



OPEN ACCESS

EDITED BY
Xu Chen,
Shaanxi Normal University, China

REVIEWED BY
Conglian Yang,
Huazhong University of Science and
Technology, China

*CORRESPONDENCE
Mingna Sun
✉ mingnas85@163.com
Lu Liang
✉ lliangaa@gzhmu.edu.cn

†These authors have contributed equally
to this work

RECEIVED 16 April 2026
REVISED 15 May 2026
ACCEPTED 19 May 2026
PUBLISHED 05 June 2026

CITATION

Chen H, Chen Z, Sun M and Liang L
(2026) Integrating *In silico* perturbation
with multilayer omics to decode
regulatory networks in cancer immunity:
a new frontier in precision oncology.
Front. Immunol. 17:1857447.
doi: 10.3389/fimmu.2026.1857447

COPYRIGHT

© 2026 Chen, Chen, Sun and Liang. This
is an open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which does
not comply with these terms.

Integrating *In silico* perturbation with multilayer omics to decode regulatory networks in cancer immunity: a new frontier in precision oncology

Haifang Chen^{1†}, Zhiyu Chen^{1,2†}, Mingna Sun^{1,2*} and Lu Liang^{1,2*}

¹Operations Management Department, The Affiliated Traditional Chinese Medicine Hospital, Guangzhou Medical University, Guangzhou, China, ²The Fifth Affiliated Hospital, Key Laboratory of Molecular Target & Clinical Pharmacology, National Medical Products Administration (NMPA), and the State Key Laboratory of Respiratory Disease, The School of Pharmaceutical Sciences, Guangzhou Medical University, Guangzhou, China

The traditional paradigm of linking single genes to individual phenotypes is being replaced by a systems-level framework to understand the complexity of the tumor microenvironment. In this context, *in silico* knockout has emerged as a powerful computational approach to predict system-wide responses to genetic or cellular perturbations. This review summarizes how multilayer regulatory information across the genome, transcriptome, proteome, and metabolome can be integrated into computational models for virtual perturbation analysis. We outline major multi-omics data sources, including bulk, single-cell, and spatial omics, and emphasize how these data are transformed into model-compatible inputs such as constraint-based matrices and latent embeddings. We then discuss the evolution of *in silico* knockout methodologies, from genome-scale metabolic models and flux balance analysis to advanced deep learning frameworks that enable the prediction of non-linear and unseen perturbations. The integration of spatial transcriptomics further extends these approaches to tissue-level modeling of cell-cell interactions. In tumor immunology, these methods facilitate the identification of immune regulatory genes, the analysis of immune evasion mechanisms, and the prioritization of therapeutic targets. Despite current challenges in multi-omics integration and biological complexity, *in silico* knockout provides a promising framework for advancing precision immunotherapy.

KEYWORDS

foundation models, *in silico* knockout, multi-omics integration, tumor immunology, tumor microenvironment

1 Introduction

The emergence of cancer immunotherapy has fundamentally transfigured the therapeutic paradigm for malignant tumors. Its primary objective is to reinvigorate the host's immune system, enabling it to recognize and eliminate malignant cells. Clinical breakthroughs in immune checkpoint blockade (ICB) and adoptive cell therapies, such as CAR-T and CAR-NK, have offered durable clinical benefits for patients with various previously intractable late-stage hematological and solid tumors (1, 2). However, a significant

proportion of patients still exhibit primary or acquired resistance, a phenomenon largely attributed to sophisticated mechanisms of immune evasion. Within the tumor microenvironment (TME), malignant cells do not exist in isolation but orchestrate a suppressive ecosystem. This immune-shield is maintained by specialized cell populations, such as immunosuppressive *SPP1+* or *TREM2+* macrophages and hypoxia-induced fibroblasts, which collectively hinder the infiltration and effector function of cytotoxic lymphocytes (3, 4).

The resilience and adaptability of this ecosystem are governed by complex, multilayered regulatory networks that bridge genetic, epigenetic, and metabolic dimensions. At the epigenetic level, specific modifications such as histone lactylation (H3K18la) have been identified as pivotal spatiotemporal regulators that link glycolytic flux to the transcriptional activation of oncogenes and immune-evasive signatures (5). Concurrently, metabolic reprogramming serves as a fundamental pillar of immune modulation; for instance, the availability of cytosolic acetyl-coenzyme A acts as a rheostat for mitophagy, while mannose metabolism can drastically reshape T cell differentiation and anti-tumor potency (6, 7). Furthermore, the spatial architecture of the TME, as exemplified by paracrine signaling loops such as the hypoxia-induced Wnt5a secretion from fibroblasts, reinforces these regulatory circuits, creating a self-sustaining niche that promotes cancer progression (8).

Despite the power of functional genomics, particularly genome-wide CRISPR screens, conventional experimental approaches face substantial bottlenecks in deconstructing these multidimensional networks. Physical knockout screens in primary human immune cells are not only resource-intensive and technically demanding but also constrained by the curse of dimensionality, where the

astronomical number of gene combinations and their context-specific effects within a dynamic TME cannot be experimentally tested (9, 10). Furthermore, traditional assays often provide only a static snapshot, failing to capture the metabolic flux or the transient regulatory states of cells during the progression of treatment.

