SMASH
machine learning for science and humanities postdoctoral program

Behind the Scenes of AI Literature Search: From Evidence to Answers

Yun Wang

UNIVERZA V LJUBLJANI
Jožef Stefan Institute
IZUM
REPUBLIC OF SLOVENIA MINISTRY OF THE ENVIRONMENT, CLIMATE AND ENERGY SLOVENIAN ENVIRONMENT AGENCY

I FEEL SLOVENIA
Co-funded by the European Union

Have you ever used them for searching literature / writing text?







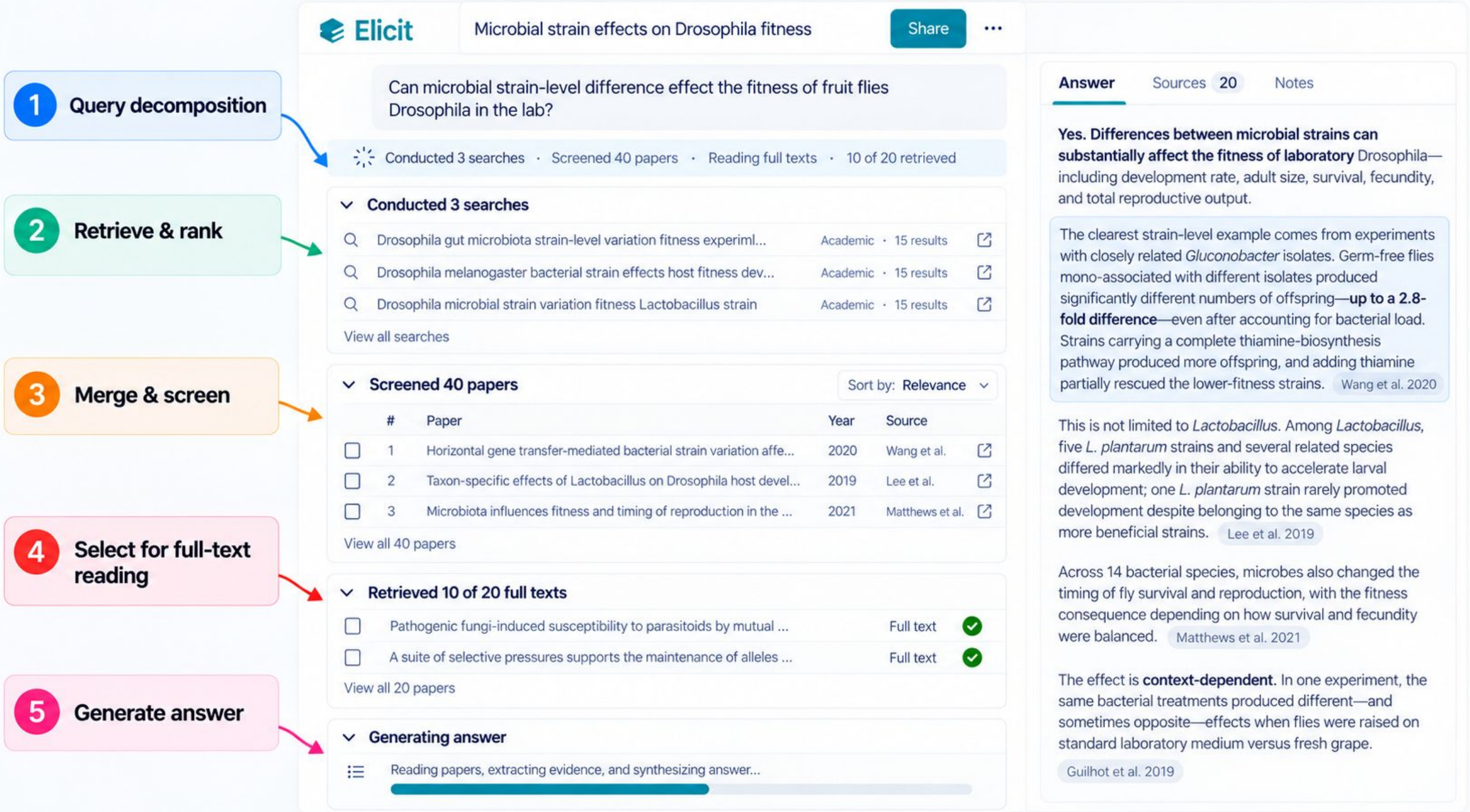











1 Query decomposition
2 Retrieve & rank
3 Merge & screen
4 Select for full-text reading
5 Generate answer
Elicit
Microbial strain effects on Drosophila fitness
Share
Can microbial strain-level difference effect the fitness of fruit flies Drosophila in the lab?
Conducted 3 searches · Screened 40 papers · Reading full texts · 10 of 20 retrieved
Conducted 3 searches
Drosophila gut microbiota strain-level variation fitness experiml... Academic · 15 results
Drosophila melanogaster bacterial strain effects host fitness dev... Academic · 15 results
Drosophila microbial strain variation fitness Lactobacillus strain Academic · 15 results
View all searches
Screened 40 papers
Sort by: Relevance
Paper Year Source
1 Horizontal gene transfer-mediated bacterial strain variation affe... 2020 Wang et al.
2 Taxon-specific effects of Lactobacillus on Drosophila host devel... 2019 Lee et al.
3 Microbiota influences fitness and timing of reproduction in the ... 2021 Matthews et al.
View all 40 papers
Retrieved 10 of 20 full texts
Pathogenic fungi-induced susceptibility to parasitoids by mutual ... Full text
A suite of selective pressures supports the maintenance of alleles ... Full text
View all 20 papers
Generating answer
Reading papers, extracting evidence, and synthesizing answer...
Answer
Sources 20
Notes
Yes. Differences between microbial strains can substantially affect the fitness of laboratory Drosophila—including development rate, adult size, survival, fecundity, and total reproductive output.
The clearest strain-level example comes from experiments with closely related Gluconobacter isolates. Germ-free flies mono-associated with different isolates produced significantly different numbers of offspring—up to a 2.8-fold difference—even after accounting for bacterial load. Strains carrying a complete thiamine-biosynthesis pathway produced more offspring, and adding thiamine partially rescued the lower-fitness strains. Wang et al. 2020
This is not limited to Lactobacillus. Among Lactobacillus, five L. plantarum strains and several related species differed markedly in their ability to accelerate larval development; one L. plantarum strain rarely promoted development despite belonging to the same species as more beneficial strains. Lee et al. 2019
Across 14 bacterial species, microbes also changed the timing of fly survival and reproduction, with the fitness consequence depending on how survival and fecundity were balanced. Matthews et al. 2021
The effect is context-dependent. In one experiment, the same bacterial treatments produced different—and sometimes opposite—effects when flies were raised on standard laboratory medium versus fresh grape.
Guilhot et al. 2019



