타깃·바이오마커 근거를 위한 새로운 오믹스-온톨로지 융합 기술

새로운 생물학 파운데이션 모델을 고도화하고 검증하여 실제 연구에 적용하기 위한 R&D 제안

Q-omics

https://qomics.ai → Try Q-omics for free



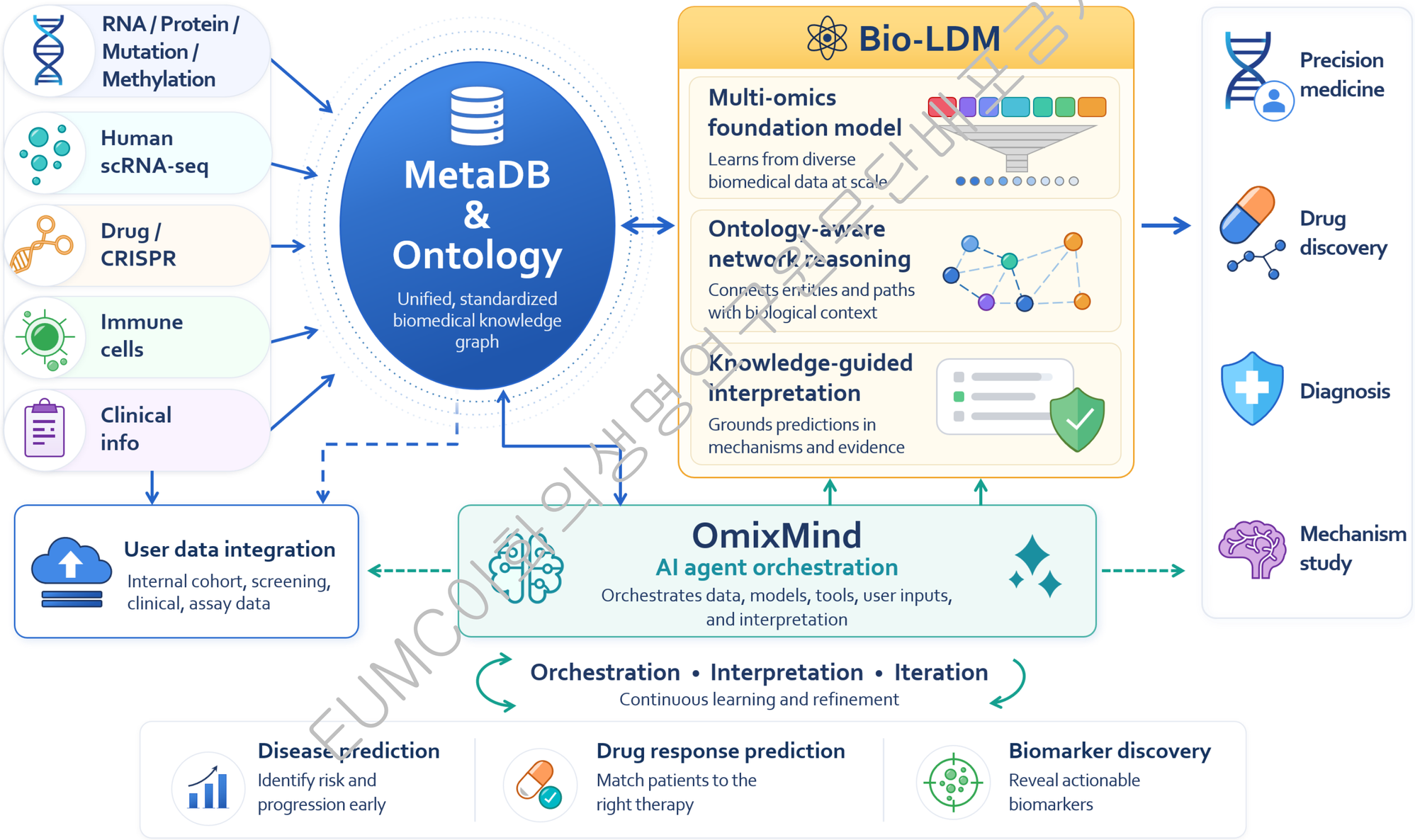
EUMC의학의생명연구원 무단배포금지

2026년 10월

㈜씨비스

AI-Native Infrastructure for Biopharma Innovation

Integrating data, knowledge, and AI to accelerate discovery, translation, and better therapies



EUMC이화의생명연구원무단배포금지

핵심 기술: Bio-LDM (Biological Large Data Model)

생물학적 맥락에 특화된 유전자 관계와 기능 모듈, 질환 기전을 도출하는
온톨로지 기반 오믹스 분석 엔진

문제: 오믹스 데이터는 수많은 분자 과정이 상호작용한 '결과'만을 보여줌

제안 : 오믹스 + 온톨로지 → 숨겨진 생물학적 맥락 학습 용이

오믹스 데이터 + 온톨로지 →

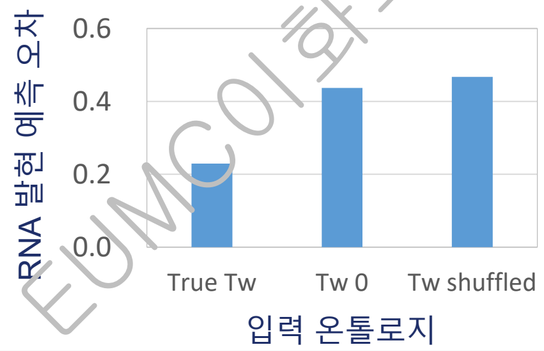
유전자–유전자 관계의 정량화

$w = -\log(P\ value)$ x α , $\boldsymbol{T_W}(A,B) = \frac{\{\sum_{x_i \in A \cap B} w(xi)\}}{\{\sum_{y_i \in A \cup B} w(yi)\}}$

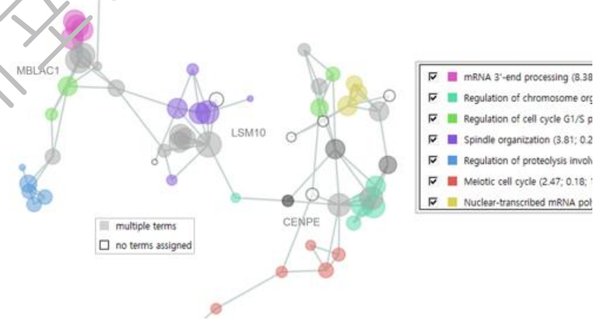
α: 유전자 기여도

$$Gene - Gene\ \text{score}_{ij} \approx \frac{Q_i K_j^T}{\sqrt{d_k}} + f(Tw_{ij})$$

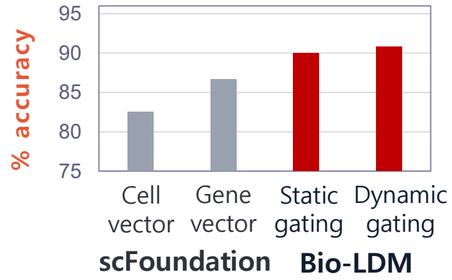
사전학습 성능:
'온톨로지 + 오믹스' 모델이
오믹스 단독 학습 모델을 능가



