

멀티에이전트 AI, 바이오의학 가설과 검증 방안 제시

- ◇ 멀티에이전트 AI인 Co-Scientist와 Robin이 치료 후보물질을 제안. 실제 세포실험에서 일부 후보물질의 효과가 확인됐으며, Robin은 후속 유전자 발현 데이터도 분석. 다만 AI의 제안은 기존 문헌 정보의 재조합으로, AI가 기존 문헌과 독립적으로 새로운 생물학적 지식을 창출할 수 있는지는 여전히 미해결

▶ 주요 출처 : Nature, "AI systems devise hypotheses and ways to test them", 2026.6.30.

여러 AI 에이전트가 협력해 바이오의학 가설을 세우고, 이를 검증할 방법까지 고안하는 멀티에이전트 AI* 연구 사례 등장

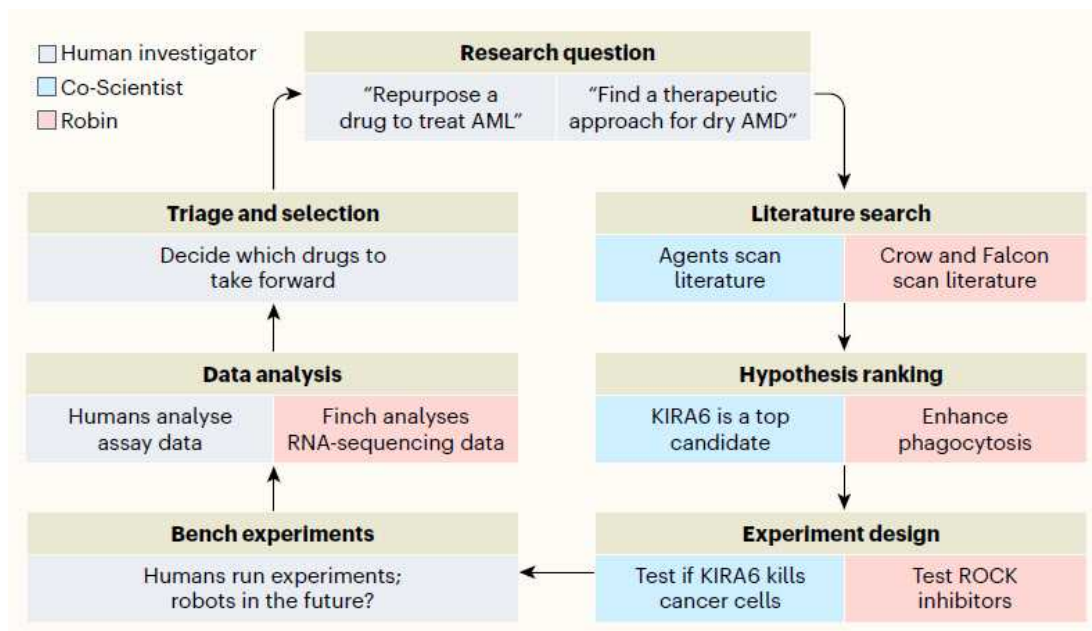
- * 여러 자율형 AI 에이전트가 역할을 나눠 복잡한 문제를 공동으로 해결하는 시스템
- 여러 에이전트가 가설 생성·검토 등의 역할을 나누고, 토론과 비교를 반복해 가설의 순위를 정하고 상위 가설을 개선하는 방식으로 작동
 - Co-Scientist와 Robin*을 실제 질환 연구에 적용해 치료 후보물질을 도출
 - * 해당 연구팀이 개발한 멀티에이전트 AI 시스템으로, Co-Scientist가 가설 생성에 초점을 둔다면, Robin은 실험 데이터 분석까지 아우르는 것이 특징
 - ※ 급성 골수성 백혈병(AML, acute myeloid leukaemia)과 건성 연령관련 황반변성(dry AMD, dry age-related macular degeneration)의 치료 후보물질 제시
 - 제안된 후보물질 중 일부는 연구진의 세포실험에서 효과가 확인됐으며, Robin은 후속 실험에서 얻은 유전자 발현 데이터도 분석
- Co-Scientist의 에이전트 간 협업을 급성 골수성 백혈병 후보물질 도출에 적용
 - AML 약물 재창출 후보로 세포 스트레스 반응에 관여하는 효소 IRE1 α 억제제 KIRA6를 제안
 - ※ 세포실험에서 KIRA6는 비백혈병 대조세포 사멸 농도보다 최대 18배 낮은 농도로 백혈병 세포를 선택적으로 사멸
- 항생제 내성 관련하여 연구진은 Co-Scientist에게 항생제 내성에 대한 가설을 세워보라고 요청
 - AI가 스스로 얼마나 새로운 아이디어를 만들어내는지 보려는 의도로, 배경 정보를 거의 주지 않음
 - Co-Scientist는 이동성 DNA 서열이 박테리오파지 꼬리를 이용해 항생제 내성 유전자를 세균 종간 전달하는 기전 제안
 - ※ 당시 미발표 실험 연구 결과와 일치하는 제안이나, 이를 구성한 개별 정보는 기존 문헌에 공개