To bridge this gap, computational modeling and in silico perturbation (virtual knockout) have emerged as indispensable frontiers in precision oncology. By leveraging large-scale foundation models (e.g., Geneformer) pre-trained on vast single-cell datasets, or utilizing Genome-Scale Metabolic Models (GEMs), researchers can now simulate the systematic deletion of genes within a digitized cellular framework (11, 12). These virtual perturbations allow for the high-throughput identification of master regulators that can revert drug-resistant states or sensitize tumors to immunotherapy by predicting the downstream impact on transcriptional meta programs and metabolic homeostasis. This review aims to synthesize recent advancements in integrating multilayered omics with in silico strategies, highlighting how these digital experiments are redefining our ability to decode and therapeutically target the regulatory essence of cancer immunity. The integrated workflow of in silico knockout, spanning from multi-omics data acquisition to virtual perturbation and clinical validation, is illustrated in Figure 1.

2 Multilayer regulatory networks and multi-omics integration

The core of systems biology lies in the shift from a reductionist view of single genes mapping to single phenotypes to a holistic

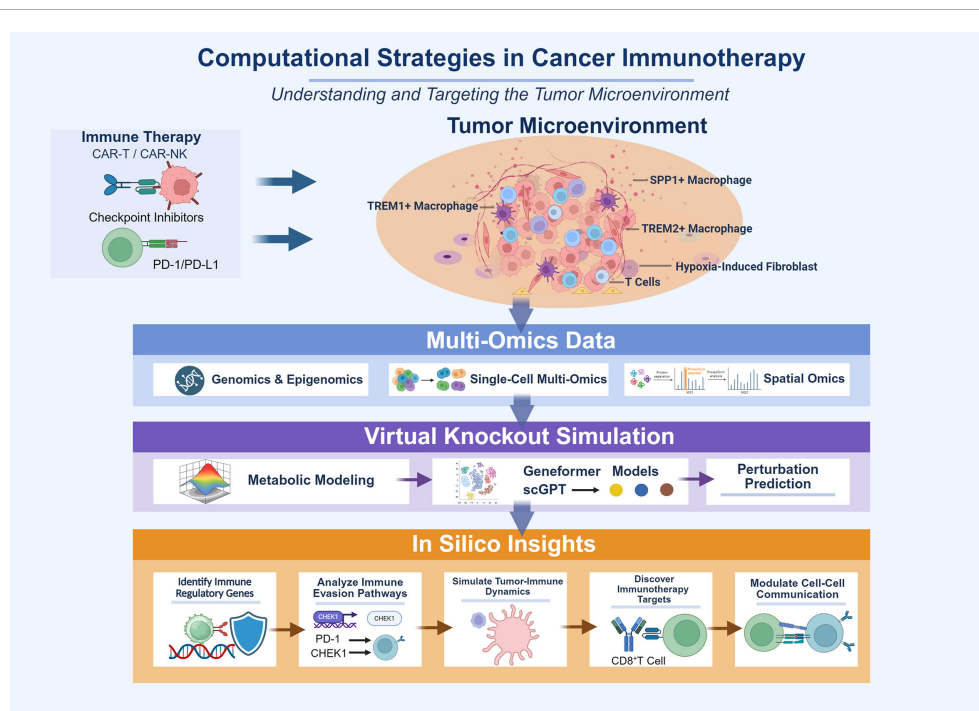


FIGURE 1
 Integrated computational workflow of in silico knockout in tumor immunology.

understanding of how biological entities interact across multiple scales. In the context of cancer immunity, decoding the landscape requires a systematic integration of information flowing through the central dogma and beyond, encompassing the genome, transcriptome, proteome, and metabolome. The comprehensive architecture of these regulatory layers, along with their respective multi-omics data sources and computational integration strategies, is systematically summarized in [Table 1](#).

2.1 The architecture of multilayer networks

Biological regulation is inherently hierarchical and interconnected. A multilayer network represents these interactions as a series of coupled layers. At the foundation lies the genomic and epigenomic layer, where structural variations and chromatin modifications, such as H3K18la lactylation, dictate the accessibility of genetic information and set the stage for subsequent regulation ([5](#), [13](#)). This is followed by the transcriptional layer, the dynamic output of the genome where transcription factors coordinate the expression of meta programs defining cell identity, such as the *EMT-inflammatory* or *Prolif-stress* states observed in malignant tissues ([3](#), [14](#)). The functional execution then occurs at the proteomic and post-translational layer, where proteins undergo modifications such as ubiquitination by E3 ligases like Cop1 in order to modulate signaling stability and immune cell infiltration ([4](#), [20](#)). Finally, the metabolic layer represents the ultimate functional state where small molecules like acetyl-coenzyme A or mannose act as both fuel and signaling metabolites, providing real-time feedback to the epigenetic and transcriptional layers ([6](#), [7](#), [21](#)).

2.2 Data sources for integrative analysis

The construction of these high-fidelity networks relies on high-dimensional data harvested from diverse repositories. Bulk omics data from sources like The Cancer Genome Atlas (TCGA) and GEO provide a pan-cancer overview of genomic alterations and clinical outcomes, which is essential for identifying broad immunosuppressive drivers like CHEK1 ([22](#), [23](#)). To resolve the inherent heterogeneity of TME, single-cell omics (e.g., scRNA-seq and scATAC-seq) are utilized to identify specific rare cell populations, such as *SPPI+* macrophages, that drive immune evasion ([15](#), [24](#)). Furthermore, the addition of spatial transcriptomics provides a crucial spatial coordinate that reveals how the proximity of various cell types, such as hypoxia-induced fibroblasts, triggers signaling loops like the Wnt5a axis that promote cancer progression ([17](#)).