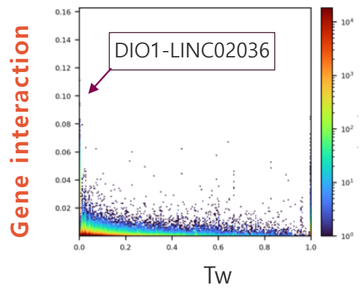
1) 신약 개발을 위한 핫스팟 유전자 시각화



2) SOTA 모델을 상회하는 암 샘플 유형 예측 정확도



3) 암 진행과 연관된 새로운 유전자–유전자 연관성 발굴



EUMC이화의생명연구원무단배포금지

Bio-LDM 활용

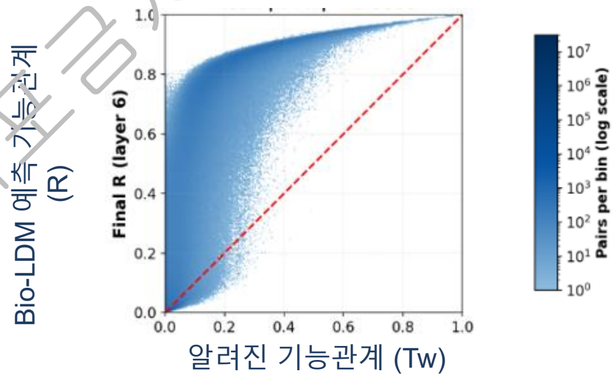
“Bulk samples + Single cells” 하이브리드 사전학습 모델:

- Bulk samples: 암 환자 샘플 및 암세포주 3만 종 : 총 33,375개 유전자 학습
- Single cells: 다양한 질환 + 정상세포 3백만 또는 3천만 종: 총 37,163 개 유전자 학습

파인튜닝 및 전이학습 모델 개발:

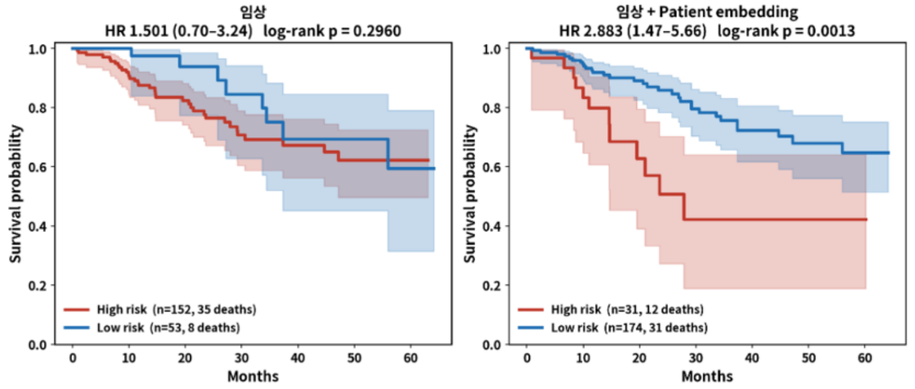
- 혈관 연구 모델 개발: Vascular Atlas 80만개 RNA exp → EndMT atlas 구축
- CRISPR / Drug 세포주 효과 데이터 >1만 5천개,
- 기전 분석 네트워크 모델
  - Normal vs.Tumor,
  - Responder vs. Non-responder
- Target hotspot prediction,
- Synthetic lethal target-drug combination

기전 연구를 위한 유전자-유전자 관계 분석



* 기존 온톨로지에 알려진 유전자 기능관계와는 다른 새로운 유전자간 연관성을 R vs Tw 분석으로 확인 가능

유전자 관계, R값 (patient embedding) 검증



* 임상 정보와 더불어 오믹스 기반 유전자 관계 정보 R이 환자 생존율 예측을 향상시킴

EUMC이화의생명연구원무단배포금지

* Bio-LDM 활용: 사용자 데이터 분석 예시

유전자-유전자 관계 정량화 (R)

Key gene-gene interactions

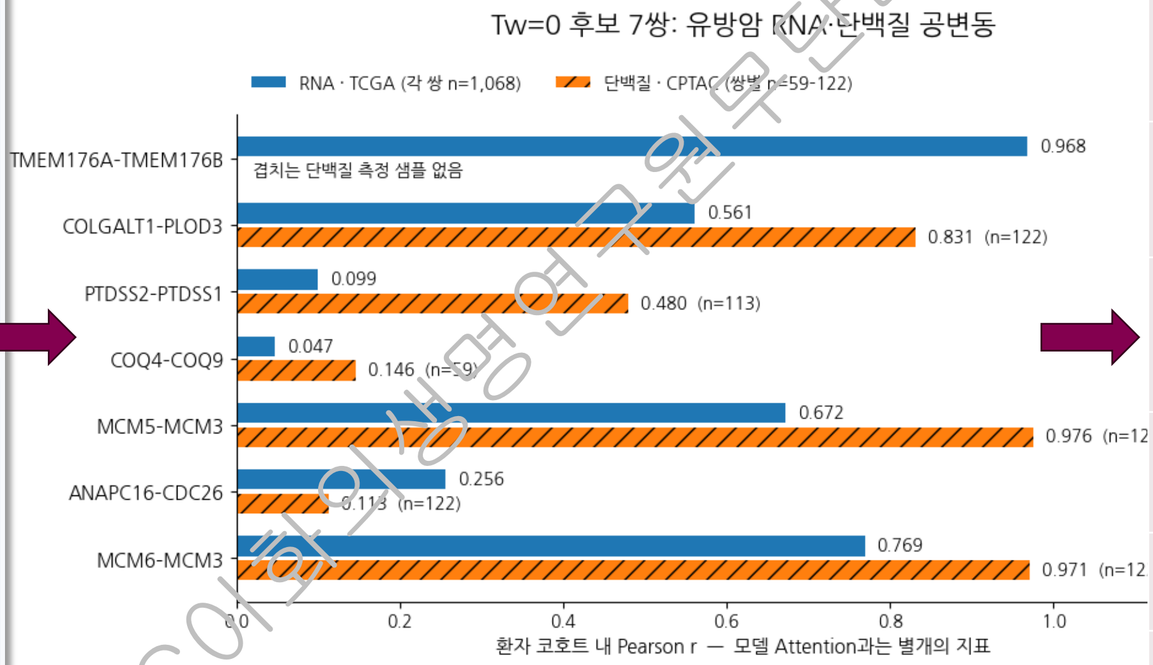
#	Gene A	Gene B	Functional similarity (Tw)	Attention
1	S100A9	S100A8	0.586	**0.2146**
2	TMEM176B	TMEM176A	0.000	**0.2021**
3	TMEM176A	TMEM176B	0.000	**0.1931**
4	MT-ND6	MT-ND5	0.737	**0.1622**
5	MZT2B	MZT2A	0.147	**0.1617**
6	C1R	C1S	0.697	**0.1583**
7	TUBA1B	TUBA1C	0.356	**0.1488**
8	APOE	APOC1	0.266	**0.1462**
9	MT1G	MT1H	0.768	**0.1406**
10	FABP5	FABP7	0.326	**0.140**

Potentially unknown gene-gene interactions

#	Gene A	Gene B	Functional similarity (Tw)	Attention
1	TMEM176B	TMEM176A	0.000	**0.2021**
2	TMEM176A	TMEM176B	0.000	**0.1931**
3	COLGALT1	PLOD3	0.000	**0.1110**
4	PTDSS2	PTDSS1	0.000	**0.0916**
5	COQ4	COQ9	0.000	**0.0797**
6	COQ9	COQ4	0.000	**0.0795**
7	MCM5	MCM3	0.000	**0.0783**
8	PLOD3	COLGALT1	0.000	**0.0782**
9	ANAPC16	CDC26	0.000	**0.0781**
10	MCM6	MCM3	0.000	**0.0770**

* 연구자의 TNBC 샘플의 RNA seq 데이터로 부터 연관성 (R, 맥락) 높은 유전자-유전자 쌍 랭킹

Q-omics 검증 분석



출처: Q-omics MCP 직접 조회 (2026-10-13). 전체 코호트 통계 사용; RNA·단백질은 서로 다른 코호트. 공변동 ≠ 인과성.

