BiomniGEM: A paradigm shift in biological AI through multimodal understanding

Jilin-AI
Jilin University
Jilin-AI@outlook.com

Abstract

We introduce **BiomniGEM**, a large language model (LLM) specifically designed for multi-omics data understanding and biological reasoning. The model is trained on **SynBioCoT**, a dataset containing **¿10k** samples with multi-trace Chain-of-Thought (CoT) annotations, targeting multi-omics reasoning tasks.

Existing general-purpose LLMs struggle to interpret raw biological data such as **DNA**, **cell expression profiles**, and **proteins**, and find it even more challenging to perform causal reasoning based on experimental evidence (Wei et al., 2022a). To address this, we propose **BiomniGEM**, a biological reasoning model built upon **Qwen3-8B** (Yang and Qwen Team, 2025), which *textifies* three omics modalities (DNA / cell profile / protein) and explicitly embeds omics information during joint training, enabling the model to *directly read* and reason over biological data. Our textification draws on prior ideas of turning single-cell profiles into "cell sentences" (Levine et al., 2024) and complements sequence-level foundation models in genomics and proteomics (Ji et al., 2021; Ferruz et al., 2022).

The **SynBioCoT** dataset spans cell profiles, DNA, and protein modalities and covers **five categories of omics tasks**, aiming to systematically teach both *omics understanding* and *biological reasoning*. Experimental results show that BiomniGEM consistently **outperforms** leading commercial and open-source models in understanding raw omics data and demonstrates **superior reasoning performance**, achieving the best results on our benchmark tasks (in line with recent evidence that explicit reasoning can boost single-cell task generalization (Fang et al., 2025)).

1 Introduction

In recent years, large language models (LLMs) have achieved remarkable progress in general natural language processing tasks such as question answering, summarization, and text generation (Wei et al., 2022a). However, when applied to specialized domains like biology—particularly in complex scenarios involving single-cell transcriptomics and multi-omics data analysis—LLMs still face significant challenges. Traditional bioinformatics approaches rely on task-specific neural architectures that, while effective within narrow scopes, often lack **generality** and **interpretability**. Moreover, they fail to leverage the extensive knowledge and linguistic priors accumulated during LLM pretraining. A central question thus arises: how can an LLM **truly understand** DNA sequences, cell expression profiles, and protein data, and perform **biological reasoning** upon them?

In this work, we present **BiomniGEM**, an LLM designed for **raw biological data understanding** and **scientific reasoning**. Specifically, we convert single-cell expression data into a **"cell sentence"** by rank-ordering gene names according to expression levels (gene-ranking textualization), formatted as `<cell>...</cell>` following the ideas of Cell2Sentence (Levine et al., 2024). Likewise, DNA and protein sequences are enclosed within `<dna>...</dna>` and `<protein>...</protein>` tags, respectively. This representation allows the LLM to **directly**

process and interpret multi-omics inputs without leaking meta-information. The **textification** strategy not only enables the model to handle multimodal inputs but also provides a unified interface for **joint multi-omics training**.

To systematically train and evaluate these capabilities, we constructed **SynBioCoT**, a dataset containing **more than 10k** samples, each annotated with multi-step reasoning traces (multi-trace CoT) suitable for reinforcement-learning (RL) supervision and feedback. **SynBioCoT** covers five classes of biological tasks—three single-modality tasks (cell / DNA / protein) and two multi-omics tasks (**multi-omics alignment** and **multi-omics integration**)—and includes instruction-alignment data that help the model interpret textified omics representations and exploit fine-grained features (e.g. rank information within cell sentences). We developed an **automatic annotation pipeline** based on asymmetric evidence and limited meta-information, incorporating **rejection sampling** to improve data quality and using incorrect reasoning traces as **negative CoT samples** for additional supervision and faster convergence. This process can be viewed as a form of **self-distillation**—rather than simply distilling answers from a stronger model, we elicit latent knowledge from the base model itself and re-teach it in a reasoning-explicit format, thereby enhancing autonomous reasoning within biological contexts (Fang et al., 2025). We find that even when applied to baseline models, this automatic labeling and training approach leads to consistent performance gains.

