

Performance

100%

80%

60%

40%

20%

0%

2024

2025

2026

2027

2028

2029

2030

- MMLU Pro Biology
- WMDP Biology
- GPQA Diamond Biology
- LAB-Bench Protocol
- LAB-Bench SeqQA
- LAB-Bench Cloning
- BioLP-bench
- LAB-Bench LitQA2