* R값 높은 유전자 조합에 대한, MetaDB에서의 단백질 및 RNA 발현 데이터에서의 유의한 상관성 분석

LLM 해석

Tw=0 유전자쌍	TCGA RNA r	CPTAC 단백질 r (샘플 수)	PubMed 근거와 해석
TMEM176A–TMEM176B	0.968	측정쌍 없음	결합 및 수지상세포 성숙 조절이 동물계에서 이미 보고됨. 강한 공발현 관계를 회수한 사례. (PubMed)
COLGALT1–PLOD3	0.561	0.831 (122)	사람의 콜라겐 가공 KOGG 복합체 구조가 cryo-EM으로 확인됨. 가장 설득력 있는 대표 사례. (PubMed)
PTDSS2–PTDSS1	0.099	0.480 (113)	같은 phosphatidylserine 합성 경로의 효소들. RNA 공발현은 약하지만 단백질 공변동은 중간 수준. (PubMed)
COQ4–COQ9	0.047	0.146 (59)	알려진 CoQ 합성 복합체 구성원. 다만 이번 유방암 코호트의 발현 공변동 근거는 약함. (PubMed)
MCM5–MCM3	0.672	0.976 (122)	알려진 MCM2–7 DNA 복제 복합체. 단백질 수준에서 매우 강하게 뒷받침됨. (PubMed)
ANAPC16–CDC26	0.256	0.113 (122)	동일 APC/C 세포주기 복합체의 구성원이나, 발현 공변동은 약함. (PubMed)
MCM6–MCM3	0.769	0.971 (122)	역시 MCM2–7 복합체. 강한 기능적 일관성과 단백질 공변동. (PubMed)

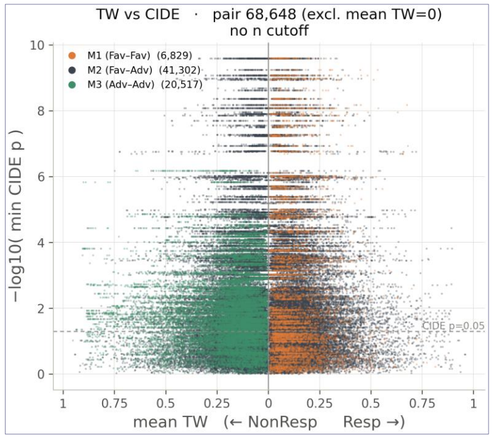
* PubMed 및 기타 Ref 기반 데이터 해석

EUMC 아주의생명연구원 무단배포금지

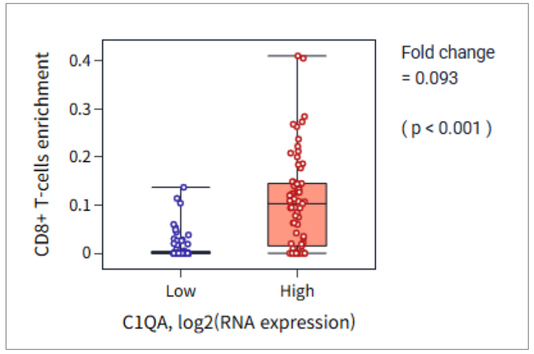
적용 예시: 펨브롤리주맙 반응 데이터에서 검증 가능한 가설까지

입력: 펨브롤리주맙 치료 환자의 전사체 프로파일 (반응군 vs. 비반응군)

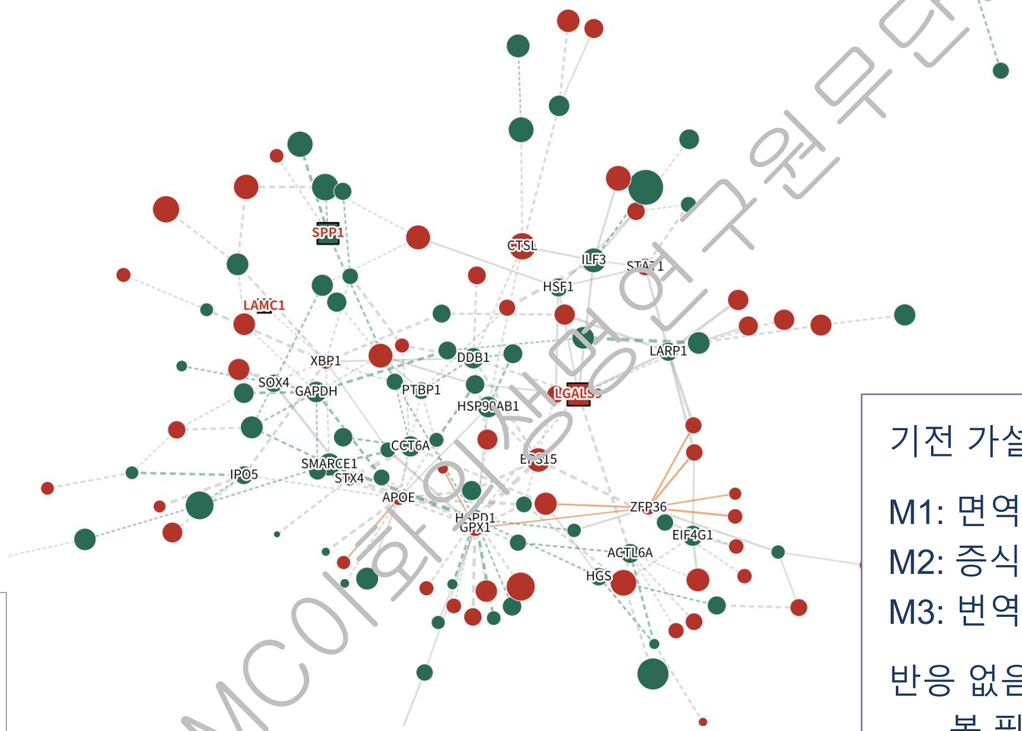
유전자 쌍 연관성:
반응군 vs 비반응군



21개 암 계통에서 C1QA 고발현군
샘플의 CD8+ T세포 농축도가 더 높음



Selected gene-pair network · node 129 edge 184 · |log2FC| ≥ 0.3



* 반응군 (붉은색), 비반응군 (녹색) 의 Bio-LDM 유전자 관계 네트워크
(노드와 엣지는 각각 반응군과 비반응군에서 유전자 중요도와 유전자간 연관성 (R값)을 시각화)

