

Artificial intelligence-driven virtual knockout in cancer research: Progress, challenges, and future perspectives

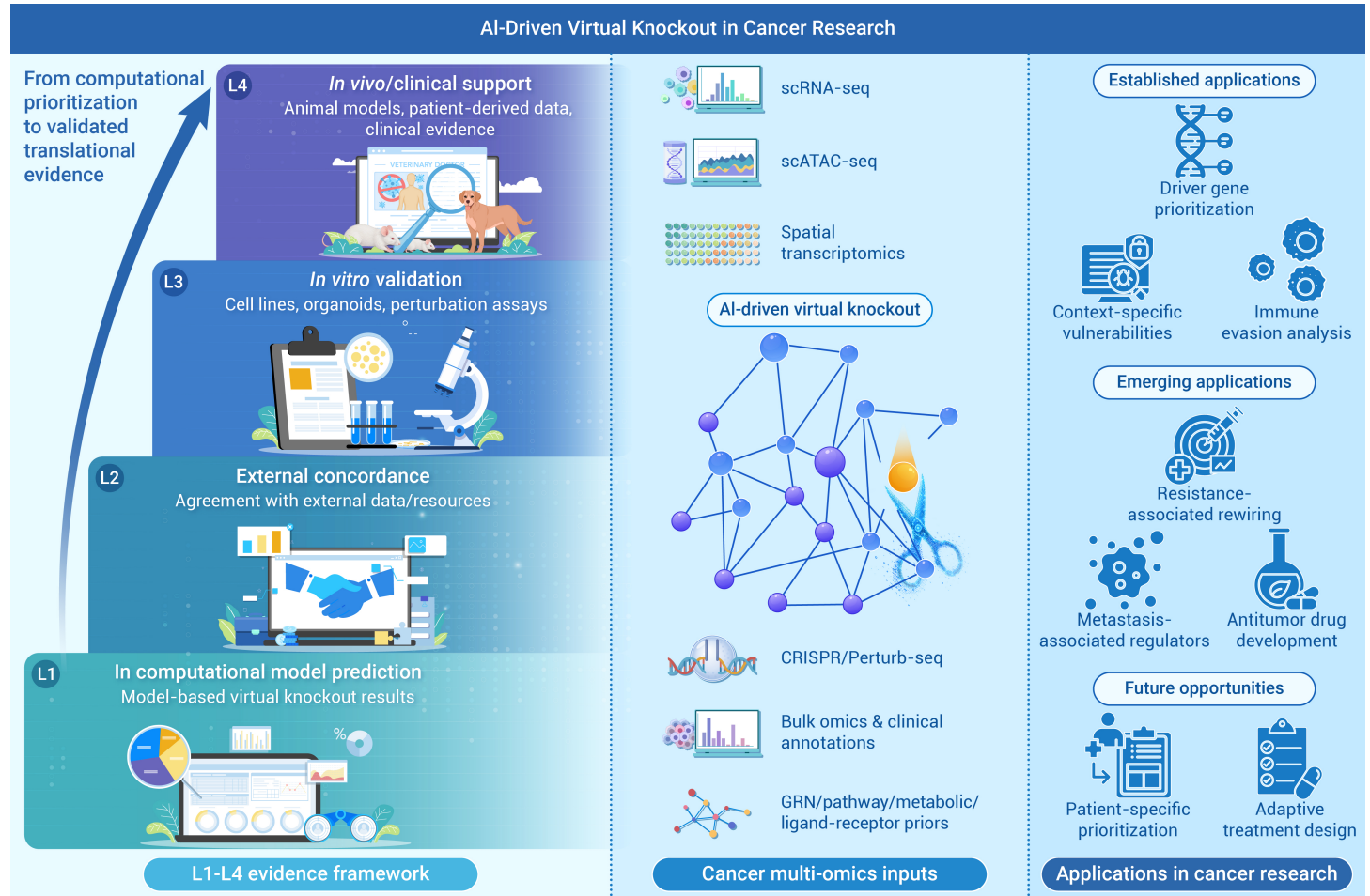
Xiangyu Wang,^{2,7,11} Min Liu,^{3,11} Meng Jia,^{6,11} Quan Wang,¹ Xiangwen Li,⁸ Qinqiao Zou,^{7,9} Xiuxiang Huang,^{7,9} Lina Zhang,¹ Xinrong Lv,¹⁰ Jie Wang,^{5,*} Lizeng Peng,^{4,*} and Wei Fang^{1,7,*}

