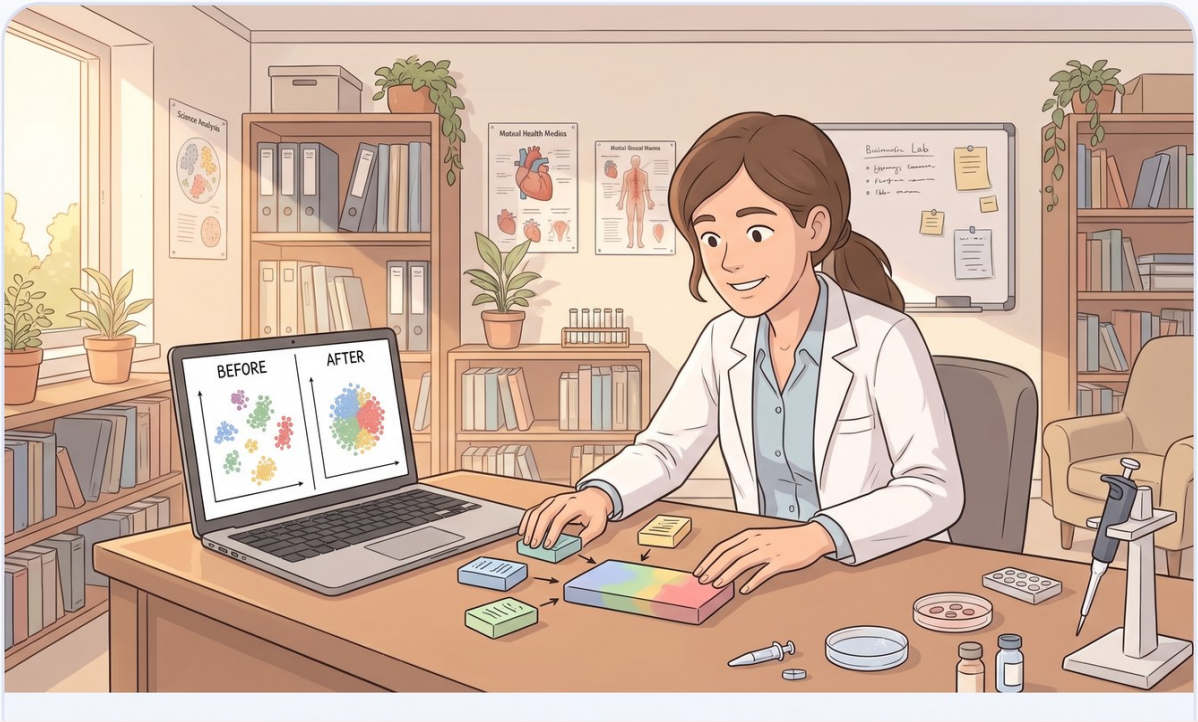
‹ CONTENTS

04 Batch effects & cross-dataset harmonization

Questions Q11–Q12 · 2 in this theme

BATCH EFFECTS & CROSS-DATASET HARMONIZATION

Q11 I want to pool several datasets but batch effects dominate the signal. How do I merge cohorts safely?



HOW BIRAGAS HELPS

Data harmonization performs batch-effect correction and cross-dataset harmonization, with PCA/UMAP diagnostics so you can verify batches actually mixed. This lets you combine cohorts to gain power without letting a technical batch masquerade as biology.

AUDIENCE
Bioinformatician

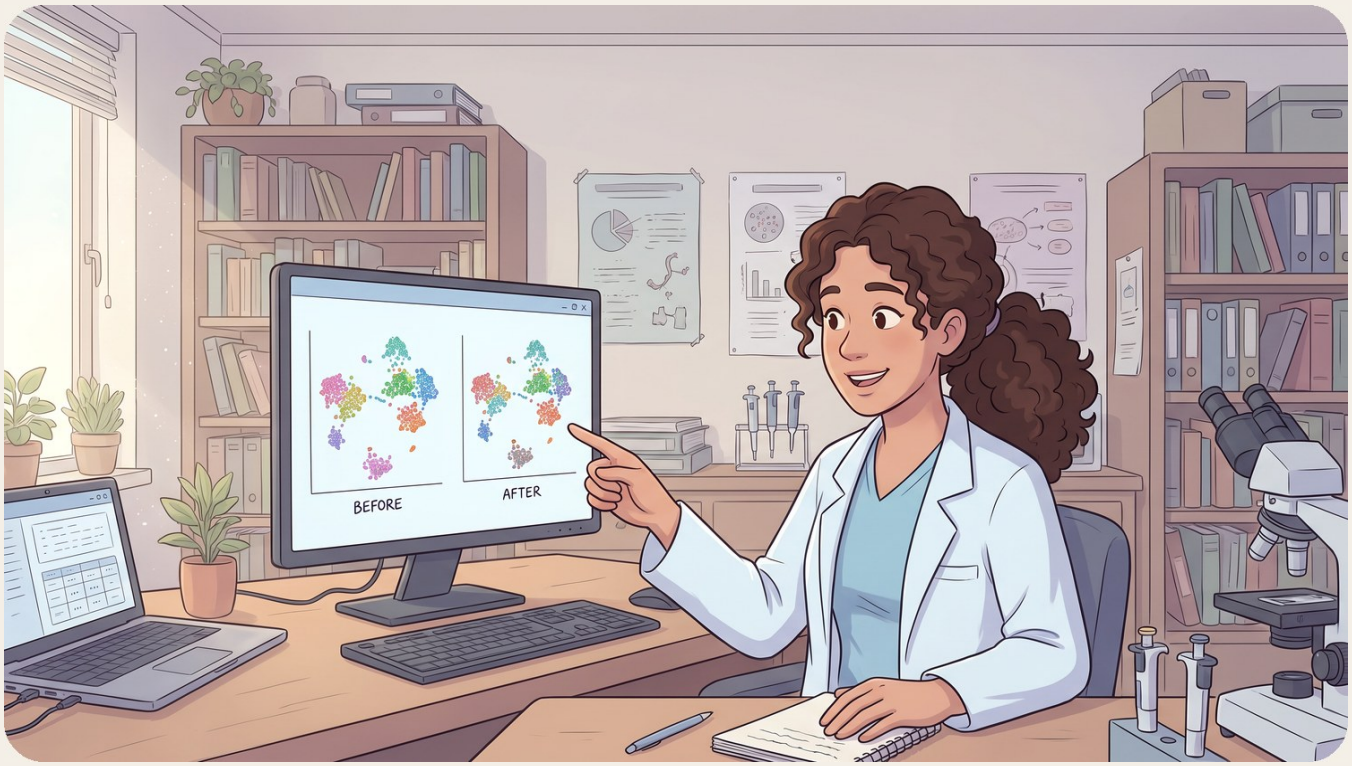
MODULES
Data harmonization

RUNS
Full engine

OUTPUT
harmonized matrix + PCA/UMAP

BATCH EFFECTS & CROSS-DATASET HARMONIZATION

Q12 How do I know harmonization didn't erase real biological differences along with the batch effect?



HOW BIRAGAS HELPS

The module returns PCA/UMAP before/after diagnostics so you can confirm samples separate by biology (condition) and not by batch after correction - making the correction auditable rather than a black box.

AUDIENCE
Senior Scientist

MODULES
Data harmonization

RUNS
Full engine

OUTPUT
diagnostic plots

THEME

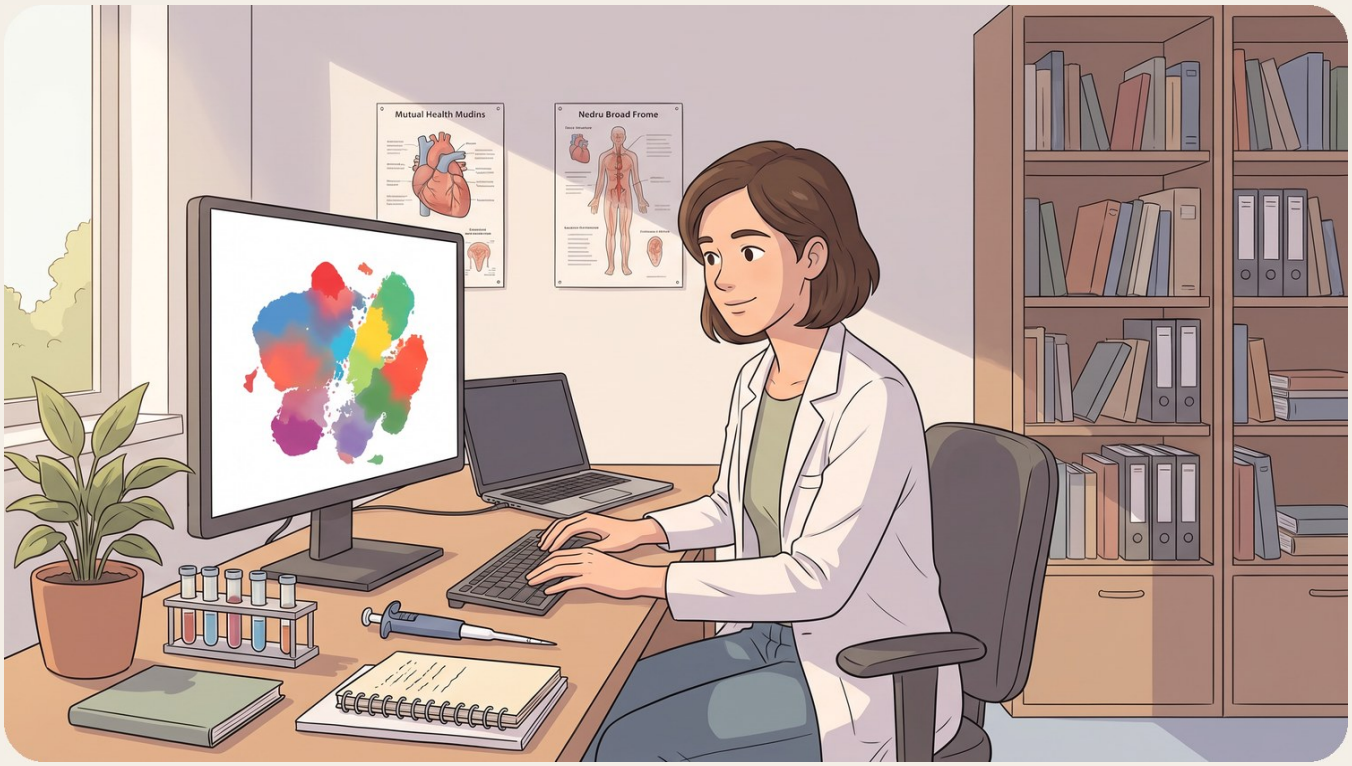
‹ CONTENTS

05 Single-cell & multi-omics

Questions Q13–Q15 · 3 in this theme

SINGLE-CELL & MULTI-OMICS

Q13 I have scRNA-seq data - how do I go from raw matrix to clusters, cell types, DE and trajectories?



HOW BIRAGAS HELPS

Single-cell analysis handles clustering, integration across samples, per-cluster differential expression, and trajectory / diffusion-pseudotime (DPT) inference in one module - the standard Scanpy/Seurat-style workflow without hand-assembling it.

AUDIENCE
Bioinformatician

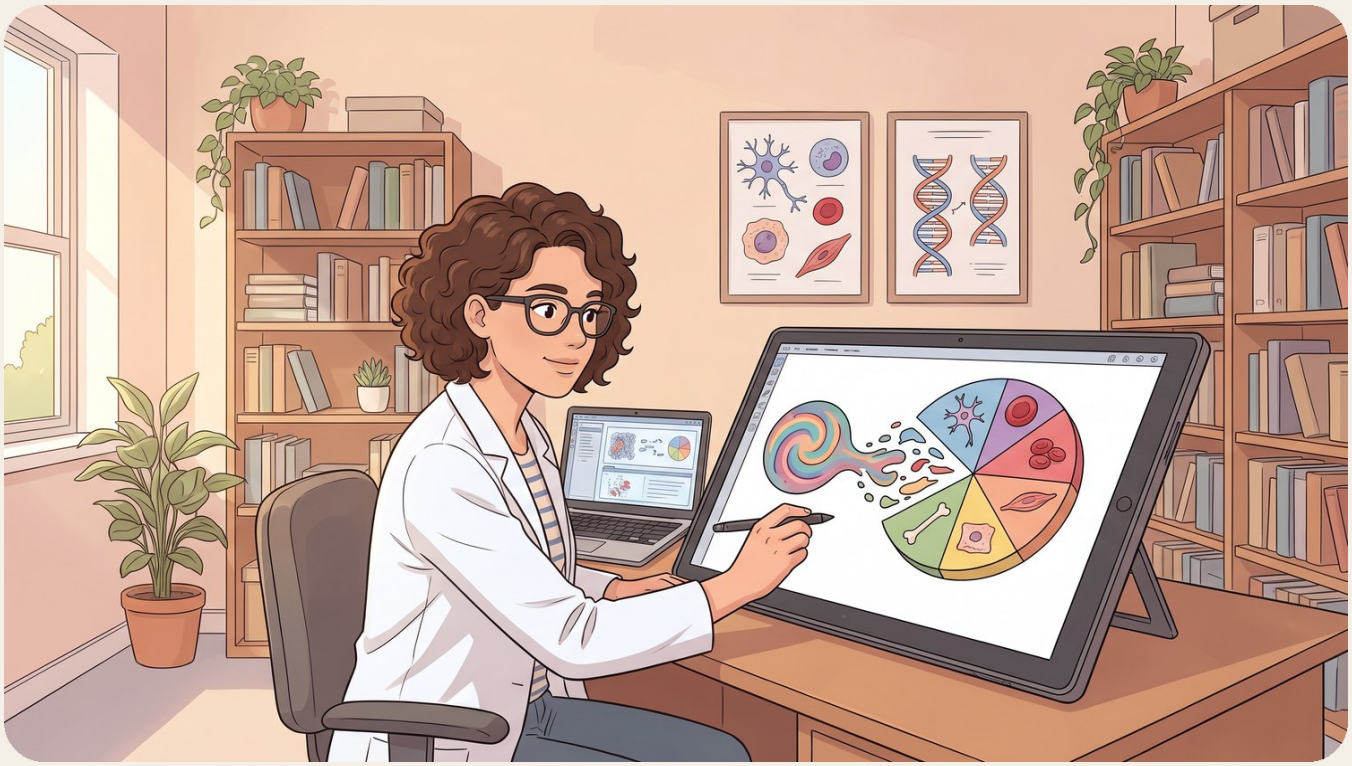
RUNS
In-app

MODULES
Single-cell analysis

OUTPUT
single-cell dir

SINGLE-CELL & MULTI-OMICS

Q14 I only have bulk RNA-seq but need cell-type proportions. Can I estimate composition without single-cell data?



HOW BIRAGAS HELPS

Deconvolution estimates cell-type abundances from bulk expression using CIBERSORTx / xCell / Bisque - so you recover composition shifts (e.g. immune infiltration) from bulk data you already have.

AUDIENCE
Bioinformatician

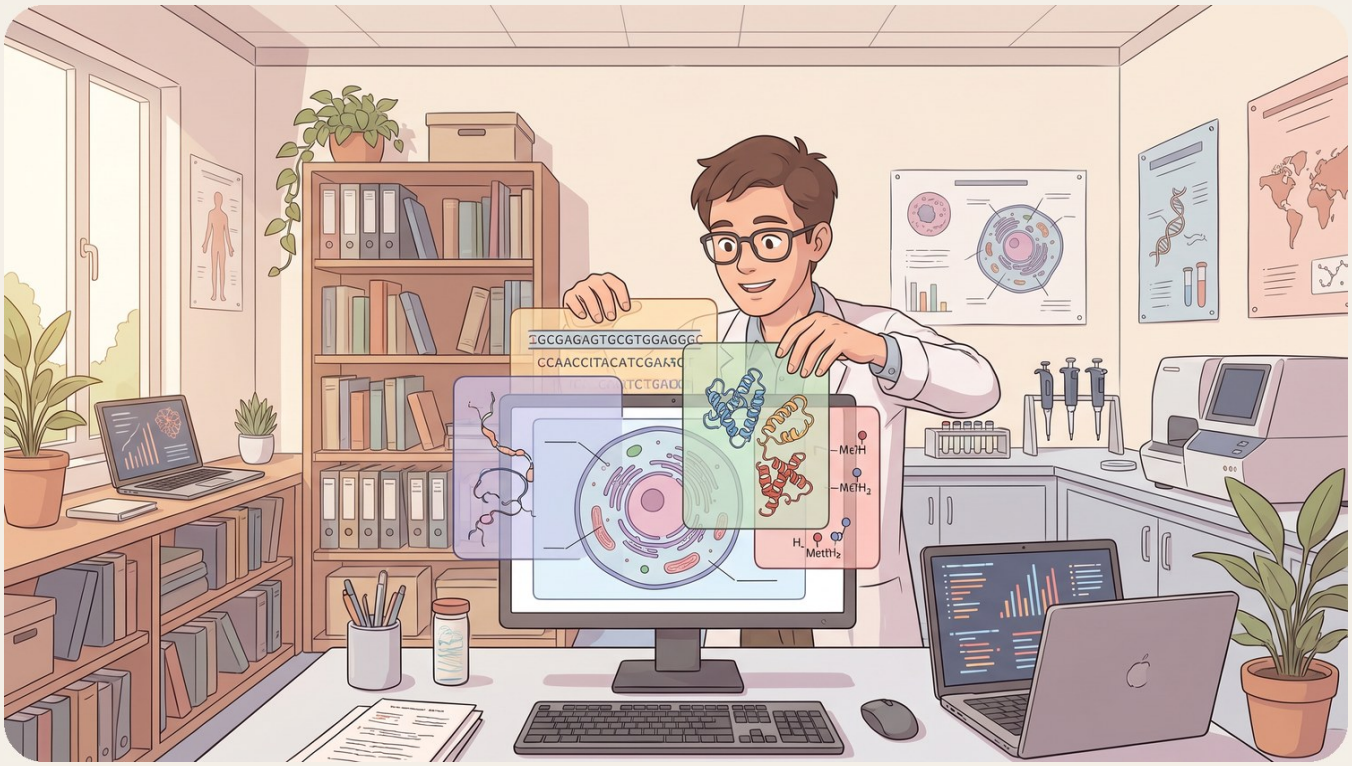
MODULES
Deconvolution

RUNS
In-app

OUTPUT
cell-type abundances

SINGLE-CELL & MULTI-OMICS

Q15 I have transcriptomics plus other omics layers. How do I integrate them instead of analyzing each alone?