BioASQ

TREC RAG

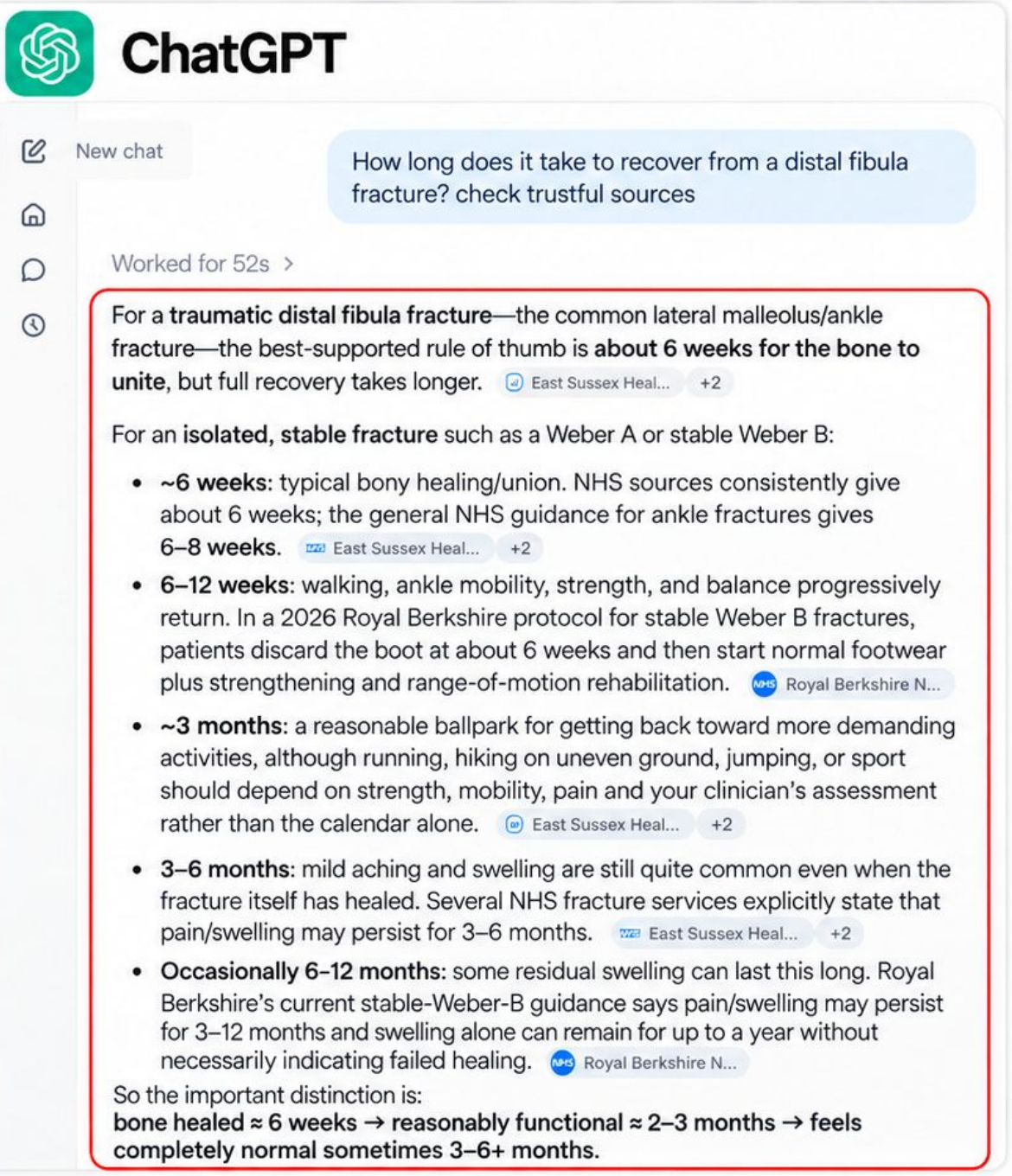
TREC 2026 · AGENT-FIRST

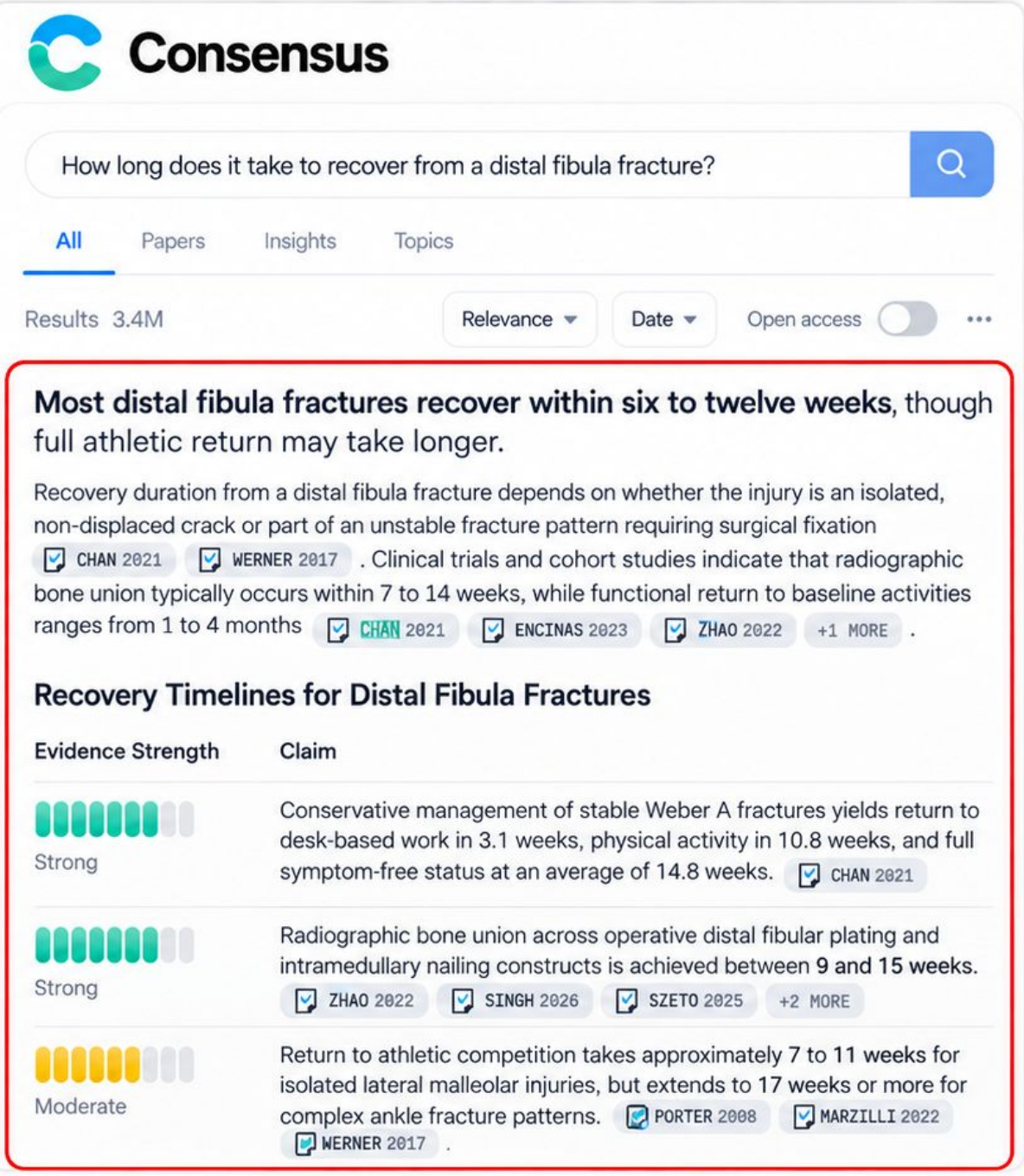
Retrieval-Augmented Generation Track

A TREC (NIST) benchmark for systems that combine retrieval with large language models.

- trec-rag.github.io

SMASH
machine learning for science and humanities postdoctoral program





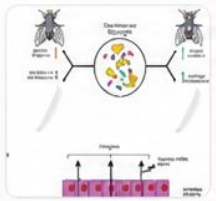
Adapted from screenshots of ChatGPT and Consensus

ChatGPT

Can microbial strain-level difference affect the fitness of fruit flies Drosophila in the lab?

Worked for 1m 37s >

Yes. Microbial strain-level variation can affect *Drosophila melanogaster* fitness in laboratory experiments, and there are some unusually clean demonstrations of this. The effects can appear in offspring number, developmental rate, body size, fecundity, lifespan, and nutritional performance.



Horizontal gene transfer-mediated bacterial strain variation affects host fitness in Drosophila

Wang et al. 2021 · BMC Biology

Yun Wang et al.

Mono-associated flies with closely related *Gluconobacter* isolates differed strongly in offspring production...

Other relevant studies

Consensus

Can microbial strain-level difference affect the fitness of fruit flies Drosophila in the lab?

All | Papers | Insights | Topics

Results 2.8M | Relevance | Date | Open access

1 **Horizontal gene transfer-mediated bacterial strain variation affects host fitness in Drosophila**

Wang et al. 2021 · BMC Biology

Yun Wang et al.

The thiamine biosynthesis pathway contributes to variation in offspring production in Drosophila, potentially enhancing host adaptation.

PDF | 4 quotes | Cross-sectional study

2 **Genetic Influences of the Microbiota on the Life Span of Drosophila melanogaster**

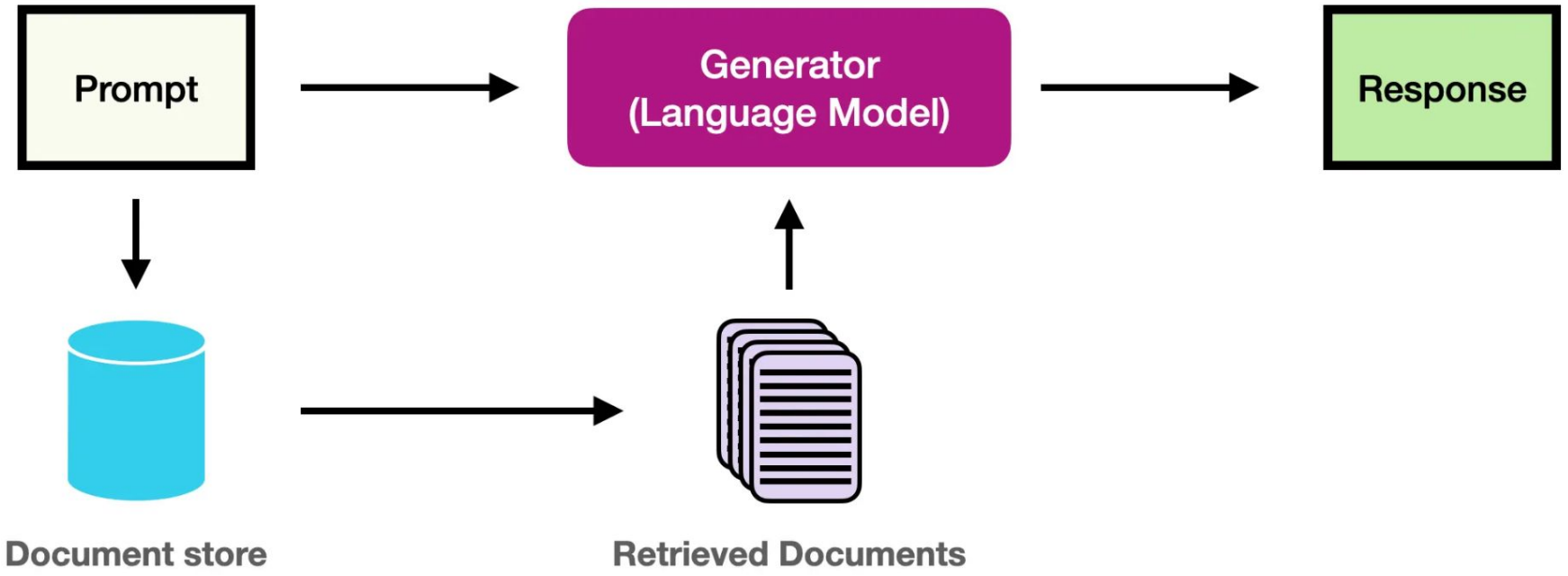
2020 · Applied and Environmental Microbiology

Bacterial genes influence fruit fly longevity, with methionine restriction and glucose and vitamin B6 metabolism playing stronger roles than methionine metabolism.

