AI-assisted data mining platform for target, biomarker and mechanism discovery

Q-omics
https://qomics.ai

Sukjoon Yoon, Professor & CEO,
Dept of Biological Sciences, Sookmyung women's university & CBiS Inc., Seoul, Korea, Contact: yoonsj@sookmyung.ac.kr

Ontology-driven omics data mining

Bio-LDM (Biological Large Data Model)

Core ontology-grounded intelligence engine for discovering context-specific gene relationships, functional modules and disease mechanisms.

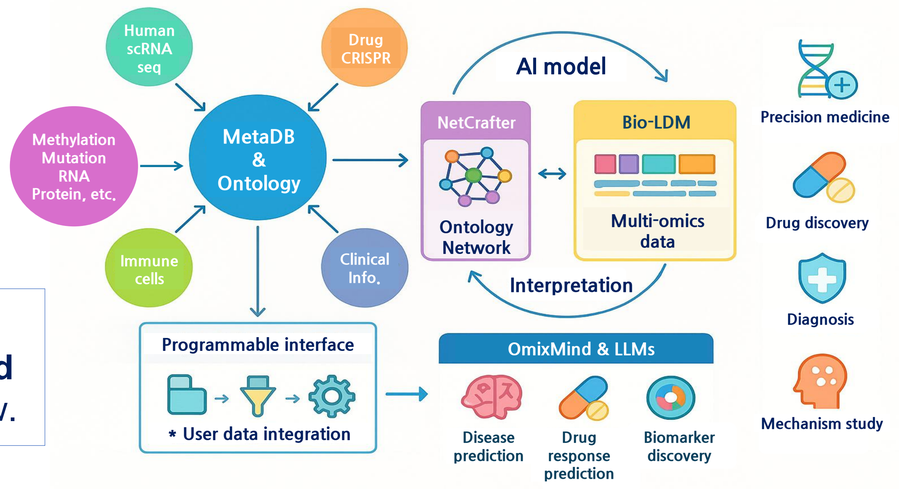
MetaDB

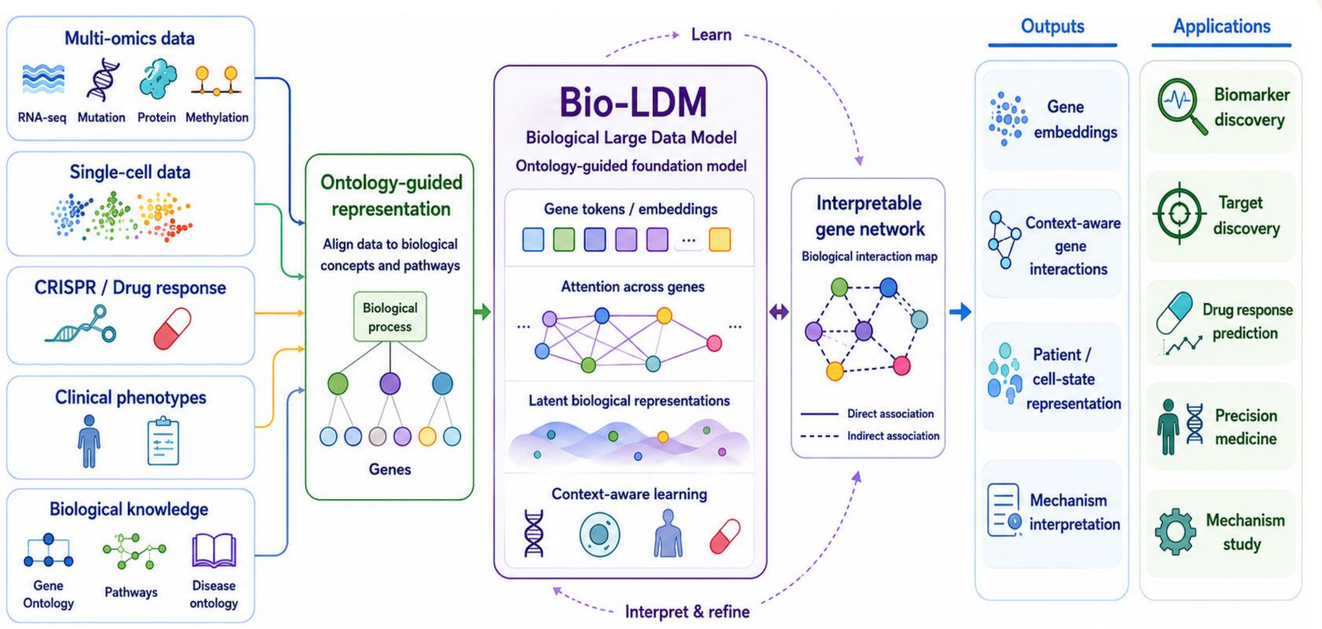
Over 1 billion multi-omics data point
30 billion cross-associations reinforced by consensus scores.