모듈	유리한 유전자	불리한 유전자
M1	C1QA, C1QC, LGALS9, LYZ, SRGN	
M2	C1QA, C1QC, IL32, LGALS9, LYZ, SRGN	LAMB1, LAMC1, POSTN, SPP1
M3		CTHRC1, LAMB1, POSTN, SPP1

기전 가설

M1: 면역 활성화 / IFN 반응
M2: 증식·스트레스 프로그램을 동반한 면역–종양 상호작용
M3: 번역, 생합성, 에너지 대사

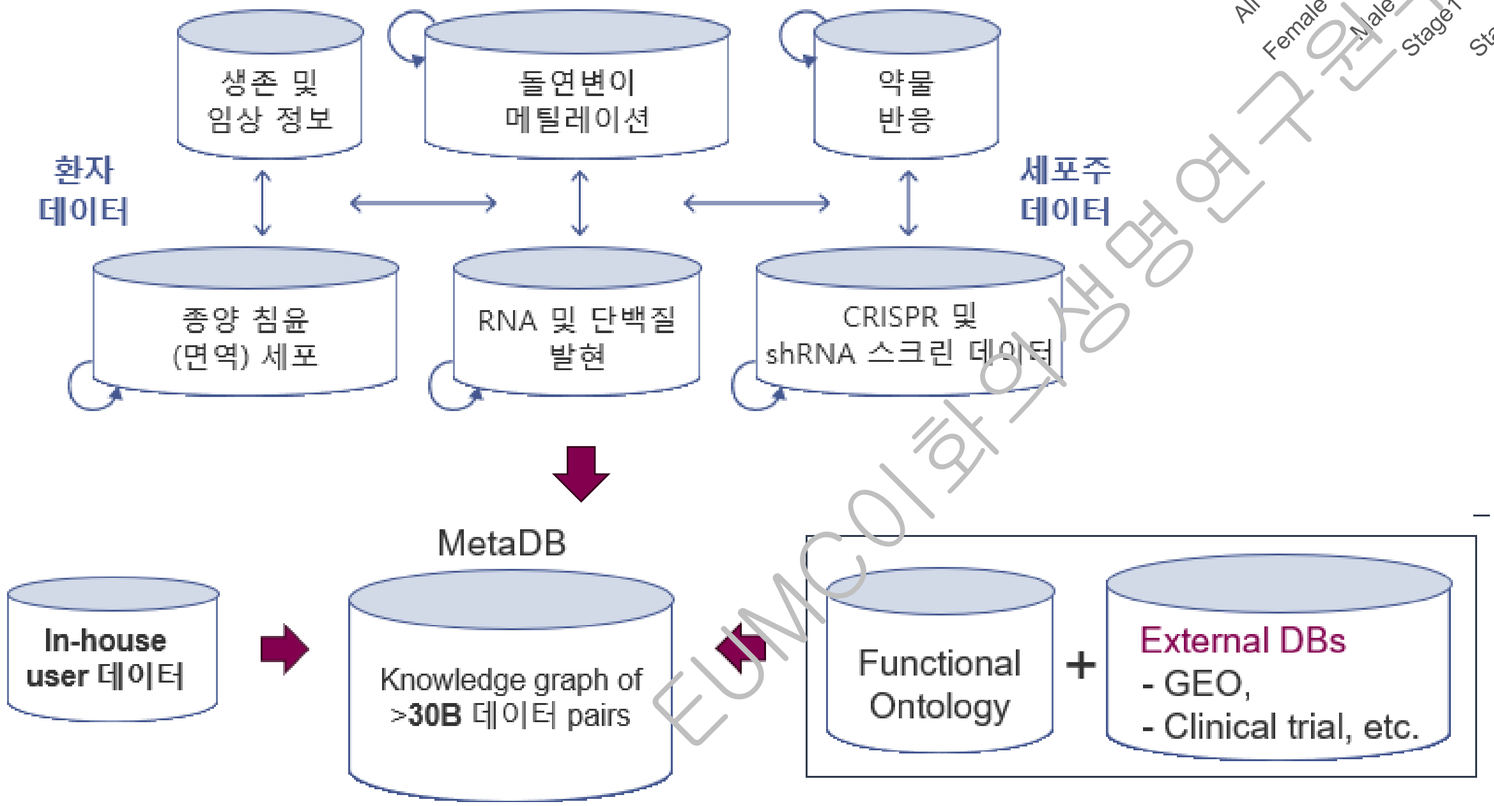
반응 없음: 생합성·대사·스트레스 적응 프로그램을 추가로 살펴볼 필요

EUMC이화의생명연구원무단배포금지

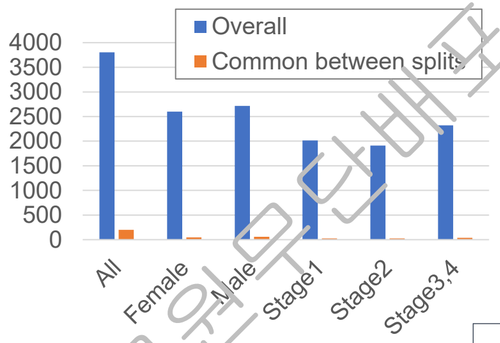
근거 기반 신뢰도: MetaDB

- 10억 건 이상의 멀티오믹스 데이터 포인트
- **컨센서스 스코어로 보강한 300억 건의 교차 연관성**

→ **컨센서스 분석** → 샘플 유형과 조건을 넘어 재현되는 신호를 우선함으로써 분석 결과의 신뢰도를 높입니다.



Prognostic RNA exp., LUAD

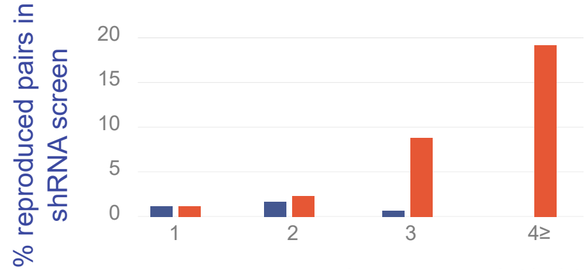


→ 특정 유전자의 RNA 발현과 환자생존율 상관성 분석

→ 동종 환자샘플 random split간에 상관성 데이터 컨센서스 매우 낮음

→ Split간 컨센서스가 있는 소수의 유전자가 중요함

특정 유전자 조합의 CRISPR 효능과 RNA 발현 간 연관성 연구



CRISPR efficacy와 RNA exp.의 상관성
유의도에 대한 암종간 컨센서스

→ 암종간 컨센서스 스코어가 높은 CRISPR efficacy - RNA exp. 유전자 조합은, 독립적인 shRNA 스크린 실험에서도 재현성을 보임 (RNA 발현에 따른 shRNA 효과 상관성이 유의함)