HOW BIRAGAS HELPS

Multi-omics analysis integrates multiple omics layers into shared latent factors / cross-omics structure, so you find coordinated signals across layers rather than reconciling separate single-omics gene lists by hand.

AUDIENCE
Senior Scientist

RUNS
In-app

MODULES
Multi-omics analysis

OUTPUT
multi-omics factors

THEME

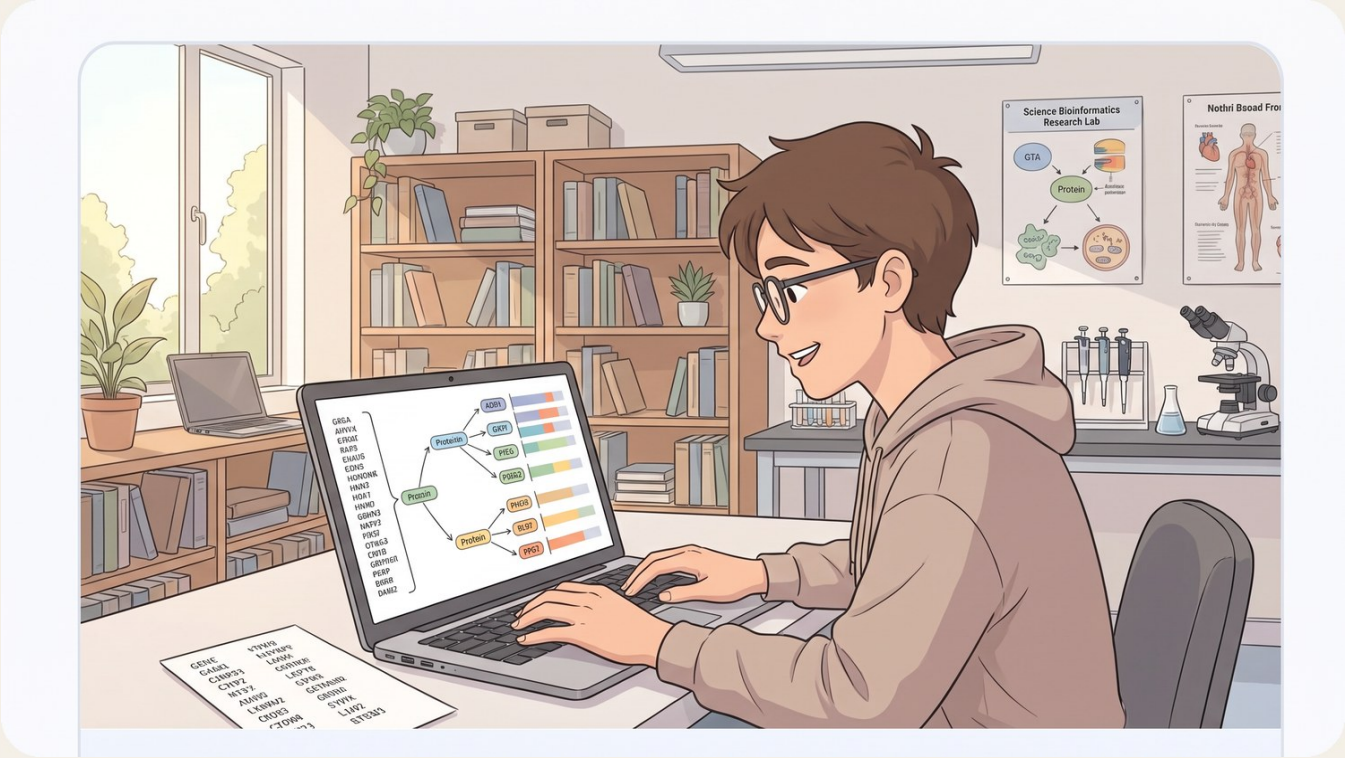
‹ CONTENTS

06 Functional enrichment & pathways

Questions Q16–Q18 · 3 in this theme

FUNCTIONAL ENRICHMENT & PATHWAYS

Q16 How do I turn my gene list into biological pathways/processes?



HOW BIRAGAS HELPS

Pathway enrichment runs overrepresentation analysis and consolidates redundant terms over your prioritized gene set; GSEA pathway enrichment runs ranked gene-set enrichment (NES) over the full DEG result so you capture coordinated shifts even without a hard significance cutoff.

AUDIENCE
Student / Trainee

RUNS
In-app

MODULES
Pathway enrichment; GSEA pathway enrichment

OUTPUT
pathway consolidation

FUNCTIONAL ENRICHMENT & PATHWAYS

Q17 ORA vs GSEA - when should I use which, and can I run both?

HOW BIRAGAS HELPS

Both are provided. Use overrepresentation (ORA / Pathway enrichment) on a defined significant set; use GSEA when you want to detect coordinated, subtle shifts across the ranked DEG list. Running both and comparing consolidated pathways is a common cross-check.

AUDIENCE
Bioinformatician

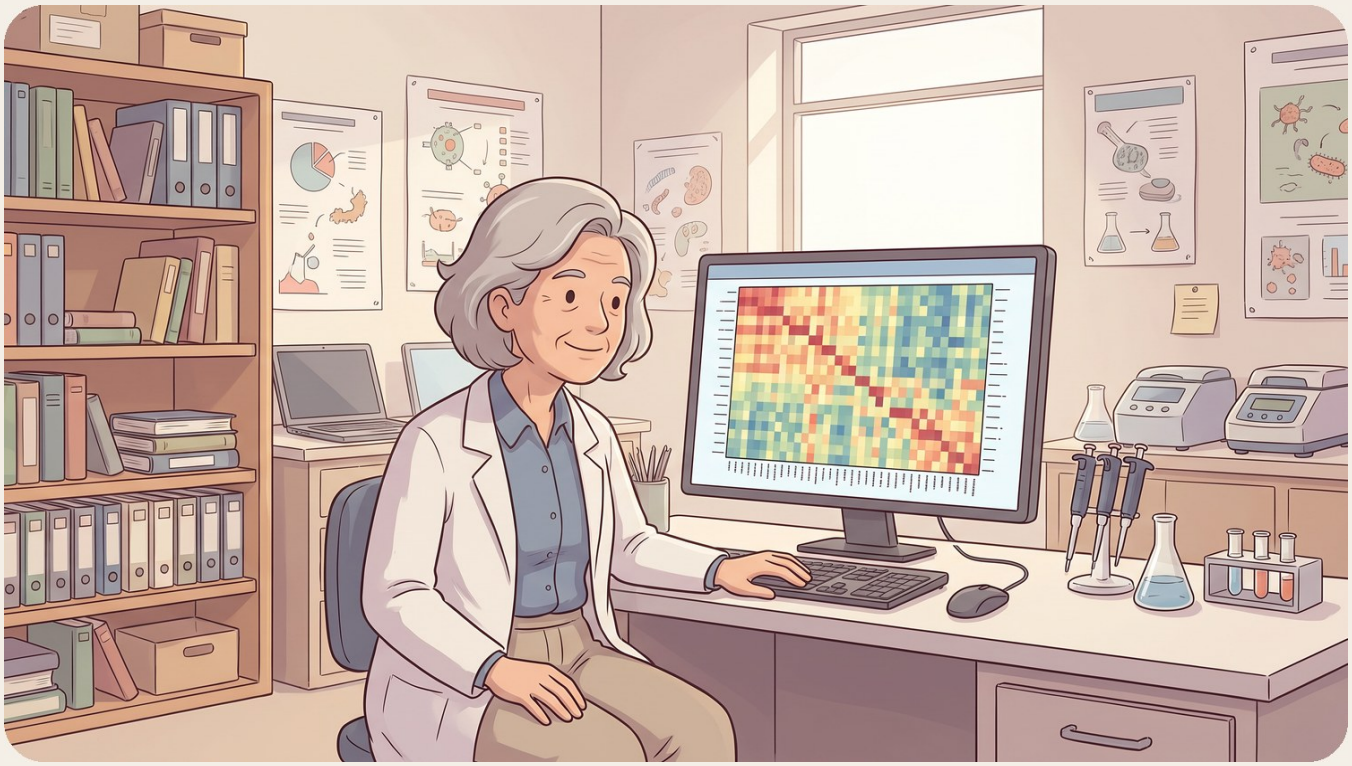
RUNS
In-app

MODULES
Pathway enrichment; GSEA pathway enrichment

OUTPUT
consolidated pathways

FUNCTIONAL ENRICHMENT & PATHWAYS

Q18 I want per-sample pathway activity to stratify patients, not just a global enrichment. Is that possible?



HOW BIRAGAS HELPS

In-silico pathway activity infers pathway activity per sample, giving you a sample x pathway matrix you can use to stratify patients or feed responder analysis - moving from one global enrichment to individual-level pathway signatures.

AUDIENCE
Senior Scientist

MODULES
In-silico pathway activity

RUNS
Full engine

OUTPUT
pathway activity dir

THEME

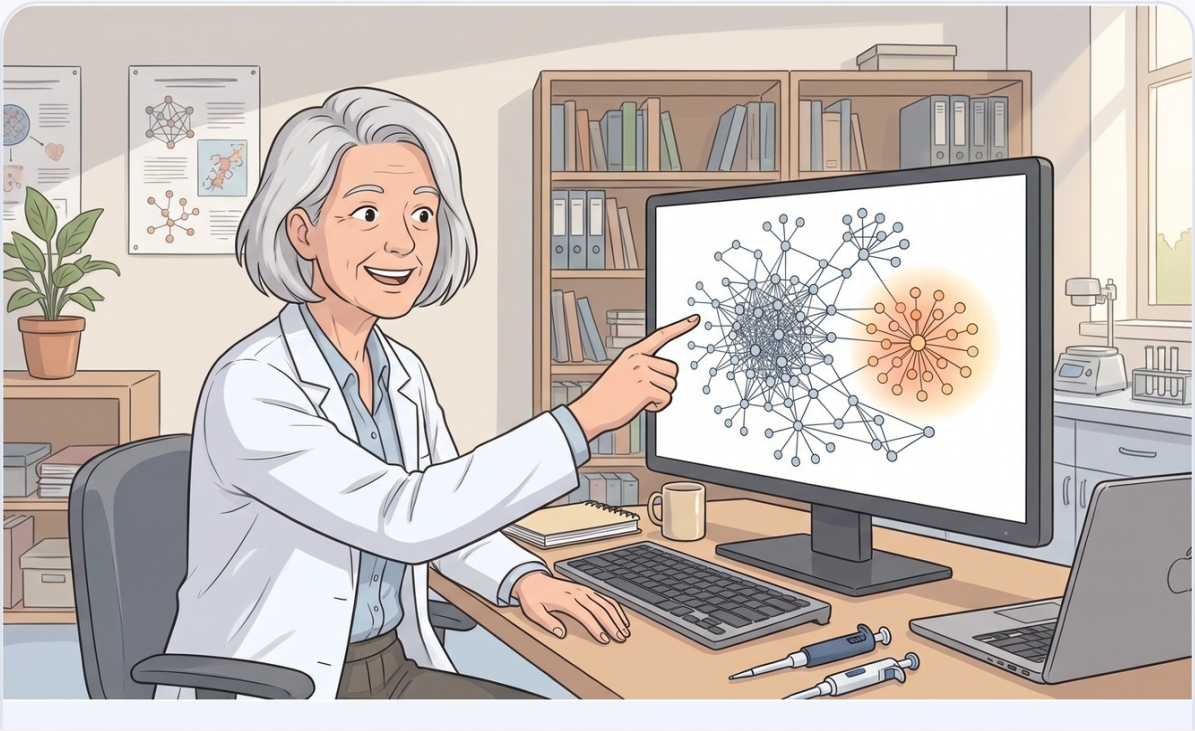
‹ CONTENTS

07 Target prioritization & disease networks

Questions Q19–Q21 · 3 in this theme

TARGET PRIORITIZATION & DISEASE NETWORKS

Q19 How do I move from 'differentially expressed' to 'likely disease-relevant module' rather than isolated genes?



HOW BIRAGAS HELPS

Disease-module detection finds enriched subnetworks / disease modules in the expression landscape, so you reason about coherent functional modules (and their hubs) instead of a flat gene list.

AUDIENCE
Senior Scientist

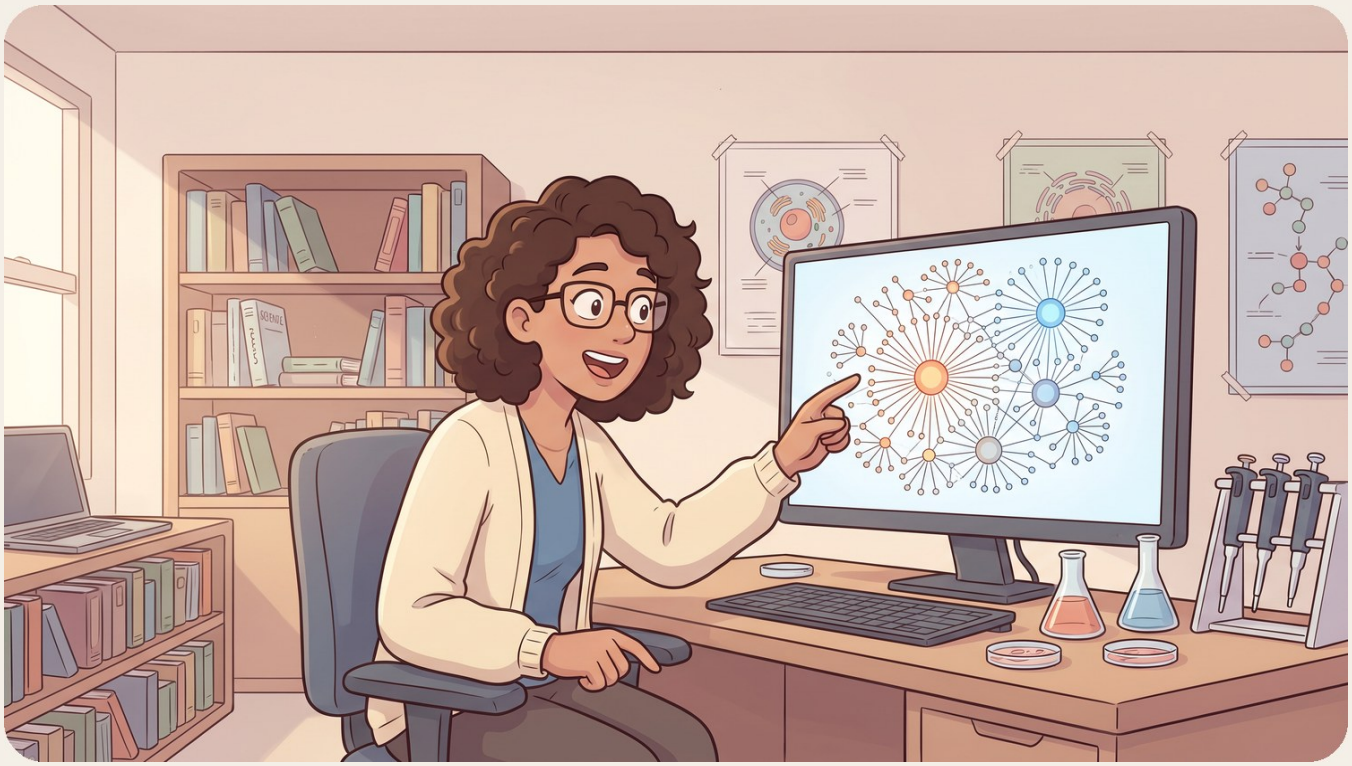
RUNS
In-app

MODULES
Disease-module detection

OUTPUT
module dir

TARGET PRIORITIZATION & DISEASE NETWORKS

Q20 In my candidate network, which genes are the structurally important control points worth targeting?



HOW BIRAGAS HELPS

Centrality calculator scores DAG nodes by betweenness, closeness, eigenvector and PageRank centrality - surfacing the structural bottlenecks/hubs that make the best intervention points, ranked on the validated causal graph rather than a correlation network.

AUDIENCE
Drug Discovery / Translational

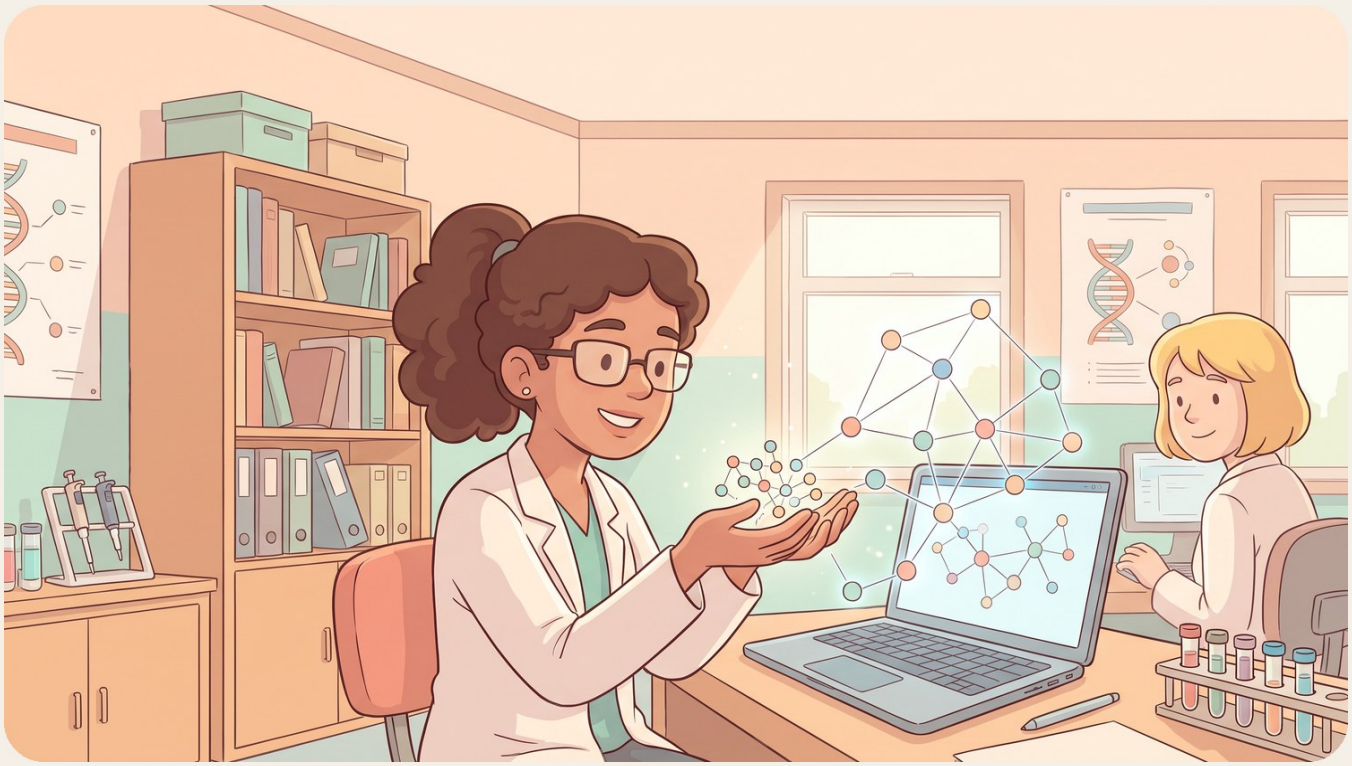
MODULES
Centrality calculator

RUNS
In-app

OUTPUT
centrality-enriched DAG

TARGET PRIORITIZATION & DISEASE NETWORKS

Q21 How do I expand a small seed target set into a fuller, evidence-backed candidate set?