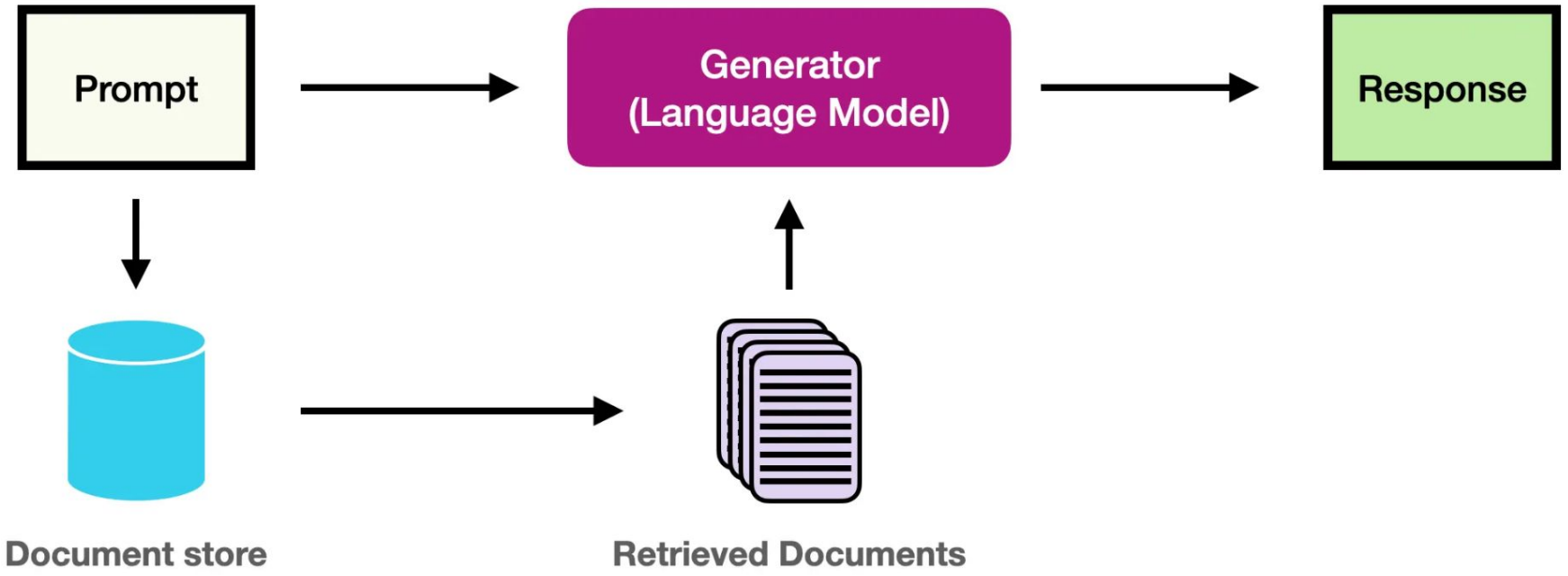
PDF | 6 quotes | Non-randomized experiment

Adapted from screenshots of ChatGPT and Consensus

Retrieval-Augmented Generation (RAG)



Retrieval-Augmented Generation (RAG)



ChatGPT Claude

consensus Elicit Asta

What is the biggest difference between chatGPT and consensus?

Where do the search?

ChatGPT

Claude

→ the broader web

---

consensus

Elicit

Asta

→ a defined literature corpus



Document store

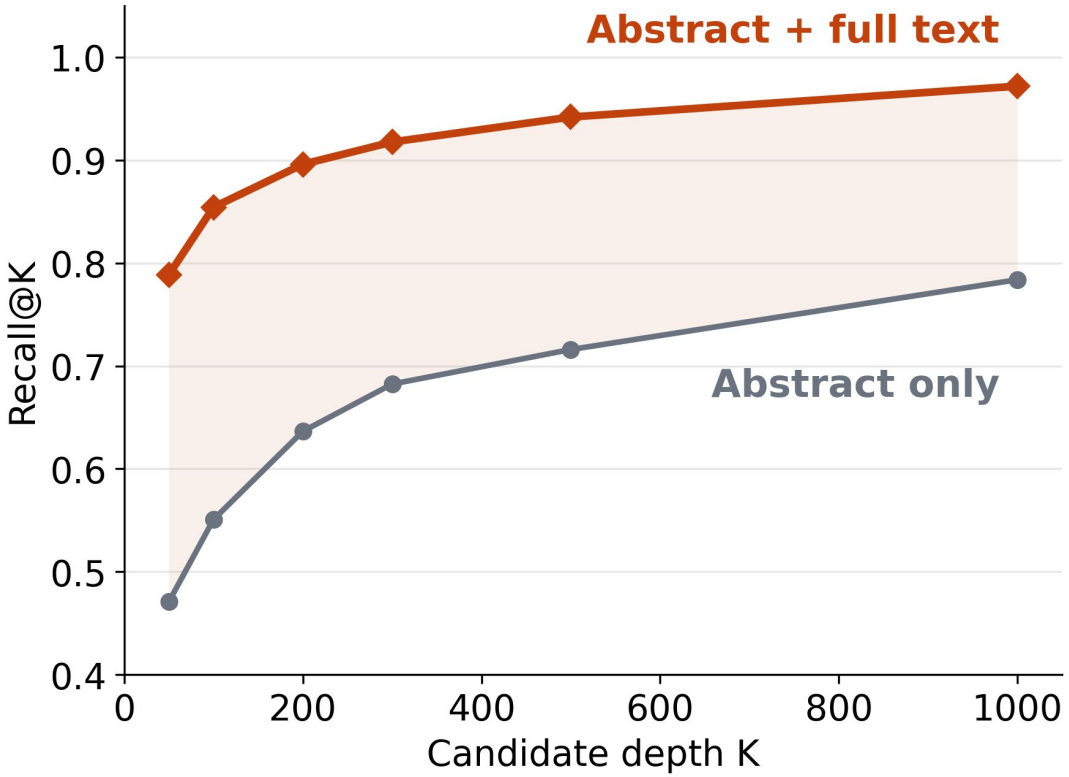
PubMed

bioRxiv

arXiv

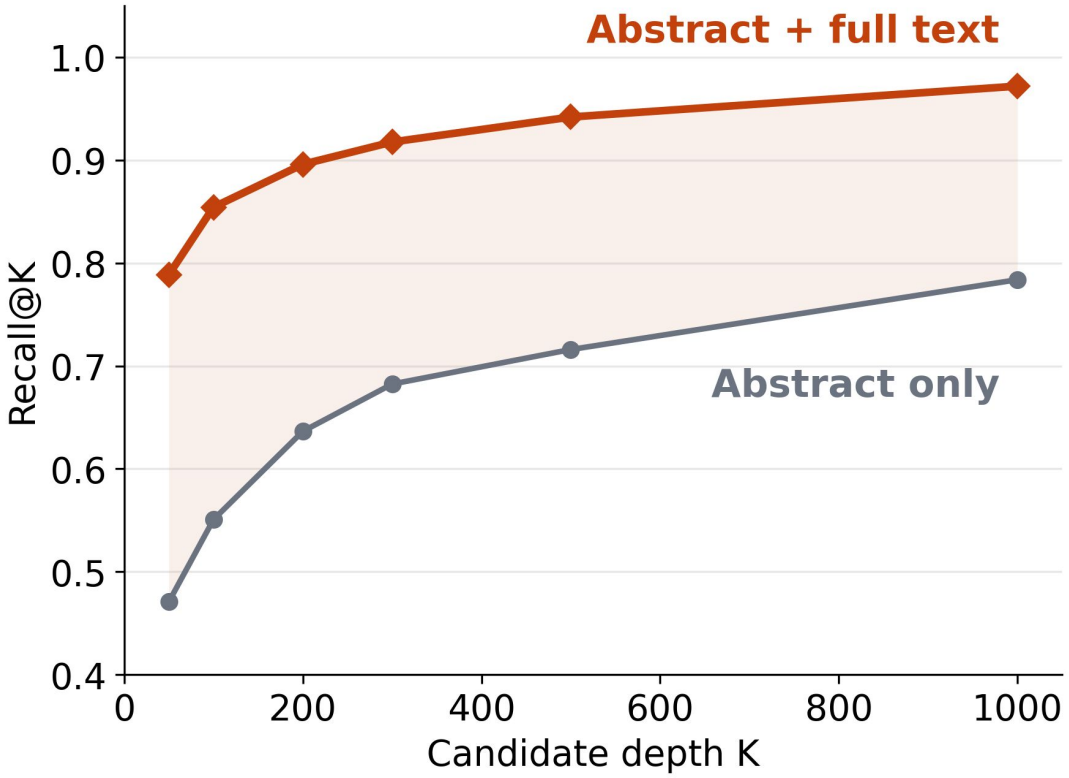
Semantic Scholar

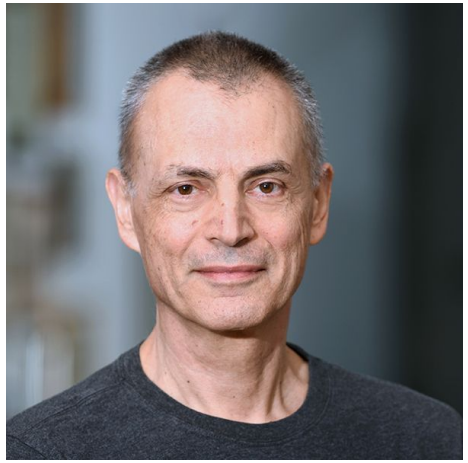
Abstracts can miss relevant evidence



Adapted from Y. Wang, G. Shaulsky, T. Curk, B. Zupan, Benchmarking Literature Retrieval for a Model Organism: A Dictyostelium Case Study, accepted at Discovery Science 2026 (LNCS, Springer).

