Ei

Basecamp Research Launches Trillion Gene Atlas to Scale AI-Designed Therapeutics

Descrizione

COMUNICATO STAMPA â?? CONTENUTO PROMOZIONALE

AUSTIN, Texas and SAN JOSE, Calif., March 18, 2026 /PRNewswire/ â?? Basecamp Research, a frontier AI lab for biological design, today announced the launch of the Trillion Gene Atlas, a landmark scientific initiative to generate and model biological data at the trillion-gene scale. Launched in collaboration with Anthropic, Ultima Genomics and PacBio, and powered by NVIDIA AI infrastructure, the Trillion Gene Atlas aims to expand known evolutionary genetic diversity 100-fold by collecting genomic data from more than 100 million species across thousands of sites worldwide.

This is possible thanks to Basecamp Researchâ??s growing network of global biodiversity partners. Its ultimate goal is to provide the vast, diverse training data required for AI systems to learn from evolution to design new medicines on demand.

â??Todayâ??s biological AI models are trained on a narrow slice of life on Earth,â?• said Glen Gowers, Co-founder and CEO of Basecamp Research, speaking at SXSW in Austin. â??The Trillion Gene Atlas expands the known genetic universe by orders of magnitude beyond what is in public databases. Training models at this scale establishes a new paradigm for programmable therapeutic design.â?•

The initiative, which is on the scale of the Human Genome Project, was unveiled during the Health Track at SXSW and the NVIDIA GTC conference in San Jose.

With huge increases in model size and computing power, diverse data is a critical enabler for progress in AI drug development and real-world benchmarks. All current sequence-based foundation models rely on variants of the same public repositories, with 80% of these trained on a public database containing fewer than 250 million sequences.

Basecamp Researchâ??s EDEN foundation models, released in January, bypass the industryâ??s evolutionary â??data wallâ?• by training entirely on BaseDataâ?¢, a proprietary genomic database that is currently more than 10 times larger than all public resources combined. By learning from an

unprecedented 10 billion new-to-science genes across 1 million newly discovered species, EDEN unlocked critical new scaling laws for AI in biology.

This massive expansion in dataset diversity moved EDEN beyond simple prediction, making it the first model capable of designing diverse therapeutics directly from a disease prompt. In wet-lab validation, EDEN demonstrated zero-shot activity in primary human T-cells without any human or clinical data needed. The model has successfully generated hits across multiple frontier modalities, notably pioneering AI-Programmable Gene Insertion (aiPGI) to insert healthy genes and designing targeted antimicrobial peptides with a 97% hit rate against priority pathogens.

The Trillion Gene Atlas builds on this approach by greatly expanding the breadth and contextual depth of genomic data in the known â??internet of biologyâ?• suitable for AI training.

â??Bigger models alone arenâ??t enough,â?• added Phil Lorenz, CTO of Basecamp Research. â??EDEN showed that performance in biological AI follows much steeper scaling trajectories with higher quality and fully contextualized data. The Trillion Gene Atlas extends that principle 100-fold.â?•

Over the past six years, Basecamp Research has built a network of scientific collaborators across 31 countries, establishing a scalable evolutionary genomics pipeline purpose-built for AI training. By pioneering a combination of new regulatory and economic frameworks and fully off-grid DNA sequencing technologies, the company collects high-quality genomic data from ecosystems beyond the reach of traditional laboratories.

These partnerships are grounded in knowledge exchange, local capacity building, and equitable Access and Benefit-Sharing agreements aligned with emerging Digital Sequence Information regulations. This framework enables responsible, large-scale, high-quality genomic data collection while investing in scientific infrastructure and training within partner regions.

As part of the Atlas launch, Basecamp is announcing new partnerships in Chile, Argentina and an expanded collaboration in Antarctica, further extending its global biodiversity network.

The Trillion Gene Atlas is enabled by advances in ultra-high-throughput short- and long-read sequencing and accelerated computing. Basecamp has partnered with Ultima Genomics and PacBio to deliver industrial-scale sequencing including data-rich, high-accuracy long reads.

Ultima is a developer of ultra-high throughput next- generation sequencing (NGS) systems. Ultimaâ??s latest sequencing system, the UG200 Series, advances the companyâ??s unique wafer-based sequencing architecture to enable high-throughput, whole-genome and multi-omics sequencing at an industrial scale at a low price point to enable initiatives like the Trillion Atlas.

â??Biology has been fundamentally data-starved when compared to other fields like language or computer vision as researchers have lacked the tools required to generate data at scaleâ?• said Gilad Almogy, Founder and CEO of Ultima Genomics. â??We strongly believe that AI will have an immense impact on our understanding of biology and human health, and the UG200 Series was designed from the ground up to enable the massive datasets required for BioAI to deliver on this promise. We are excited our technology can enable Basecamp in their vision and advance innovative initiatives like the Trillion Gene Atlas.â?•

â??PacBio HiFi sequencing delivers highly accurate long reads that preserve full genomic context and enables subspecies and even strain-level resolution in complex samples.â?• said Christian Henry, President and CEO of PacBio. â??HiFi data provides the reliable, information-rich foundation biological AI models need to learn from nature at scale and power initiatives like the Trillion Gene Atlas.â?•

The Trillion Gene Atlas will be powered by NVIDIAâ??s accelerated computing infrastructure to process vast quantities of genetic data at the petabase scale. As part of this effort, Basecamp plans to leverage NVIDIA Parabricks to significantly accelerate metagenomic assembly. This collaboration focuses on both advanced engineering and the development of new algorithmic methods to optimize how complex environmental samples are reconstructed. Thanks to this acceleration, processing quadrillions of DNA base pairs, a task that would have previously taken over 20 years, is expected to take less than two years to complete.

Through parallelized data processing, automated annotation, and large-scale model training, the partners expect to compress a task that previously would have required more than 20 years of processing time to less than two years. This compression of sequencing, assembly, annotation and model training is intended to expand the performance and scope of biological foundation models across therapeutic development.

Anthropic joins as part of its broader effort to add new capabilities for life sciences: connecting Claude to more scientific platforms. Working with the Claude for Life Sciences team, the aim is to harness the Trillion Gene Atlas and EDEN to make Claude an even more productive research partner for scientists and clinicians, and support organisations bringing new scientific advancements to the public.

By combining Claudeâ??s advanced reasoning capabilities, EDENâ??s therapeutic design capabilities, and
NVIDIAâ??s CUDA-X Libraries to process unstructured data, the initiative aims to create an integrated workflow for interpreting complex clinical data and translating it directly into therapeutic design.The Trillion Gene Atlas is built on three pillars: large-scale DNA sequencing, global data supply partnerships and advanced computing. Together with AI systems capable of reasoning across complex data, these foundations can help turn vast datasets into therapeutic discoveries. By increasing the evolutionary data available to AI by another 100x, Basecamp Research aims to make drug design faster and more systematic â?? extending EDENâ??s earlier advances in fields like gene therapy and the fight against antibiotic-resistant bacteria.

View original content:https://www.prnewswire.co.uk/news-releases/basecamp-research-launches-trillion-gene-atlas-to-scale-ai-designed-therapeutics-302716632.html

COMUNICATO STAMPA â?? CONTENUTO PROMOZIONALE: Immediapress Ã¨ un servizio di diffusione di comunicati stampa in testo originale redatto direttamente dallâ??ente che lo emette. Lâ??Adnkronos e Immediapress non sono responsabili per i contenuti dei comunicati trasmessi

â??

immediapress/pr-newswire

Categoria

1. Comunicati

Tag

1. ImmediaPress

Data di creazione
Marzo 18, 2026
Autore
redazione