HOW BIRAGAS HELPS

Gene-set builder expands your target set by combining the prioritized DEGs with genetic evidence, so the expansion is grounded in causal/genetic support rather than guilt-by-association co-expression alone.

AUDIENCE
Drug Discovery / Translational

MODULES
Gene-set builder; Gene prioritization

RUNS
Full engine

OUTPUT
expanded gene set

THEME

‹ CONTENTS

08 Temporal & dynamic analysis

Questions Q22–Q23 · 2 in this theme

TEMPORAL & DYNAMIC ANALYSIS

Q22 I have a time-course / treatment series. How do I model expression trajectories over time?

HOW BIRAGAS HELPS

Temporal analysis (bulk) fits bulk time-course models of expression trajectories across treatment/time, so you characterize dynamic response patterns instead of collapsing timepoints into a single contrast.

AUDIENCE
Bioinformatician

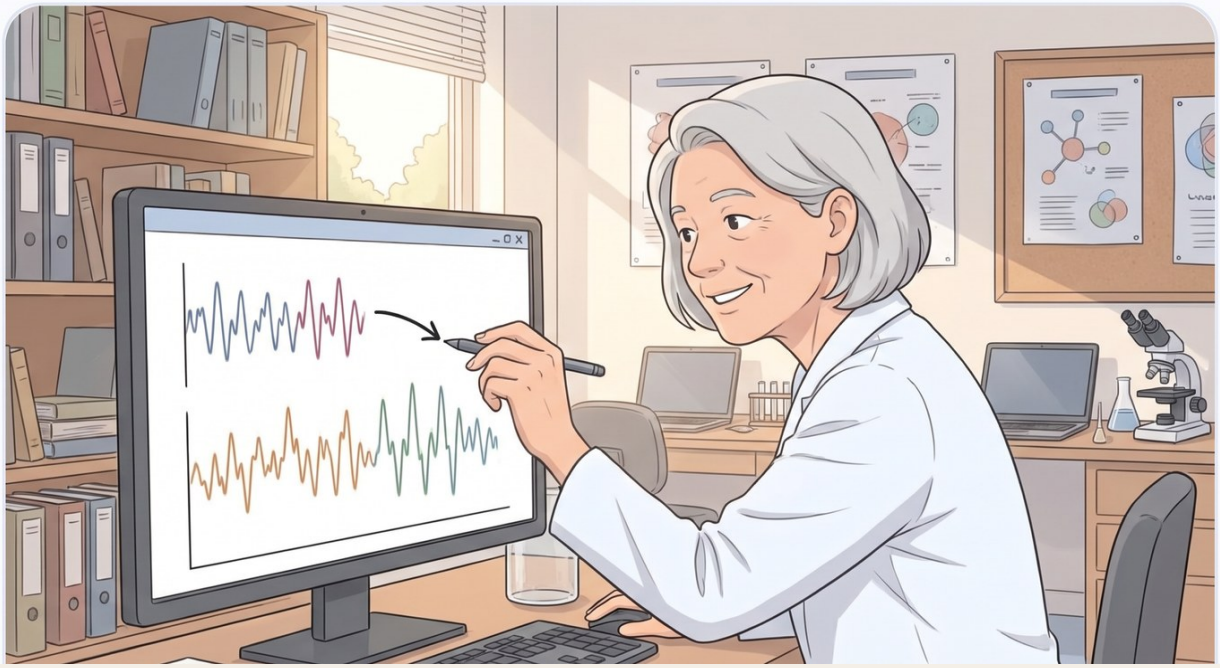
RUNS
Full engine

MODULES
Temporal analysis (bulk)

OUTPUT
temporal dir

TEMPORAL & DYNAMIC ANALYSIS

Q23 Can I infer directional, lead-lag relationships between genes from time-series data?



HOW BIRAGAS HELPS

Temporal Granger analysis derives Granger-causality edges from time-series expression - i.e. which genes' past values predict others' future values - giving you a temporal directionality layer that pure correlation can't provide. This is one of the four evidence streams feeding the causal graph.

AUDIENCE
Senior Scientist

MODULES
Temporal Granger analysis

RUNS
In-app

OUTPUT
granger edges

THEME

‹ CONTENTS

09 CRISPR & functional perturbation

Questions Q24–Q27 · 4 in this theme

CRISPR & FUNCTIONAL PERTURBATION

Q24 I ran a pooled CRISPR screen. How do I call hits rigorously?

HOW BIRAGAS HELPS

CRISPR screening performs pooled-screen hit-calling with MAGeCK (RRA and MLE) and BAGEL2 - the field-standard essentiality/enrichment callers - so your hit list is reproducible and method-cross-checked rather than a single ad-hoc ranking.

AUDIENCE
Bioinformatician

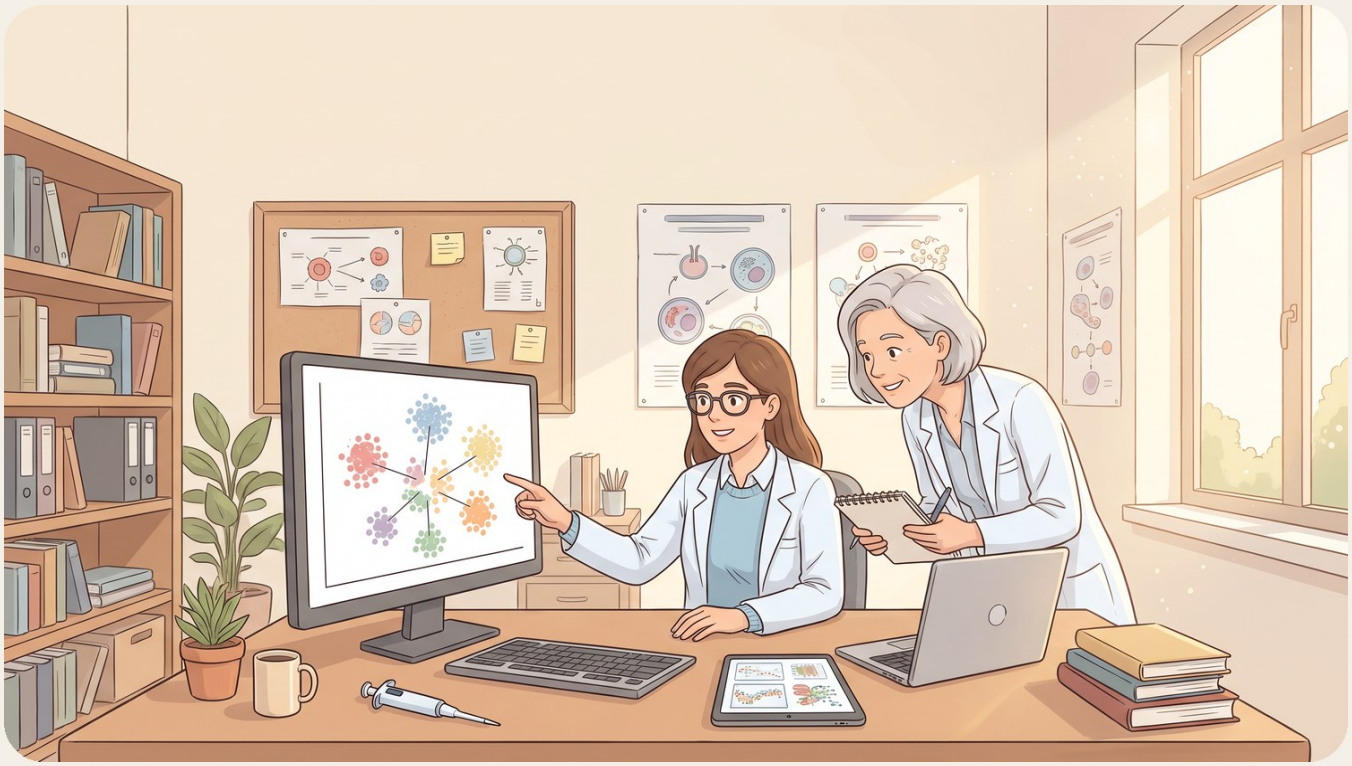
MODULES
CRISPR screening

RUNS
Full engine

OUTPUT
screening dir

CRISPR & FUNCTIONAL PERTURBATION

Q25 I have single-cell CRISPR (Perturb-seq) data. How do I connect guides to transcriptional phenotypes?



HOW BIRAGAS HELPS

CRISPR Perturb-seq analyzes sgRNA-based single-cell screens with hit calling, linking each perturbation to its single-cell transcriptional consequence - so you read out mechanism, not just fitness.

AUDIENCE
Bioinformatician

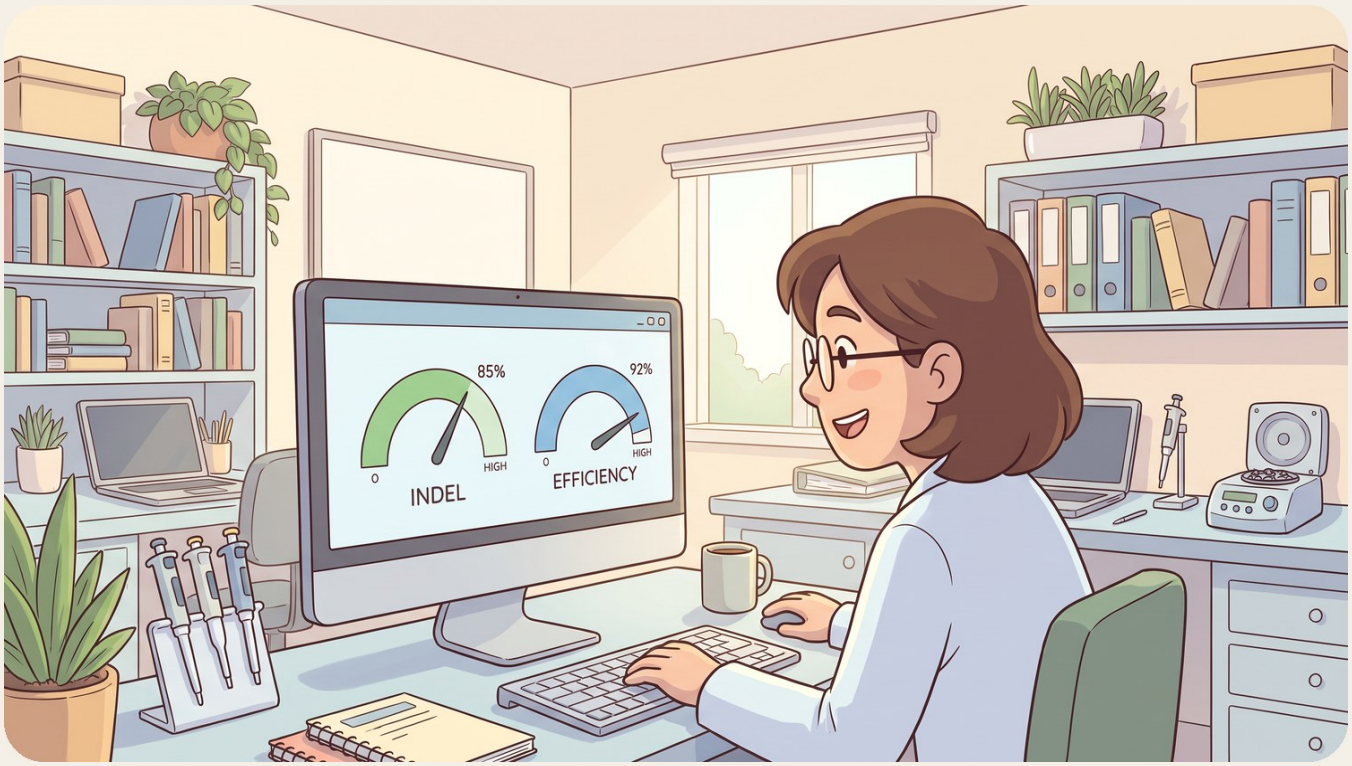
MODULES
CRISPR Perturb-seq

RUNS
Full engine

OUTPUT
crispr dir

CRISPR & FUNCTIONAL PERTURBATION

Q26 How do I profile editing outcomes (indels, efficiency) from targeted / amplicon CRISPR?



HOW BIRAGAS HELPS

CRISPR targeted / amplicon performs deep targeted/amplicon editing profiling - indel spectra and editing efficiency - so you QC and quantify your edits precisely at each locus.

AUDIENCE
Bioinformatician

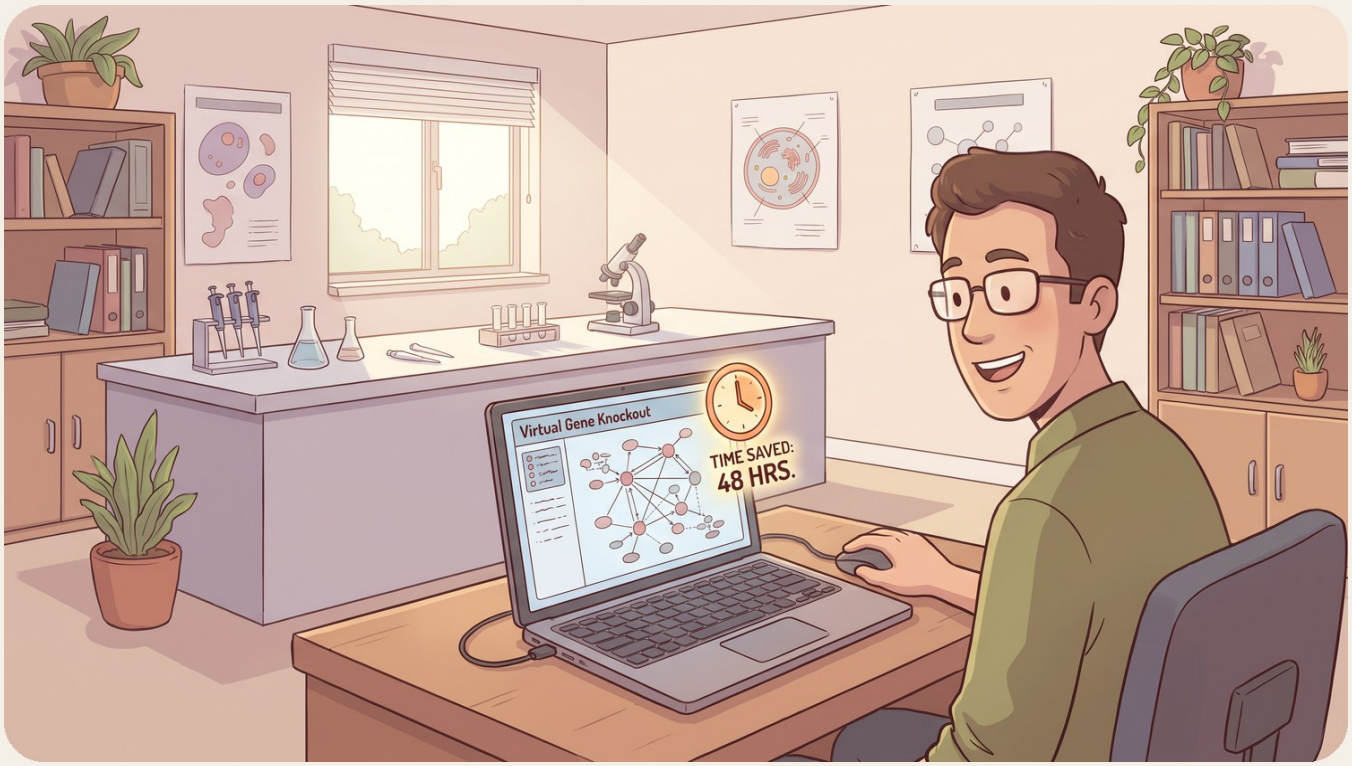
MODULES
CRISPR targeted / amplicon

RUNS
Full engine

OUTPUT
targeted dir

CRISPR & FUNCTIONAL PERTURBATION

Q27 Wet-lab CRISPR is slow and expensive. Can I predict knockout consequences before touching a cell?



HOW BIRAGAS HELPS

Perturbation analysis runs in-silico knockouts over your prioritized genes using DepMap dependency and L1000 drug-response evidence. Its MultiKnockout engine uses a 7-method ensemble (topological do-calculus, Bayesian noisy-OR, Monte-Carlo n=500, pathway-specific, feedback-adjusted, ODE-dynamic, mutual-information); the MegaKnockout engine computes any knockout's disease effect in ~O(1) via a sparse resolvent [I-(alpha)W]^-1 (alpha=0.85), scaling to 22.2 billion gene-pair predictions. It filters inactive/paradoxical predictions before you commit wet-lab resources.