2.3 Methodologies for multi-omics integration

Translating these disparate data types into a unified regulatory map requires advanced computational frameworks. Network modeling techniques, particularly GEMs, utilize biochemical constraints to link genomic data with metabolic flux, enabling the simulation of how gene knockouts perturb the entire system's equilibrium ([25](#)). In parallel, machine learning and deep learning are increasingly employed to extract non-linear features from multi-omics data. Interpretable AI models, such as PRIME, and transformer-based

architectures like Geneformer allow researchers to prioritize mutational intolerance or predict treatment responses by learning the latent grammar of gene regulation across different biological layers ([26–28](#)).

Crucially, these integration methodologies serve as the functional bridge to perturbation analysis. By converting raw multi-omics datasets into structured inputs, including metabolite-reaction stoichiometric matrices for mechanistic models and high-dimensional latent embeddings for deep learning architectures, these frameworks provide the necessary starting configurations for the virtual knockout simulations described in the following section.

3 In silico knockout: from metabolic flux to AI-driven perturbation simulation

The field of *in silico* knockout has evolved from simple logic-based gate simulations to a sophisticated multi-dimensional framework capable of predicting complex non-linear responses within the tumor ecosystem. By leveraging computational models to simulate the absence of genes or cellular components, researchers can perform high-throughput screening and identify therapeutic vulnerabilities without the immediate need for exhaustive wet-lab experimentation.

The efficacy of these simulations depends on the precision of the integrated data inputs derived from the aforementioned omics layers, which define the baseline state of the digital twin prior to perturbation.

3.1 Model-based virtual knockout via GEMs

GEMs represent a cornerstone in understanding the metabolic reprogramming that drives cancer progression and drug resistance. This approach utilizes integrated transcriptomic and proteomic profiles as regulatory constraints to tailor generic metabolic templates into patient-specific models. In the context of breast cancer research, Flux Balance Analysis (FBA) serves as the primary engine for model-based virtual knockouts. FBA translates biochemical reactions into mathematical constraints, allowing for the calculation of steady-state flux distributions across the entire metabolic network. Recent frameworks have integrated single-gene knockout simulations with clustering algorithms like UMAP and k-means to identify metabolic targets that can revert the flux state of drug-resistant cells back to a drug-sensitive parental phenotype ([29](#)). Utilizing tools such as the COBRA toolbox or RAVEN, researchers can integrate patient-specific transcriptomic data to build personalized models. The hallmark of this approach is its ability to predict the impact of a knockout on the biomass production rate, directly identifying metabolic essential genes that can be prioritized as potent drug targets to sensitize resistant cells.

3.2 Deep learning and foundation models for perturbation prediction

As biology moves beyond linear modeling, the focus has shifted toward data-driven approaches exemplified by generative foundation models. Unlike mechanistic models that require explicit biochemical

TABLE 1 Multilayer regulatory architecture, data sources, and computational integration strategies.

Layer / component	Biological features	Data type & source	Integration strategy	Output for downstream models	References
Genomic & Epigenomic Layer	DNA mutations, structural variations, chromatin accessibility (e.g., lactylation)	Bulk sequencing (TCGA, GEO), epigenomics	Feature extraction, mutation profiling	Gene-level constraints, regulatory priors	(5)
Transcriptional Layer	Gene expression programs (e.g., EMT-inflammatory, Prolif-stress)	Bulk RNA-seq, scRNA-seq	Dimensionality reduction, clustering	Expression matrices, cell states, embeddings	(3)
Proteomic & Post-translational Layer	Protein abundance, ubiquitination (e.g., Cop1), signaling modulation	Proteomics, inferred from transcriptomics	Network inference, activity scoring	Signaling activity scores, regulatory networks	(4)
Metabolic Layer	Metabolites (e.g., acetyl-CoA, mannose), metabolic flux	Metabolomics, inferred flux models	Constraint-based modeling (GEMs)	Stoichiometric matrices, flux distributions	(6, 7)
Bulk Omics Data	Population-level genomic & clinical patterns	TCGA, GEO	Statistical modeling, biomarker discovery	Gene signatures, risk models	(13, 14)
Single-cell Omics	Cellular heterogeneity (e.g., SPPI ⁺ macrophages)	scRNA-seq, scATAC-seq	Clustering, trajectory inference	Cell-type-specific features	(15, 16)
Spatial Omics	Cell-cell proximity, niche interactions (e.g., Wnt5a signaling)	Spatial transcriptomics	Spatial mapping, neighborhood analysis	Spatial graphs, interaction networks	(17–19)
Network Modeling (GEMs)	Gene–metabolism linkage	Multi-omics integrated	Constraint-based reconstruction	Flux balance models (FBA-ready)	(2)
Machine Learning Models	Non-linear feature extraction	Multi-omics datasets	Representation learning	Latent embeddings	(13)
Deep Learning / Transformers	Gene regulation grammar, perturbation prediction	Large-scale omics data	Pretraining + fine-tuning	Predictive models for knockout simulation	(13, 14)
Bridging to In Silico Knockout	Conversion to model-ready inputs	Integrated multi-omics	Encoding into matrices / embeddings	Inputs for GEMs & AI perturbation models	(2, 13, 14)

rules, these architectures ingest massive single-cell multi-omics datasets to construct universal cell embeddings. While early deep learning frameworks provided a starting point, current trends involve models like scGPT and GearNet, which utilize pre-trained cell embeddings to learn the language of gene expression.