EUMC이화의생명연구원무단배포금지

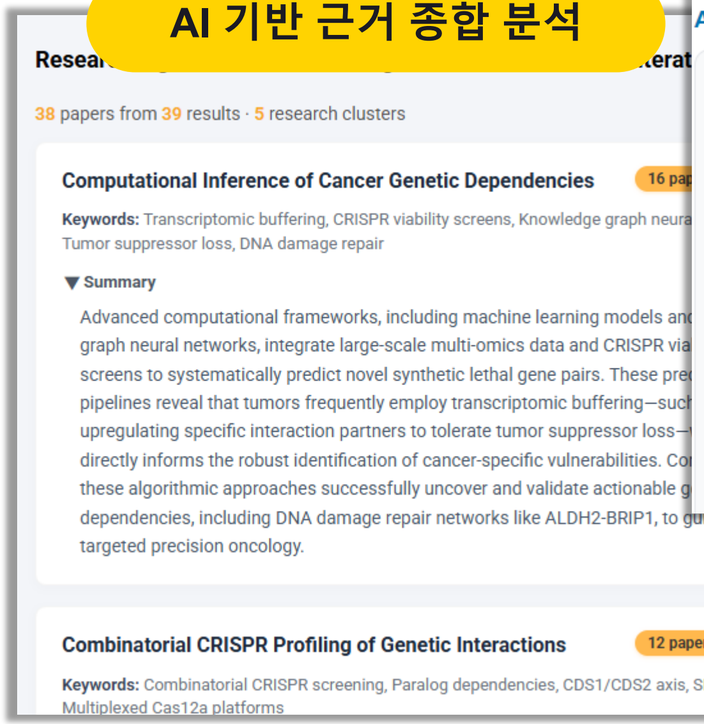
접근성 극대화: OmixMind

질문 + 데이터 → 문헌·데이터베이스 검색 → Q-omics 과제 생성 → Bio-LDM·MetaDB 오케스트레이션 → 근거 해석

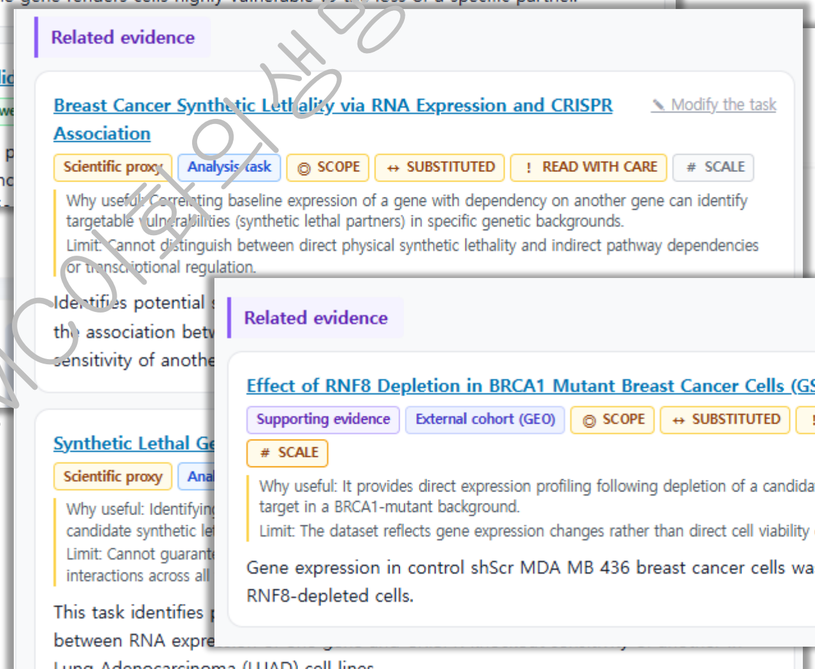
AI 기반 근거 종합 분석

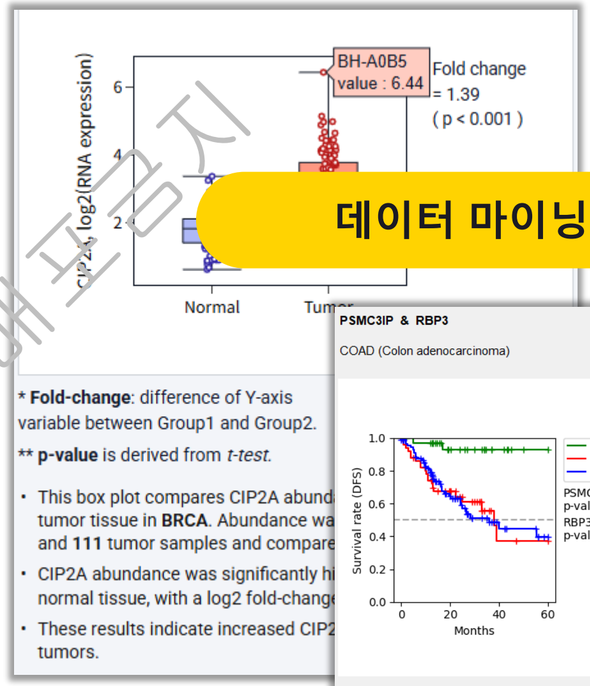
Q-omics 분석 태스크 도출

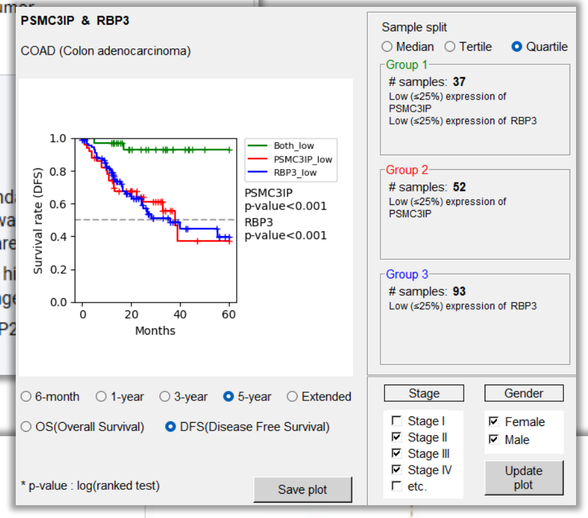
데이터 마이닝 및 해석













EUMC의생명연구원무단배포금지

* LLM을 Q-omics에 연결:

연구·신약 개발을 위한 AI 에이전트

- 기존 연구 워크플로우에 Q-omics 분석을 그대로 추가
- 실시간 데이터 마이닝 및 분석 기능 제공
- Claude와 ChatGPT 지원

핵심은 인터페이스 챗봇이 아니라, 그 아래에 구현된 **원시데이터, 메타데이터 및 분석 아키텍처 및 Bio-LDM을 제공**하는 데이터 마이닝 모델임.

실제 Q-omics 근거로 답하는 Claude

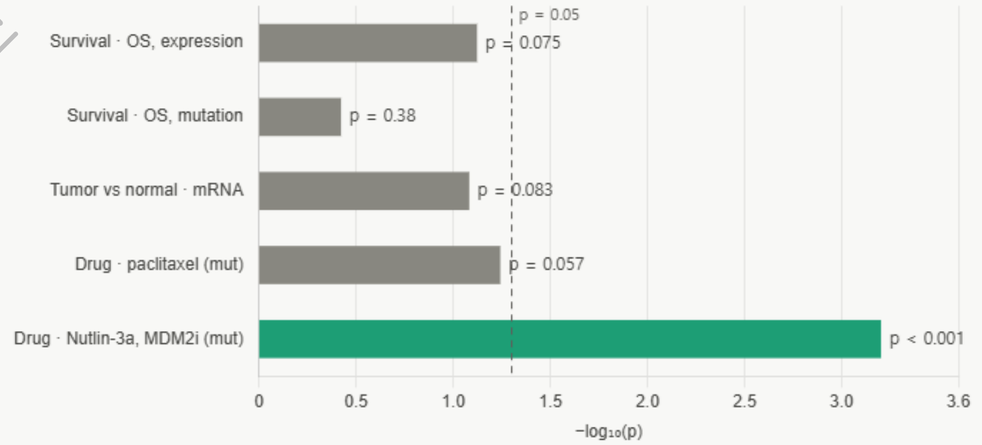
Is TP53 a consistent hit in BRCA across survival, tumor-vs-normal expression, and drug response?