OmixMind & LLMs

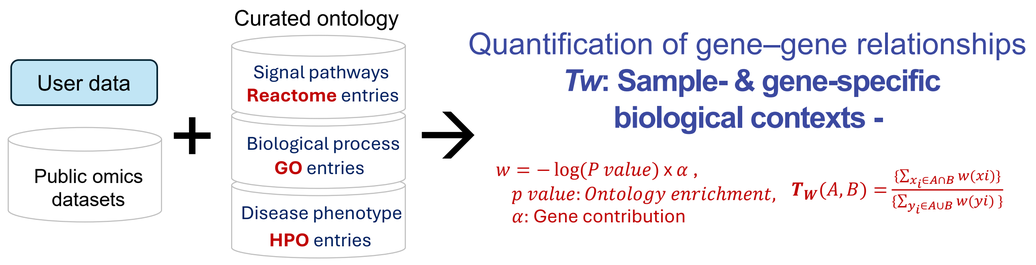
Programmable LLM interface orchestrating Bio-LDM, MetaDB and NetCrafter for fully customized, project-specific analyses.

Bio-LDM discovers the hidden biology; **MetaDB** validates it, **NetCrafter** explains it and **OmixMind** turns it into a fully customized discovery workflow.

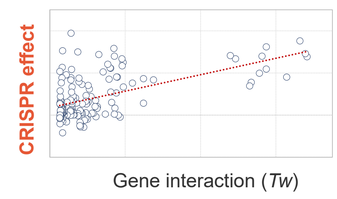




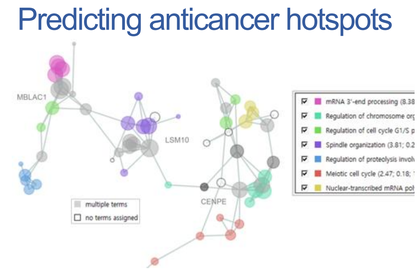
NetCrafter: Ontology-derived quantitative gene network modeling



Significant correlations in gene interaction with anticancer CRISPR efficacy



Predicting anticancer hotspots



Meta DB prioritizes reproducible targets and biomarkers

- **Consensus scoring** on >30 billion associated data pairs in **MetaDB**
  → **Reliable targets and biomarkers across lineages and subtypes**

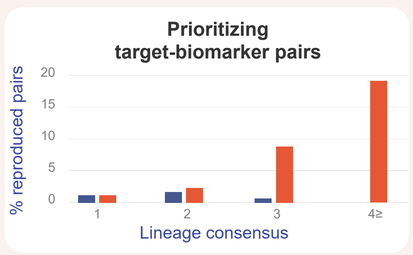
Lineage and sampling consensus analysis of PD-1 expression associated with patient survival

- **PDCD1** expression was significantly associated with patient survival in **28 of 34** TCGA cancer types (p-values < 0.05). Each lineage's result is itself backed by many sampling conditions.
- Higher **PDCD1** was **favorable** across these lineages, holding up best in **SKCM** (**135 of 216** conditions), then **UCEC** at **92** and **UVM** at **88**.

Next step: Open the Consensus view to see where the signal holds, or draw a Kaplan–Meier plot for a specific gene.