A defining advantage of these foundation models is their dual capability for zero-shot and fine-tuned perturbation predictions. In a zero-shot setting, models such as scGPT leverage their vast pre-training on diverse cell atlases to predict the effects of unseen perturbations exemplified by the knockout of genes not explicitly included in the training data by navigating the learned latent space of gene regulation (30). Conversely, fine-tuned prediction allows these architectures to be specialized for specific biological contexts, a process exemplified by Geneformer, which can be refined with context-specific datasets to identify dosage-sensitive genes or predict emergent resistance mechanisms in a particular tumor subtype (31). By refining the pre-trained weights with smaller, high-fidelity datasets, fine-tuning significantly enhances the accuracy of predicting localized signaling responses compared to generic models.

For instance, transformer-based perturbation frameworks have been employed to reveal ligand-centered activation mechanisms in T cells within the colorectal cancer microenvironment (32). Furthermore, deep learning architectures are now being used for large-scale *in silico* screening of anticancer drugs at the single-cell level, identifying specific vulnerabilities such as the EZH2 or PLK1 pathways (33). By disentangling baseline cellular states from perturbation-induced effects, these AI frameworks capture complex non-linear compensatory mechanisms, predicting how bypass pathways might be activated following a primary gene knockout.

3.3 Spatial *in silico* perturbations and microenvironmental modeling

The integration of spatial transcriptomics has expanded the definition of *in silico* knockout from the intracellular level to the tissue architecture level. This approach allows for the simulation of perturbations within specific anatomical niches, such as the tumor-stroma interface. Rather than just removing a gene, researchers can simulate the removal of specific cell populations or the blockade of cell-cell communication pathways. For example, large-scale single-cell analysis combined with *in silico* perturbation has been used to dissect the dynamic evolution of hepatocellular carcinoma, identifying specific metabolic and inflammatory meta-programs that drive metastasis (16). Similarly, by virtually knocking out the secretion of ligands like Wnt5a from hypoxia-induced inflammatory fibroblasts (18) or targeting the ADAM12-mediated myofibroblast program (34), one can observe the subsequent state changes in adjacent malignant cells. This reveals the niche-dependency of tumor progression, providing a theoretical roadmap for identifying checkpoints that impede anti-tumor immunity.

4 Applications of *In silico* knockout in tumor immunology

The integration of *in silico* knockout technology into tumor immunology has revolutionized our ability to dissect the complex interactions within TME. By leveraging computational models to

simulate the loss of specific genetic components, researchers can predict immune responses and identify therapeutic vulnerabilities with unprecedented precision.

4.1 Identification of key immune regulatory genes

In silico perturbation serves as a high-throughput discovery engine for identifying genes that govern the balance between immune activation and suppression. To identify immune suppressor genes, researchers utilize deep learning frameworks like scGPT or SCENIC+ to simulate the deletion of myeloid-specific factors. For instance, simulating the knockout of the E3 ubiquitin ligase Cop1 has been shown to alter the secretion of Ccl2 and Ccl5, thereby reducing the infiltration of pro-tumorigenic macrophages and enhancing anti-tumor immunity (35). This computational prediction was subsequently validated through *in vivo* CRISPR-Cas9 screens and bone marrow chimeras, confirming that Cop1 deficiency significantly inhibits tumor growth by remodeling the myeloid compartment. Conversely, the search for immune activator genes often involves perturbing metabolic or signaling checkpoints. Recent simulations of mannose metabolism pathways revealed that disrupting specific metabolic nodes can reshape T cell differentiation, effectively turning exhausted populations into potent effectors (36). These findings were physically corroborated by dietary mannose supplementation in tumor-bearing mice, which demonstrated a marked increase in CD8⁺ T cell effector function and improved response to anti-PD-1 therapy.

4.2 Analysis of immune evasion mechanisms

Virtual knockout techniques allow for the systematic deconstruction of the pathways tumors use to bypass immune surveillance. In the context of the PD-1/PD-L1 pathway, *in silico* models can predict the downstream transcriptional shifts that occur when these checkpoints are blocked, helping to identify resistance signatures. This approach is particularly effective for studying T cell exhaustion, where simulating the knockout of transcription factors like TOX or specific kinases such as CHEK1 reveals how the epigenetic landscape of a T cell can be reprogrammed from a dysfunctional state back to a functional one (37). Specifically, the predicted role of CHEK1 as an immunosuppressive driver was clinically reinforced by the observation that high CHEK1 expression in lung adenocarcinoma patients correlates with poor prognosis and reduced immune cell infiltration.

4.3 Discovery of immunotherapy targets

The transition from discovery to clinical application is accelerated by using virtual perturbations to prioritize checkpoint inhibitors and combination therapies. Beyond classic PD-1/CTLA-4 targets, *in silico* screens have identified stromal checkpoints such as ADAM12 in fibroblasts; simulating its loss predicts a significant increase in CD8⁺ T cell infiltration into previously cold tumors (34, 38). This stromal dependency was physically validated using patient-derived organoids and mouse models, where the

pharmacological inhibition of the ADAM12-mediated myofibroblast program successfully restored T cell access to the tumor core. Furthermore, virtual knockout facilitates the prediction of synergistic effects in combination therapy. By simulating the simultaneous inhibition of a metabolic regulator and a signaling receptor, models can identify which dual-targeting strategies most effectively overcome the immunosuppressive barriers of the TME without the exhaustive cost of physical combinatorial libraries.