In summary, our contributions are:

- We propose **BiomniGEM**, an LLM tailored for raw biological data understanding and reasoning.
- We construct **SynBioCoT**, a >10k-sample multi-omics CoT dataset covering both single- and cross-omics tasks.
- We introduce an **asymmetric-information-based automatic labeling** and **rejection-sampling** strategy, establishing a reusable CoT data-construction paradigm for AI-for-Science.

2 Related Works

2.1 Open-Weight General Models and Biological Adaptation

Recently, OpenAI released the **GPT-OSS** series (gpt-oss-20B / 120B) as the first open-weight models supporting **Chain-of-Thought (CoT)** reasoning and structured output (Wei et al., 2022a). These models adopt a **Mixture-of-Experts (MoE)** architecture, achieving performance comparable to o4-mini on standard reasoning benchmarks and exhibiting strong tool-use and step-by-step reasoning capabilities.

The biological variant, **Bio-GPT-OSS-20B**, further fine-tunes GPT-OSS-20B on PubMed literature, life-science textbooks, and biomedical QA datasets, outperforming general-purpose LLMs in biological text comprehension and scientific QA, and supporting gradual reasoning and evidence extraction.

However, these models are still **trained primarily on natural-language corpora**, without explicitly embedding **raw omics data layers** (e.g. DNA sequences, protein sequences, or single-cell profiles). As a result, their ability in **multi-omics reasoning** and **data-driven causal interpretation** remains limited.

2.2 Textualization of Omics Data

Several studies have attempted to embed omics data into textual sequences so that LLMs can process biological data within a unified linguistic space. A representative example is

Cell2Sentence (Levine et al., 2024), which converts cell expression profiles into *cell sentences* by sorting gene names in descending order of expression levels, training an LLM for classification or generative tasks.

The approach achieved performance comparable to traditional representation learning foundation models. Nevertheless, it mainly relies on an **instant-response** training paradigm, lacking explicit **CoT reasoning traces**, and thus demonstrates poor generalization across downstream biological tasks. This suggests that data textualization alone—without reasoning supervision—fails to unlock the full potential of LLMs in scientific problem solving.

2.3 Reasoning-Driven Biological Tasks and Cross-Task Generalization

Another research direction emphasizes **explicit reasoning** to enable LLMs to perform logical inference over biological data. **Cell-o1** (Fang et al., 2025) was inspired by expert-driven batch-level cell annotation and introduced the *cell puzzle* task, where the model must perform one-to-one matching among multiple cell sentences and utilize contextual metadata to predict cell types.

This design compels the model to perform cross-sample reasoning rather than single-instance classification, achieving significantly higher **batch-level accuracy** than conventional LLMs. Unlike Cell2Sentence, which exhibits almost no generalization across tasks, Cell-o1 demonstrates strong robustness and transferability, outperforming its baseline model **Qwen3-8B** (Yang and Qwen Team, 2025) in multiple biological tasks. These findings collectively highlight that **biological reasoning** remains a key deficiency of current general-purpose LLMs, and incorporating **explicit reasoning processes** is crucial to improving cross-task generalization and scientific interpretability.

However, **systematic multi-omics integration and reasoning-oriented CoT training** remain largely unexplored areas. Against this backdrop, we propose **BiomniGEM** to address these limitations.

3 SynBioCoT: A Large-Scale Multi-omics CoT Dataset

3.1 Dataset Overview

To enable large language models (LLMs) to comprehend raw biological data, we introduce **SynBioCoT**—a large-scale multi-omics *Chain-of-Thought (CoT)* reasoning dataset specifically designed for synthetic biology. SynBioCoT is not merely a corpus of biological text; rather, it aims to teach models how to *reason* over multiple omics modalities—**DNA**, **cell expression profiles**, and **proteins**—through the design of cross-modality tasks and multi-step reasoning traces Wei et al. (2022b).

Unlike traditional biological datasets that contain only single-turn question–answer pairs or static annotations, each SynBioCoT sample consists of three components: an **omics-based query** that explicitly includes `<cell>`, `<dna>`, and/or `<protein>` tags to represent distinct biological modalities; a **reasoning trace (CoT)** providing a multi-step logical chain that demonstrates how the model arrives at its conclusion; and a **final answer** giving a concise conclusion, often describing a biological function, mechanism, or causal relationship.