Abstracts can miss relevant evidence



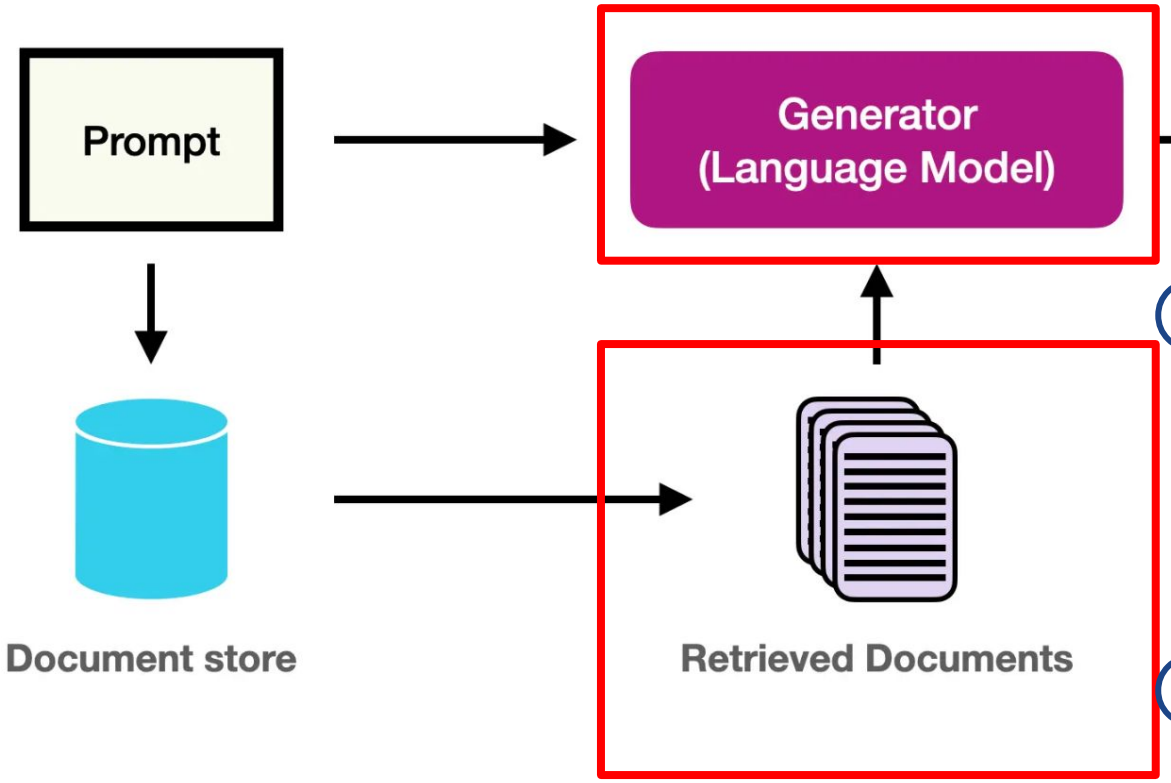


Prof. Gad Shaulsky

Tomorrow 9:45 - Deciphering Dictyostelium Biology: From Genetic Pathways to AI Systems for Knowledge and Discovery

Adapted from Y. Wang, G. Shaulsky, T. Curk, B. Zupan, Benchmarking Literature Retrieval for a Model Organism: A Dictyostelium Case Study, accepted at Discovery Science 2026 (LNCS, Springer).

Two key stages of RAG

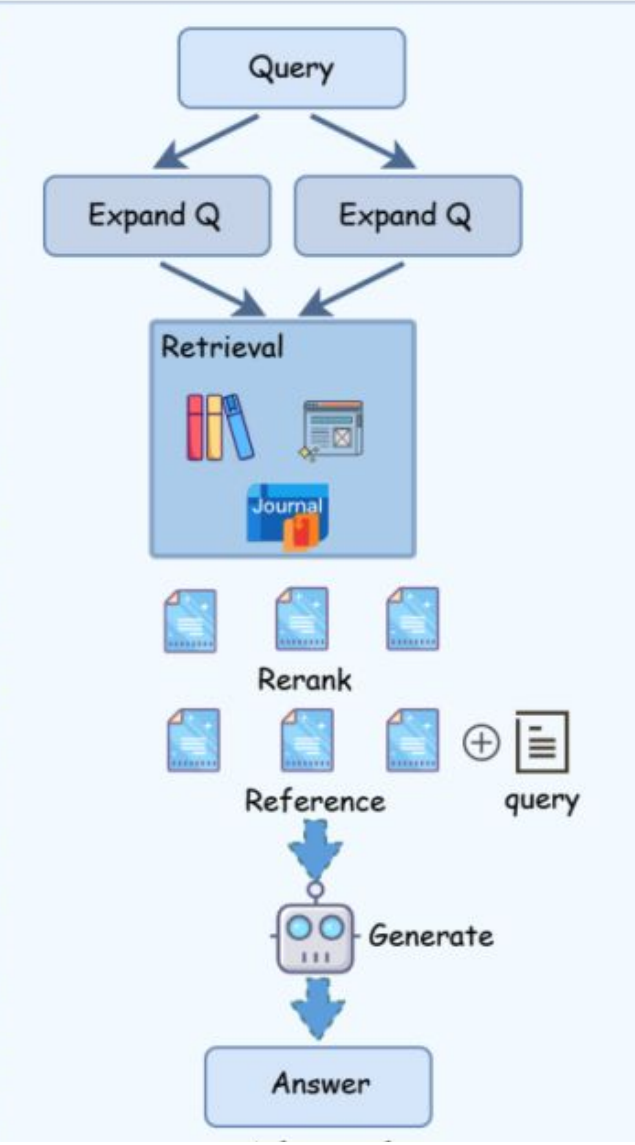


② Generate answer

① Find evidence:

Retrieval + Reranking

① Find evidence: retrieval and reranking



EXAMPLE

200 candidate papers

Found by retrieval

Rerank

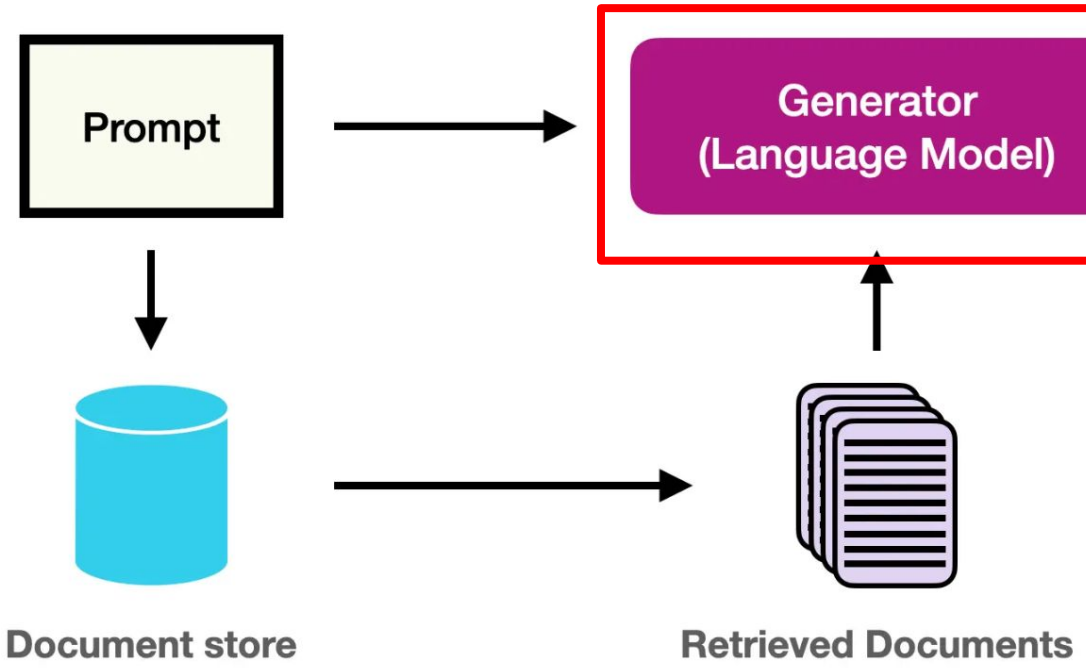
Assess relevance more closely

Top 20 for generation

Illustrative numbers

He et al.,arXiv:2505.01146 (2025).

② Generation: Turning Evidence into Answers

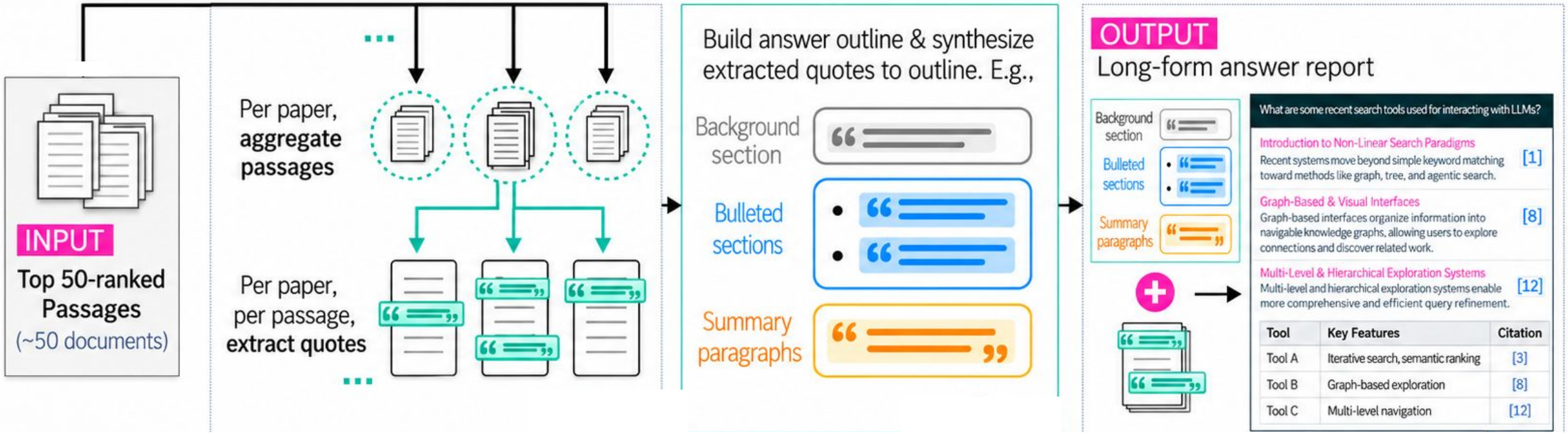


Turning evidence into an answer

- Which LLM?
- What instructions?
- One step or multiple steps?