4.4 Regulation of cell-cell communication at the single-cell level

At the resolution of single-cell transcriptomics, virtual knockout is used to perturb ligand-receptor networks to understand microenvironment remodeling. Tools like CellChat, when combined with perturbation modules, allow researchers to delete a specific ligand (such as Wnt5a or MIF) from a tumor cell and observe the ripple effect across the entire communication interactome (19). For example, the virtual removal of THBS2⁺ cancer-associated fibroblasts (CAFs) or SPP1⁺ macrophages has demonstrated how these specific sub-populations act as hubs for immunosuppression. The functional importance of these predictions was substantiated by clinical spatial profiling, which revealed that the physical co-localization of SPP1⁺ macrophages and ITGA5⁺ fibroblasts is a deterministic feature of metastasis in hepatocellular carcinoma (39). Their removal from the model results in a complete restructuring of the spatial and chemical signals that normally exclude T cells, providing a theoretical roadmap for precision fibroblast-targeting strategies.

5 Discussion

The transition from descriptive omics to predictive *in silico* perturbation represents a paradigm shift in cancer immunotherapy. However, the maturation of this field requires addressing fundamental tensions between computational abstraction and biological complexity.

5.1 Schools of thought: mechanistic rigor vs. data-driven scalability

A significant divergence exists in how the field approaches the digital twin of the TME. One school of thought prioritizes mechanistic rigor, utilizing GEMs and FBA to simulate perturbations within biologically defined stoichiometric constraints. Proponents argue that these models provide causal transparency that black-box AI cannot match. Conversely, an emerging school of thought champions data-driven scalability, leveraging transformer-based foundation models like scGPT to learn regulatory grammars directly from massive single-cell repositories. The controversy lies in whether the latent spaces of AI truly capture biological causality or merely sophisticated correlations. Resolving this tension likely requires hybrid architectures where mechanistic laws act as stabilizers for AI-driven predictions.

5.2 Current research gaps: beyond transcriptional hegemony

Despite the proliferation of single-cell datasets, a critical research gap persists in the cross-layer fidelity of virtual knockouts. Most current models are anchored in transcriptional abundance, yet mRNA levels often correlate poorly with protein activity or metabolic flux due to post-translational modifications. This domain gap means that a virtual knockout of a transcript might ignore the functional resilience provided by protein stability or metabolic bypasses. To address this deficiency, the integration of multimodal technologies such as Cellular Indexing of Transcriptomes and Epitopes by Sequencing (CITE-seq) into existing frameworks is essential (40). By providing simultaneous quantification of surface protein expression and RNA abundance, CITE-seq data allow virtual knockout models to anchor transcriptional perturbations in concrete protein-level changes, thereby refining the predictive accuracy of signaling state transitions. Furthermore, incorporating emerging single-cell proteomic signatures enables the modeling of post-translational regulatory logic, a process facilitated by frontier architectures like SCN- β that leverage multi-scale embeddings to bridge the gap between gene expression and functional protein activity (41). Future models must transition from purely transcriptomic transformers to multimodal architectures capable of cross-referencing these diverse data streams to ensure that a simulated genetic loss reflects true biochemical ablation. Furthermore, while we can simulate the loss of a single gene, we currently lack the computational capacity to model the combinatorial curse of dimensionality, specifically how the simultaneous perturbation of multiple checkpoints within diverse spatial niches, including the interaction between ADAM12+ fibroblasts and SPP1 + macrophages, reshapes the global immune landscape. Accompanying this biological complexity is a significant escalation in hardware requirements, as the traditional curse of dimensionality has effectively transitioned into a curse of GPU memory. Running large-scale foundation models like Geneformer or scGPT for extensive perturbation screens necessitates substantial computational infrastructure, often requiring high-performance A100 or H100 GPU clusters to handle the quadratic scaling of self-attention mechanisms (31). This creates a practical barrier for many academic laboratories, where the memory intensity of processing millions of cells across thousands of simulated gene knockouts can exceed available local resources. Addressing this requires not only more powerful hardware but also the development of memory-efficient algorithms, such as FlashAttention or quantized low-rank adaptation, to democratize access to frontier *in silico* knockout frameworks (42).

5.3 Bridging the discrepancy: from digital simulations to biological reality

A fundamental debate centers on the equivalence of virtual and physical ablation. Computational models typically simulate a perfect loss of function, whereas real-world CRISPR-Cas9 or pharmacological interventions involve varying degrees of efficiency, off-target effects, and systemic toxicity. Current simulations often operate in a

fluid digital environment, largely failing to account for the biophysical constraints of the 3D tumor architecture, such as interstitial pressure and oxygen gradients. Bridging this gap requires the integration of 4D spatiotemporal modeling, where the digital twin evolves not just in its molecular state, but also within the physical and mechanical pressures of the evolving tumor ecosystem. Beyond physical constraints, a critical frontier involves transitioning from static snapshots to models that capture the temporal dynamics of clonal evolution and the emergence of drug resistance. Current efforts are beginning to integrate longitudinal single-cell sequencing data with mathematical modeling of evolutionary fitness, a process exemplified by frameworks like CloneAlign that link chromosomal copy number alterations to transcriptional shifts (43). Future digital twins must incorporate these longitudinal trajectories to simulate how specific perturbations, such as the knockout of a primary oncogenic driver, might inadvertently select for resistant sub-clones or activate latent compensatory pathways over time. By utilizing multi-time-point data to parameterize these models, researchers can move toward 4D simulations that predict not only the immediate response to therapy but also the long-term evolutionary bottlenecks that lead to clinical relapse (44).