AUDIENCE
Drug Discovery / Translational

MODULES
Perturbation analysis

RUNS
Full engine

OUTPUT
perturbation dir

THEME

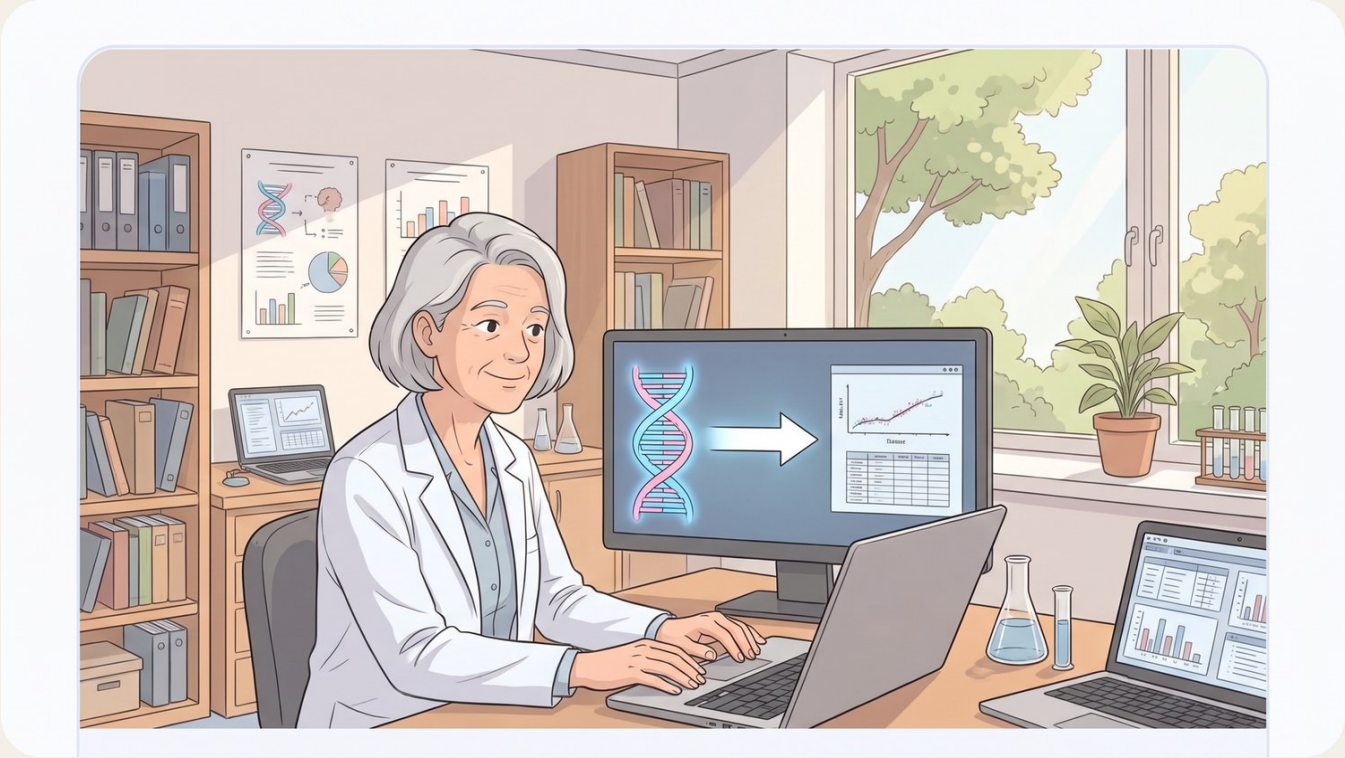
‹ CONTENTS

10 Genetic causal evidence (GWAS / MR)

Questions Q28–Q30 · 3 in this theme

GENETIC CAUSAL EVIDENCE (GWAS / MR)

Q28 How can I tell whether a gene is genuinely causal for the disease and not just correlated in expression?



HOW BIRAGAS HELPS

GWAS / Mendelian randomization uses natural genetic variants as instruments to estimate the causal effect on outcome, running 5 MR methods (IVW, MR-Egger, Weighted Median, MR-PRESSO, CAUSE) with 4 sensitivity tests (Cochran Q, Steiger directionality, leave-one-out, Egger intercept) and requiring F-statistic > 10 to reject weak instruments, plus eQTL colocalization to confirm a shared causal variant. This orients edges causally and guards against confounding and reverse causation.

AUDIENCE
Senior Scientist

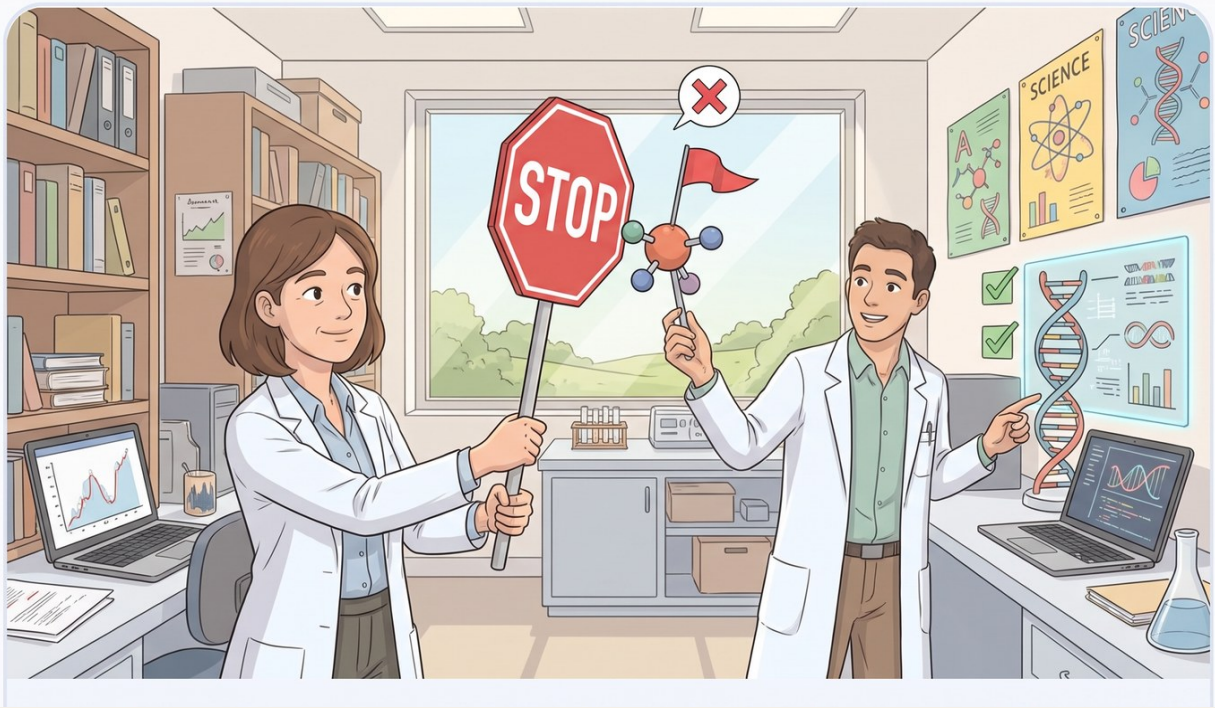
RUNS
Full engine

MODULES
GWAS / Mendelian randomization

OUTPUT
MR + genetic evidence

GENETIC CAUSAL EVIDENCE (GWAS / MR)

Q29 How do I avoid designing a therapeutic against the wrong biomarker (a classic, billion-dollar mistake)?



HOW BIRAGAS HELPS

Mendelian randomization is specifically used to prevent false leads from confounding or reverse causation. By requiring genetic-instrument support (and colocalization) before a target is elevated, BiRAGAS screens out associations that would otherwise lead you to build a drug for a passenger biomarker.

AUDIENCE
Drug Discovery / Translational

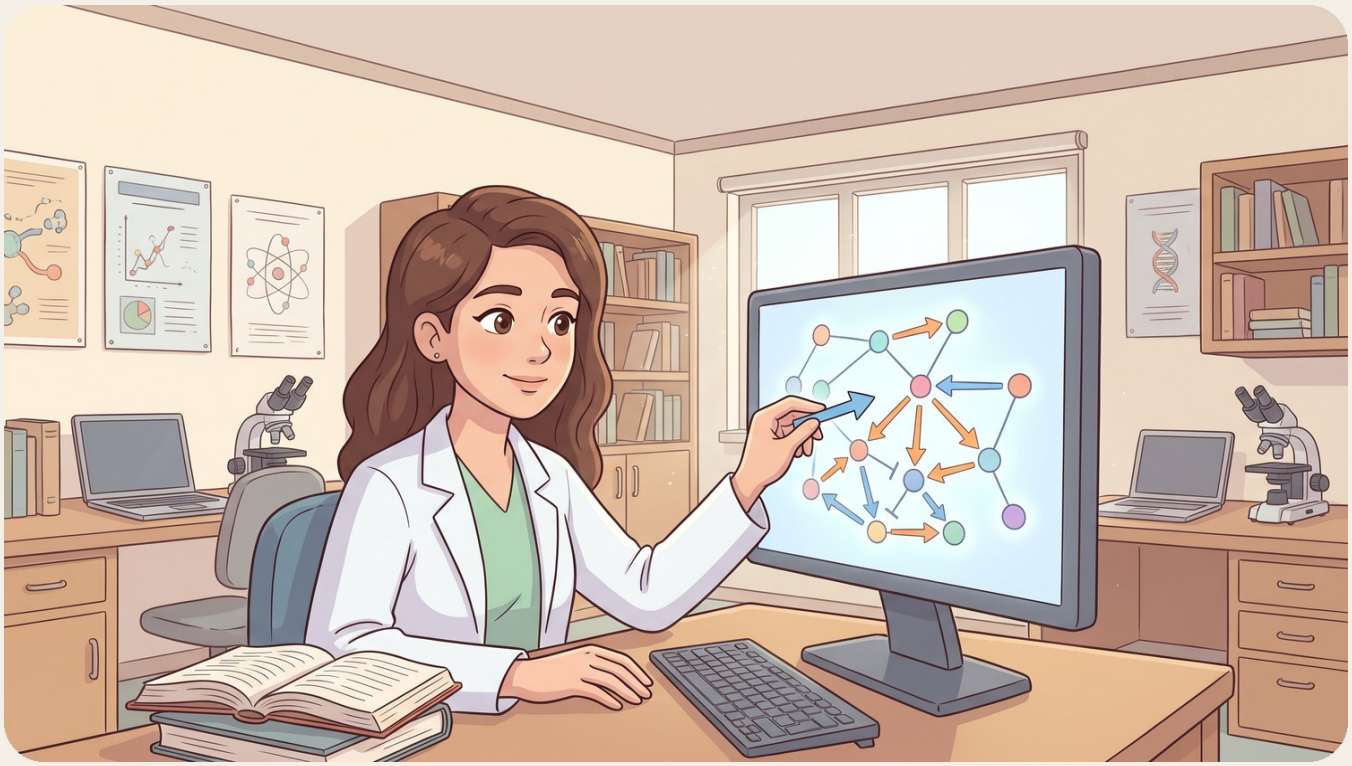
MODULES
GWAS / Mendelian randomization; Causal calculus

RUNS
Full engine

OUTPUT
genetic causal evidence

GENETIC CAUSAL EVIDENCE (GWAS / MR)

Q30 Can I anchor literature-curated signaling relationships to my own data's genetic evidence?



HOW BIRAGAS HELPS

SIGNOR causality overlays literature-curated signed causal edges (activation/inhibition) from SIGNOR, anchored by the genetic evidence in your analysis - so mechanistic prior knowledge is fused with, and constrained by, your data rather than assumed.

AUDIENCE
Senior Scientist

MODULES
SIGNOR causality

RUNS
Full engine

OUTPUT
SIGNOR subnetwork

THEME

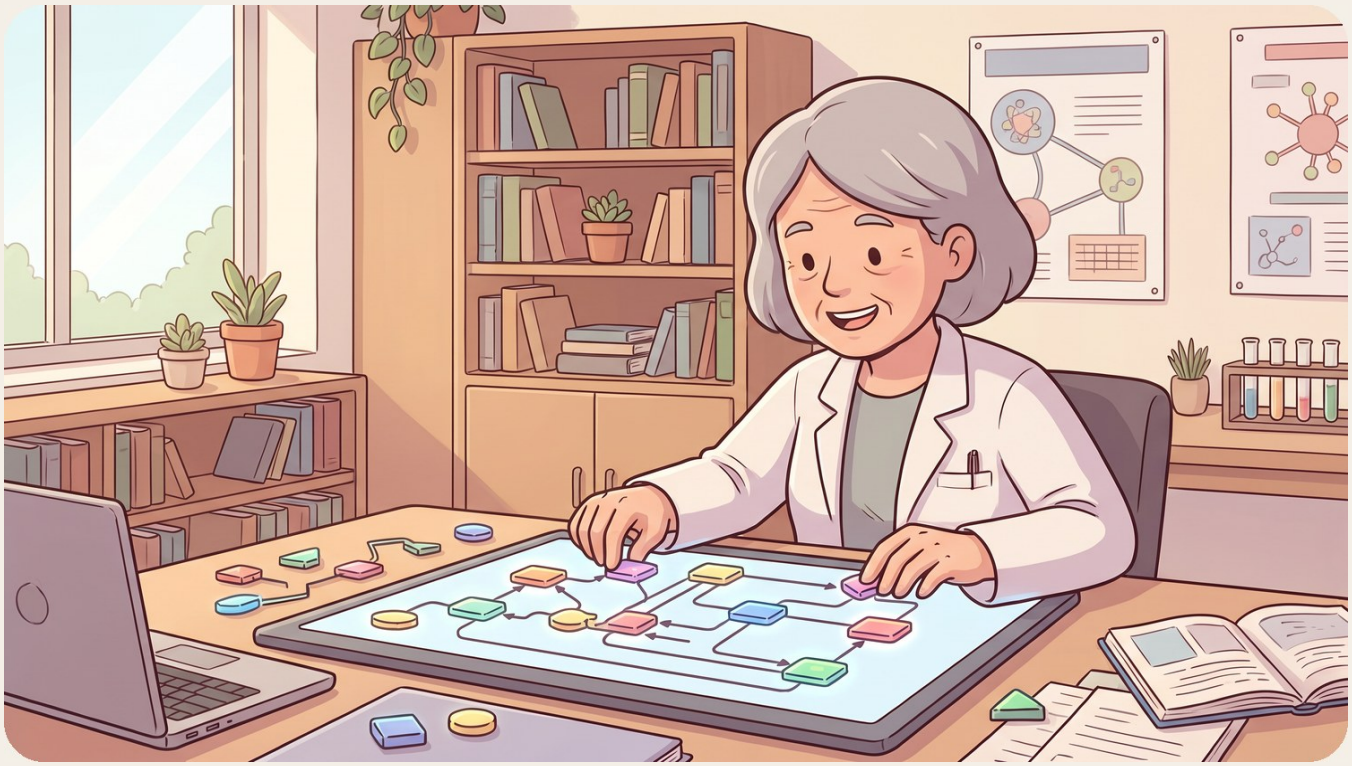
‹ CONTENTS

11 Causal inference & DAGs

Questions Q31–Q33 · 3 in this theme

CAUSAL INFERENCE & DAGS

Q31 How do I actually build a cause-and-effect graph of my disease rather than a co-expression network?



HOW BIRAGAS HELPS

Causal DAG builder learns a consensus directed acyclic graph with the PC-stable algorithm (gaussCItest, bootstrap n=100) and fuses 9 evidence streams via noisy-OR confidence aggregation [C = 1-(1-C_old)(1-C_new)], with streams weighted MR 0.95 > GWAS/SIGNOR 0.90 > CRISPR/eQTL 0.85 > database 0.80 > temporal 0.65 > statistical 0.35. The cascade reduces ~20,000 genes to a handful of Tier-1 drivers, each edge carrying its data lineage and rationale - not an undirected correlation map.

AUDIENCE
Senior Scientist

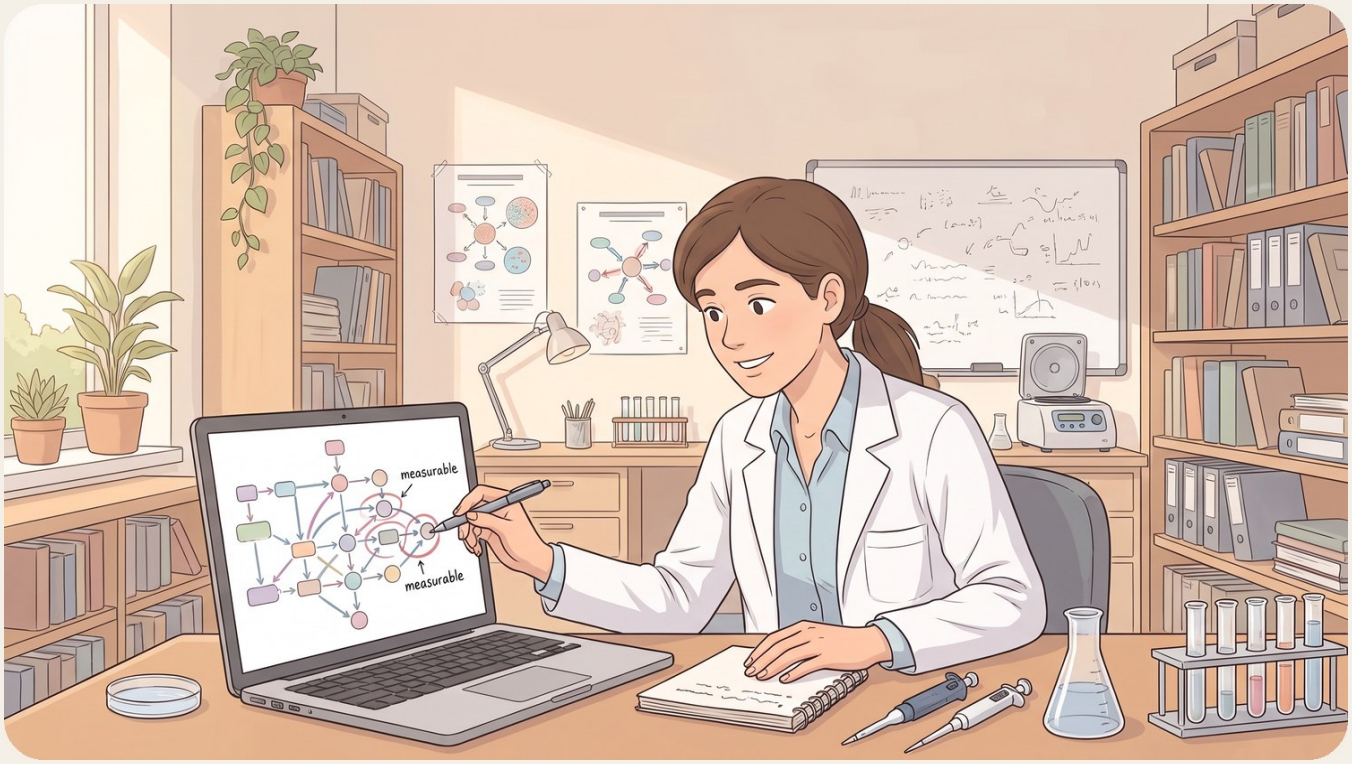
MODULES
Causal DAG builder

RUNS
In-app

OUTPUT
consensus + validated DAG

CAUSAL INFERENCE & DAGS

Q32 Given my causal graph, which effects are actually identifiable, and what are the confounders/back-door paths?