Asta

ScholarQA
A Allen AI project



1 **Extract evidence**

2 **Organize content**

3 **Write the report**

Key idea: generation can be a pipeline, not just one prompt.

Evaluation: Important, but Often Overlooked



Evaluation: Important, but Often Overlooked



Did we find the right evidence?

Did we generate a good answer?

Retrieval evaluation

Established metrics

Recall

Did we find the relevant papers?

Precision

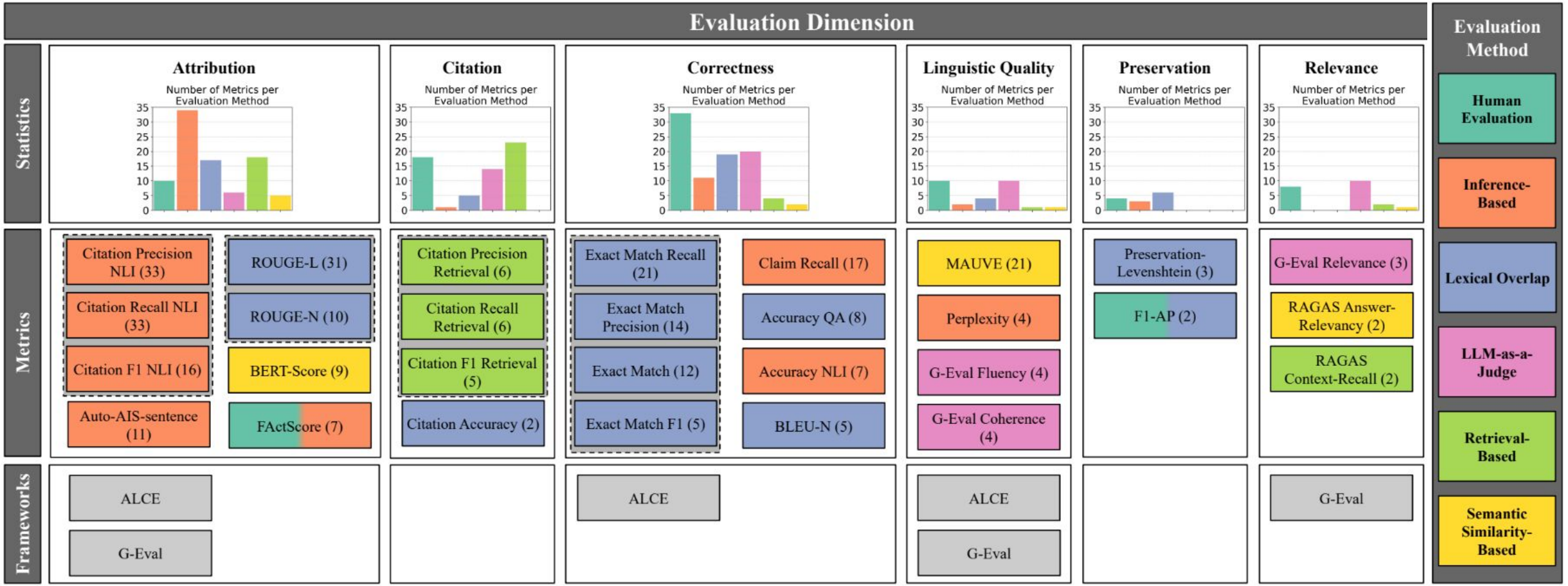
How many results are relevant?

Ranking

Are the best papers near the top?

Generation evaluation

Different methods capture different aspects of answer quality



Schreieder et al., ACL 2026.