5.4 Potential future developments: toward autonomous discovery

The future of *in silico* perturbations lies in the move toward autonomous, closed-loop discovery systems. We anticipate the development of self-correcting models that utilize active learning to suggest the most informative wet-lab experiments, which in turn refine the model's predictive accuracy. Furthermore, the integration of multi-modal foundation models incorporating spatial transcriptomics, digital pathology, and clinical longitudinal data will allow for the simulation of personalized therapeutic journeys. By transitioning from static snapshots to dynamic, high-fidelity simulations that capture non-linear stochasticity and phenotypic plasticity, *in silico* frameworks will eventually move beyond being mere screening tools to becoming the central engine for precision oncology and the design of synergistic combination immunotherapies.

Author contributions

HC: Visualization, Formal analysis, Conceptualization, Data curation, Writing – original draft. ZC: Visualization, Writing – original draft, Software, Methodology, Conceptualization. MS: Supervision, Writing – review & editing, Project administration, Funding acquisition. LL: Writing – review & editing, Resources, Funding acquisition, Project administration, Supervision.

Funding

The author(s) declared that financial support was received for this work and/or its publication. We acknowledged the National

Natural Science Foundation of China (82570120) and Natural Science Foundation of Guangdong Province (2024A1515012290). This study was also supported by Key Project in Key Areas of Guangdong Provincial Universities (2025ZSZX2054), Open Research Project of the National Key Laboratory of New Drug Targets Discovery and Novel Drug Development for Major Diseases (SKLD2025M02), and Featured Innovation Project of Ordinary Higher Education Institutions in Guangdong Province (2023KTSCX106).

Acknowledgments

The authors would like to thank all the participants involved in this study.

Conflict of interest

The remaining author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of Interest.