Architected multi-pronged analysis strategy for TP53 significance across three cancer dimensions >

Orchestrated multi-dimensional analysis of TP53 across BRCA datasets >

All three pillars are now measured, and the picture is clear. Let me pull the visualization guidance so I can summarize cleanly.

No — TP53 is **not** a consistent hit across these three axes in BRCA. It's a strong, mechanistically clean hit in drug response, but it falls short of significance in both survival and tumor-vs-normal expression in this cohort. Here's the whole picture at a glance:



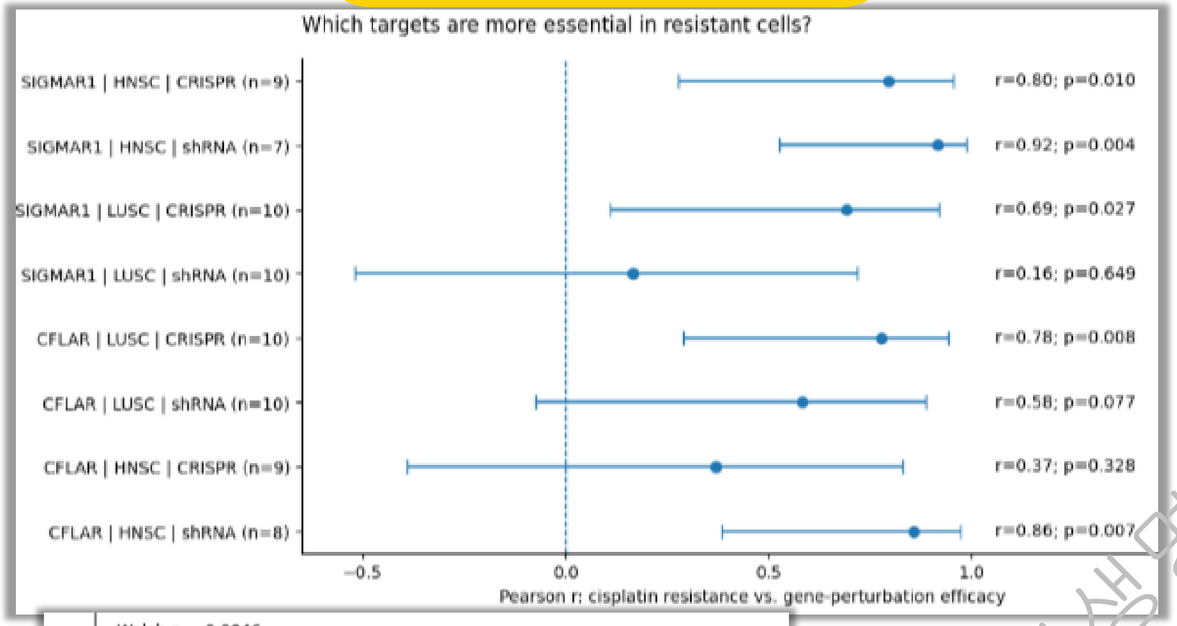
What each axis actually showed (TCGA BRCA tissue for the first two; GDSC2 breast cell lines for drug response):

Survival is the weakest. Splitting patients by TP53 mRNA at the median gives no significant overall-survival separation (p = 0.075), and splitting by mutation status — mutant vs wild-type — is weaker still (p = 0.38). Tumor-vs-normal is similar: TP53 mRNA is only ~1.13×

EUMC이화의생명연구원 무단배포금지

예시) Cisplatin 저항성 극복 연구:

Q-omics 데이터 마이닝 I



SIGMAR1에 대한 Q-omics 메타DB 분석

SIGMAR1 data type	Analysis	Lineage consensus	Lineage of highest sampling consensus
▼ Patient survival			
Methylation (SIGMAR1_ma...	Kaplan-Meier plot	21	UVM (144)
Methylation (SIGMAR1_me...	Kaplan-Meier plot	21	STAD (70)
Mutation	Kaplan-Meier plot	1	HNSC (14)
Protein (mass-spec)	Kaplan-Meier plot	7	HNSC (31)

Patient Samples | Cell Line

Associated data type	Strength (# associated data)	Lineage of highest associated data
▼ Mutation		
Infiltrating cells	1	UCEC (1)
Methylation	24	UCEC (24)
RNA	50	UCEC (21)
▼ Protein (mass-spec)		
Function (RNA)	938	LUAD (287)
Function (mass-spec)	1230	UCEC (380)
Mutation	4	
Protein (mass-spec)	6130	
RNA	345	
▼ RNA		
Drug	1	

메타 데이터 해석

핵심 결과

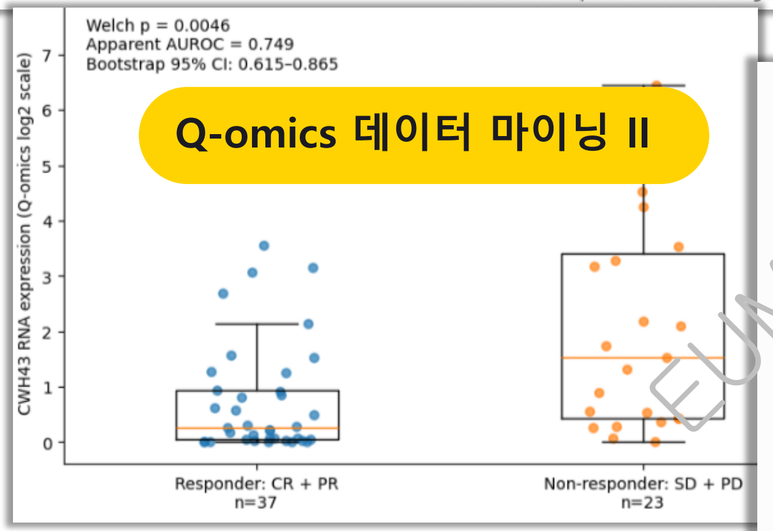
시스플라틴 임상 반응과 연관: 암종 consensus 3, 샘플링 **6/48**, 대표 결과 p=0.012–0.027.