Generation evaluation

No single agreed standard

Accuracy

Factual correctness

Coverage

Key information included

Citations

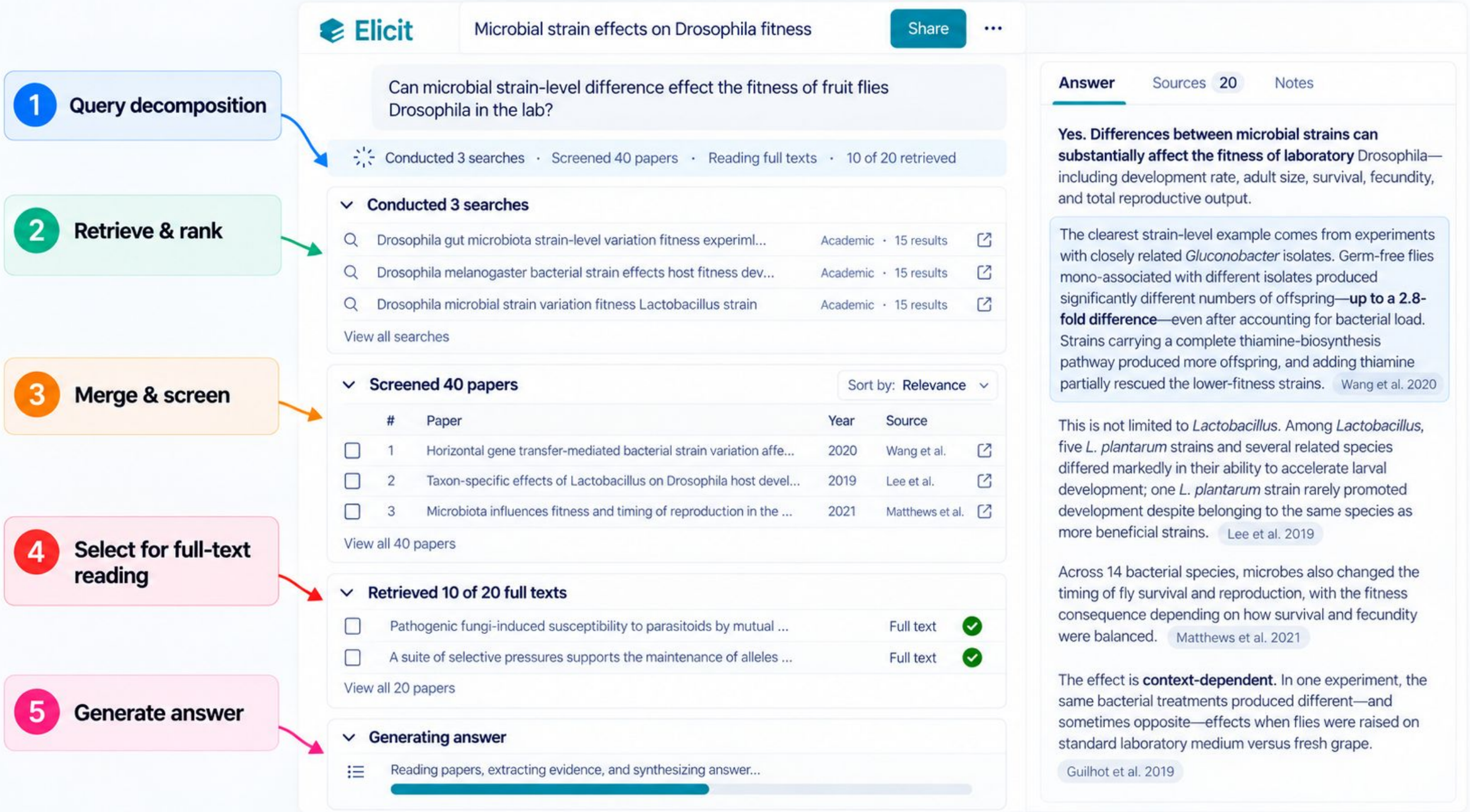
Claims supported by evidence

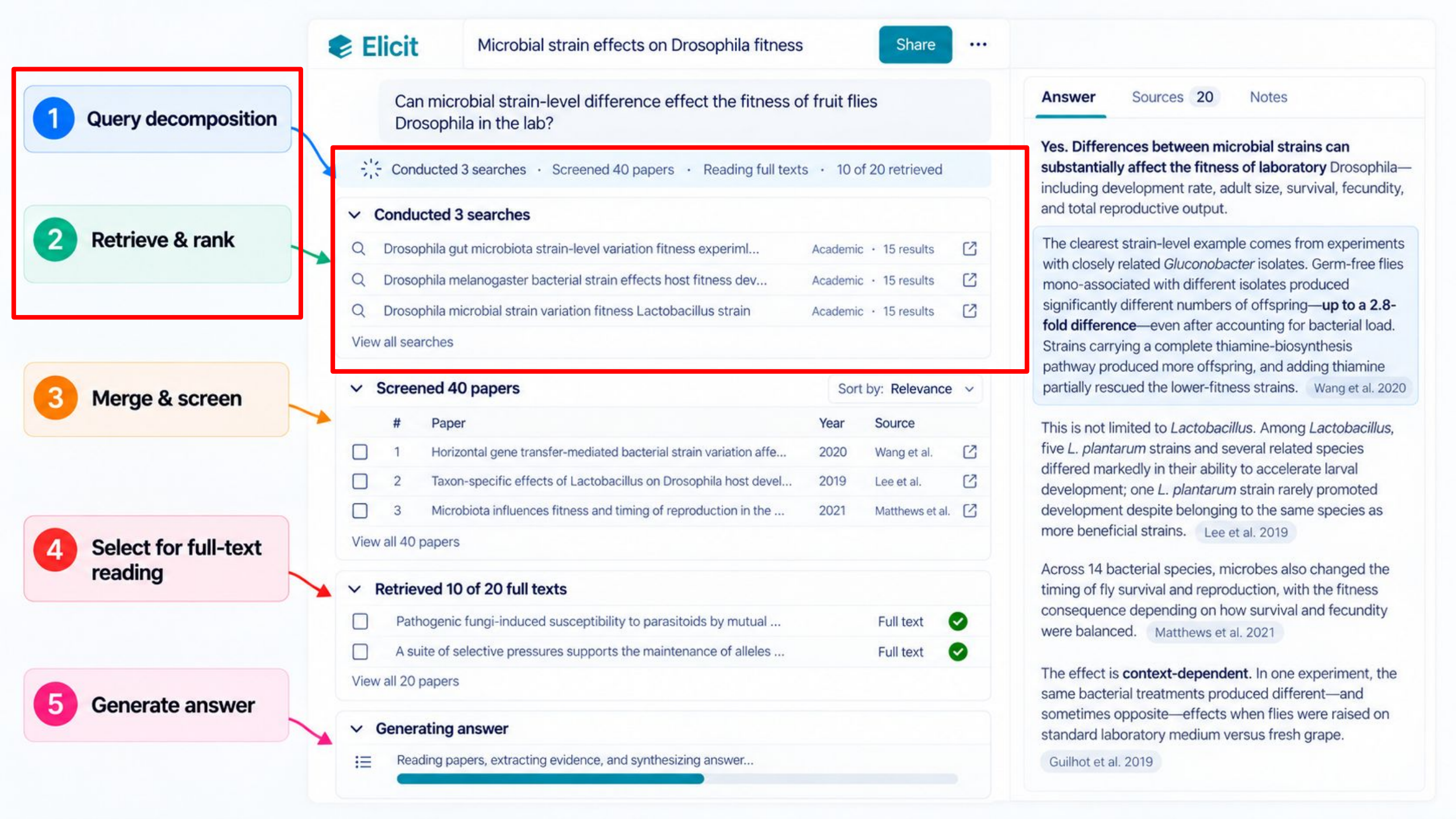
AI **LLM judges**
Increasingly used

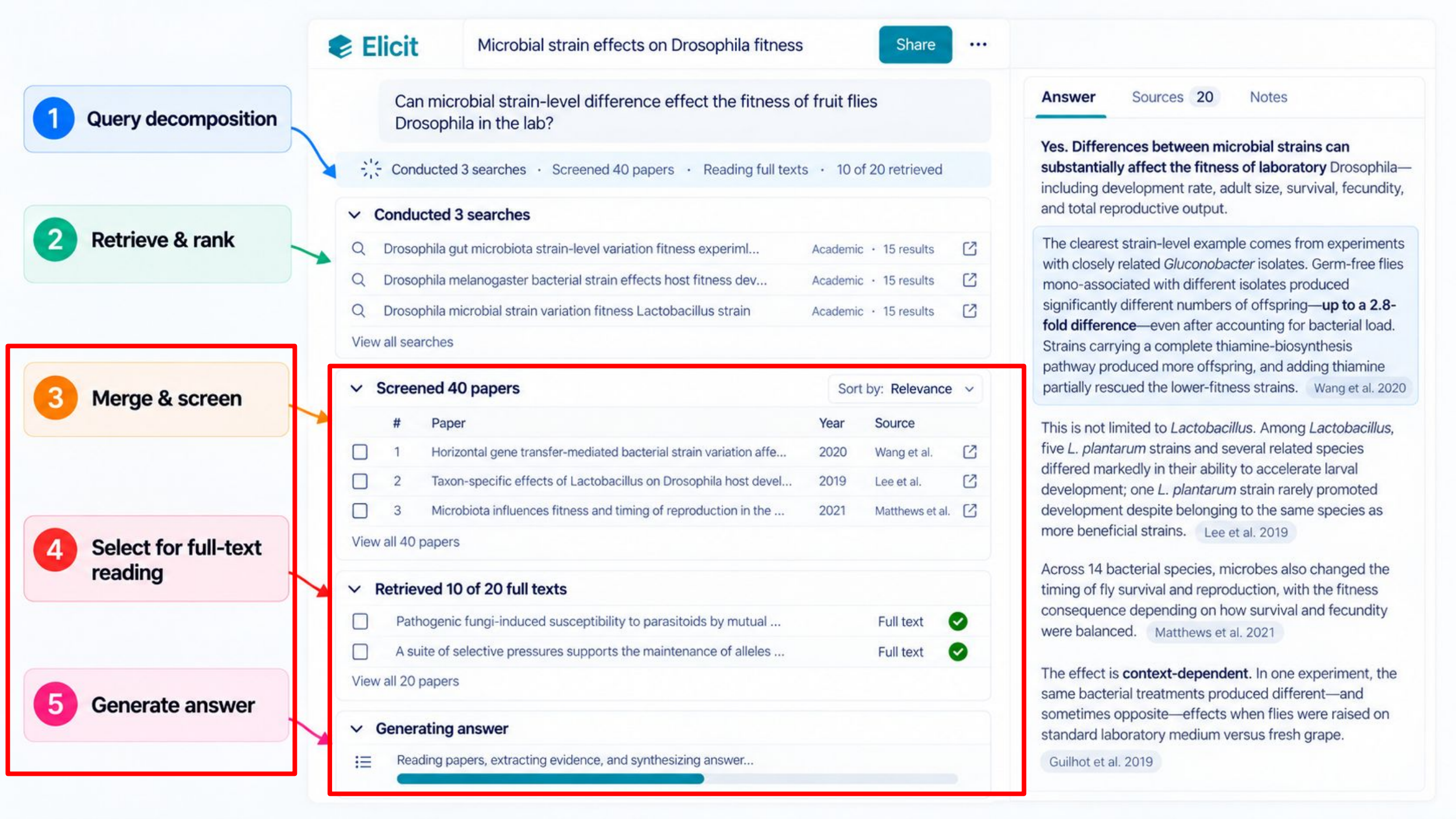
Human evaluation
Still the gold standard

Inside AI literature tools




1 Query decomposition
2 Retrieve & rank
3 Merge & screen
4 Select for full-text reading
5 Generate answer
Elicit
Microbial strain effects on Drosophila fitness
Share
Can microbial strain-level difference effect the fitness of fruit flies Drosophila in the lab?
Conducted 3 searches · Screened 40 papers · Reading full texts · 10 of 20 retrieved
Conducted 3 searches
Drosophila gut microbiota strain-level variation fitness experiml... Academic · 15 results
Drosophila melanogaster bacterial strain effects host fitness dev... Academic · 15 results
Drosophila microbial strain variation fitness Lactobacillus strain Academic · 15 results
View all searches
Screened 40 papers
Sort by: Relevance
Paper Year Source
1 Horizontal gene transfer-mediated bacterial strain variation affe... 2020 Wang et al.
2 Taxon-specific effects of Lactobacillus on Drosophila host devel... 2019 Lee et al.
3 Microbiota influences fitness and timing of reproduction in the ... 2021 Matthews et al.
View all 40 papers
Retrieved 10 of 20 full texts
Pathogenic fungi-induced susceptibility to parasitoids by mutual ... Full text
A suite of selective pressures supports the maintenance of alleles ... Full text
View all 20 papers
Generating answer
Reading papers, extracting evidence, and synthesizing answer...
Answer
Sources 20
Notes
Yes. Differences between microbial strains can substantially affect the fitness of laboratory Drosophila—including development rate, adult size, survival, fecundity, and total reproductive output.
The clearest strain-level example comes from experiments with closely related Gluconobacter isolates. Germ-free flies mono-associated with different isolates produced significantly different numbers of offspring—up to a 2.8-fold difference—even after accounting for bacterial load. Strains carrying a complete thiamine-biosynthesis pathway produced more offspring, and adding thiamine partially rescued the lower-fitness strains. Wang et al. 2020
This is not limited to Lactobacillus. Among Lactobacillus, five L. plantarum strains and several related species differed markedly in their ability to accelerate larval development; one L. plantarum strain rarely promoted development despite belonging to the same species as more beneficial strains. Lee et al. 2019
Across 14 bacterial species, microbes also changed the timing of fly survival and reproduction, with the fitness consequence depending on how survival and fecundity were balanced. Matthews et al. 2021
The effect is context-dependent. In one experiment, the same bacterial treatments produced different—and sometimes opposite—effects when flies were raised on standard laboratory medium versus fresh grape.
Guilhot et al. 2019