*Correspondence: Fangwei635@163.com (W.F.); penglizeng@sdnu.edu.cn (L.P.); JieW629@163.com (J.W.)

Received: March 16, 2026; Accepted: July 12, 2026; Published Online: July 17, 2026; <https://doi.org/10.59717/j.xinn-oncol.2026.100027>

© 2026 The Author(s). This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

GRAPHICAL ABSTRACT



PUBLIC SUMMARY

- AI-driven virtual knockout simulates gene loss before experiments.
- This review covers gene regulatory network, generative, metabolic, and foundation model tools.
- An L1-L4 framework is established to assess the reliability of virtual knockout findings.
- Virtual knockout prioritizes cancer targets and context-specific vulnerabilities.
- Virtual knockout links computational prediction with experimental validation.

Artificial intelligence-driven virtual knockout in cancer research: Progress, challenges, and future perspectives

Xiangyu Wang,^{2,7,11} Min Liu,^{3,11} Meng Jia,^{6,11} Quan Wang,¹ Xiangwen Li,⁸ Qinbiao Zou,^{7,9} Xiuxiang Huang,^{7,9} Lina Zhang,¹ Xinrong Lv,¹⁰ Jie Wang,^{5,*} Lizeng Peng,^{4,*} and Wei Fang^{1,7,*}

¹Department of Pharmacy, Weihai Municipal Hospital, Cheeloo College of Medicine, Shandong University, Weihai 262400, China

²Department of Hepatobiliary Surgery, Central Hospital Affiliated to Shandong First Medical University (Jinan Central Hospital), Jinan 250013, China

³Department of Pharmacy, Qilu Hospital of Shandong University, Jinan 250012, China

⁴Key Laboratory of Novel Food Resources Processing Ministry of Agriculture, Institute of Food & Nutrition Science and Technology, Shandong Academy of Agricultural Sciences, Jinan 250100, China

⁵Department of Pharmacy, The First Affiliated Hospital of Xinjiang Medical University, Wulumuqi 830054, China

⁶Department of Microbiology, Perelman School of Medicine, University of Pennsylvania, PA 19104, USA

⁷Department of Public Health, International College, Krirk University, Bangkok 10220, Thailand

⁸School of Pharmaceutical Sciences Shandong University, Jinan 250012, China

⁹Beihai Vocational College of Wellness, Beihai 536000, China

¹⁰Department of Neurobiology, School of Basic Medical Sciences, Nanjing Medical University, Nanjing 211166, China

¹¹These authors have contributed equally to this work and share first authorship

*Correspondence: Fangwei635@163.com (W.F.); penglizeng@sdnu.edu.cn (L.P.); JieW629@163.com (J.W.)

Received: March 16, 2026; Accepted: July 12, 2026; Published Online: July 17, 2026; <https://doi.org/10.59717/j.xinn-oncol.2026.100027>

© 2026 The Author(s). This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

Citation: Wang X., Liu M., Jia M., et al. (2026). Artificial intelligence-driven virtual knockout in cancer research: Progress, challenges, and future perspectives. *The Innovation Oncology* 1:100027.