This design allows the model to learn not only *what* the correct answer is, but also *how* the reasoning process unfolds—providing a structured paradigm for scientific reasoning within LLMs.

3.2 Omics Textification

To bridge the gap between symbolic biological data and language models, we adopt a **textification** interface that converts omics modalities into structured textual representations. For

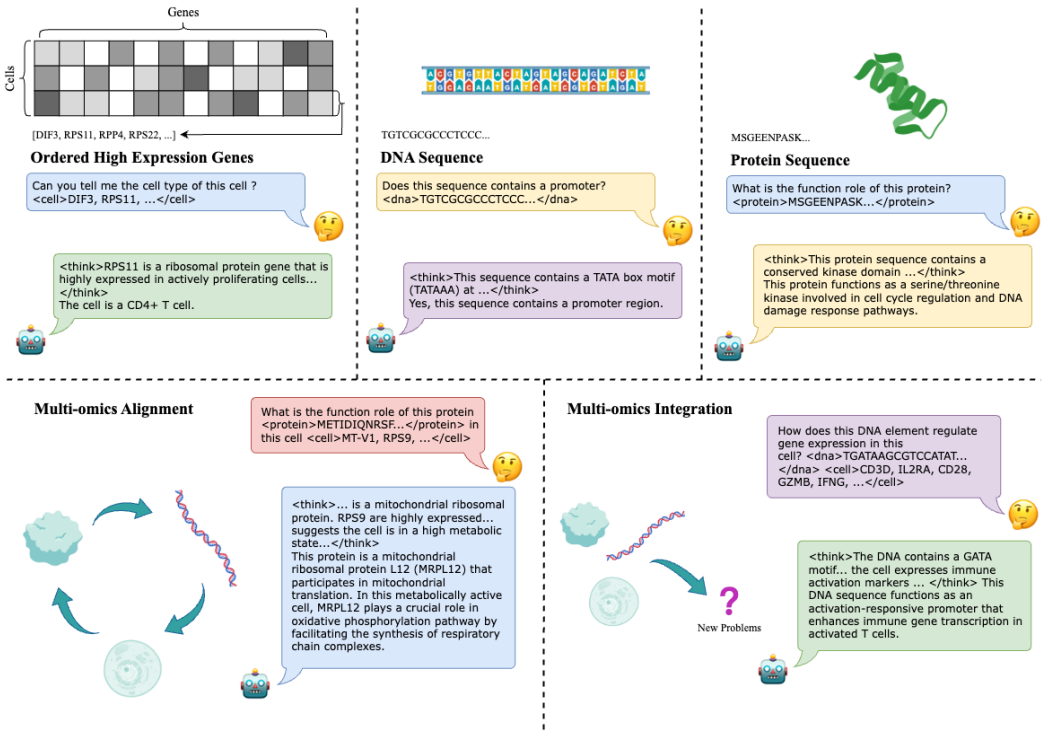

Figure 1: Overview of SynBioCoT task categories and structure.

cell data, single-cell gene expression profiles are transformed into tokenized sequences by ranking genes in descending order of expression and enclosing them in `<cell>...</cell>`. This preserves rank information, enabling LLMs to interpret expression intensity through token order. Although the rank-based representation introduces mild distortion relative to the original UMI space, prior work has shown that a lightweight encoder can reconstruct UMI distributions with ∼85% accuracy from ranked gene lists, indicating that the "cell sentence" representation retains most biological information while being compatible with language modeling. For **DNA** and **protein** modalities, nucleotide sequences are directly encoded as `<dna>ATG...TAA</dna>`, and amino acid sequences are represented as `<protein>MSDSEK...</protein>`, both preserving raw sequence information without auxiliary metadata.

Two key mechanisms support SynBioCoT's design: **textification**, which maps raw omics data into a unified symbolic format (`<dna>`, `<protein>`, `<cell>`), allowing models to process sequence-level evidence within a linguistic space; and **reasoning-chain annotation**, which records intermediate logical steps, teaching the model how to traverse from empirical evidence to causal explanations.