1 Query decomposition
2 Retrieve & rank
3 Merge & screen
4 Select for full-text reading
5 Generate answer
Elicit
Microbial strain effects on Drosophila fitness
Share
Can microbial strain-level difference effect the fitness of fruit flies Drosophila in the lab?
Conducted 3 searches · Screened 40 papers · Reading full texts · 10 of 20 retrieved
Conducted 3 searches
Drosophila gut microbiota strain-level variation fitness experiml... Academic · 15 results
Drosophila melanogaster bacterial strain effects host fitness dev... Academic · 15 results
Drosophila microbial strain variation fitness Lactobacillus strain Academic · 15 results
View all searches
Screened 40 papers
Sort by: Relevance
Paper Year Source
1 Horizontal gene transfer-mediated bacterial strain variation affe... 2020 Wang et al.
2 Taxon-specific effects of Lactobacillus on Drosophila host devel... 2019 Lee et al.
3 Microbiota influences fitness and timing of reproduction in the ... 2021 Matthews et al.
View all 40 papers
Retrieved 10 of 20 full texts
Pathogenic fungi-induced susceptibility to parasitoids by mutual ... Full text
A suite of selective pressures supports the maintenance of alleles ... Full text
View all 20 papers
Generating answer
Reading papers, extracting evidence, and synthesizing answer...
Answer
Sources 20
Notes
Yes. Differences between microbial strains can substantially affect the fitness of laboratory Drosophila—including development rate, adult size, survival, fecundity, and total reproductive output.
The clearest strain-level example comes from experiments with closely related Gluconobacter isolates. Germ-free flies mono-associated with different isolates produced significantly different numbers of offspring—up to a 2.8-fold difference—even after accounting for bacterial load. Strains carrying a complete thiamine-biosynthesis pathway produced more offspring, and adding thiamine partially rescued the lower-fitness strains. Wang et al. 2020
This is not limited to Lactobacillus. Among Lactobacillus, five L. plantarum strains and several related species differed markedly in their ability to accelerate larval development; one L. plantarum strain rarely promoted development despite belonging to the same species as more beneficial strains. Lee et al. 2019
Across 14 bacterial species, microbes also changed the timing of fly survival and reproduction, with the fitness consequence depending on how survival and fecundity were balanced. Matthews et al. 2021
The effect is context-dependent. In one experiment, the same bacterial treatments produced different—and sometimes opposite—effects when flies were raised on standard laboratory medium versus fresh grape.
Guilhot et al. 2019


1 Query decomposition
2 Retrieve & rank
3 Merge & screen
4 Select for full-text reading
5 Generate answer
Elicit
Microbial strain effects on Drosophila fitness
Share
Can microbial strain-level difference effect the fitness of fruit flies Drosophila in the lab?
Conducted 3 searches · Screened 40 papers · Reading full texts · 10 of 20 retrieved
Conducted 3 searches
Drosophila gut microbiota strain-level variation fitness experiml... Academic · 15 results
Drosophila melanogaster bacterial strain effects host fitness dev... Academic · 15 results
Drosophila microbial strain variation fitness Lactobacillus strain Academic · 15 results
View all searches
Screened 40 papers
Sort by: Relevance
Paper Year Source
1 Horizontal gene transfer-mediated bacterial strain variation affe... 2020 Wang et al.
2 Taxon-specific effects of Lactobacillus on Drosophila host devel... 2019 Lee et al.
3 Microbiota influences fitness and timing of reproduction in the ... 2021 Matthews et al.
View all 40 papers
Retrieved 10 of 20 full texts
Pathogenic fungi-induced susceptibility to parasitoids by mutual ... Full text
A suite of selective pressures supports the maintenance of alleles ... Full text
View all 20 papers
Generating answer
Reading papers, extracting evidence, and synthesizing answer...
Answer
Sources 20
Notes
Yes. Differences between microbial strains can substantially affect the fitness of laboratory Drosophila—including development rate, adult size, survival, fecundity, and total reproductive output.
The clearest strain-level example comes from experiments with closely related Gluconobacter isolates. Germ-free flies mono-associated with different isolates produced significantly different numbers of offspring—up to a 2.8-fold difference—even after accounting for bacterial load. Strains carrying a complete thiamine-biosynthesis pathway produced more offspring, and adding thiamine partially rescued the lower-fitness strains. Wang et al. 2020
This is not limited to Lactobacillus. Among Lactobacillus, five L. plantarum strains and several related species differed markedly in their ability to accelerate larval development; one L. plantarum strain rarely promoted development despite belonging to the same species as more beneficial strains. Lee et al. 2019
Across 14 bacterial species, microbes also changed the timing of fly survival and reproduction, with the fitness consequence depending on how survival and fecundity were balanced. Matthews et al. 2021
The effect is context-dependent. In one experiment, the same bacterial treatments produced different—and sometimes opposite—effects when flies were raised on standard laboratory medium versus fresh grape.
Guilhot et al. 2019

Summary



Corpus

Coverage matters



Retrieval & reranking

Find the right evidence



Generation

No single recipe



Evaluation

Check claims & citations

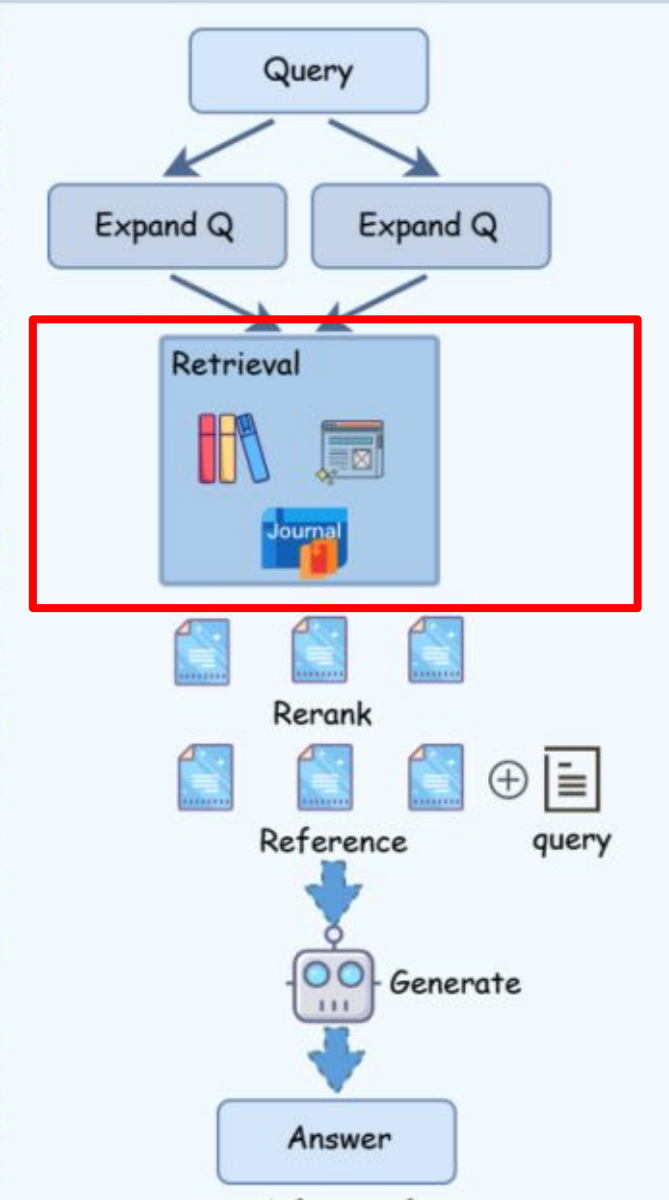
Thank you!

I FEEL SLOVENIA

Co-funded by the European Union

SMASH
machine learning for science and humanities postdoctoral program

①ⓐ Retrieval: finding candidate papers through prebuilt indexes



Hybrid Retrieval combines

- **Keyword search:** matches words

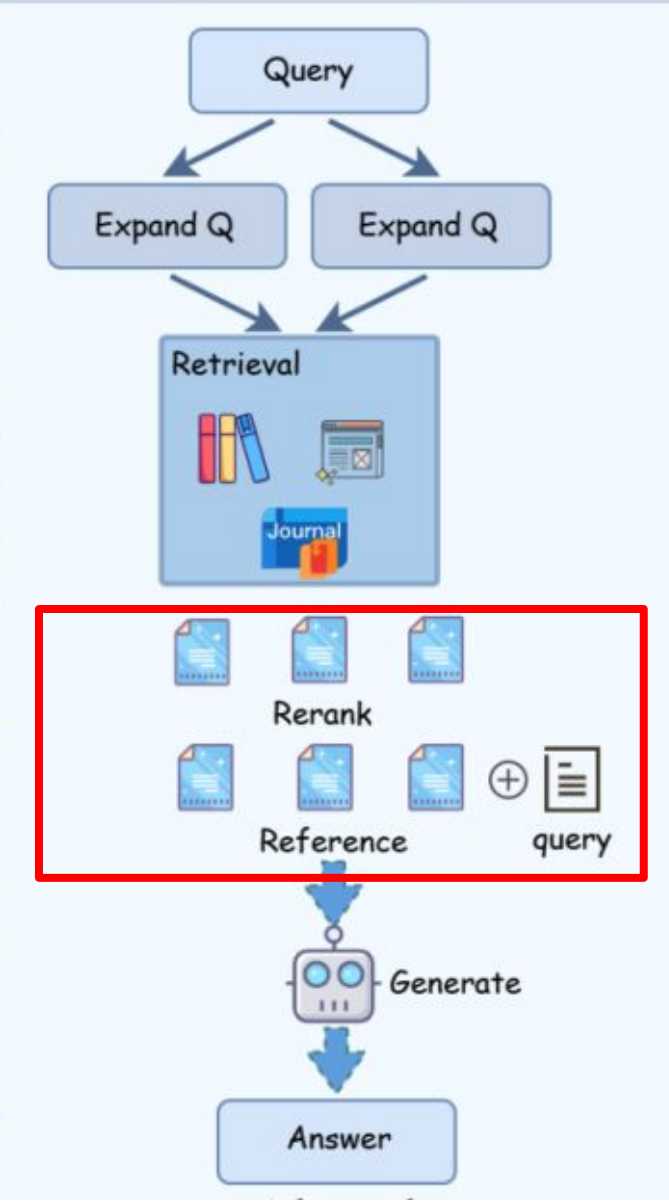
  e.g acute myocardial infarction -> acute myocardial infarction

- **Semantic search:** matches meaning

  e.g acute myocardial infarction <-> heart attack

He et al.,arXiv:2505.01146 (2025).