Artificial intelligence-driven virtual knockout is emerging as an intervention-oriented computational paradigm that helps move cancer research from descriptive association toward perturbation-informed hypothesis generation and model-derived counterfactual prediction. Virtual knockout simulates biologically meaningful loss-of-function perturbations in genes, pathways, or regulatory nodes. This approach facilitates target prioritization, model-based mechanistic interpretation, screening of combinatorial strategies, and optimization of downstream validation before experimental intervention. In this review, we define virtual knockout as a class of computational frameworks that model the consequences of loss-of-function perturbations and estimate model-derived post-perturbation biological states, while distinguishing strict virtual knockout models from related perturbation-inference paradigms. We summarize major modeling strategies, including gene regulatory network perturbation, trajectory and transcription factor program modeling, generative model-derived counterfactual prediction, constraint-based metabolic simulation, and communication-informed perturbation analysis, and review their applications in driver gene prioritization, context-specific vulnerability discovery, resistance-associated rewiring, tumor microenvironment analysis, metastasis-associated regulator prioritization, drug development, and patient-oriented therapeutic prioritization. We further propose a four-level evidence framework (L1-L4) and discuss major challenges, including data sparsity, tumor heterogeneity, context dependence, and cross-platform inconsistency. Overall, virtual knockout represents a promising bridge between computational inference, experimental validation, and precision oncology, although its current role is best viewed as one of prioritization and decision support rather than direct clinical replacement.

INTRODUCTION

With ongoing global population aging, the incidence and mortality of cancer continue to rise, causing nearly 10 million deaths annually and making it the second leading cause of death worldwide.¹ The occurrence of cancer is closely associated with compromised genomic stability within cells.^{2,3} The core of anticancer drug development lies in identifying key driver mutations that promote tumor initiation and progression, and designing specific targeted therapies against them.^{4,5} However, preclinical pharmacodynamic studies for a given target rely heavily on extensive cell-based assays and animal models that recapitulate human tumor features, which are extremely time- and resource-intensive,^{6,7} while only about 15%-20% of drug candidates entering clinical trials eventually achieve market approval.⁸ Consequently, bioinformatics has become an indispensable component of modern cancer research.⁹

Compared with correlation-based approaches, such as differential expression analysis or pathway enrichment analysis, virtual knockout places greater emphasis on counterfactual perturbation. Rather than asking only which genes are associated with a phenotype, virtual knockout asks how a biological system is predicted to change after a defined loss-of-function event. In this way, virtual knockout provides a bridge from descriptive association to perturbation-informed, testable hypotheses and is increasingly used to prioritize candidate targets, generate falsifiable mechanistic hypotheses, and guide downstream validation strategies.¹⁰ In cancer research, virtual knockout enables *in silico* evaluation of candidate loss-of-function perturbations before experimental intervention, supporting target prioritization, mechanistic investigation, and prediction of context-specific vulnerabilities.¹¹ These capabilities are particularly useful for dissecting tumor heterogeneity and identifying vulnerable targets in specific cellular subsets (Figure 1),^{12,13} including tumor-intrinsic dependency (Figure 1A), tumor state transition and plasticity (Figure 1B), tumor microenvironment and immune evasion (Figure 1C), and therapeutic target and combination discovery (Figure 1D).

THE CONCEPT DEFINITION AND PERTURBATION PREDICTION OF VIRTUAL KNOCKOUT-RELATED FOUNDATION MODELS

The concept definition of virtual knockout

Virtual knockout generally refers to the *in silico* deletion or inhibition of a specific gene, protein, pathway, or regulatory node within a computational framework, followed by simulation of the resulting system-level response to estimate the functional contribution of the perturbed entity in the original biological system. A method is considered strict virtual knockout when it meets four minimal criteria:

- (i) the perturbation is explicitly encoded as loss-of-function, such as knockout, knockdown, inhibition, deletion, or constraint-based inactivation;
- (ii) the perturbation target is clearly defined, such as a gene, protein, regulatory node, metabolic reaction, ligand-receptor pair, or pathway component;
- (iii) the model generates downstream responses beyond simple association;
- (iv) the predicted effect is biologically interpretable and, in principle, experimentally or externally testable.