References

- Li Y, Lu B, Zhao Z, Gong M, Lyu D, Zhang H, et al. Integrative pan-cancer analysis of dipeptidyl peptidase 4 with clinical and *in vitro* validation in prostate cancer. *Front Immunol.* (2026) 17:1616889. doi: 10.3389/fimmu.2026.1616889
- Biederstädt A, Basar R, Park JM, Uprety N, Shrestha R, Reyes Silva F, et al. Genome-wide CRISPR screens identify critical targets to enhance CAR-NK cell antitumor potency. *Cancer Cell.* (2025) 43:2069–2088.e11. doi: 10.1016/j.ccell.2025.07.021
- Xia P, Shuang S, Fu D, Liu L, Yang D, Guo Y, et al. Large-scale single-cell analysis and *in silico* perturbation reveal dynamic evolution of HCC: from initiation to therapeutic targeting. *NPJ Precis Oncol.* (2026) 10:100. doi: 10.1038/s41698-026-01307-2
- Wang X, Tokheim C, Gu SS, Wang B, Tang Q, Li Y, et al. *In vivo* CRISPR screens identify the E3 ligase Cop1 as a modulator of macrophage infiltration and cancer immunotherapy target. *Cell.* (2021) 184:5357–5374.e22. doi: 10.1016/j.cell.2021.09.006
- Shen A, Zheng S, Tang X, Yao Y, Yin E, Sun N, et al. Integrated multi-omics analyses identified the H3K18la-based spatiotemporal characteristics and risk-stratified treatment strategy in lung adenocarcinoma. *Cancer Lett.* (2026) 638:218158. doi: 10.1016/j.canlet.2025.218158
- Zhang Y, Shen X, Shen Y, Wang C, Yu C, Han J, et al. Cytosolic acetyl-coenzyme A is a signalling metabolite to control mitophagy. *Nature.* (2026) 649:1022–31. doi: 10.1038/s41586-025-09745-x
- Qiu Y, Su Y, Xie E, Cheng H, Du J, Xu Y, et al. Mannose metabolism reshapes T cell differentiation to enhance anti-tumor immunity. *Cancer Cell.* (2025) 43:103–121.e8. doi: 10.1016/j.ccell.2024.11.003
- Harada A, Yasumizu Y, Harada T, Fumoto K, Sato A, Maehara N, et al. Hypoxia-induced Wnt5a-secreting fibroblasts promote colon cancer progression. *Nat Commun.* (2025) 16:3653. doi: 10.1038/s41467-025-58748-9
- Shifrut E, Carnevale J, Tobin V, Roth TL, Woo JM, Bui CT, et al. Genome-wide CRISPR screens in primary human T cells reveal key regulators of immune function. *Cell.* (2018) 175:1958–1971.e15. doi: 10.1016/j.cell.2018.10.024
- Biederstädt A, Rezvani K. Engineered natural killer cells for cancer therapy. *Cancer Cell.* (2025) 43:1987–2013. doi: 10.1016/j.ccell.2025.09.013
- Lim J, Jung HD, Park SY, Jeon M, Kim DS, Cho R, et al. Genome-scale knockout simulation and clustering analysis of drug-resistant breast cancer cells reveal drug sensitization targets. *Proc Natl Acad Sci USA.* (2025) 122:e2425384122. doi: 10.1073/pnas.2425384122
- Wang Y, Yue B, Ni H, Chen J, Shi R, Wang Z, et al. POSTN+ cancer-associated fibroblast-CCL3+ macrophage crosstalk defines the immune-excluded tumor micro-environment in clear cell renal cell carcinoma. *Transl Oncol.* (2026) 65:102682. doi: 10.1016/j.tranon.2026.102682
- Xiang H, Kasajima R, Azuma K, Tagami T, Hagiwara A, Nakahara Y, et al. Multi-omics analysis-based clinical and functional significance of a novel prognostic and immunotherapeutic gene signature derived from amino acid metabolism pathways in lung adenocarcinoma. *Front Immunol.* (2024) 15:1361992. doi: 10.3389/fimmu.2024.1361992
- Huang R, Ma J, Yao J, Pang J, Zhang M, Wen L, et al. Integrated multi-omics and network toxicology elucidate the multi-target mechanisms of environmental hormones in driving hepatocellular carcinoma. *Ecotoxicol Environ Saf.* (2026) 309:119519. doi: 10.1016/j.ecoenv.2025.119519
- Tang T, Li Y, Xiyun N, Wu H, Fan L, Zhang X, et al. Crosstalk between SPP1+ macrophages and ITGA5+ fibroblasts promotes hepatocellular carcinoma metastasis. *Hepatol Commun.* (2026) 10:e00907. doi: 10.1097/HC9.0000000000000907
- Chen J, Lu W, Lou Y, Liu J, Liao X, Bai Y, et al. Integrating single-cell- and spatial-resolved transcriptomics unravels the inter-tumor heterogeneity and immunosuppressive landscape in HBV- and Clonorchis sinensis-associated hepatocellular carcinoma. *Mol Cancer.* (2026) 25:3. doi: 10.1186/s12943-025-02381-z
- Mehdawi LM, Prasad CP, Ehrnström R, Andersson T, Sjölander A. Non-canonical WNT5A signaling up-regulates the expression of the tumor suppressor 15-PGDH and induces differentiation of colon cancer cells. *Mol Oncol.* (2016) 10:1415–29. doi: 10.1016/j.molonc.2016.07.011
- Cheng R, Sun B, Liu Z, Zhao X, Qi L, Li Y, et al. Wnt5a suppresses colon cancer by inhibiting cell proliferation and epithelial-mesenchymal transition. *J Cell Physiol.* (2014) 229:1908–17. doi: 10.1002/jcp.24566
- Chen T, Zhang F, Liu J, Huang Z, Zheng Y, Deng S, et al. Dual role of WNT5A in promoting endothelial differentiation of glioma stem cells and angiogenesis of glioma derived endothelial cells. *Oncogene.* (2021) 40:5081–94. doi: 10.1038/s41388-021-01922-2
- Ji P, Gong Y, Jin ML, Wu HL, Guo LW, Pei YC, et al. *In vivo* multidimensional CRISPR screens identify Lgals2 as an immunotherapy target in triple-negative breast cancer. *Sci Adv.* (2022) 8:eabl8247. doi: 10.1126/sciadv.abl8247
- Vysochan A, Sengupta A, Weljie AM, Alwine JC, Yu Y. ACS2-mediated acetyl-CoA synthesis from acetate is necessary for human cytomegalovirus infection. *Proc Natl Acad Sci USA.* (2017) 114:E1528–35. doi: 10.1073/pnas.1614268114
- Wang T, Zhao H, Sun X, Ding Y, Zhu Z, Yu X, et al. Pancancer fine-mapping of mutational intolerance identifies CHEK1 as an immunosuppressive driver in lung adenocarcinoma. *Adv Sci (Weinh).* (2026) 13:e21265. doi: 10.1002/adv.202521265

Generative AI statement

The author(s) declared that generative AI was used in the creation of this manuscript. AI was used to optimize the schematic diagrams and to refine the linguistic quality of the manuscript. All AI outputs were rigorously reviewed and verified by the authors, who take full responsibility for the final content.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

