Fauna Brain™: Agents for Accelerated Drug Target Discovery

Deek Guruge, Michelle M. Li, Poornima Neela, Evelyn Tran, Manashvi Borad, Natalie DeForest, Gennady Gorin, Phil McNamara, Katharine Grabek, Linda Goodman
Fauna Bio

1. Introduction

The traditional drug discovery pipeline is slow, expensive, and prone to failure (Sertkaya et al., 2024; BIO et al., 2021; IQVIA Institute, 2024). A key bottleneck is the evaluation of targets; it is a complex workflow that requires synthesizing diverse biological and pharmacological data and literature evidence, and cannot be solved by a single machine learning model (Buniello et al., 2025; Minikel et al., 2024; Guo et al., 2024; Wu et al., 2023; Li et al., 2024). We introduce Fauna Brain™, a multi-agent system that automates this drug target discovery workflow (Guo et al., 2024; Wu et al., 2023; Li et al., 2024).

Fauna Brain™ decomposes drug target discovery into specialized tasks handled by dedicated agents. It harnesses insights from proprietary data contained in our Convergence™ platform that details the phenomenon of **animal disease resistance (ADR)**, analyzing the adaptations that protect species against conditions like obesity, heart disease, and kidney disease. Fauna Brain™ additionally leverages public human genomic databases and scientific literature (Fauna Bio, 2025; Ferris et al., 2025; Buniello et al., 2025). Altogether, Fauna Brain™ rapidly screens and scores candidate obesity targets, thereby improving the speed, scale, and translational potential of target selection.

2. Methods

Fauna Brain™ is optimized through an iterative process of prompt and software engineering such that each agent's capabilities and reasoning logic are refined specifically for drug target discovery (Wu et al., 2023; Guo et al., 2024). The core orchestration uses a StateFlow-based framework that models the workflow as a finite state machine (FSM) (Wu et al., 2025). In this FSM, each major stage is a 'state' that constrains agent actions and only allows advancement via predefined transitions. This state-based control

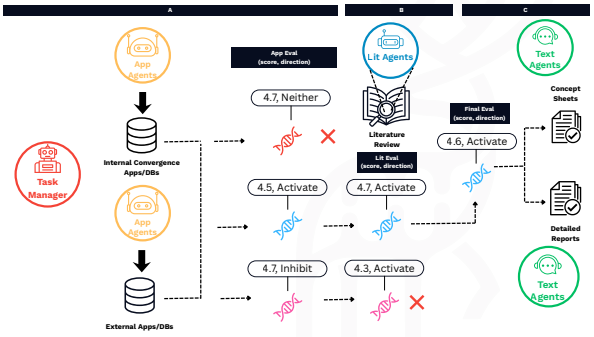

Figure 1: Overview of Fauna Brain™. **(A)** Task Manager deploys App Agents and **(B)** Lit Agents to gather evidence on candidate targets. **(C)** Text Agents generate concept sheets and reports for human evaluation.

provides strong guardrails, enforcing a logical progression that ensures every evaluation is consistent, accurate, and repeatable.

Fauna Brain™ is composed of four main types of AI agents that work together to automate drug discovery; it uses Gemini-2.5-Pro as its backbone (Google DeepMind, 2025) (Fig. 1). The **Task Manager** coordinates specialized agents to automate the research workflow. **App Agents** interact with various software and databases to execute specific tasks, while **Lit Agents** analyze scientific papers to evaluate the candidate target. This framework is supported by various **Generic Agents** that handle general-purpose functions involving vision and text-based analysis.

Fauna Brain™ executes a workflow that begins with **Target Screening**, where App Agents use detailed, role-specific prompts in a zero-shot manner to gather evidence from proprietary and public databases, assigning an initial weighted score and therapeutic direction to each gene (Wu et al., 2023; Fauna Bio, 2025). Targets that meet a predefined threshold and

therapeutic direction advance (Buniello et al., 2025) to the **Literature Review** stage. Lit Agents use a retrieval-augmented generation (RAG) approach to process scientific papers automatically downloaded and embedded per target, extracting key data on safety, druggability, mouse models, etc to generate a score and therapeutic direction. Targets that have congruent therapeutic direction from App and Lit Agents and meet specific thresholds are selected by Fauna Brain™ as promising candidates, and concept sheets and detailed reports can be generated for them (Lewis et al., 2020; Priem et al., 2022). The entire high-throughput system is deployed on Google Cloud Platform (GCP), where a **parallel architecture** allows dozens of agents to operate simultaneously for rapid, large-scale screening.

3. Results

3.1. Benchmark Performance and Outcomes

To validate Fauna Brain's ability to solve the "needle-in-a-haystack" challenge of target discovery, we created a deceptively difficult "haystack" of 549 potential obesity targets, all of which were pre-selected for a plausible biological link to the disease based on differential expression data. Of them, three targets met a comprehensive matrix of criteria for druggability, safety, and therapeutic potential (Emmerich et al., 2021). The remaining 546 candidates formed the negative set, as they failed to meet these standards during multiple rounds of target selection. This benchmark therefore evaluated whether the AI could distinguish a truly viable drug target from a field of plausible but flawed candidates.

Table 1: Benchmarking Fauna Brain™ on nominating targets for obesity.