3.3 Task Design

Building upon prior text-based representations, SynBioCoT addresses two core challenges: (1) How can models fully interpret omics information? (2) How can models jointly reason over multiple omics modalities? Hasin et al. (2017)

3.3.1 Single-omics Tasks

For each modality—**cell**, **DNA**, and **protein**—we design two types of tasks: **annotation tasks** that require concise factual outputs (e.g., gene function or regulatory role), and **explanation tasks** that require biologically meaningful reasoning describing *why* or *how* the observed phenomenon occurs.

A major motivation for CoT-based supervision arises from the nature of biological data: high noise, sparse markers, weak direct correlations, and long reasoning chains. Conventional instruction-tuned models tend to exhibit overly conservative behavior—for example, in differential expression prediction, models frequently default to "no change" predictions (recall of "no" $\approx 100\%$) due to reliance on prior probabilities rather than biological inference. Interestingly, when the *correct answer* is provided as context, the same model can often articulate accurate biological reasoning—suggesting that the limitation lies not in knowledge, but in the absence of biological-style reasoning paths Chen et al. (2024).

To address this, we propose an **asymmetric-information automatic annotation framework** for CoT construction:

1. **Guided generation** – a labeling model receives the correct answer or partial metadata and is asked to *explain* or *justify* it. A secondary model then restructures the response into a standardized, multi-step CoT schema.

2. **Reverse verification** – the CoT (without the answer) is fed back to the original labeling model, which must infer the final answer solely from the reasoning trace.

3. **Consistency filtering (rejection sampling)** – a lightweight evaluator model compares the generated answer against the ground truth. Correct pairs are retained as *positive CoT samples*, while inconsistent ones are labeled as *negative samples*.

This yields a dataset containing both positive and negative reasoning traces, where negative samples provide strong contrastive supervision and accelerate convergence during training.

3.3.2 Multi-omics Tasks

Simply concatenating multiple omics modalities in the input proved ineffective. Empirical tests show that generic LLMs fail to leverage additional omics information—sometimes even performing worse when given more modalities. This degradation arises from biological data's inherent noise and cross-modality interference, where conflicting markers obscure reasoning Subramanian et al. (2023).

To address this, SynBioCoT introduces a **two-stage multi-omics design**. In the first stage, **alignment tasks** teach the model to map relationships between modalities—for example, explaining why a given DNA sequence encodes a particular protein, or describing the biological role of a protein in a specific cell type. These tasks promote the formation of *semantic bridges* across omics layers. In the second stage, **integration tasks** require joint inference using multiple omics sources to answer a third-level question (e.g., "Given a DNA variant, its protein conformation, and a cell expression profile, what phenotype change will occur and why?"). These tasks encourage the model to resolve noise and redundancy across sources, fostering true multi-omics causal reasoning.

The multi-omics dataset construction follows the same asymmetric annotation framework, but with explicit instructions compelling the labeling model to integrate evidence from all modalities. Task-specific guidance (e.g., "align before infer") is injected to ensure biologically meaningful CoT traces and stable cross-modal reasoning.

3.4 Instruction Alignment

Before fine-tuning on reasoning tasks, models must first learn the semantics of the structured input format. We therefore introduce an **instruction alignment** stage to align model understanding with the `<cell>`, `<dna>`, and `<protein>` representations Wang et al. (2023). In this stage, the model is trained on **semantic interpretation tasks**, where it is asked to describe what each data segment represents and which biological information it encodes. Prompts include explicit contextual hints (e.g., how the cell sentence is constructed) and require the model

to interpret rank information explicitly. Concurrently, **format and protocol tuning** ensures that models follow unified input–output protocols, producing CoT-style reasoning followed by concise answers. This alignment stage ensures that the model develops an intuitive understanding of the omics-text interface, leading to more stable and semantically consistent downstream training.

4 Training

We fine-tune and optimize the **Qwen3-8B** model Qwen Team (2024) on the **SynBioCoT** dataset to enable multi-omics comprehension and biological reasoning. The overall training pipeline comprises four stages: **BioCorpus Pretraining**, **Instruction Alignment SFT**, **Cold-start SFT**, and **GRPO Optimization (Group-based Reinforcement Preference Optimization)**.