23. Sun L, Ma Y, Geng C, Gao X, Li X, Ru Q, et al. DPP4, a potential tumor biomarker, and tumor therapeutic target: review. *Mol Biol Rep.* (2025) 52:126. doi: 10.1007/s11033-025-10235-6
24. Zhang L, Yu X, Zheng L, Zhang Y, Li Y, Fang Q, et al. Lineage tracking reveals dynamic relationships of T cells in colorectal cancer. *Nature.* (2018) 564:268–72. doi: 10.1038/s41586-018-0694-x
25. Villeneuve DJ, Hembruff SL, Veitch Z, Cecchetto M, Dew WA, Parissenti AM. cDNA microarray analysis of isogenic paclitaxel- and doxorubicin-resistant breast tumor cell lines reveals distinct drug-specific genetic signatures of resistance. *Breast Cancer Res Treat.* (2006) 96:17–39. doi: 10.1007/s10549-005-9026-6
26. Li Y, Huan C, Sun H, Zhang W, Guo Z, Li C, et al. Spatial transcriptomics and snRNA-seq expose CAF niches orchestrating dual stromal-immune barriers in hepatocellular carcinoma. *Adv Sci (Weinh).* (2025) 12:e14661. doi: 10.1002/advs.202514661
27. Wang Y, Xiang YB, Chen XW, Zhang T, Wang JY, Liu WY, et al. PRIME: an interpretable artificial intelligence model based on liquid biopsy improves prediction of progression risk in non-small cell lung cancer. *Mil Med Res.* (2026) 12:94. doi: 10.1186/s40779-025-00679-z
28. Ran R, Brubaker DK. A ligand-centered framework for $\gamma\delta$ T cell activation in colorectal cancer revealed by single-cell and transformer-based perturbation. *Front Immunol.* (2026) 16:1715827. doi: 10.3389/fimmu.2025.1715827
29. Adham SA, Al Kalbani A, Al Zeheimi N, Al Dalali M, Al Kharusi N, Siddiqi A, et al. Glycemic load impacts the response of acquired resistance in breast cancer cells to chemotherapeutic drugs *in vitro*. *PLoS One.* (2024) 19:e0311345. doi: 10.1371/journal.pone.0311345
30. Cui H, Wang C, Maan H, Pang K, Luo F, Duan N, et al. scGPT: toward building a foundation model for single-cell multi-omics using generative AI. *Nat Methods.* (2024) 21:1470–80. doi: 10.1038/s41592-024-02201-0
31. Theodoris CV, Xiao L, Chopra A, Chaffin MD, Al Sayed ZR, Hill MC, et al. Transfer learning enables predictions in network biology. *Nature.* (2023) 618:616–24. doi: 10.1038/s41586-023-06139-9
32. Ran R, Trapecar M, Brubaker DK. Systematic analysis of human colorectal cancer scRNA-seq revealed limited pro-tumoral IL-17 production potential in gamma delta T cells. *Neoplasia.* (2024) 58:101072. doi: 10.1016/j.neo.2024.101072
33. Zhang P, Wang X, Cen X, Zhang Q, Fu Y, Mei Y, et al. A deep learning framework for *in silico* screening of anticancer drugs at the single-cell level. *Natl Sci Rev.* (2024) 12:nwae451. doi: 10.1093/nsr/nwae451
34. Li J, Liu H, Guo Q, Zhang Y, Li J, Diao T, et al. Single-cell screens identify ADAM12 as a fibroblast checkpoint impeding anti-tumor immunity. *Cancer Cell.* (2026) 44:424–442.e14. doi: 10.1016/j.ccell.2025.12.018
35. Li DQ, Ohshiro K, Reddy SD, Pakala SB, Lee MH, Zhang Y, et al. E3 ubiquitin ligase COP1 regulates the stability and functions of MTA1. *Proc Natl Acad Sci USA.* (2009) 106:17493–8. doi: 10.1073/pnas.0908027106
36. Scheffler CM, Beavis PA, Darcy PK. A metabolic pathway for improving adoptive cellular therapy. *Cancer Cell.* (2025) 43:8–10. doi: 10.1016/j.ccell.2024.11.014
37. Qiu M, Lan C, Xia Z, Wu H, Wang Z, Wei J, et al. GPX2 induces macrophage M2 polarization through the MIF signaling pathway to promote colorectal cancer progression. *Int J Biol Macromol.* (2025) 331:148341. doi: 10.1016/j.ijbiomac.2025.148341
38. Chen F, Bai G, Liu Q, He G, Ding Z, Liang J, et al. Integrated multi-omics identifies a CD54+ iCAF-ITGAL+ macrophage niche driving immunosuppression via CXCL8-PDL1 axis in cervical cancer. *Mol Cancer.* (2025) 24:262. doi: 10.1186/s12943-025-02471-y
39. Zhou X, Han J, Zuo A, Ba Y, Liu S, Xu H, et al. THBS2+ cancer-associated fibroblasts promote EMT leading to oxaliplatin resistance via COL8A1-mediated PI3K/AKT activation in colorectal cancer. *Mol Cancer.* (2024) 23:282. doi: 10.1186/s12943-024-02180-y
40. Stoeckius M, Hafemeister C, Stephenson W, Houck-Loomis B, Chattopadhyay PK, Swerdlow H, et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods.* (2017) 14:865–8. doi: 10.1038/nmeth.4380
41. Xu J, Huang DS, Zhang X. scmFormer integrates large-scale single-cell proteomics and transcriptomics data by multi-task transformer. *Adv Sci (Weinh).* (2024) 11:e2307835. doi: 10.1002/advs.202307835
42. Zou Z, Liu Y, Bai Y, Luo J, Zhang Z. scTrans: Sparse attention powers fast and accurate cell type annotation in single-cell RNA-seq data. *PLoS Comput Biol.* (2025) 21:e1012904. doi: 10.1371/journal.pcbi.1012904
43. Campbell KR, Steif A, Laks E, Zahn H, Lai D, McPherson A, et al. clonealign: statistical integration of independent single-cell RNA and DNA sequencing data from human cancers. *Genome Biol.* (2019) 20:54. doi: 10.1186/s13059-019-1645-z
44. Sun R, Hu Z, Sottoriva A, Graham TA, Harpak A, Ma Z, et al. Between-region genetic divergence reflects the mode and tempo of tumor evolution. *Nat Genet.* (2017) 49:1015–24. doi: 10.1038/ng.3891