

A Landmark STATE Model Publication in Cell Positions Myllia Biotechnology as the Go-To Partner for Perturb-Seq Datasets in Primary Human Cells

Myllia Biotechnology, a leader in high-content CRISPR screening in primary human cell models, today highlighted the landmark publication of the STATE virtual cell model by Arc Institute in *Cell* ([https://www.cell.com/cell/fulltext/S0092-8674\(26\)00921-9](https://www.cell.com/cell/fulltext/S0092-8674(26)00921-9)) as a defining moment for the field of AI-driven cell biology - and reaffirmed its commitment to serving as the industry's premier source of Perturb-seq datasets generated in disease-relevant, primary human cells.

The STATE model, introduced by Arc Institute, is a machine learning architecture that predicts how stem cells, cancer cells, and immune cells respond to drugs, cytokines, or genetic perturbations. STATE was trained on observational data from nearly 170 million cells and perturbational data from more than 100 million cells across 70 cell lines, and consists of two interlocking modules - the State Embedding (SE) model and the State Transition (ST) model - that together predict transcriptional shifts following perturbation with unprecedented accuracy. In benchmarks, STATE demonstrated a 50% improvement in distinguishing perturbation effects and twice the accuracy in identifying true differentially expressed genes compared to existing state-of-the-art methods.

The publication underscores a critical bottleneck now facing the entire single-cell foundation model (scFM) community: while observational scRNA-seq data is abundant, large-scale, high-quality perturbation data in physiologically relevant primary human cells remains scarce - and it is precisely this type of data that is required to teach AI models the causal biology of human disease.

Myllia: Built for the Single-Cell Perturbomics Era

Myllia's CROP-Seq (Perturb-Seq) platform combines pooled CRISPR perturbation with single-cell RNA sequencing in context-specific, primary human cell models to deliver deeply mechanistic, causal datasets at scale. Unlike screens performed in immortalized cancer lines, Myllia's datasets capture the true biology of the human cells that drugs must ultimately act upon.

"The STATE publication is a watershed moment. It shows, in the most rigorous way yet, that virtual cell models can learn causal biology - but only if they are fed with high-quality, large-scale perturbation data," said Dr. Adam Krejci, VP of Computational Biology at Myllia Biotechnology.

"The bottleneck for the next generation of scFMs is no longer compute or architecture. It is data - specifically, perturbation data in human cell models. That is exactly what we generate at Myllia and we're engaging in discussions with partners who have appetite for training or benchmarking an AI model of the human cell. Such models are in need of Perturb-seq data that reflects real human biology - either in cancer lines as a proxy, or, ideally, in primary human cells."

"For years, we have been building the industrial-scale Perturb-seq engine in primary human cell models that this moment demands," added Dr. Tilmann Buerckstuemmer, Co-Founder and CSO of Myllia Biotechnology.

"Our CROP-Seq platform routinely delivers single-cell perturbomics datasets in primary human cells and other disease-relevant models - the very contexts where target discovery succeeds or fails. Following the STATE paper, we are actively expanding single-cell perturbomics collaborations with AI/ML groups, scFM developers, and pharma partners who recognize that the quality of a virtual cell model is only as good as the quality of the perturbation data behind it. Myllia is open for functional genomics, target discovery and scFM data generation collaborations - come talk to us."

Single-Cell Perturbation Data for the AI/ML and scFM Community

With STATE now setting a new bar for predicting cellular perturbation phenotypes, Myllia invites AI/ML teams, foundation model developers, and biopharma partners working on virtual cells, perturbation prediction, or in silico target discovery to engage on single-cell perturbomics collaborations. Myllia's technology platform at a glance:

- Custom Perturb-seq campaigns in disease-relevant primary human cells, including therapeutic areas such as immunology and inflammation, kidney biology, oncology and metabolic disease
- Genome-scale and focused screens producing causal, mechanism-resolved single-cell readouts suitable for scFM training data and evaluation benchmarks
- Off-the-shelf CROP-Seq perturbation datasets available for licensing, generated in primary human cell models
- Bespoke library design and bioinformatic support tailored to model training and benchmarking needs – an end-to-end offering enabling advanced functional genomics

About Myllia Biotechnology

Myllia Biotechnology, headquartered in Vienna, Austria, is redefining functional genomics by uniting CRISPR-based perturbations, single-cell resolution, and advanced computational modeling. Our platform integrates high-content CRISPR screens with single-cell RNA sequencing in primary human cells to deliver deeply physiologic, mechanistic insights. By generating large-scale, custom perturbation datasets, we enable AI-driven foundation models of virtual cells that infer causal relationships, reveal mechanisms of action, and accelerate unbiased target discovery. This end-to-end approach produces genetically validated, high-value targets for early-stage drug development and generates AI-ready data at a scale and fidelity previously impossible.

Media & Partnership Contact

Dr. Henrik Schmidt | VP, Commercial – Europe & APAC

Myllia Biotechnology GmbH

Am Kanal 27, 1110 Vienna, Austria

Web: www.myllia.com

Email: info@myllia.com

For partnership inquiries: www.myllia.com/contact