HOW BIRAGAS HELPS

Causal calculus applies Pearl's do-calculus to the validated DAG: P(Y|do(X)) is evaluated by graph surgery (removing X's incoming edges, setting do(X), propagating). It determines identifiability, derives the correct back-door adjustment sets via the back-door criterion and d-separation, and detects forks/colliders and pleiotropy (MR-Egger intercept) - so estimates are the ones the graph actually supports.

AUDIENCE
Drug Discovery / Translational

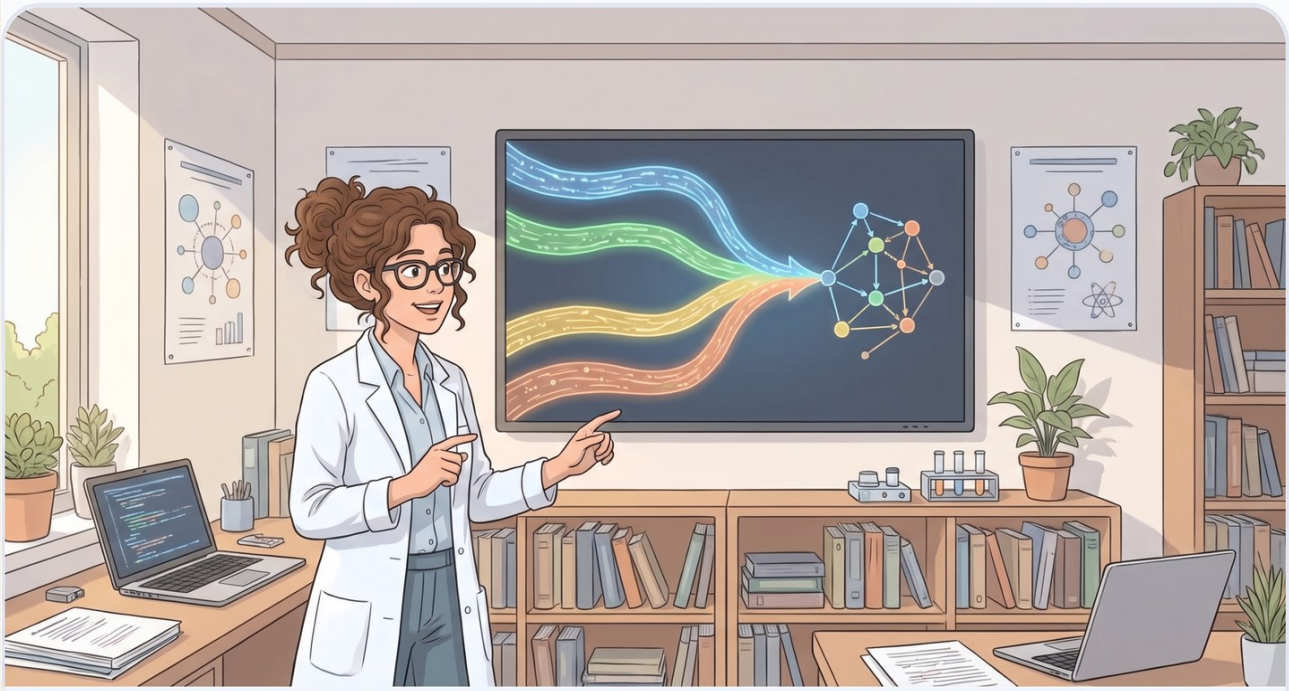
MODULES
Causal calculus

RUNS
In-app

OUTPUT
identified causal structure

CAUSAL INFERENCE & DAGS

Q33 How does the platform combine four different kinds of evidence into one trustworthy answer?



HOW BIRAGAS HELPS

BiRAGAS converges four independent streams - Observational (DE/enrichment), Temporal (Granger), Genetic (5-method MR + eQTL colocalization), and Interventional (do-calculus + in-silico CRISPR) - into one causal graph. Agreement across independent modalities (multi-stream convergence) is what raises confidence, and each finding is scored accordingly.

AUDIENCE
Senior Scientist

RUNS
Mixed

MODULES
Causal DAG builder; Causal calculus; GWAS / Mendelian randomization; Temporal Granger analysis; Perturbation analysis

OUTPUT
validated multi-evidence DAG

THEME

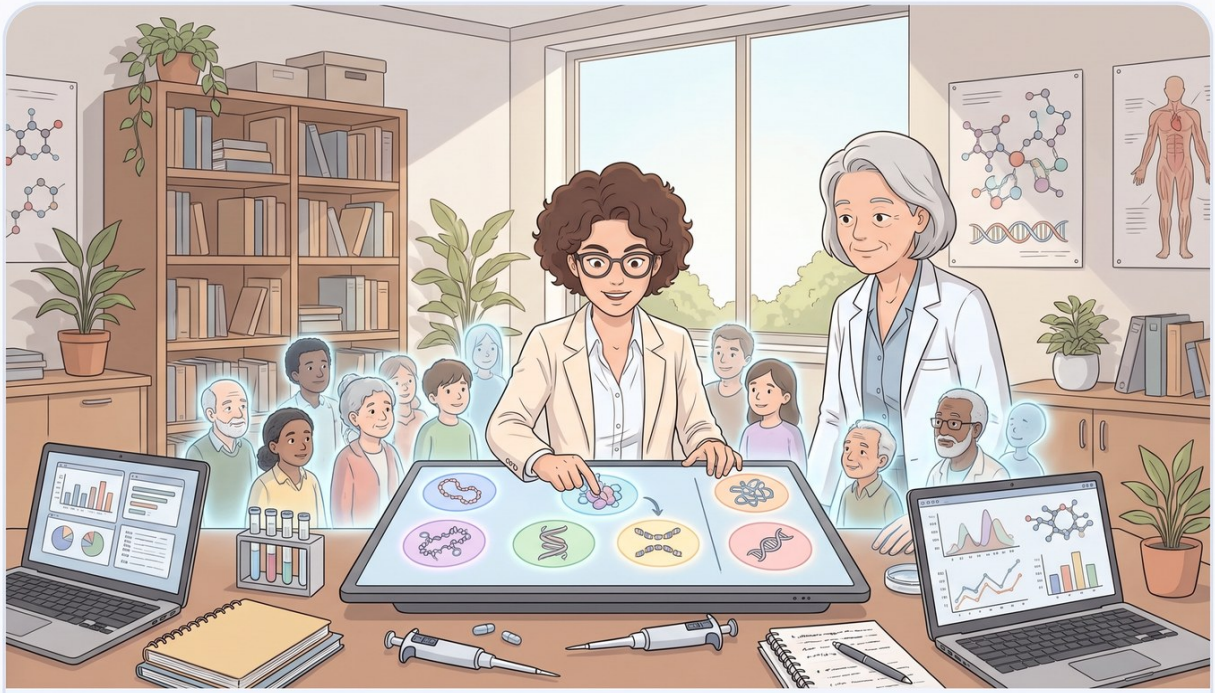
‹ CONTENTS

12 In-silico trials & patient stratification

Questions Q34–Q36 · 3 in this theme

IN-SILICO TRIALS & PATIENT STRATIFICATION

Q34 Can I rank therapeutic targets at the population level and see where resistance would emerge?



HOW BIRAGAS HELPS

In silico perturbation (population) ranks targets on a 7-dimension score (causal 0.25, centrality 0.20, druggability 0.15, efficacy 0.15, safety 0.10, resistance 0.10, translatability 0.05) and runs explicit resistance analysis across 5 mechanisms (feedback, bypass, redundancy, compensation, variability) - then recommends 2-3 target combinations to block the escape routes it predicts.

AUDIENCE
Drug Discovery / Translational

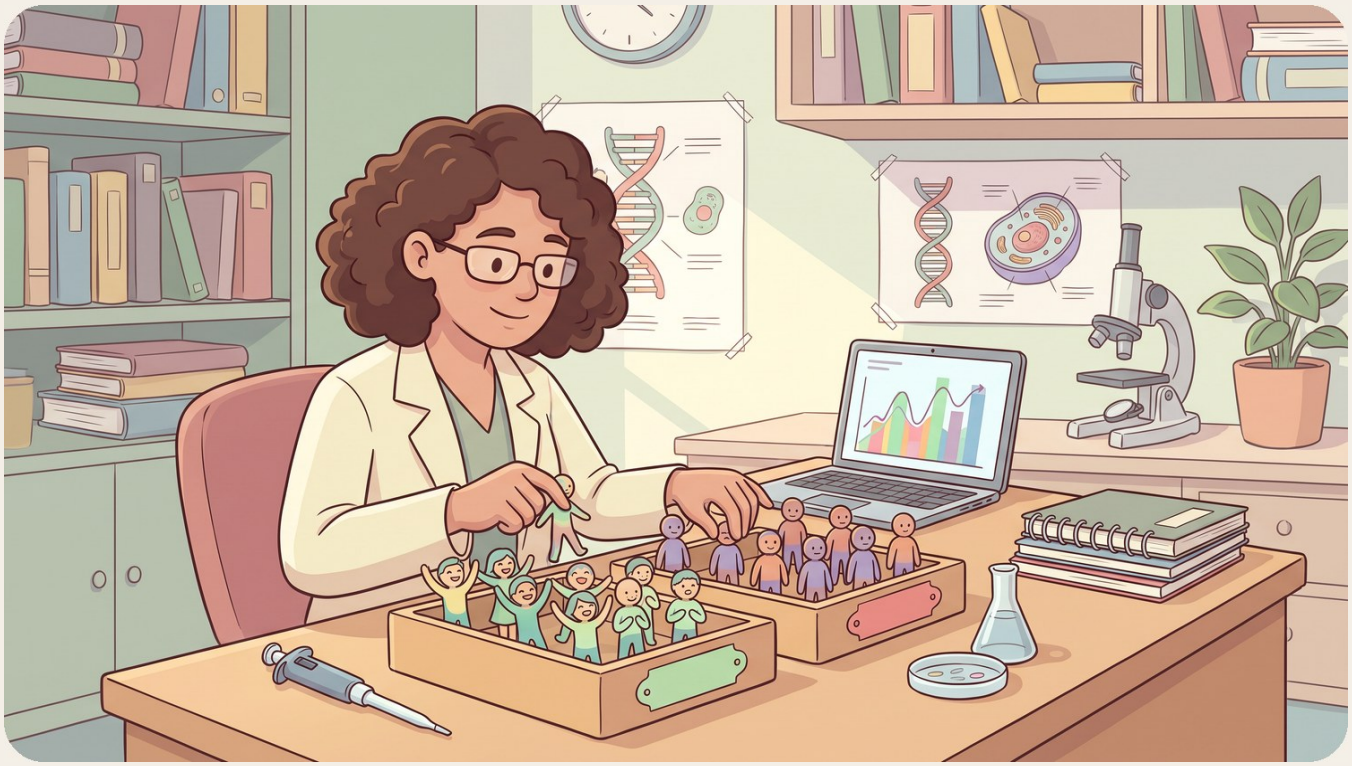
RUNS
In-app

MODULES
In silico perturbation (population)

OUTPUT
ranked targets

IN-SILICO TRIALS & PATIENT STRATIFICATION

Q35 How do I figure out which patients will respond to a candidate drug (and which won't)?



HOW BIRAGAS HELPS

In silico perturbation (patient) does patient-stratified target ranking and assigns responder clusters - moving therapy from population guesswork toward individual-level precision by simulating perturbation response per patient subgroup.

AUDIENCE
Drug Discovery / Translational

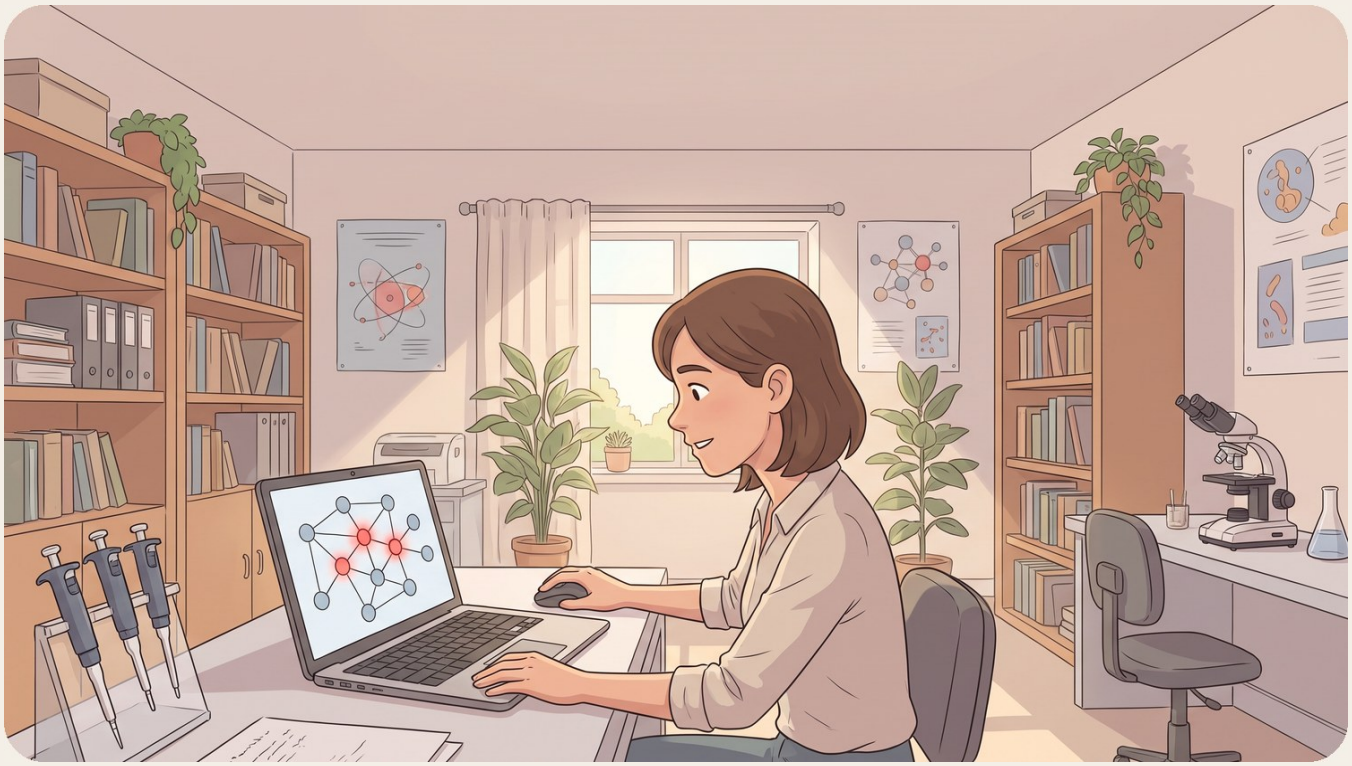
MODULES
In silico perturbation (patient)

RUNS
Full engine

OUTPUT
patient-stratified targets

IN-SILICO TRIALS & PATIENT STRATIFICATION

Q36 Why are some patients resistant to an experimental drug? Can BiRAGAS explain the mechanism?



HOW BIRAGAS HELPS

By simulating perturbations on patient-stratified causal graphs and analyzing compensation/resistance nodes, BiRAGAS surfaces the mechanistic bypass routes (e.g. a compensatory pathway) that make a subgroup resistant - the kind of 'why' that a validated AML case study demonstrated (bypass calls matched patient observations).