Lineage	Measure	Sample split	Gender	Stage	Period	AUC ...	AUC ...	p-value	Sampling consensus
SKCM ...	OS	Median	All	All	Extend...	0.433	0.260	< 0.001	135
UCEC (...	DFS	Median	All	All	Extend...	0.756	0.507	< 0.001	92
UVM (...	OS	Median	All	All	Extend...	0.424	0.759	< 0.001	88
HNSC ...	DFS	Median	Male	All	Extend...	0.464	0.261	< 0.001	74
KIRP (...	OS	Quartile	All	All	Extend...	0.591	0.836	0.001	61
STAD (...	OS	Quartile	All	III,IV	Extend...	0.609	0.203	0.007	58
LGG (B...	OS	Median	All	All	Extend...	0.380	0.508	< 0.001	51
LAML (...	DFS	Tertile	All	All	3-year	0.430	0.730	< 0.001	40
CESC (...	OS	Tertile	All	All	5-year	0.870	0.721	0.004	32
KIRC (...	DFS	Quartile	All	All	5-year	0.750	0.855	0.005	32

Hit selection: p-values < 0.05 & min. sample number = 11
What is AUC (Area Under Curve)?
Kaplan Meier plot
Consensus samplings



Lineage consensus of associations between **CRISPR** efficacy and RNA expression

→ Gene pairs with high consensus showed stronger reproducibility in independent **shRNA** screens.

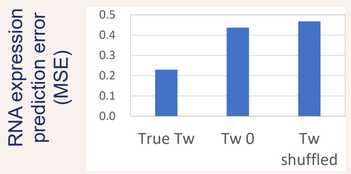
Bio-LDM: Ontology-guided Bio-AI archiecture

Problem: Sparse omics data and incomplete ontology context limit conventional Bio-AI

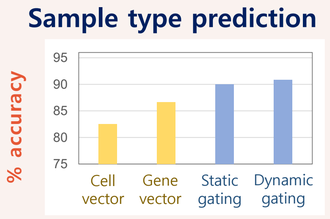
Solution: Bio-LDM™ learns structured biological context through NetCrafter-derived *Tw*-gated attention

$$Gene-Gene\ \text{score}_{ij} = \frac{Q_i K_j^T}{\sqrt{d_k}} + f(Tw_{ij}),\quad Ontology\ gating:\ f_{ij} \approx Tw_{ij}(Q_i \cdot w_{Tw} + K_j \cdot w_{Tw})$$

Omics data prediction



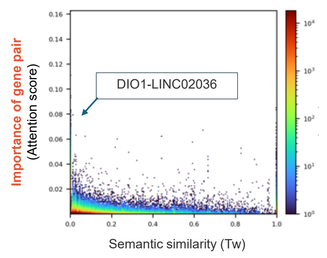
Sample type prediction



- *Paper published (Briefings in Bioinformatics, 2026)*
- *Patent application filed in Korea, USA & China*

Finding novel, tumor-specific gene-gene interaction

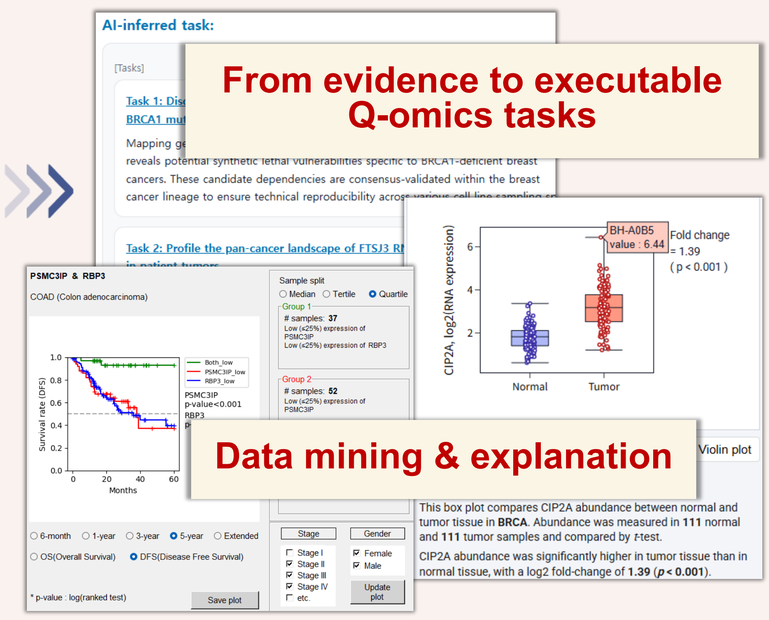
* Comparative analysis of attention score with Tw (functional association)