- Crow(간략 문헌 탐색)·Falcon(심층 문헌 탐색)·Finch(실험데이터 분석) 3개 에이전트로 구성된 Robin을 dry AMD 치료 후보물질 도출에 활용
 - dry AMD 치료 후보로 망막 색소 상피세포의 식세포 작용(phagocytosis) 저하를 제시
 - 상위 5개 후보물질의 세포실험 결과, ROCK 억제제 Y-27632의 식세포 작용 촉진 효과가 가장 큰 것으로 확인
- ※ Finch는 후속 유전자 발현 데이터를 분석해 Y-27632 처리 세포의 ABCA1 발현이 대조군보다 3배 높다는 점을 확인

■ 문헌 탐색·가설 생성·실험 설계까지는 두 AI가 담당하고 있으며, Robin은 여기에 더해 데이터 분석(Finch) 까지 커버하며 범위 확장

- 두 시스템은 수천 편의 논문과 인접 분야의 정보를 종합해 한 명의 연구자가 감당하기 어려운 가설·후보물질 탐색을 지원
 - 웹 검색·논문 검색·코드 실행 등 외부 도구를 활용해 텍스트 입출력을 넘어 탐색 범위 확대

< 바이오의학 발견 사이클에 진입하는 AI 에이전트 >



¶ 연구 질문 → 문헌 탐색 → 가설 순위화 → 실험 설계 → 벤치 실험 → 데이터 분석 → 선별 → 다시 연구 질문으로 이어지는 하나의 순환 고리에서 각 단계마다 인간 과학자, Co-Scientist(AML 사례)와 Robin(dry AMD 사례)의 역할을 제시

출처 : Nature, "AI systems devise hypotheses and ways to test them", 2026.6.30.

- 실제 데이터 기반 생물정보 분석 벤치마크 BixBench*에서 Robin의 Finch가 비에이전트** 방식의 Claude 3.7 Sonnet보다 높은 성능 기록

* 실제 데이터를 이용해 생물정보 분석 과제의 수행 능력을 평가하는 벤치마크

** 자율적인 의사결정 기능 없이 모델을 사용하는 방식

- 다만 최근 등장한 Claude Opus 4.8과 같은 범용 코딩 에이전트가 이 격차를 좁힐 수 있어, 전문 과학 에이전트의 우위가 지속될지는 불확실

< Co-Scientist와 Robin 비교 >

구분	Co-Scientist	Robin
에이전트 구성·작동 방식	토론·비교·순위화·개선	Crow·Falcon·Finch 협업
적용 분야·과제	급성 골수성 백혈병, 항생제 내성	건성 연령관련 황반변성
기반 LLM	Google Gemini	OpenAI o4-mini, Anthropic Claude
공개 여부	현재 비공개	오픈소스

출처 : Nature, "AI systems devise hypotheses and ways to test them", 2026.6.30.

■ AI 제안 가설은 기존 지식의 재조합 수준으로, 새로운 생물학적 지식 창출을 위한 세 방향의 기술적 확장 필요

- KIRA6와 ROCK 억제제는 기존에 알려졌거나 임상에서 사용 중이며, 항생제 내성 기전의 구체적인 결론은 당시 미발표였지만 이를 구성한 개별 정보는 기존 문헌에 공개
 - 기존 문헌과 독립적으로 새로운 생물학적 지식을 신뢰성 있게 창출할 수 있는지는 여전히 미해결
- 향후 과제로 실험 자동화, 데이터 우선 접근, 생물학적 내부 모델 구축 등 세 가지 방향 제시
 - (실험 자동화) 자율실험실과 가설 생성 시스템을 연계해 여러 가설을 동시 검증
 - ※ 가설 생성 시스템과 자율실험실을 결합한 발표 사례는 아직 없음
 - (데이터 우선 접근) 유전자 교란 스크리닝, 대규모 인간 유전정보, 단일 세포 지도, 연구실 자체 데이터 등을 가설 생성에 활용
 - (생물학적 내부 모델 구축) 미분방정식·에이전트 기반 모델·세포데이터 파운데이션 모델을 활용해 생물학적 근거에 기반한 가설 도출
- AI 성능이 향상되더라도 연구 질문 설정, 신규성·파급효과·임상적 관련성 평가, 전략적·윤리적 판단은 인간 연구자의 역할로 남을 전망
 - 예상치 못한 데이터에 따라 연구 질문 자체를 바꾸는 데 필요한 직관과 용기도 인간 과학자의 역할일 것으로 예측