4.1 BioCorpus Pretraining

To enhance the model's representation of biological knowledge, we conduct domain-specific pretraining on a self-constructed corpus, **BioCorpus**. The dataset combines three sources: *NCBI Biology Books* and *OpenStax Biology* textbooks (providing systematic biological foundations), *iGEM community wikis and component descriptions* (offering experimental and applied contexts of synthetic biology), and a subset of *PubMedQA* Jin et al. (2019) (supplying structured QA pairs with verified biological reasoning).

Textbooks contribute hierarchical conceptual structure, iGEM data capture experimental semantics, and PubMedQA ensures QA quality and linguistic clarity. Compared with full research papers, these sources present higher information density and reduced structural noise, thus yielding more stable convergence for mid-sized models ($\approx$ 8 B parameters). The resulting corpus ($\approx$ 1 B tokens) supports lightweight in-domain pretraining on Qwen3-8B to "refresh" biological knowledge and acquire the discourse patterns of scientific writing.

4.2 Instruction Alignment SFT

Following pretraining, we conduct supervised fine-tuning on an **instruction alignment** dataset to align the model with the structured input–output protocol of multi-omics tasks. Inputs employ modality tags `<cell>...</cell>`, `<dna>...</dna>`, and `<protein>...</protein>`, whereas outputs follow a unified CoT-style format.

The loss function is defined as:

$$\mathcal{L}_{\text{IA}} = \mathbb{E}_{(x,y)}[-\log \pi_\theta(y_{\text{UOP}} \mid x)] + \lambda_{\text{fmt}} \cdot \mathbf{1}_{\text{format violation}}, \tag{1}$$

where π_θ denotes the model distribution, and λ_{fmt} is a penalty coefficient assigned when output format deviates from the defined protocol. This stage ensures robust modality recognition and structural compliance before downstream reasoning tasks.

4.3 Cold-start SFT

Subsequently, we fine-tune the aligned model on single-modality tasks (cell, DNA, and protein) from **SynBioCoT**, enabling it to internalize CoT reasoning schemas and biological semantics.

The objective function is:

$$\mathcal{L}_{\text{SFT}} = \mathbb{E}_{(x,r^*,a^*)}[-\log \pi_\theta(r^* \mid x) - \log \pi_\theta(a^* \mid x, r^*)], \tag{2}$$

where r^* and a^* represent the ground-truth reasoning trace and final answer, respectively.

4.4 GRPO Optimization

To further refine reasoning quality and structural coherence, we adopt **Group-based Reinforcement Preference Optimization (GRPO)** DeepSeek-AI (2025), a reinforcement-learning approach that balances RLHF stability with sample efficiency by comparing multiple candidate responses within each group.

The optimization objective is:

$$\max_{\theta} \mathbb{E}_x \left[\mathbb{E}_{y \sim \pi_\theta(\cdot|x)}[R(x,y)] - \beta \, \mathrm{KL}(\pi_\theta(\cdot \mid x) \parallel \pi_{\mathrm{ref}}(\cdot \mid x)) \right], \tag{3}$$

where π_{ref} denotes the SFT reference policy and β is the KL regularization coefficient. Candidate rewards for each prompt are normalized via group statistics:

$$\tilde{r}_i = \frac{r_i - \mu_x}{\sigma_x + \epsilon}, \qquad q_i = \frac{\exp(\alpha \tilde{r}_i)}{\sum_j \exp(\alpha \tilde{r}_j)}. \tag{4}$$

The resulting loss for GRPO training is:

$$\mathcal{L}_{\mathrm{GRPO}} = -\,\mathbb{E}_x \left[\sum_i q_i \log \pi_\theta(y_i \mid x) \right] + \beta \, \mathrm{KL}(\pi_\theta(\cdot \mid x) \parallel \pi_{\mathrm{ref}}(\cdot \mid x)). \tag{5}$$