By contrast, methods that predict cell states, drug responses, or counterfactual expression profiles without a clearly specified loss-of-function target are considered adjacent but distinct perturbation-inference approaches.^{14,15}

The perturbation prediction of virtual knockout-related foundation models

In addition to the task-specific gene regulatory network (GRN), metabolic, communication-informed, trajectory-based, and generative perturbation

Conceptual Overview of Virtual Knockout Applications in Cancer Research

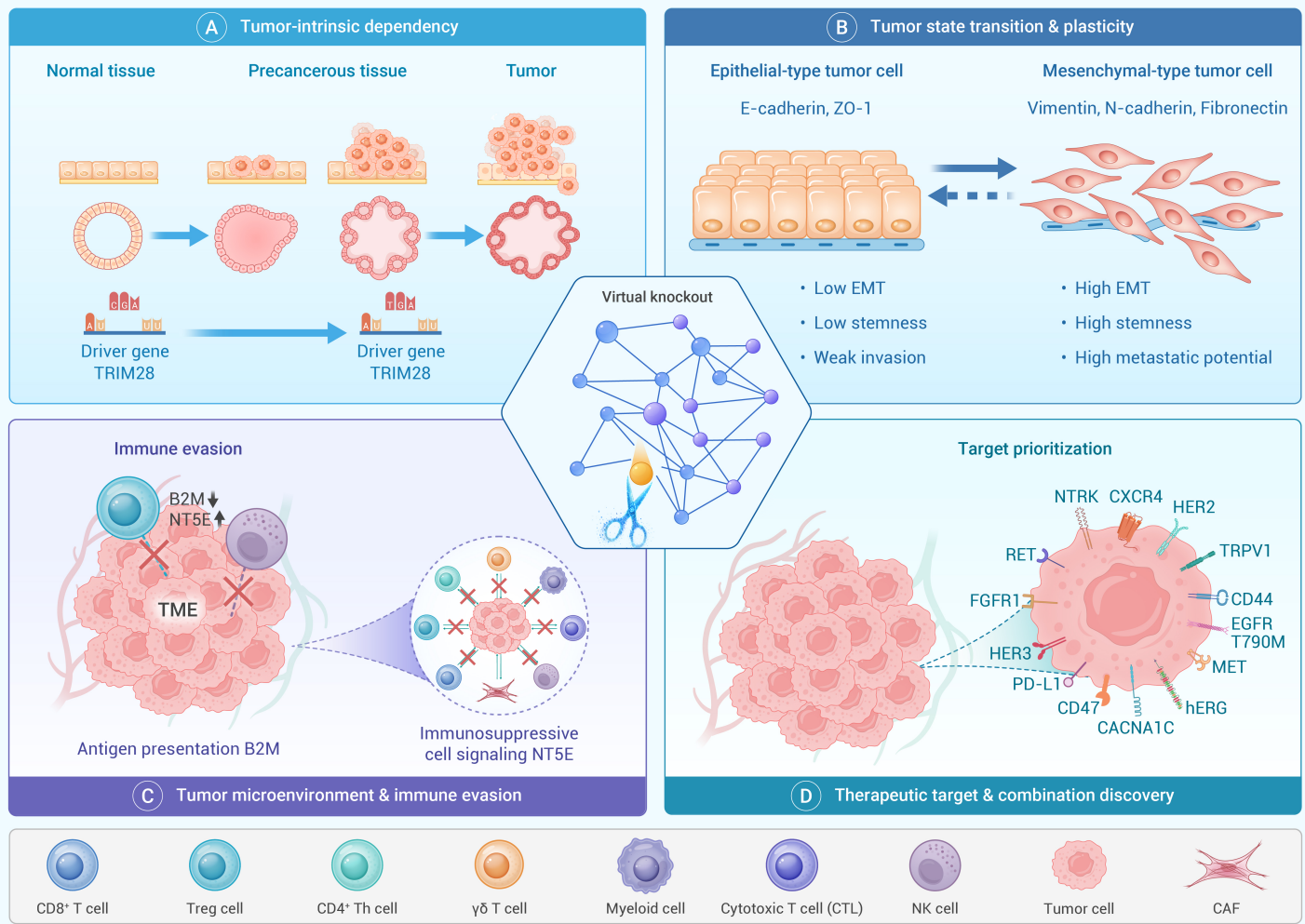


Figure 1. Conceptual overview of virtual knockout applications in cancer research Virtual knockout is an *in silico* perturbation framework for simulating loss-of-function events and inferring their consequences in cancer systems. The figure summarizes four major application domains: (A) Tumor-intrinsic dependency, (B) Tumor state transition and plasticity, (C) Tumor microenvironment and immune evasion, and (D) Therapeutic target and combination discovery. Together, these domains highlight virtual knockout as a scalable framework for mechanistic inference, hypothesis generation, and context-specific target prioritization in cancer research. **Abbreviation:** EMT, epithelial-mesenchymal transition.

models, single-cell foundation models have recently emerged as a distinct class of virtual knockout-related perturbation prediction frameworks.¹⁶ Models such as Geneformer,¹⁷ scGPT,¹⁸ scFoundation¹⁹ and Scouter²⁰ are pretrained on large-scale single-cell transcriptomic corpora and learn transferable representations of genes, cells, and cellular states. These models can support *in silico* perturbation screens through gene masking, token deletion, embedding perturbation, or learned perturbation representations, thereby predicting cell-state shifts or transcriptional responses after candidate perturbations. The scope and boundary of virtual knockout and related perturbation-inference frameworks are summarized in Table 1.

CORE ANALYTICAL ROLES OF VIRTUAL KNOCKOUT IN CANCER RESEARCH