OmixMind: Text-to-Data mining

AI-assisted research synthesis

From evidence to executable Q-omics tasks

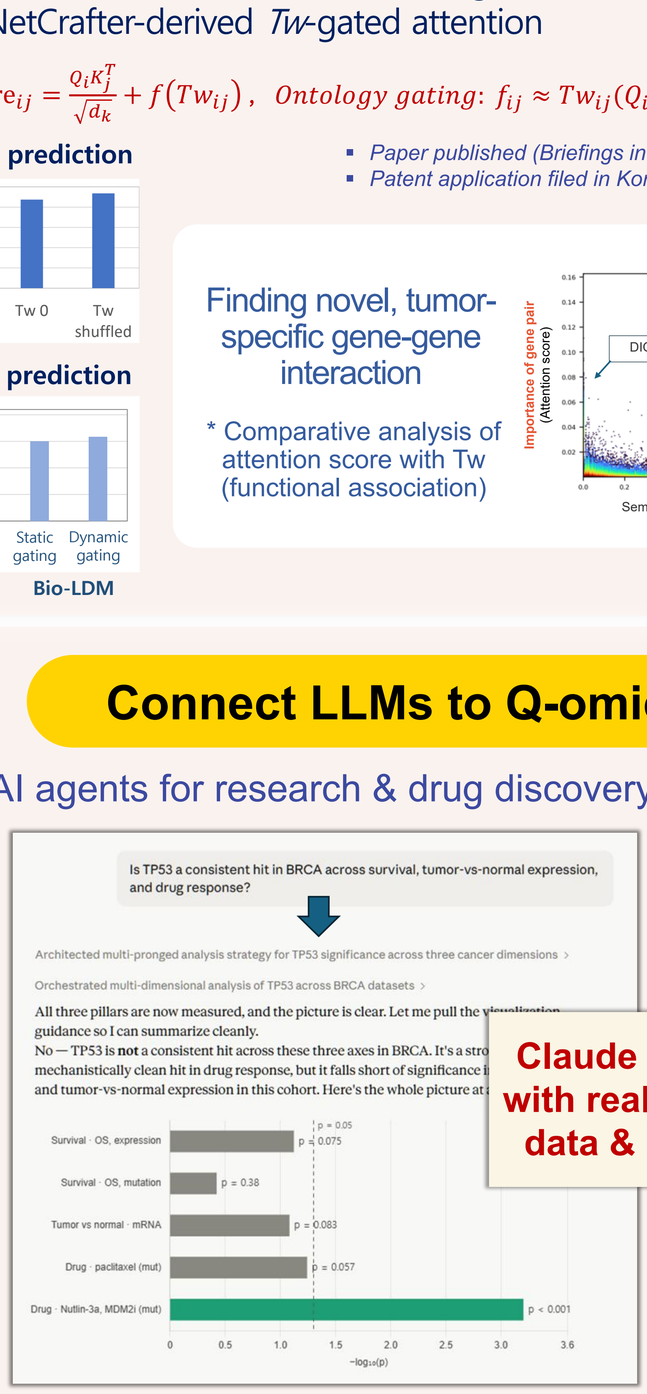


Data mining & explanation

Connect LLMs to Q-omics

AI agents for research & drug discovery pipelines

Claude answers with real Q-omics data & insights



Diverse applications including:

- Drug target and biomarker discovery
- Functional interpretation for underlying mechanisms and disease phenotypes
- Biomarkers in immunotherapy response
- Synthetic lethal gene pairs
- Tumor- antigen discovery
- Undruggable targets and new modalities in drug discovery
- AI-assisted analysis of user data

Availability & Collaboration

- Explore Q-omics **SaaS** at https://qomics.ai, AWS Cloud and via **MCP** on Claude and ChatGPT
- For collaboration, we offer partner-specific **Evidence Sprint**, **cloud/API integration** and confidential **PoC** options