4.5 Reward Design and Multi-trace Construction

4.5.1 Reward Function

Each output sample consists of a **Reasoning** section and a final **Answer**. Because these pairs are generated via controlled annotation rather than online reinforcement, the reward depends purely on structural quality:

$$R(x,y) = w_{\mathrm{len}} R_{\mathrm{len}} - w_{\mathrm{rep}} P_{\mathrm{rept}}, \tag{6}$$

where R_{len} encourages concise yet sufficient reasoning, and P_{rept} penalizes redundant text. Specifically,

$$R_{\mathrm{len}} = \exp\left(-\frac{|\ell(y_{\mathrm{cot}}) - \ell^*|}{\tau} \right), \tag{7}$$

where ℓ^* is the target reasoning length and τ a smoothing factor.

4.5.2 Multi-trace Sample Construction

For each query x, we construct five controlled reasoning variants with predefined rewards:

Type	**Description**	**Reward**
Gold	Correct CoT reasoning (from SFT)	1.0
Short	Retains 60% of steps, correct answer	0.5
Long	Inserts 30% extraneous steps, partially noisy	0.5
Redundant	Duplicates reasoning fragments, inducing verbosity	0.7
No-reason	Retains only minimal reasoning (20%)	0.2
Negative	Incorrect reasoning or answer	-1.0

Table 1: Controlled multi-trace variants for GRPO training.

This contrastive design enhances sensitivity to reasoning quality and enables GRPO to effectively distinguish high- and low-quality CoT samples, improving both logical consistency and biological interpretability.

5 Conclusion

In this paper, we present **BiomniGEM**—a large language model (LLM) capable of understanding raw omics data and performing biological reasoning. We constructed the **SynBio-CoT** dataset, which systematically embeds multi-omics data—including **DNA**, **cell expression profiles**, and **proteins**—into textual representations. Using an asymmetric-information automatic annotation mechanism, we generated **multi-trace reasoning samples** that capture structured biological logic.

Building upon this dataset, we fine-tuned the **Qwen3-8B** model through a four-stage training pipeline: *BioCorpus pretraining*, *instruction alignment*, *cold-start supervised fine-tuning*, and *GRPO reinforcement optimization*. This multi-stage process enables BiomniGEM to perform interpretable and biologically consistent multi-omics reasoning.

Experimental results demonstrate that BiomniGEM significantly outperforms both general-purpose and commercial LLMs across a range of biological reasoning tasks. These findings validate the necessity and effectiveness of **Chain-of-Thought (CoT)** reasoning in biological contexts. The model not only comprehends raw omics information but also produces causal explanations and functional analyses consistent with biological principles, establishing a new paradigm for applying AI in life science research.

To promote community research and reproducibility, we have **open-sourced** the model weights, dataset, and construction framework for use and improvement by researchers and the synthetic biology community. We believe that BiomniGEM is not merely a model for biological research, but an **open platform** at the intersection of AI-for-Science and synthetic biology, laying the foundation for explainable and reproducible scientific AI.

Acknowledgements

We would like to express our sincere gratitude to the **iGEM community** for providing open resources, documentation, and inspiration that made this work possible. We thank the developers of the **Qwen model** for their contributions to open and high-quality large language models, which served as the foundation of BiomniGEM. We also acknowledge **OpenAI** for offering annotation and automation tools that facilitated the large-scale construction of reasoning datasets.

This work was supported by **Jilin University** and the **Jilin-AI iGEM 2025 Team**. We deeply appreciate the efforts, creativity, and collaboration of all team members who contributed to the design, data curation, model training, and experimental validation of BiomniGEM.

References

Qingyu Chen, Jingcheng Du, Sun Kim, W. John Wilbur, and Zhiyong Lu. Large language models in biomedical natural language processing: benchmarks, baselines, and recommendations. *npj Digital Medicine*, 7(1):8, 2024. URL `https://www.nature.com/articles/s41746-023-00982-9`. doi: 10.1038/s41746-023-00982-9.

Haotian Cui, Chloe Wang, Hassaan Maan, Kuan Pang, Fengning Luo, Nan Duan, and Bo Wang. scgpt: toward building a foundation model for single-cell multi-omics using generative ai. *Nature Methods*, 21(8):1470–1480, 2024. URL `https://www.nature.com/articles/s41592-024-02201-0`. doi: 10.1038/s41592-024-02201-0.