Virtual knockout plays several core analytical roles in cancer research. First, it enables the genome-wide evaluation of perturbation strength induced by single-gene loss-of-function on system outputs, thereby facilitating target prioritization and substantially narrowing the scope of downstream experimental validation.^{21,22} This capability can be further extended to the *in silico* evaluation of double-gene or multi-gene perturbations, allowing early identification of combinations more likely to exhibit synergistic or synthetic-lethal effects.^{23,24} Second, virtual knockout can transform differential expression or enrichment signals into testable mechanistic hypotheses. By introducing perturbations within computational models and tracing their predicted down-

stream consequences, it enables the analysis of pathway-level responses, state transitions, and context-specific regulatory effects.^{11,22} When applied across distinct cell types, cellular states, or molecular subpopulations defined by single-cell and related datasets, virtual knockout can reveal context-dependent vulnerabilities that are often obscured in bulk or mixed-cohort analyses.^{25,26} In addition, by incorporating environmental constraints, such as nutrient availability or exchange flux limits, virtual knockout can be used to evaluate condition-dependent essentiality and identify microenvironment-dependent vulnerabilities.^{27,28} Third, the predictions generated by virtual knockout can be aligned with measurable experimental and translational endpoints. Computational readouts can be linked to proliferation, drug response, migration and invasion, immune phenotypes, pathway activity, and molecular markers, thereby helping refine validation design and improve translational feasibility.²⁹⁻³¹ In parallel, essential genes, key enzymes, or metabolic bottleneck reactions identified through virtual knockout can be systematically linked to drug-target resources to generate druggable candidates and support drug repurposing or combination design.²⁴ In this sense, virtual knockout serves not only as a tool for hypothesis generation, but also as a practical bridge between computational inference and downstream experimental or therapeutic decision-making. The core analytical capabilities of virtual knockouts and their connections to the major application dimensions discussed in the following sections are summarized in Table 2.