요로계 세포주에서 SIGMAR1 CRISPR 억제효능↑ ↔ 시스플라틴 감수성↓. 다만 샘플링 consensus **1/3.**

골암계에서는 **RNA 고발현–감수성 증가**, 대장암계에서는 **단백질 고발현–저항성 증가.**

두경부암에서 **RNA 고발현–OS 불량**, 반면 **단백질 고발현–DFS 양호.**

Q-omics 데이터 마이닝 II



데이터 마이닝 해석

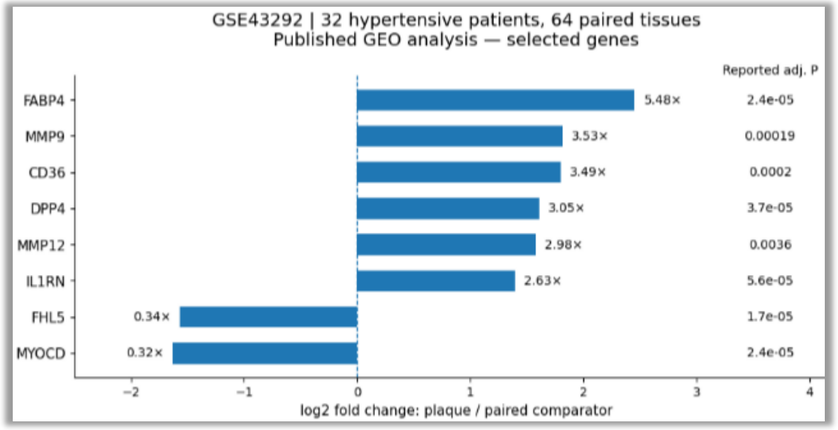
1. 실제 데이터가 지지하는 후보

후보·암종	Q-omics 원자료 재계산 결과	실험 의사결정
SIGMAR1 — 두경부암	Cisplatin 저항성과 **CRISPR 억제 효능**의 상관 **r=0.795**, **p=0.0104**, **n=9**, shRNA에서도 **r=0.917**, **p=**	**1순위 병용 검증.** 저항성 세포가 SIGMAR1 억제에 더 취약하다는 방향 품에서 일치
CFLAR/c-FLIP — 폐편평상피암	CRISPR **r=0.777**, **p=0.0082**, **n=10**. shRNA **r=0.583**, **p=0.0768**, **n=10**	**2순위.** 방향은 같지만 shRNA 전체 표본 분석은 유의하지 않아 근거가 약함
CWH43 — 방광암	반응군 **37명**, 비반응군 **23명**. 비반응군 발현이 평균 **1.405 log2 단위 높음**, Welch **p=0.0046**. 동일 코호트 **AUROC=0.749**, 95% CI **0.615–0.865**	**비반응 표지자로 독립 검증.** CWH43를 억제하면 저항성이 해소된다는 근거는 아님

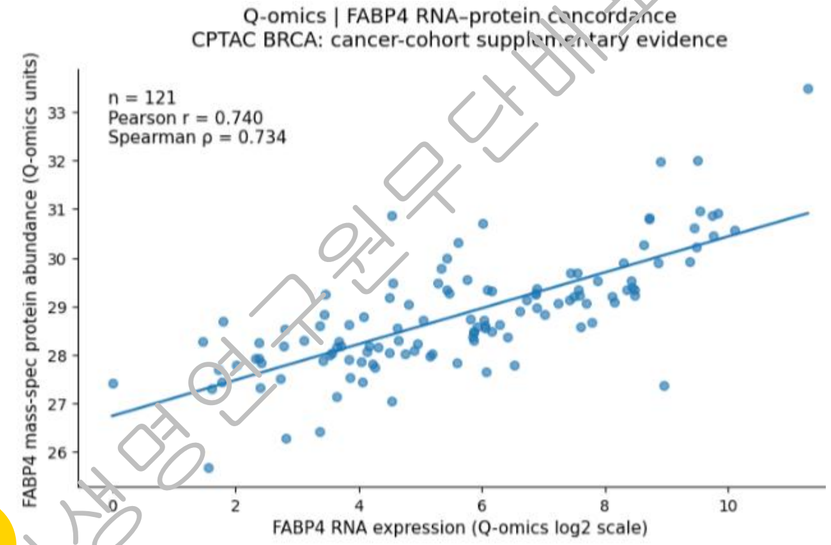
EUMC이화의생명연구원 무단배포금지

예시) 고혈압 또는 뇌혈관질환 연구:

GEO 데이터 마이닝



Q-omics 데이터 마이닝



데이터 마이닝 해석

후보·조절 방향	GSE43292 플라크/대조 발현비[1]	Q-omics 실조회 보조 근거[2]	우선 판단
MYOCD 회복	**0.32배** — 약 68% 감소 조정 P=2.39×10⁻⁵	평활근 xCell 점수와 **28개** 암종에서 유의한 연관	**기전 검증 1순위.** 평활근 기능·지질배출 회복 방향
FABP4 억제	**5.48배 증가** 조정 P=2.39×10⁻⁵	내피세포 점수 **29개**, 대식세포 점수 **16개 암종**에서 연관. RNA–단백질 **r=0.740, n=121**	**세포·병변 상태를 한정한 조건부 후보**
MMP12	**2.98배 증가** 조정 P=0.00359	대식세포 점수와 **31개 암종**에서 유의한 연관	기질손상 기전 연구 후보. **전신 억제 타깃 지명은 보류**

MYOCD, FABP4, MMP12에 대한 Q-omics 메타DB 분석

메타 데이터 해석

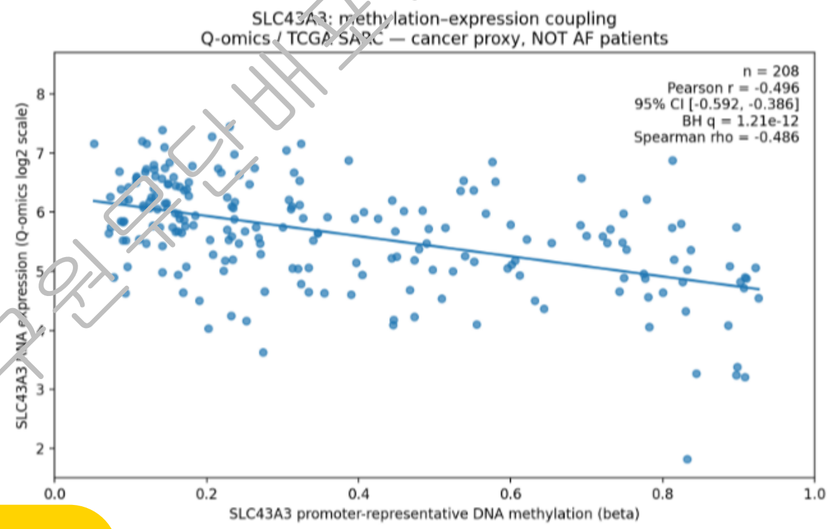
EUMC이화의생명연구원 무단배포금지

예시) 심방세동 타겟, 바이오마커 연구:

GEO / Q-omics 데이터 분석

결과	SLC43A3	ANXA4	RUNX1
표본 30개 이상인 암종 수	31	31	31
메틸레이션–RNA가 음의 상관인 암종	**28/31**	21/31	16/31
음의 상관이면서 **보정 q<0.05**인 암종	**16/31**	2/31	0/31
SARC에서의 상관	**−0.496; n=208**	−0.133; n=255	−0.066; n=256