DeepSeek-AI. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. *arXiv preprint arXiv:2501.12948*, 2025. URL `https://arxiv.org/abs/2501.12948`. Introduces Group Relative Policy Optimization (GRPO).

Y. Fang et al. Cell-o1: Training LLMs to solve single-cell reasoning tasks with batch-level rewards. *arXiv preprint arXiv:2506.02911*, 2025. URL https://arxiv.org/abs/2506.02911.

Noelia Ferruz, Steffen Schmidt, and Birte Höcker. ProtGPT2 is a deep unsupervised language model for protein design. *Nature Communications*, 13(1):4348, 2022. URL https://www.nature.com/articles/s41467-022-32007-7. doi: 10.1038/s41467-022-32007-7.

Minsheng Hao, Jia Gong, Xin Zeng, Chiming Liu, Yixin Guo, Yu Zhang, and Chen Cao. Transformers for single-cell analysis. *Nature Methods*, 21(11):1985–1987, 2024. URL https://www.nature.com/articles/s41592-024-02353-z. doi: 10.1038/s41592-024-02353-z.

Yehudit Hasin, Marcus Seldin, and Aldons Lusis. Multi-omics integration in the age of million single-cell data. *Nature Reviews Genetics*, 18(8):710–717, 2017. URL https://www.nature.com/articles/nrg.2017.54. doi: 10.1038/nrg.2017.54.

Yanrong Ji, Zhihan Zhou, Han Liu, and Ramana V. Davuluri. DNABERT: pre-trained bidirectional encoder representations from transformers for DNA-language in genome. *Bioinformatics*, 37(15):2112–2120, 2021. URL https://academic.oup.com/bioinformatics/article/37/15/2112/6128680. doi: 10.1093/bioinformatics/btab083.

Qiao Jin, Bhuwan Dhingra, Zhengping Liu, William W. Cohen, and Xinghua Lu. Pubmedqa: A dataset for biomedical research question answering. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 2567–2577, Hong Kong, China, 2019. Association for Computational Linguistics. URL https://aclanthology.org/D19-1259. doi: 10.18653/v1/D19-1259.

Daniel Levine, David van Dijk, et al. Cell2sentence: Teaching large language models the language of single cells. *bioRxiv*, 2024. URL https://www.biorxiv.org/content/10.1101/2023.09.11.557287v3. doi: 10.1101/2023.09.11.557287.

Qwen Team. Qwen2.5: A party of foundation models. *arXiv preprint arXiv:2412.15115*, 2024. URL https://arxiv.org/abs/2412.15115.

Iman Subramanian, Srikant Verma, Shantanu Kumar, Abhishek Jere, and Kulandaisamy Anamika. Multi-omics integration in biomedical research – a metabolomics-centric review. *Analytica Chimica Acta*, 1237:340533, 2023. URL https://www.sciencedirect.com/science/article/pii/S0003267022012174. doi: 10.1016/j.aca.2022.340533.

Yizhong Wang, Yeganeh Kordi, Swaroop Mishra, Alisa Liu, Noah A. Smith, Daniel Khashabi, and Hannaneh Hajishirzi. Self-instruct: Aligning language models with self-generated instructions. *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 13484–13508, 2023. URL https://aclanthology.org/2023.acl-long.754. doi: 10.18653/v1/2023.acl-long.754.

Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed H. Chi, Quoc Le, and Denny Zhou. Chain-of-thought prompting elicits reasoning in large language models. *arXiv preprint arXiv:2201.11903*, 2022a. URL https://arxiv.org/abs/2201.11903.

Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed H. Chi, Quoc V. Le, and Denny Zhou. Chain-of-thought prompting elicits reasoning in large language models. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 35, pages 24824–24837, 2022b. URL https://proceedings.neurips.cc/paper_files/paper/2022/hash/9d5609613524ecf4f15af0f7b31abca4-Abstract-Conference.html.

A. Yang and Qwen Team. Qwen3 technical report. *arXiv preprint arXiv:2505.09388*, 2025. URL https://arxiv.org/abs/2505.09388.