①ⓑ Reranking: bringing the most relevant papers to the top



EXAMPLE

100 candidate papers

Found by retrieval

Rerank

Assess relevance more closely

Top 20 for generation

He et al.,arXiv:2505.01146 (2025).

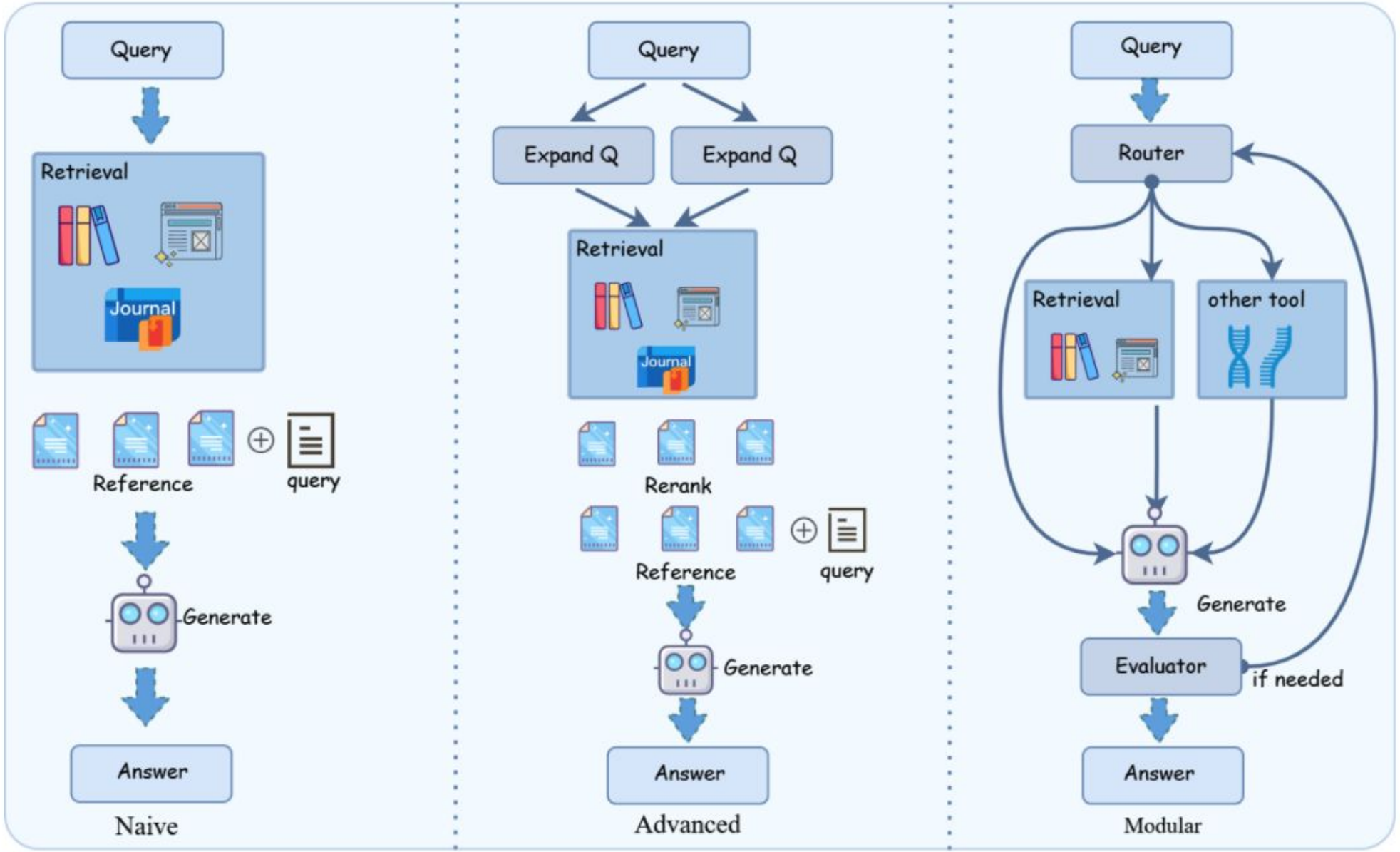
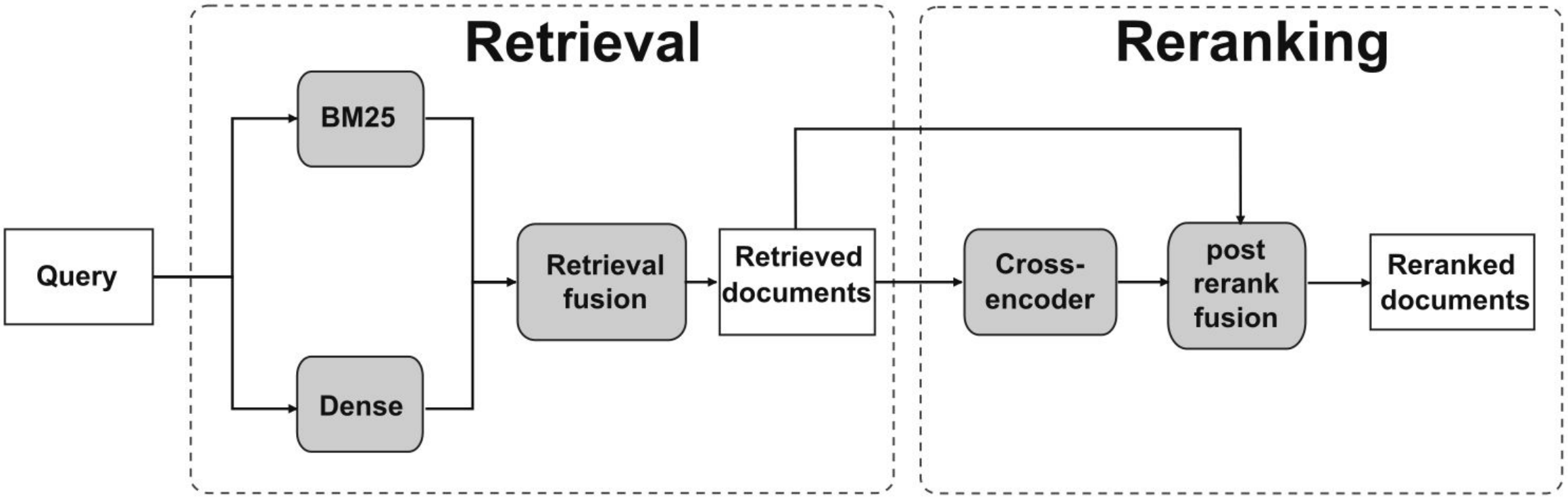


Figure 1: Basic biomedical RAG framework.

SMASH
machine learning for science and humanities postdoctoral program

Two-stage retrieval: retrieval and reranking

Our pipeline follows this design



github.com/fulaibaowang/RAG-scripts


consensus


Asta


Claude


Elicit


ChatGPT


bioRχiv


PubMed


arXiv


Semantic Scholar


SMASH
machine learning for science and humanities
postdoctoral program

1. **Query rewriting**
2. **Retrieve at scale**
3. **Rank + eligibility filter**
4. **Read included evidence**
5. **Extract and synthesize**

Consensus

References

Can microbial strain-level difference affect the fitness of fruit flies Drosophila in the lab?

Pro · 4 steps

1.4M	50	20
Retrieved	Eligible	Included

microbial strain-level differences affect the fitness of fruit flies Drosophila in the lab 225.9K

microbial strain differences affect Drosophila melanogaster fitness under laboratory conditions 1.2M

Read Abstracts and PDFs 20

Microbial Community Dynamics in Natural Drosophila melanogaster Populatio…	LEMOINE	2025
Host Species and Environmental Effects on Bacterial Communities Associate…	STAUBACH	2012
Different Long-Term Nutritional Regimens of Drosophila melanogaster Shape I…	JELENA	2025
How gut transcriptional function of Drosophila melanogaster varies with the …	BOST	2018
Interspecies Interactions Determine the Impact of the Gut Microbiota on Nutril…	NENELL	2013
Environment and diet shape the geography-specific Drosophila melanogaster …	GALE	2024
Acetic acid bacteria promote Drosophila fitness under nutritional stress	MATTHEWS	2020
Host genotype and diet interact to shape microbiome function in fruit flies	WANG	2021
Strain-specific effects of Acetobacter on Drosophila development and survival	KEEBAUGH	2018
Microbiome composition rather than taxonomy predicts Drosophila lifespan	GOULD	2018

Generated answer (excerpt)

Strain-Level Variation in Host Fitness

A microbial GWAS using *Gluconobacter* isolates from wild-caught flies found significant variation in offspring number conferred by closely related strains, tracing the difference to horizontal gene transfer of the thiamine biosynthesis pathway WANG 2021 .

Notably, the two strains differing in fitness benefit had identical 16S rRNA sequences, demonstrating that strain-level functional variation is invisible to standard taxonomic profiling WANG 2021 .

Species Interactions and Community Context

Beyond individual strains, microbe-microbe interactions shape host fitness as strongly as any single species. A combinatorial study of five core gut bacteria found that higher-order interactions (three to five species) occurred in 13–44% of possible cases and were often as important as individual species effects GOULD 2018 .

Acetobacter and Lactobacillus species differ in their nutritional contributions. Acetobacter species drove the largest reductions in development time and triglyceride levels, while most Lactobacillus species had no effect on these traits in monoculture NENELL 2013 .

Summary

Corpus

What can the tool search? Coverage and full-text access matter.

Retrieval and reranking

Well-established methods for finding and prioritising relevant evidence.

Generation

Different ways to turn evidence into answers; no single recipe or agreed measure of quality.