AUDIENCE
Drug Discovery / Translational

RUNS
Full engine

MODULES
In silico perturbation (patient); Perturbation analysis; Causal calculus

OUTPUT
resistance mechanism report

THEME

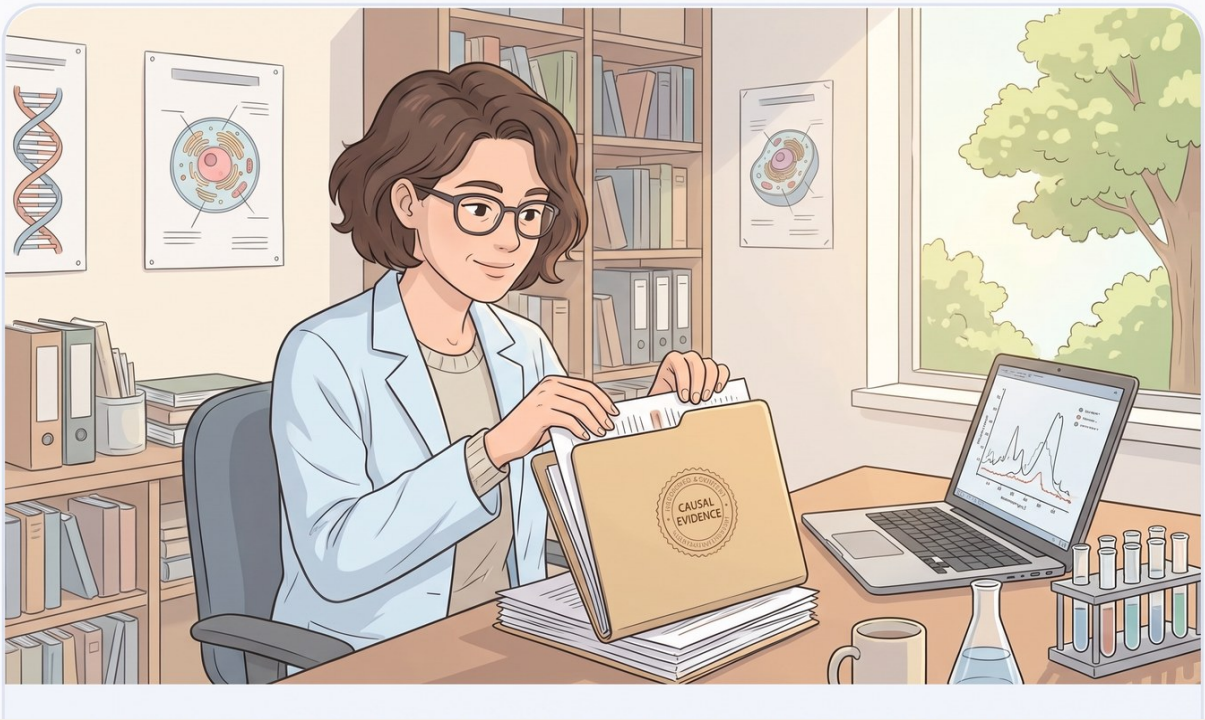
‹ CONTENTS

13 Regulatory & FDA-grade evidence

Questions Q37–Q39 · 3 in this theme

REGULATORY & FDA-GRADE EVIDENCE

Q37 The FDA now wants mechanistic causal evidence for a biomarker's role. How do I produce that?



HOW BIRAGAS HELPS

Since the 2018 Biomarker Qualification Guidance, the FDA requires evidence of a biomarker's role in the causal/outcome pathway. BiRAGAS produces this natively: every output carries Causal Confidence Scoring and a source-linked, FDA-aligned causal target dossier - it's integral to the architecture, not a bolted-on service.

AUDIENCE
Drug Discovery / Translational

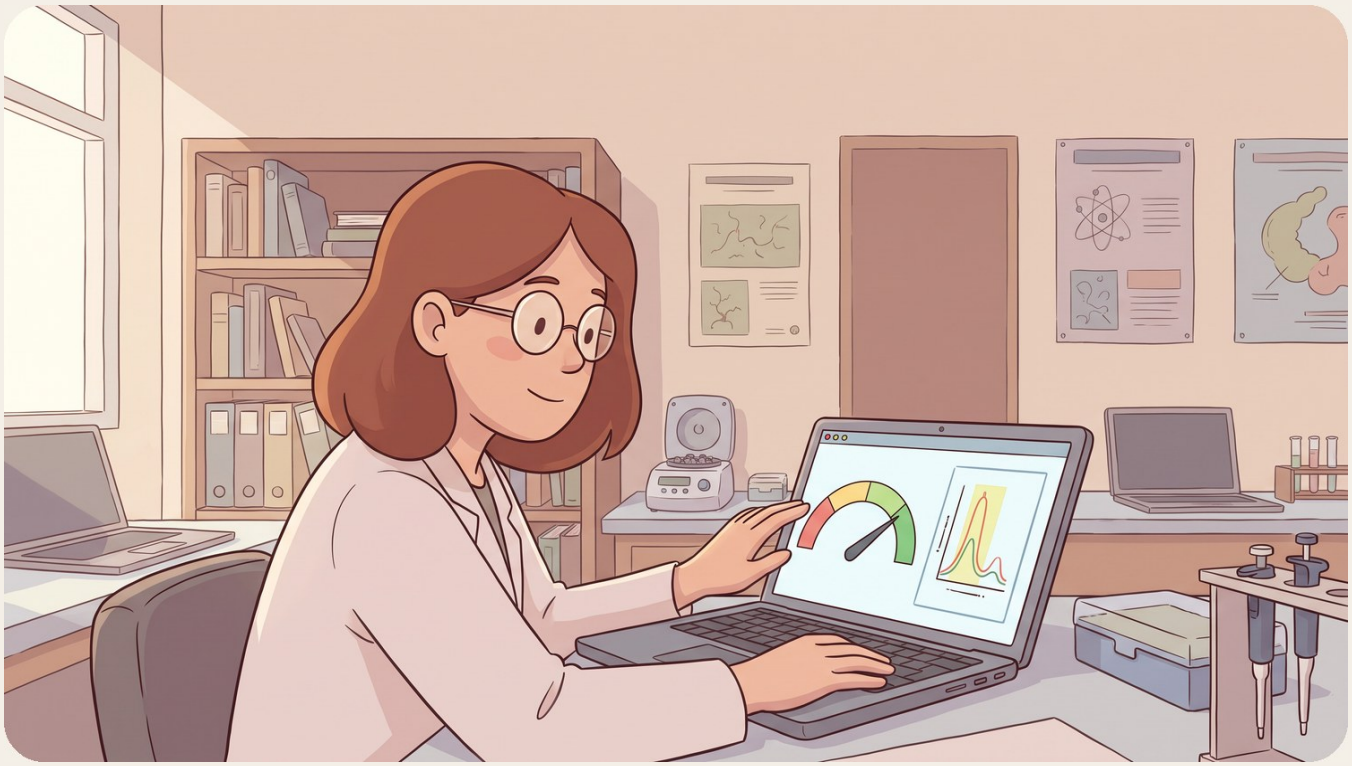
RUNS
Full engine

MODULES
Molecular report / dossier; Causal calculus

OUTPUT
causal target dossier

REGULATORY & FDA-GRADE EVIDENCE

Q38 How trustworthy is any single BiRAGAS call - can I defend it to a regulator or reviewer?



HOW BIRAGAS HELPS

Each finding gets a Causal Confidence Score = Bradford-Hill criteria met / 8 (strength, consistency, specificity, temporality, gradient, plausibility, coherence, experiment). Tiers: Tier 1 Driver = CCS >= 0.8 with predicted effect > 50% (FDA-submission grade); Tier 2 Strong Candidate >= 0.6 (2+ independent streams); Tier 3 Candidate >= 0.5. The 7-step Bayesian composite (statistical 30%, RAG-grounding 25%, MR proof 25%, multi-agent 15%, plausibility 5%) benchmarked at AUROC 0.947 against clinical outcomes. Every step is source-linked with retained provenance - no black boxes.

AUDIENCE
Drug Discovery / Translational

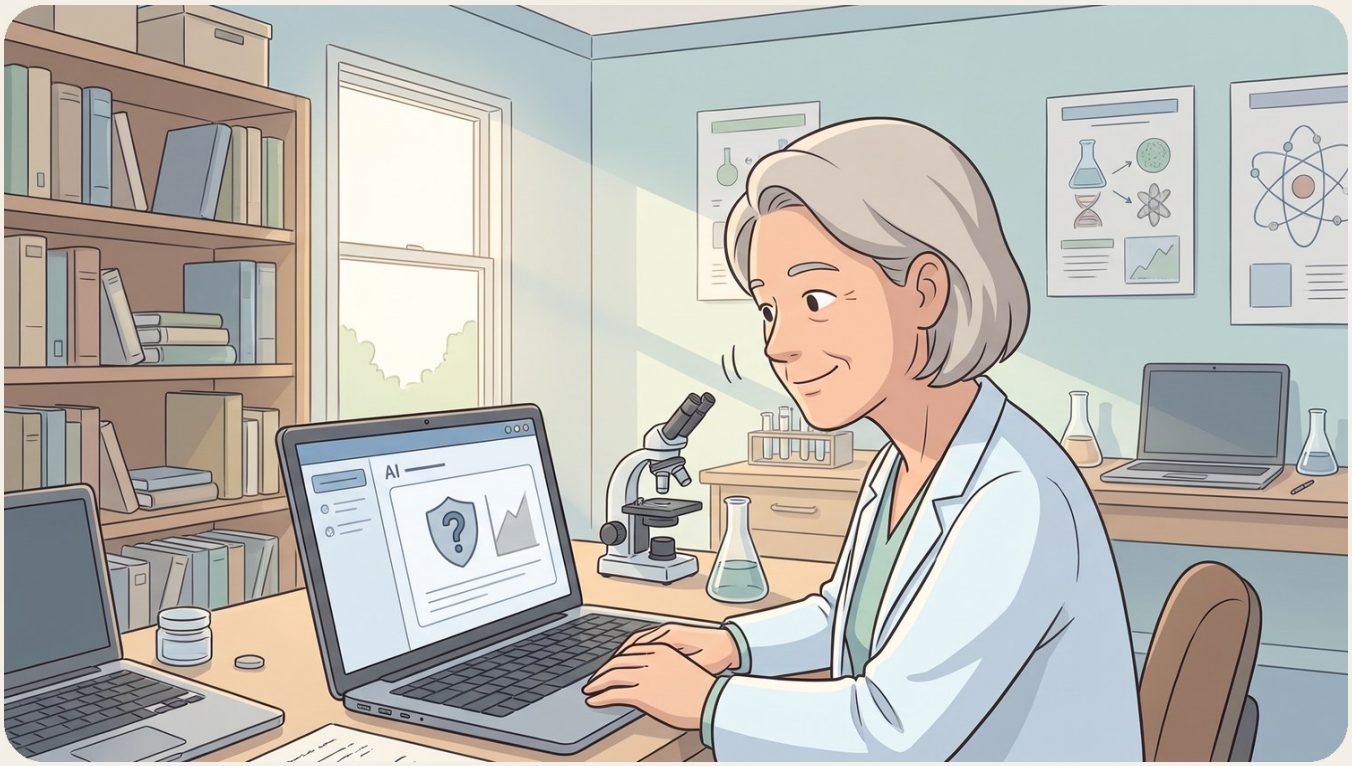
RUNS
Full engine

MODULES
Molecular report / dossier; Causal DAG builder

OUTPUT
CCS-ranked, source-linked output

REGULATORY & FDA-GRADE EVIDENCE

Q39 Does the platform ever admit uncertainty, or does it always give a confident answer?



HOW BIRAGAS HELPS

It abstains when appropriate. In the BeatAML validation, on the candidates it could not support, BiRAGAS reported low confidence or abstained rather than emit a false positive - the CCS framework is designed to know what it does not know, which is what makes the confident calls credible.

AUDIENCE
Senior Scientist

MODULES
Molecular report / dossier

RUNS
Full engine

OUTPUT
CCS with abstention

THEME

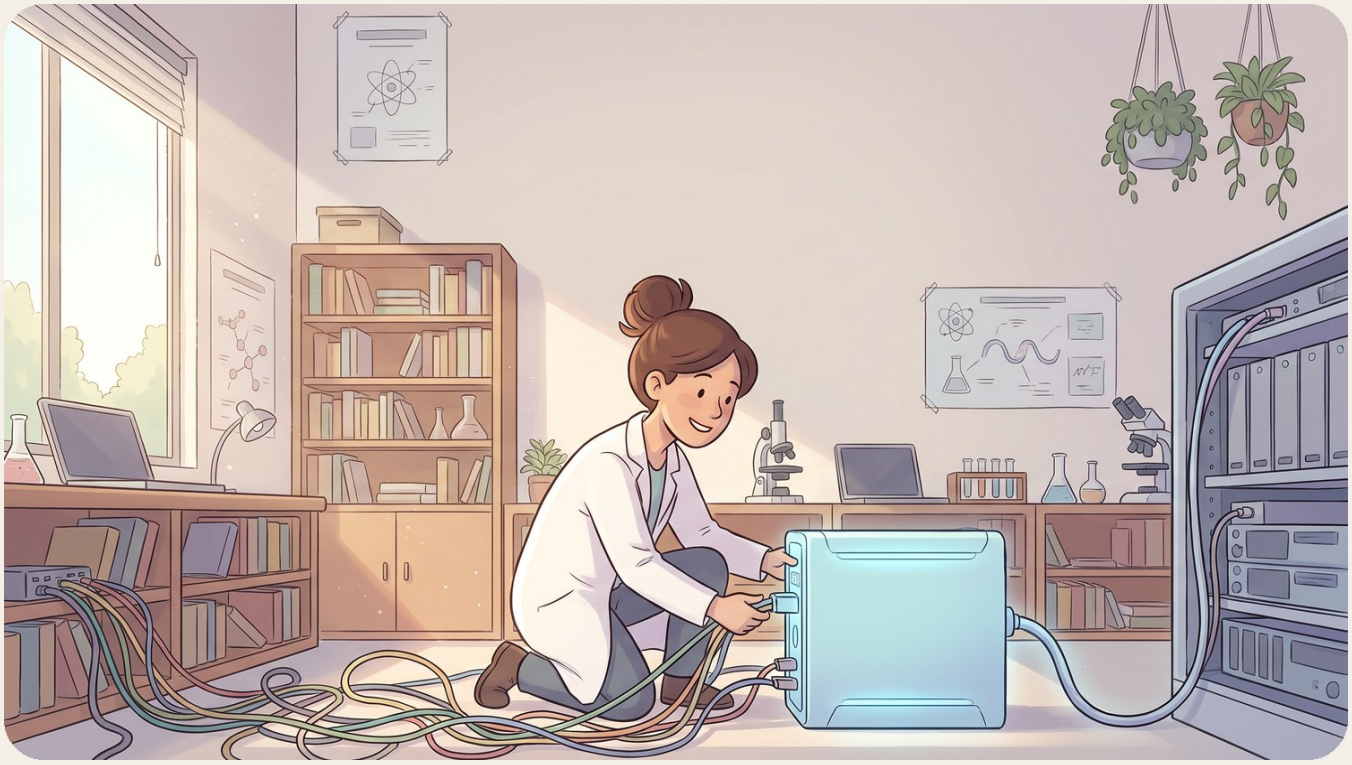
‹ CONTENTS

14 Reproducibility, provenance & automation

Questions Q40–Q42 · 3 in this theme

REPRODUCIBILITY, PROVENANCE & AUTOMATION

Q40 Our discovery pipeline is stitched from many vendors' tools with no shared provenance. How does BiRAGAS fix that?



HOW BIRAGAS HELPS

BiRAGAS is one end-to-end chained engine: raw sequences -> harmonization -> prioritization -> causal graph -> genetic validation -> in-silico trials -> dossier, with zero manual handoffs and full provenance at every step. It replaces the fragmented, non-auditable multi-vendor chain where 'evidence dies' at each handoff.

AUDIENCE
Bioinformatician

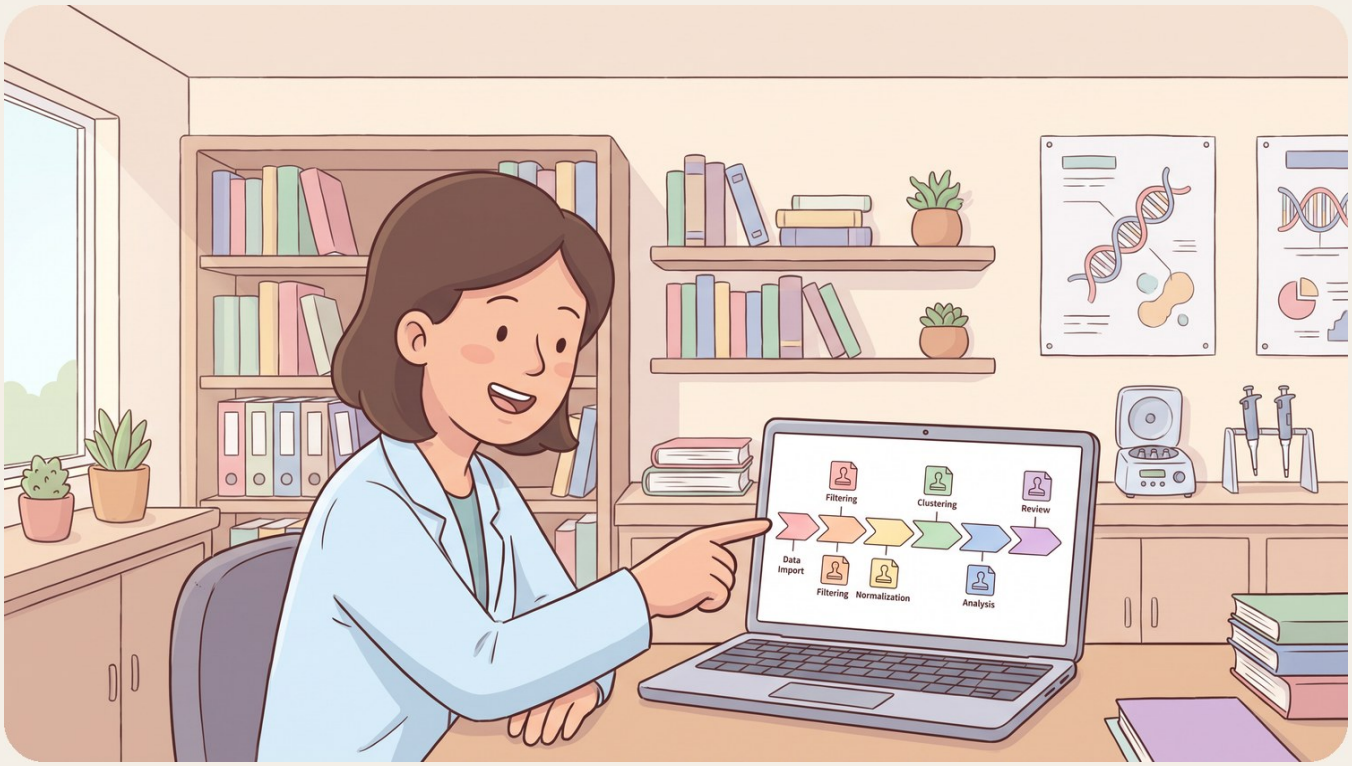
RUNS
Full chain

MODULES
All modules (chained pipeline)

OUTPUT
auditable end-to-end run

REPRODUCIBILITY, PROVENANCE & AUTOMATION

Q41 How do I make an analysis reproducible and auditable months later?



HOW BIRAGAS HELPS

Every step is fully auditable with retained data lineage and rationale (the DAG records provenance behind each connection), and the pipeline is a single engine rather than one-off scripts - so a run can be traced and re-executed rather than reverse-engineered from a folder of intermediate files.