Table 1. Scope and boundary of virtual knockout and related perturbation-inference frameworks.

Model class	Representative tools	Required inputs	Perturbation representation	Main outputs/readouts	Interpretability	Common failure modes	Recommended safeguards/suitable validation	Common leakage risks	Recommended mitigation strategies
GRN perturbation	scTenifoldKnk, CellOracle, GenKI	scRNA-seq, scATAC-seq, TF motif info	TF knockout/KD, edge removal	Downstream gene modules, network rewiring	Medium-high; network edges interpretable	Erroneous GRN edges; correlation causation; sparsity or dropout	Bootstrap/resampling, multiple network reconstructions, validation with orthogonal datasets	Same cells/patients for GRN inference and validation; non-independent edge validation	Patient/sample-level split; independent dataset validation; bootstrap GRN stability
Generative perturbation prediction	scGen, GEARS	scRNA-seq (training), baseline cell states	Latent space knockout/shift, simulated perturbation	Predicted gene expression post-knockout, cell state transitions	Medium; latent space shifts less interpretable	Model collapse; OOD cell types poorly predicted	Cross-validation, latent space regularization, perturbation reconstruction accuracy checks	Train-test overlap in perturbations, cell states, or batches; global similarity masking failure	Leave-one-perturbation/cell-type/dataset-out testing; replicate-level split; DE/pathway recovery evaluation
Constraint-based metabolic models	GSMM/FBA, ngMCS	Metabolomics, expression data, reaction network	Flux constraints, gene knockout	Essential genes, flux distributions, synthetic lethal pairs	High; flux/constraints mechanistically interpretable	Incorrect objective function; incomplete metabolic network; low predictive power for rare reactions	Multiple objective functions, validation with knockout datasets, sensitivity analysis, flux variability analysis	Outcome or drug-response leakage into constraints/objectives; mixed tumor-normal constraints	Predefined constraints/objectives; tumor-normal model separation; held-out validation; FVA/sensitivity analysis
Communication-informed perturbation	NicheNet, CellChat	scRNA-seq, ligand-receptor database, signaling prior	Ligand/receptor deletion, perturbation inference	Predicted changes in receiver gene programs, altered interactions	Medium; pathway-level interpretability	Incomplete LR database; missing cell types; false positive interactions	Validation with orthogonal LR datasets, replicate cell types, experimental perturbation (e.g., ligand blocking)	Same data for LR inference and receiver validation; incomplete LR priors; abundance-signaling confusion	Independent spatial/proteomic validation; LR database comparison
Trajectory/TF program perturbation	dynGENIE3, Scribe	scRNA-seq, pseudotime trajectories, TF targets	TF knockout/KD along trajectory	Predicted trajectory shifts, altered differentiation states	Medium-high; trajectory branch changes interpretable	Pseudotime misalignment; spurious TF-target links	Pseudotime robustness check, multiple trajectory inference methods, orthogonal TF perturbation validation	Same-data pseudotime validation; batch/cell-cycle-driven branches; correlation-based TF links	Independent/time-course validation; multiple trajectory methods; batch/cell-cycle adjustment; motif/scATAC support
AI/foundation-model-based virtual knockout-related perturbation prediction	Geneformer, scGPT, scFoundation, UCE, Scouter, GenKI	Large-scale scRNA token deletion, seq corpora; query embedding; cell states; optional perturbation data	Gene masking, token deletion, embedding perturbation, learned perturbation representation	Predicted post-perturbation transcriptional states, candidate regulators, cell-state shifts	Medium; representation-based and often indirect	Corpus bias; zero-shot unreliability; OOD failure; metric dependent overestimation; limited mechanistic interpretability	Benchmark against Perturb-seq/CRISPR screens; compare with simple baselines; evaluate DE-gene/pathway recovery; test unseen genes/cell types; report uncertainty	Pretraining-benchmark overlap; "unseen" genes/cells not truly unseen; corpus or batch bias comparison	Data-provenance checks; dataset/perturbation/cell-type holdouts; independent Perturb-seq/CRISPR benchmarks; baseline comparison