Q-omics 데이터 마이닝



데이터 마이닝 해석

심방세동 신약 타깃 및 바이오마커 발굴 연구 요약

우선순위 / 후보	GEO·Q-omics 핵심 분석 결과	기전 및 해석	후속 검증 실험
1. SLC43A3 치료 타깃 후보	AF 심방조직에서 DNA 저메틸화·H3K27ac↑·RNA↑. Q-omics에서 메틸레이션–RNA 역상관 28/31 암종, SARC r=−0.496	후성유전적 조절에 의한 발현 증가 가능성. AF 관련 새로운 조절 기전 후보	AF 심방조직에서 CpG–RNA 검증 → 심방세포별 발현 확인 → CRISPR 억제·과발현 후 전기생리·섬유화 평가
2. ANXA4 / cg11942603 임상 바이오마커	AF에서 메틸레이션 감소 Δβ=−0.08, p=0.006. 같은 코호트에서 ANXA4 RNA–좌심방 직경 r=0.36, NT-proBNP r=0.38	특정 CpG 메틸레이션과 ANXA4 발현이 심방 리모델링 및 질환 표현형을 반영할 가능성	혈액·심방조직 CpG 메틸레이션 및 RNA 측정 → 좌심방 크기·AF 부담·재발과 연관성 검증
3. RUNX1–YTHDF1–CCN2 섬유화 타깃 축	AF에서 RNA m6A–YTHDF1–RUNX1–CCN2 섬유화 기전 보고. Q-omics RUNX1–CCN2 RNA 상관 r=0.335, n=256	RUNX1 번역 조절을 통한 심방 섬유화 촉진 기전. DNA 메틸레이션과 RNA m6A는 구분 필요	심방 섬유아세포에서 RUNX1/YTHDF1 억제 → CCN2·COL1A1·α-SMA 및 콜라겐 생성 변화 확인

EUMCI 한의생명연구원 무단배포금지

왜 Q-omics인가?

생물학적 지식이 해석 단계가 아니라 학습 과정 안으로 들어옵니다.

새로운 핵심 기술

Bio-LDM

정량 온톨로지를 학습 구조 내부에 융합
— 다양한 파인튜닝 및 전이학습 가능

근거 기반

MetaDB

컨센서스 분석으로 검증된 환자·CRISPR·
약물반응 근거의 대규모 집적

전달 경로

OmixMind

연구 질문 → 분석 과제 → Bio-LDM /
MetaDB 오케스트레이션

학습된 관계에서 검증 가능한 R&D 의사결정으로

타깃	바이오마커	기전
신규 타깃–바이오마커 조합	약물반응 시그니처	기능적 유전자 모듈
합성치사 파트너 유전자	정밀 의료 근거	작용기전(MoA) 및 저항성 가설
맥락 특이적 취약점	예후 바이오마커 후보	기전 핫스팟

EUMC 이화의생명연구원 무단배포금지

왜 Q-omics인가?

공개 오믹스·기존 유전체 근거에 관계 분석을 더하고, 환자 데이터·실험으로 검증하는 연구

01 공개·내부 데이터 연계

공개 오믹스 (GEO 등)
bulk·단일세포 RNA-seq
연구 질문에 맞는 데이터 선정

기존 유전체 근거
GWAS·QTL 분석 결과
보유 타깃·바이오마커 후보

선택적 내부 데이터
환자 오믹스·임상·실험 정보
가용 데이터와 연구 범위 확인

02 관계 분석·근거 통합

컨센서스
코호트 간 재현성과
환자 재표집 안정성을 구분 평가

Tw / Bio-LDM
유전자 관계·기능 모듈 분석
타깃·기전 가설 도출

OmixMind
질문 → 분석 → 문헌·근거 통합
출처와 미확인 근거를 함께 정리

03 후보 선정·검증

후보 우선순위화
유전·오믹스·세포·기전 근거 종합
타깃·바이오마커 후보 선정

임상·기능 검증
환자 코호트·세포·검체 활용
후보별 검증 실험 공동 설계

관계 정보의 추가가치 평가
발현·유전체 단독 분석과 비교
검증할 후보 선정에 기여하는가?

산출물 | 우선순위 후보 목록 + 기능 모듈·기전 가설 + 후보별 근거표·검증 계획

EUMC이뮤노믹스연구원 무단배포금지

Q-omics

Q-omics 엔터프라이즈

데이터를 Q-omics로 옮기지 않고, Q-omics를 데이터가 있는 곳으로 옮겨갑니다.

고객사 VPC / 온프레미스 환경

데이터도, 통제권도 고객사에

내부 데이터

오믹스 · 임상 · 연구

연구자 및 내부 애플리케이션

분석하고 탐색하여 발견을 이끌어냅니다

Q-omics 프라이빗 인스턴스

Bio-LDM 엔진 · 큐레이션 MetaDB · REST API • MCP

통찰과 결과

새로운 생물학, 더 빠른 의사결정.

소프트웨어 제공 방식

Q-omics / CBIS

- Deployment package
- Onboarding
- Optional updates & support

기밀 데이터와 질의, 결과가 네트워크 밖으로 나가지 않습니다.

설계부터 프라이빗
민감 데이터는 고객사 환경 안에 그대로 남습니다

운영 환경에 그대로 배포
VPC, 프라이빗 클라우드, 온프레미스 모두 지원

연동 준비 완료
기존 연구 파이프라인에 그대로 맞물립니다.

EUMC이화의생명연구원 무단배포금지

특허 및 논문

01 핵심 기술을 보호하는 특허와 상표

- ❖ 온톨로지 기반 오믹스 데이터 기능 분석·해석 자동화 시스템 (10-2024-0119064)
- ❖ 오믹스 데이터 마이닝 버티컬 AI 자동화 시스템 및 방법 (10-2025-0072993)
- ❖ 바이오 파운데이션 모델 기반 오믹스 데이터 마이닝 자동 추론 방법 및 시스템 (10-2025-0062962), 한국·미국·중국 출원
- ❖ Bio-LDM™ 상표 출원

02 피어리뷰 논문 실적

- ❖ Current advances in comprehensive omics data mining for oncology and cancer research, Biochim Biophys Acta Rev Cancer, 2023 (상위 6% 저널)
- ❖ Q-omics: Smart software for assisting oncology and cancer research, Mol Cells, 2021
- ❖ NetCrafter: Ontology-derived gene network modeling and interpretation, Briefings in Bioinfo. 2026 (상위 2.7% 저널)
- ❖ Bridging literature and omics: Q-omics 3 integrates LLM-guided text-to-data workflows and pan-cancer multi-omics consensus scoring (2026, 심사 진행 중)

EUMC이화의생명연구원 무단배포금지

Q-omics 퀵 데모

https://qomics.ai → 웹 브라우저에서 간단히 질의 및 데이터 분석 가능

https://qomics.ai/mcp-docs/index.html → 챗지피티 또는 클로드에 커넥터로 Q-omics 설치 후 사용

Features	Standard (무료)	Premium	Enterprise
Meta DB analysis	Essential	Advanced	Advanced + Private data
Consensus Summary	Basic	Comprehensive	Integrated with private data
User data analysis	Basic	Advanced	Custom pipeline
Data download	-	Standard export	Full export, LLM-ready
OmixMind	Base model	Pro model	Fine-tuned model
External DBs (GEO, PubMed, BioMCP, etc.)	Core access	Full access	Full + Custom integration
MCP (LLM connector)	Limited usage & data access	Full access	Dedicated integration
Security	Standard	Standard	Enterprise-grade

EUMC이화의생명연구원 무단배포금지

감사합니다

기존 AI를 넘어

Q-omics로 파이프라인의 혁신을 이루십시오

CBiS

회사: https://cbis.co.kr | 제품: https://qomics.ai

EUMC이화의생명연구원 무단배포금지