AUDIENCE
Bioinformatician

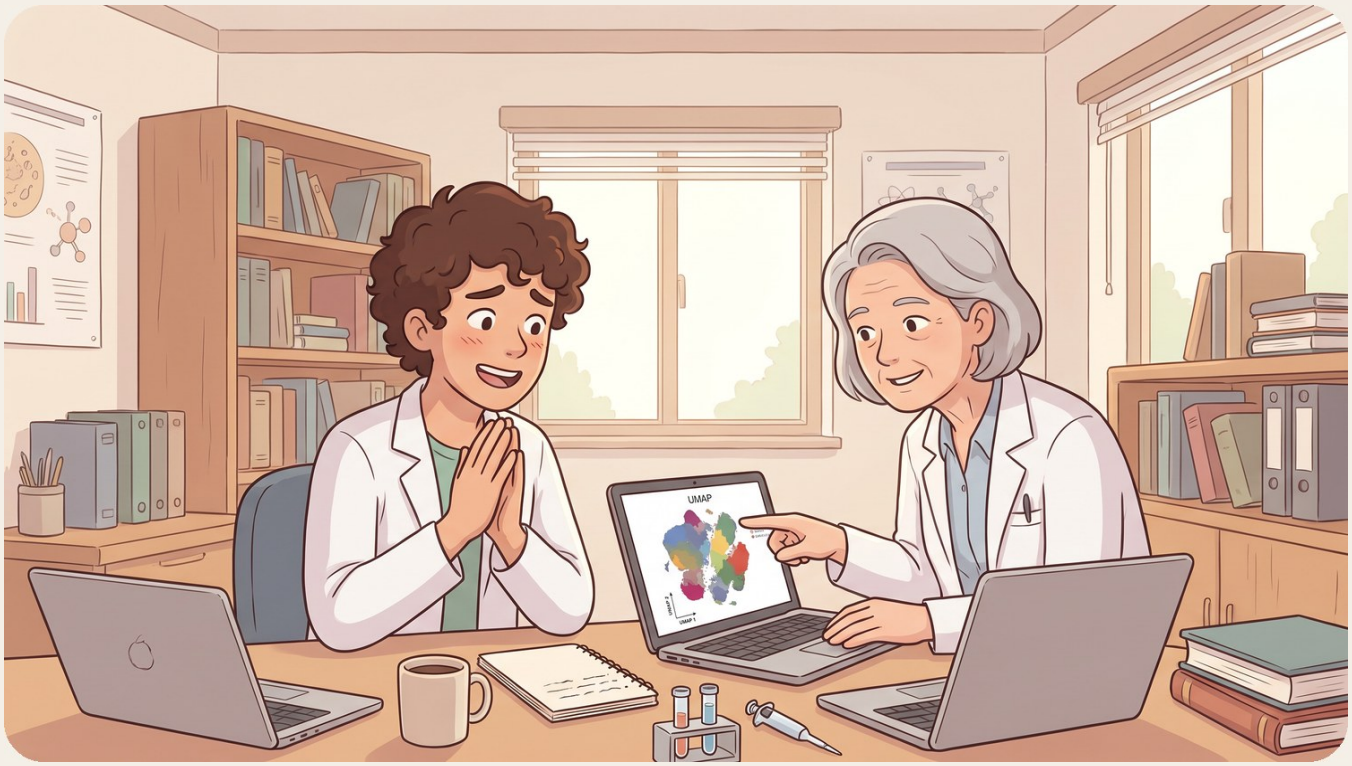
MODULES
Molecular report / dossier; Causal DAG builder

RUNS
Full chain

OUTPUT
provenance-tracked run

REPRODUCIBILITY, PROVENANCE & AUTOMATION

Q42 I don't remember the exact parameters or which module to use. Can I just describe what I want?



HOW BIRAGAS HELPS

Yes - Causion is a natural-language agent (LangGraph supervisor, 16 agents) over all 28 modules, wired to 13 external knowledge APIs (ChEMBL, KEGG, Reactome, STRING, PubMed, OpenFDA, Ensembl, DGIdb...). Say 'Map the causal genetic nodes for condition X' or 'Simulate a patient perturbation trial' and it interprets intent, routes to the right module, sets parameters and returns the result. Every claim it makes is grounded against an 8.6M-vector Bio-RAG index and passed through a hallucination shield, so you don't need to memorize the pipeline - or worry it invents edges.

AUDIENCE
Student / Trainee

RUNS
In-app

MODULES
Causion co-pilot (all modules)

OUTPUT
guided execution

THEME

‹ CONTENTS

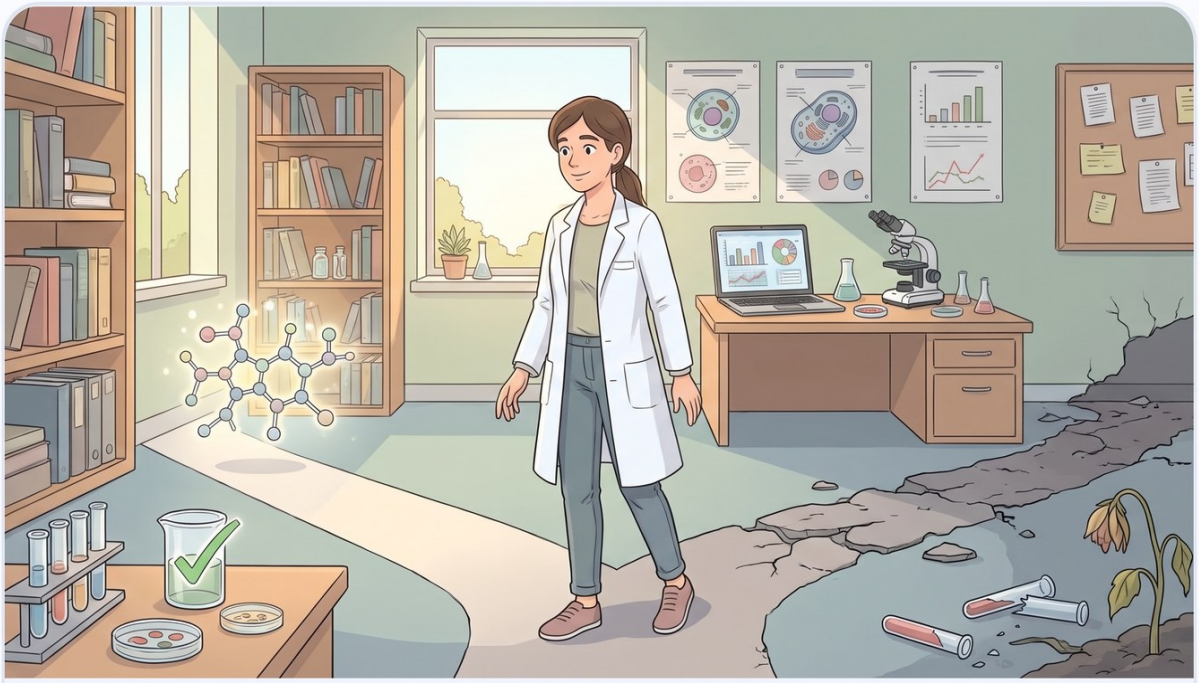
15

Strategic drug-discovery questions

Questions Q43–Q46 · 4 in this theme

STRATEGIC DRUG-DISCOVERY QUESTIONS

Q43 90% of trials fail, often from a wrong target. How does BiRAGAS reduce that risk upstream?



HOW BIRAGAS HELPS

BiRAGAS operates upstream of molecular design - at causal target identification and validation, the highest-stakes step. By requiring convergent causal evidence (genetic + temporal + interventional + observational) and scoring drivers vs passengers before a target advances, it compresses upstream timelines and reduces the late-stage failures that come from chasing the wrong target.

AUDIENCE
Drug Discovery / Translational

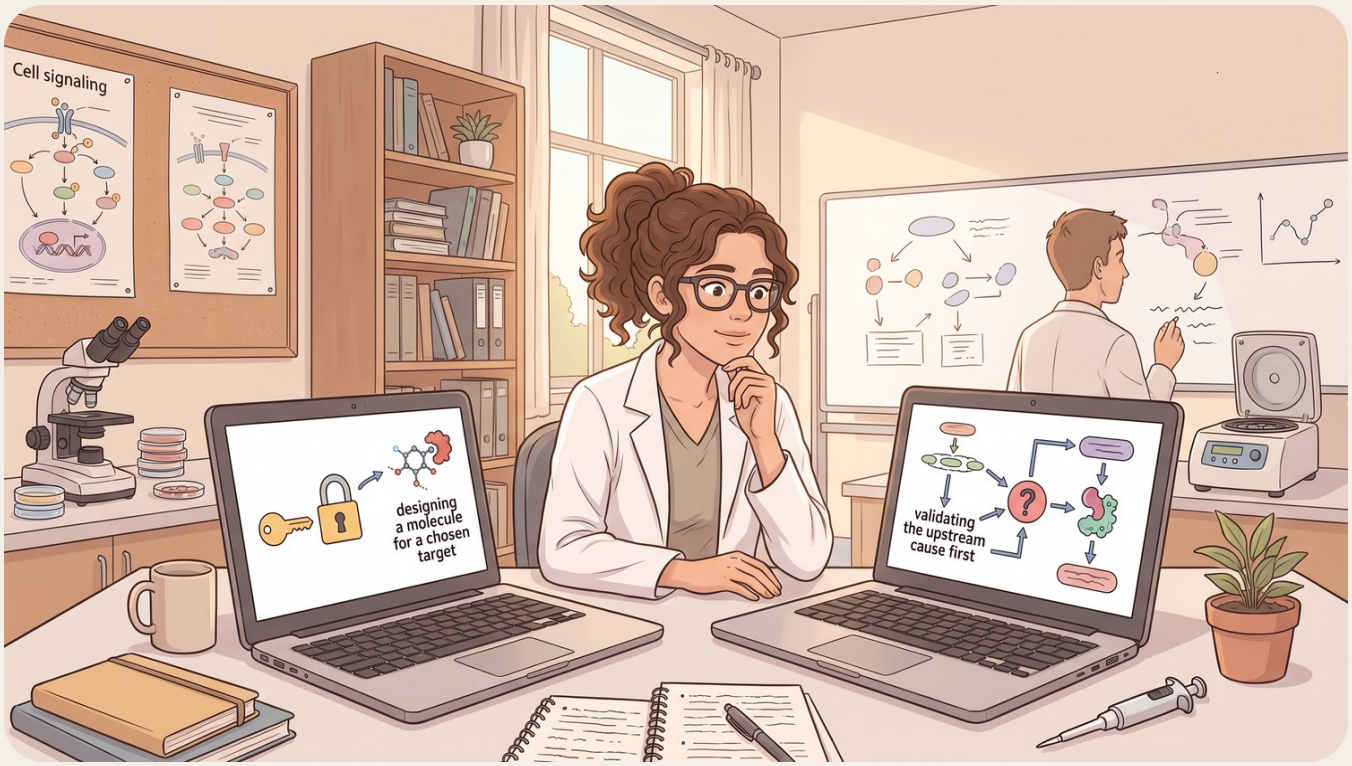
RUNS
Mixed

MODULES
Causal DAG builder; GWAS / Mendelian randomization; In silico perturbation (population)

OUTPUT
validated causal targets

STRATEGIC DRUG-DISCOVERY QUESTIONS

Q44 Competitors design drugs for a target - what does BiRAGAS do that they don't?



HOW BIRAGAS HELPS

Tools like Recursion, Insilico and Schrodinger design/optimize molecules for a given target; genomics platforms (Tempus, Foundation) produce correlation. BiRAGAS validates whether the target itself is correct, screens which patients benefit, and outputs the mechanistic causal evidence the FDA requires - the upstream causal layer the others leave to the customer to build.

AUDIENCE
Drug Discovery / Translational

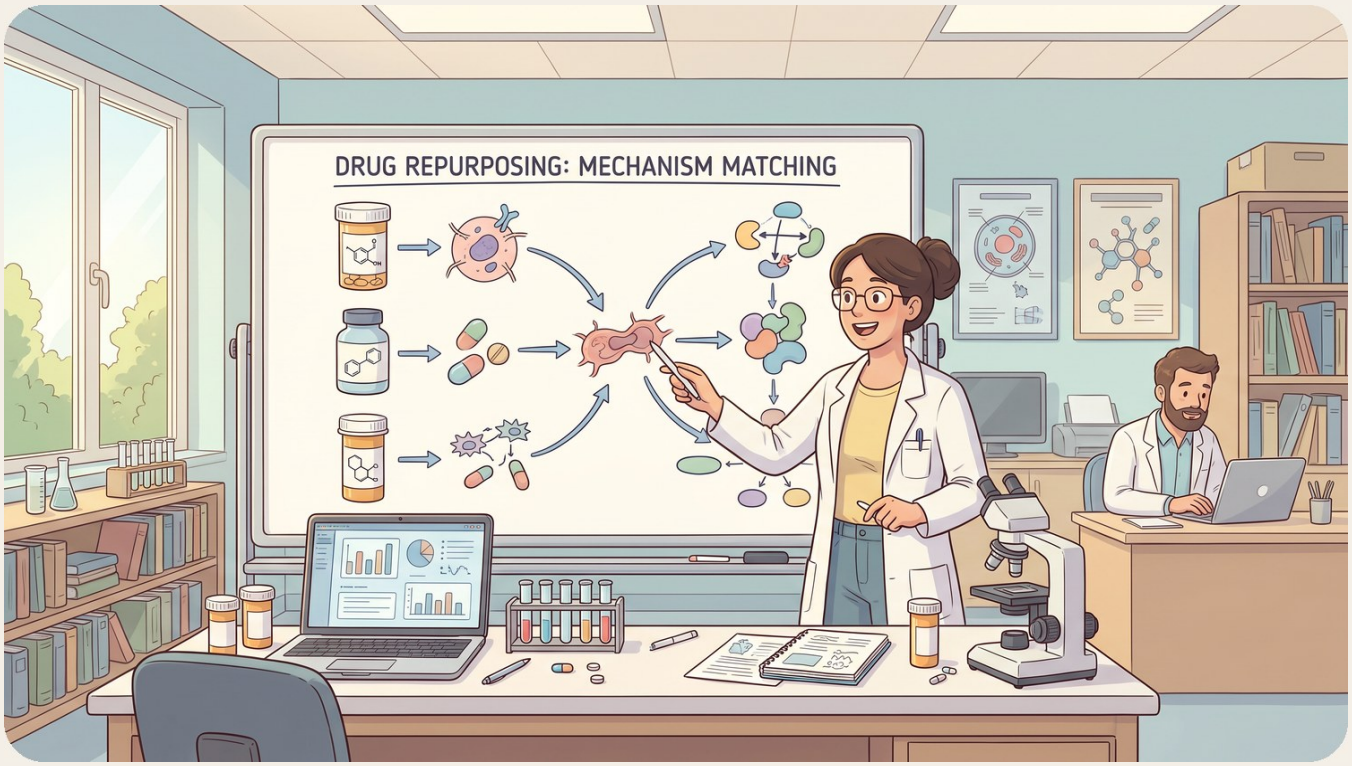
RUNS
Mixed

MODULES
Causal DAG builder; Causal calculus; Molecular report / dossier

OUTPUT
upstream causal validation

STRATEGIC DRUG-DISCOVERY QUESTIONS

Q45 Can BiRAGAS repurpose or reposition existing drugs by mechanism rather than trial-and-error?



HOW BIRAGAS HELPS

Yes - Perturbation analysis leverages DepMap and L1000 (drug-response) evidence over your causal targets, so you can connect a validated causal node to compounds that perturb it, prioritizing repositioning hypotheses by mechanism rather than screening blindly.

AUDIENCE
Drug Discovery / Translational

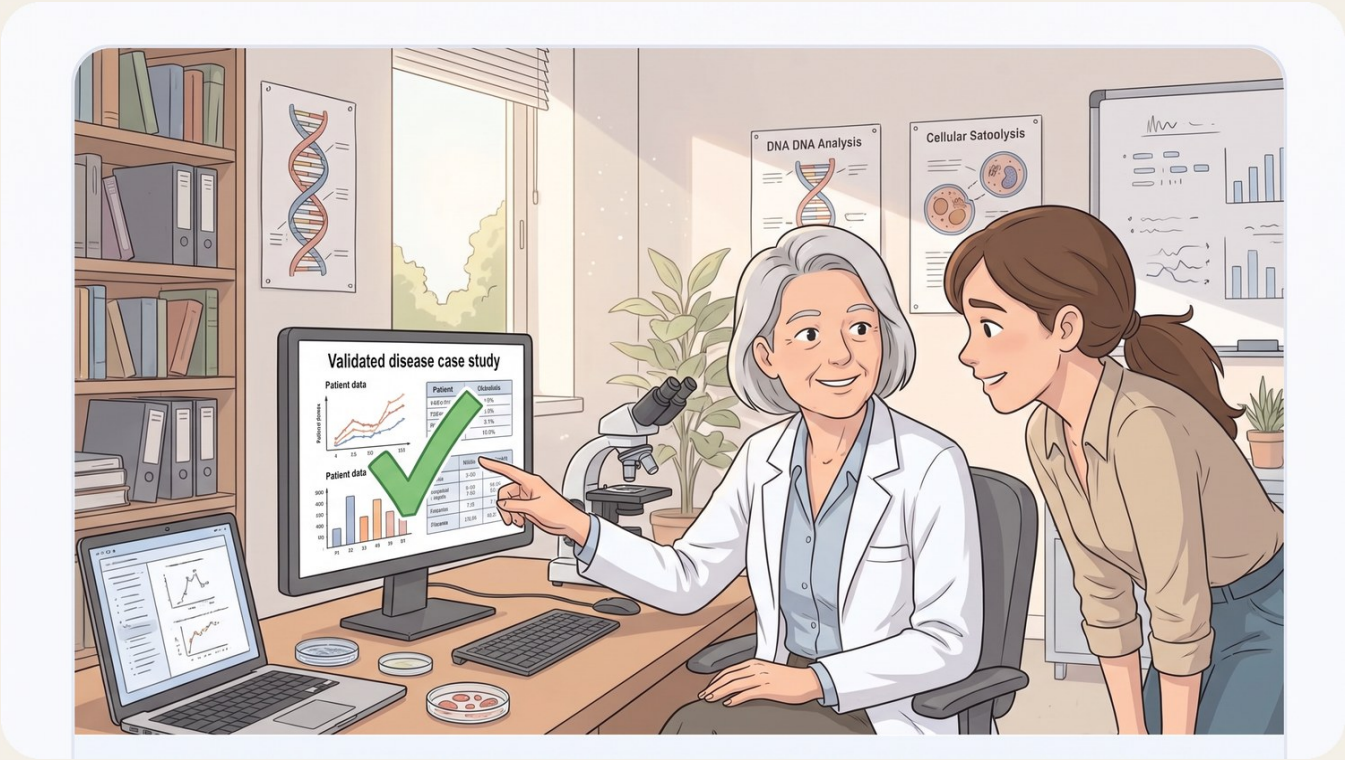
MODULES
Perturbation analysis; In silico perturbation (population)

RUNS
Full engine

OUTPUT
mechanism-ranked compounds

STRATEGIC DRUG-DISCOVERY QUESTIONS

Q46 I work on a specific disease (e.g. AML, melanoma, diabetic nephropathy). Has this been demonstrated?