Model	Bal. Acc	Prec.	Recall	F1
Transformer	0.5339	0.0059	1.0000	0.0117
Gemini (2.5 Pro)	0.4185	0.0000	0.0000	0.0000
Fauna Brain™	0.9982	0.6000	1.0000	0.7500

Fauna Brain™ outperforms both the transformer and Gemini models in accurately identifying drug targets for obesity (Tab. 1). The transformer-based model suffers from low precision, and Gemini fails to identify any of the positive cases. These results underscore the value of our multi-agent approach and unique data foundation. Although Gemini is the backbone of Fauna Brain™, it struggles to perform a comprehensive evaluation, indicating that an LLM is insufficient for drug discovery. In contrast, Fauna Brain™ achieves perfect recall and high precision.

3.2. Deployment, Usage, and Impact

Actively deployed in pharmaceutical partnerships, Fauna Brain™ accelerates our research workflows, rapidly scaling to new disease areas. The system operates at scale on the cloud, increasing the throughput and scope of our target discovery campaigns. The deployment has improved efficiency. A manual target discovery process involving six researchers typically takes two weeks at a cost of over $30k. Fauna Brain™ can screen several thousand targets in 24 hours for about $0.01 per target, making it over 400× cheaper and 14× faster while screening a vastly larger set of candidates with high accuracy and reliability, **and has already identified novel targets for obesity now under further investigation.**

4. Discussion

Developing Fauna Brain™ required overcoming three core challenges: achieving operational scale, ensuring reliable access to scientific literature, and guaranteeing the accuracy of our AI agents. To this end, we engineered the system to run parallel agents on Google Cloud Platform (Google Cloud, 2025a,b), created a custom retrieval system and tools to access scientific papers reliably (Lewis et al., 2020), and designed detailed prompts for structured execution as guardrails.

Future work will focus on automating and sophisticating the system at every level: systematically optimizing prompts and models using a framework like DSPy (Khattab et al., 2023), preprocessing a corpus of literature and evolving our RAG pipeline with more advanced techniques like hybrid search (Lewis et al., 2020; Thakur et al., 2021), and enabling greater agent autonomy (Guo et al., 2024). These improvements will allow Fauna Brain™ to tackle novel scientific challenges with even more creativity and flexibility.

Fauna Brain™ represents a significant advancement in drug discovery. By combining a sophisticated agent architecture with unique insights from ADR data, it accelerates the research pipeline while enhancing our ability to discover more effective medicines with greater translational potential.

References

BIO, Informa (Citeline), and QLS Advisors. Clinical Development Success Rates 2011–2020. Technical report, Biotechnology Innovation Organization, 2021. Authoritative 2010s–2020 snapshot of phase transition and LOA.

A. Buniello et al. Open Targets Platform 2025: facilitating therapeutic hypothesis building in drug discovery. *Nucleic Acids Research*, 53(D1):D1467–D1480, 2025.

C.H. Emmerich et al. Improving target assessment in biomedical research: the GOT-IT recommendations. *Nature Reviews Drug Discovery*, 20:64–81, 2021.

Fauna Bio. Approach / Convergence™. https://www.faunabio.com/approach, 2025. Company technical overview.

Elliott Ferris, Josue D. Gonzalez Murcia, Adriana Cristina Rodriguez, Susan Steinwand, Cornelia Stacher Hörndli, Dimitri Traenkner, Pablo J. Maldonado-Catala, and Christopher Gregg. Genomic convergence in hibernating mammals elucidates the genetics of metabolic regulation in the hypothalamus. *Science*, 389(6759):494–500, 2025. doi: 10.1126/science.adp4025.

Google Cloud. Cloud batch: Overview. https://cloud.google.com/batch/docs/overview, 2025a. Accessed 2025-09-09.

Google Cloud. Gke best practices: Batch workloads. https://cloud.google.com/kubernetes-engine/docs/best-practices/batch, 2025b. Accessed 2025-09-09.

Google DeepMind. Gemini 2.5 — model update blog. https://blog.google/, Mar 2025. Accessed September 2025.

T. Guo et al. Large Language Model-Based Multi-Agents: A Survey. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, 2024.

IQVIA Institute. Global Trends in R&D 2024: Activity, Productivity and Enablers. Technical report, IQVIA, 2024. Recent composite success-rate analysis; still low overall success.

Omar Khattab, Jesse Hall, Keshav Santhanam, et al. Dspy: Compiling declarative language model calls into self-improving pipelines. *arXiv*, 2023. URL https://arxiv.org/abs/2310.03714.

P. Lewis et al. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2020.

X. Li et al. A survey on LLM-based multi-agent systems: workflow design and applications, 2024. Published by SpringerLink.

E.V. Minikel et al. Refining the impact of genetic evidence on clinical success. *Nature*, 2024.

Jason Priem, Heather Piwowar, and Richard Orr. OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. *arXiv*, 2022. URL https://arxiv.org/abs/2205.01833.

A. Sertkaya et al. Costs of Drug Development and Research and Development Intensity in the US, 2000-2018. *JAMA Network Open*, 2024. Systematic review summarizing capitalized R&D cost ranges.

Nandan Thakur, Nils Reimers, Andreas Rücklé, Abhishek Srivastava, and Iryna Gurevych. Beir: A heterogeneous benchmark for zero-shot evaluation of information retrieval models. In *NeurIPS Datasets and Benchmarks Track*, 2021. URL https://arxiv.org/abs/2104.08663.

Q. Wu et al. AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversation. arXiv preprint arXiv:2308.08155, 2023. Updated 2024.

Yiran Wu, Tianwei Yue, Shaokun Zhang, Chi Wang, and Qingyun Wu. Stateflow: Enhancing llm task-solving through state-driven workflows. *arXiv*, 2025. URL https://arxiv.org/abs/2403.11322. First posted 2024; revised 2025.