HOW BIRAGAS HELPS

The methodology was validated retroactively on the BeatAML2 cohort (800+ patients): BiRAGAS called 8 of 11 AML drugs correctly and in the right context (bypass calls for venetoclax, FLT3, PLK1 matched patient observations), and surfaced novel targets that DE analysis alone missed. Active pilots/collaborations span melanoma (drug resistance), diabetic nephropathy, computational pathology and neurodegeneration.

AUDIENCE
Senior Scientist

RUNS
Full engine

MODULES
In silico perturbation (patient); Perturbation analysis

OUTPUT
validated case evidence

THEME

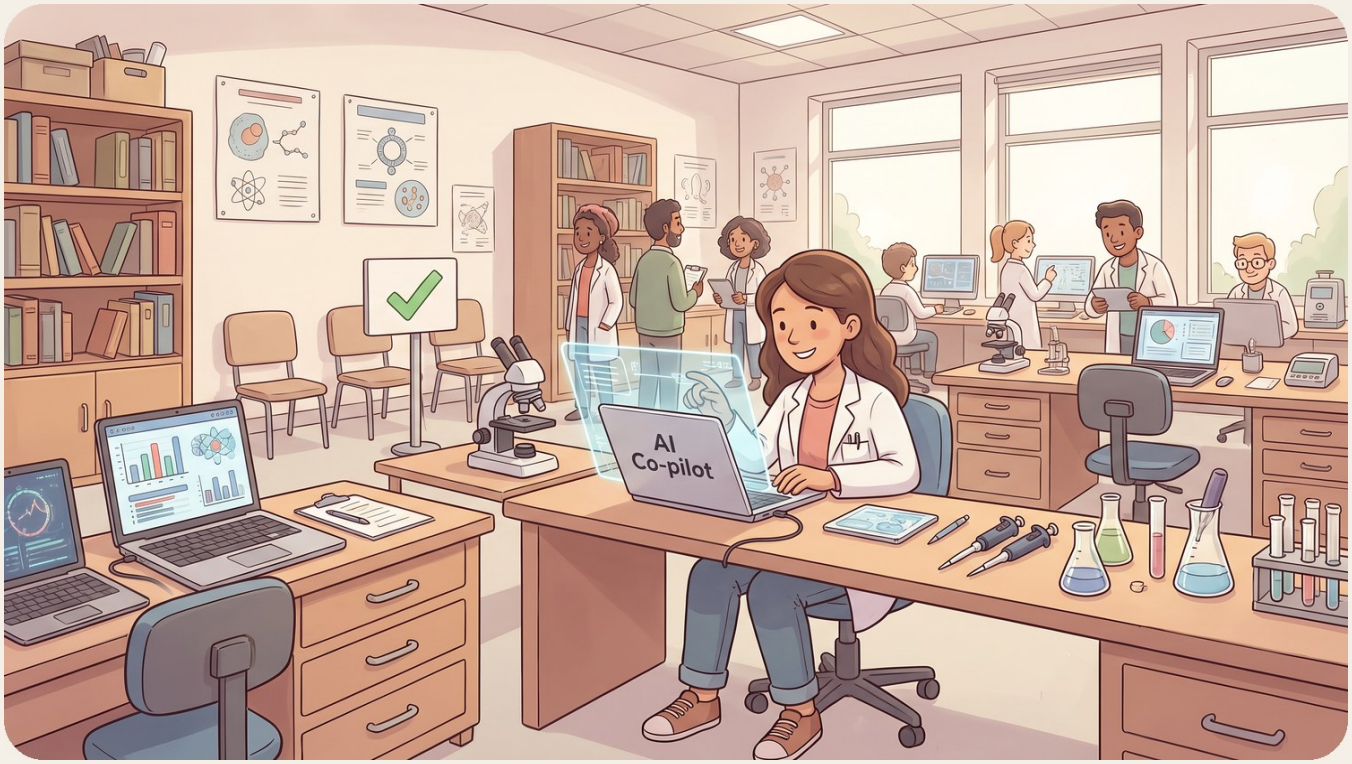
‹ CONTENTS

16 Common lab bottlenecks -> BiRAGAS solution

Questions Q47–Q51 · 5 in this theme

COMMON LAB BOTTLENECKS -> BIRAGAS SOLUTION

Q47 Bottleneck: every analysis needs a trained bioinformatician, so science waits on a queue of a few techs.



HOW BIRAGAS HELPS

Causion removes the operator bottleneck - any translational scientist can run the 28 modules in plain English, so throughput is limited by scientific questions, not by who can run a pipeline.

AUDIENCE
Bioinformatician

RUNS
In-app

MODULES
Causion co-pilot

OUTPUT
self-serve analysis

COMMON LAB BOTTLENECKS -> BIRAGAS SOLUTION

Q48 Bottleneck: results are association-only; I can't tell drivers from passengers.

HOW BIRAGAS HELPS

The causal layer (DAG + do-calculus + MR + in-silico perturbation) and Causal Confidence Scoring explicitly rank drivers vs passengers - the exact gap correlation tooling can't close.

AUDIENCE
Bioinformatician

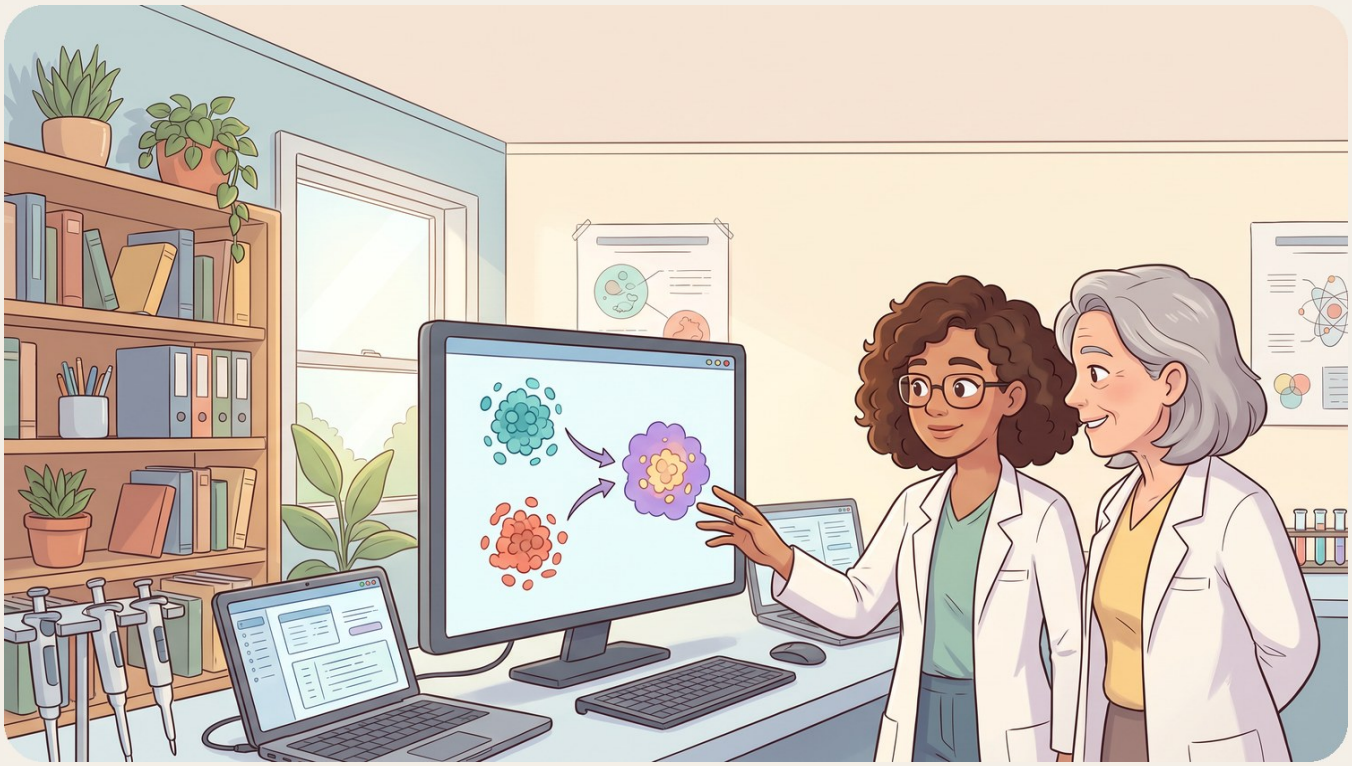
RUNS
Mixed

MODULES
Causal DAG builder; Causal calculus; GWAS / Mendelian randomization

OUTPUT
driver vs passenger ranking

COMMON LAB BOTTLENECKS -> BIRAGAS SOLUTION

Q49 Bottleneck: I can't reconcile conflicting single-tool results (DEG says one thing, network another).



HOW BIRAGAS HELPS

Multi-stream convergence fuses observational, temporal, genetic and interventional evidence into one graph and scores agreement - so conflicting single-tool outputs are reconciled into a single confidence-weighted answer instead of leaving you to arbitrate.

AUDIENCE
Senior Scientist

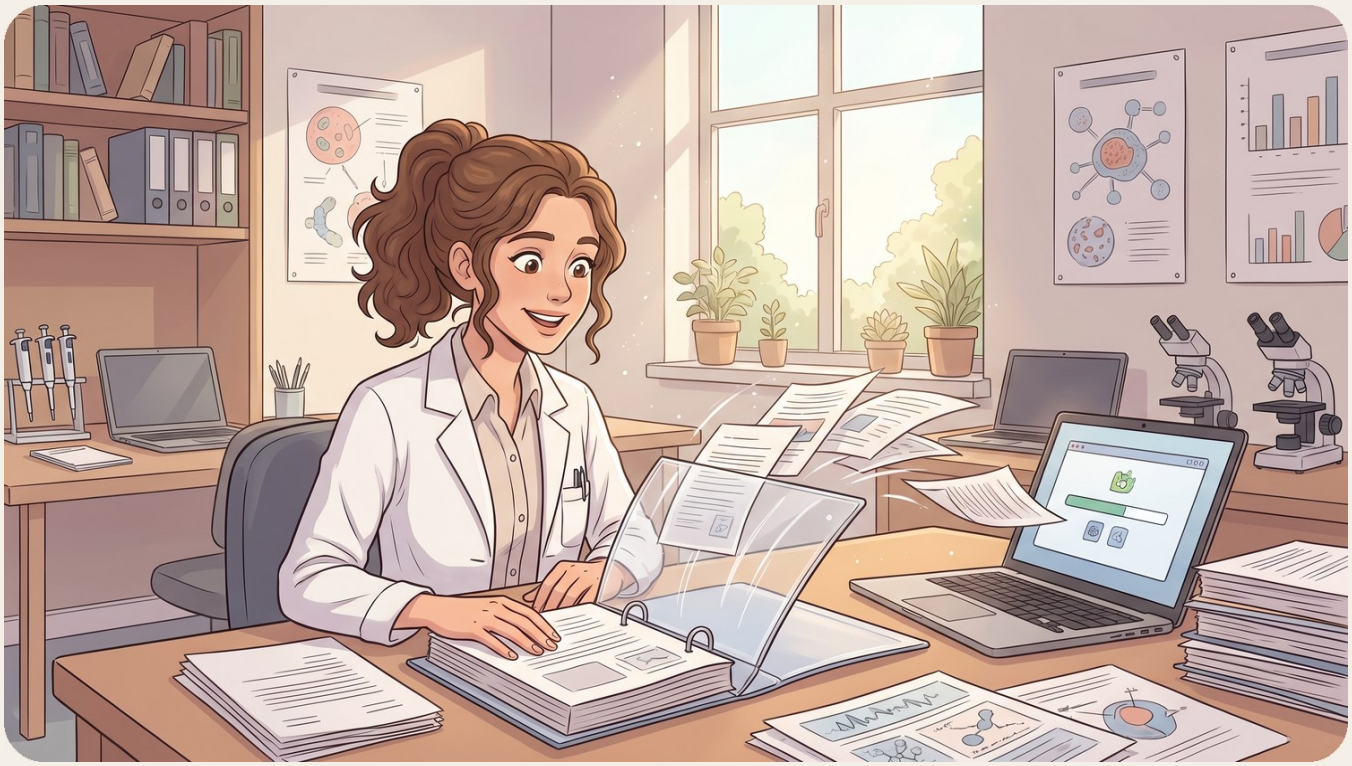
MODULES
Causal DAG builder

RUNS
In-app

OUTPUT
reconciled consensus DAG

COMMON LAB BOTTLENECKS -> BIRAGAS SOLUTION

Q50 Bottleneck: building the FDA regulatory-evidence layer by hand takes months per target.



HOW BIRAGAS HELPS

BiRAGAS emits the source-linked, FDA-aligned dossier automatically the moment a target is identified - turning a slow, manual regulatory-writing process into a software feature.

AUDIENCE
Drug Discovery / Translational

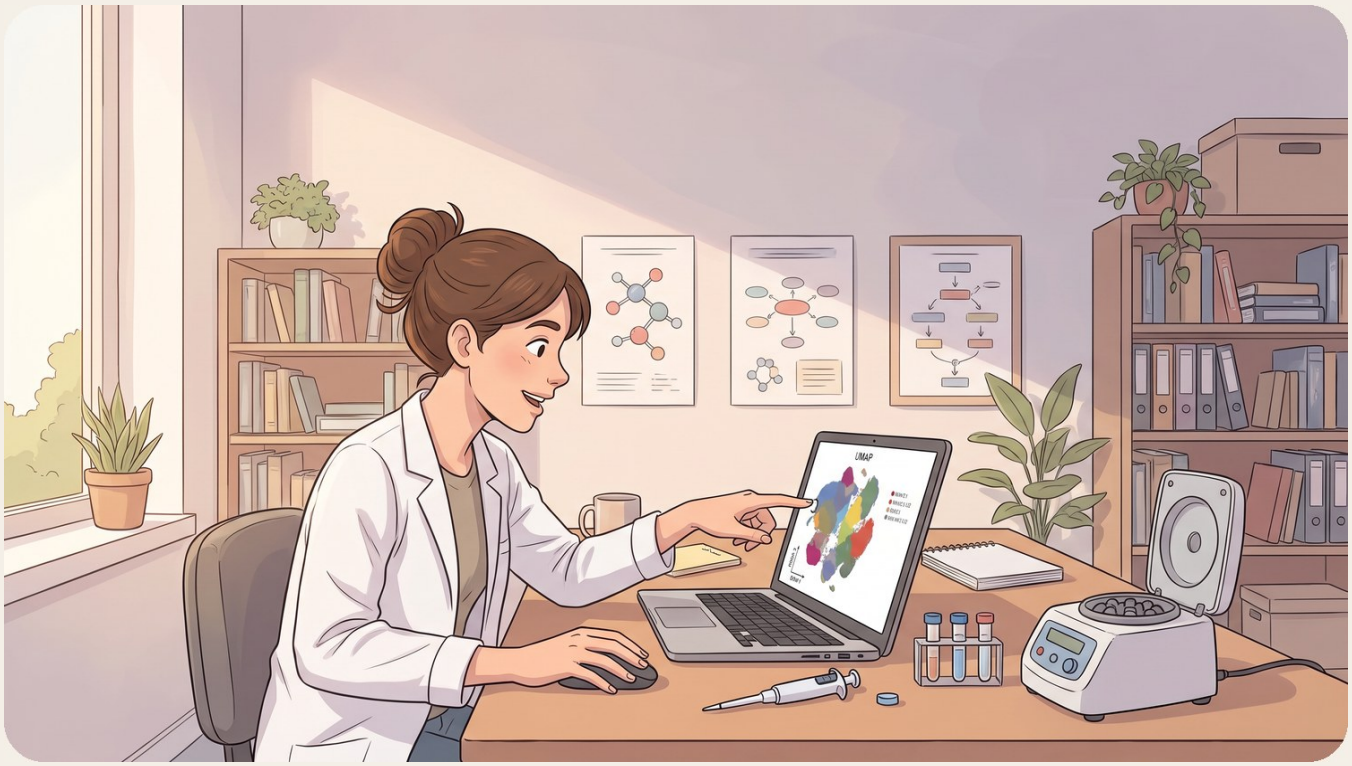
RUNS
Full engine

MODULES
Molecular report / dossier

OUTPUT
auto-generated dossier

COMMON LAB BOTTLENECKS -> BIRAGAS SOLUTION

Q51 Bottleneck: wet-lab validation (CRISPR, patient trials) is too slow and expensive to explore broadly.



HOW BIRAGAS HELPS

In-silico perturbation and in-silico CRISPR simulate knockout and patient-response 'what-if' trials in minutes, filtering paradoxical/inactive hypotheses so only the strongest candidates consume wet-lab and clinical resources.

AUDIENCE	RUNS
Senior Scientist	Full engine
MODULES	**OUTPUT**
Perturbation analysis; In silico perturbation (patient)	in-silico pre-screen

51 questions.
One causal engine.

From raw reads to validated causal targets — BiRAGAS answers the whole question, and tells you how confident it is.