Taken together, this comparison emphasizes that model selection should be driven by the biological question and validation plan rather than by tool availability alone. GRN-based and trajectory-based approaches are most useful when the goal is to identify upstream regulators and interpretable regulatory programs. Generative perturbation models are well suited for large-scale candidate screening and prediction of transcriptional responses, particularly when perturbation training data are available. Constraint-based metabolic models are appropriate for metabolic dependency, synthetic lethality, and drug-repurposing analyses. Communication-informed perturbation models are particularly useful for tumor microenvironment studies, especially when the perturbation target involves ligand-receptor pairs, cytokine signaling, or cell-cell communication axes. Foundation-model-based perturbation predictors may further expand scalability, but their outputs should be interpreted cautiously unless benchmarked against independent perturbation datasets or experimentally validated. **Abbreviations:** TF, transcription factor; CRISPR, clustered regularly interspaced short palindromic repeats; GRN, gene regulatory network; ATAC, assay for transposase-accessible chromatin; KD, knockdown; FBA, flux balance analysis; UCE, universal cell embedding; LR, ligand-receptor. Interpretability indicates how mechanistically understandable the model outputs are; "high" means outputs can be mapped to clear biological mechanisms; "medium" indicates partial interpretability; "low" indicates latent or black-box outputs. Common failure modes are frequently observed limitations or risks for each model class, which may affect prediction accuracy or generalizability. Recommended safeguards/suitable validation include suggested computational or experimental strategies to mitigate the failure modes. Tools listed are representative, not exhaustive, and new tools (e.g., GenKI, *In Silico* Perturbation Oracle) are emerging. The leakage-risk and mitigation columns are intended as a prospective checklist for applying the L1-L4 framework across different virtual knockout-related modeling paradigms. They highlight common sources of training-testing information leakage and recommended safeguards, but the listed strategies are not exhaustive.

Table 2. Core capabilities of virtual knockout.

Dimension	Capability	Main function
Prioritization power	Genome-wide prioritization	Identifies high-confidence candidate targets from large gene sets.
Prioritization power	Combination optimization	Prioritizes multi-gene or multi-target perturbation strategies.
Mechanistic interpretability	Mechanistic hypothesis generation	Predicts how perturbation effects propagate through regulatory or functional networks.
Mechanistic interpretability	Context-specific stratification	Compares perturbation effects across cell types, states, or molecular subgroups.
Mechanistic interpretability	Microenvironment-aware modeling	Evaluates environment-dependent vulnerabilities under hypoxia, nutrient limitation, or stromal signaling.
Experimental and translational linkage	Readout alignment	Links predictions to measurable endpoints such as proliferation, drug response, or immune phenotypes.
Experimental and translational linkage	Target-to-drug translation	Connects predicted dependencies to drug resources and candidate compounds.
Experimental and translational linkage	Validation planning	Guides follow-up validation based on predicted effect size, uncertainty, and context specificity.

Table 3. Evidence levels (L1-L4) for virtual knockout studies.

Evidence level	Definition	What it supports	Typical outputs	QC/safeguards	Validation/assays
L1	In-model virtual knockout only	Prioritization	Targets; pathways/modules; state shift; SL pairs; drug leads	Bootstrap; sensitivity; permutation; no leakage	None
L2	L1 + external concordance	Reproducibility	L1 + replicated trends	Cross-dataset; FDR	Data: DepMap, Perturb-seq, LINCS. KB: STRING, Reactome, KEGG
L3	<i>In vitro</i> perturbation	Perturbation-specific effect (<i>in vitro</i>)	Phenotype + mechanism	≥2 sgRNAs; rescue; orthogonal; replicate	CRISPR knockout/CRISPRi; drugs (dose-response); target engagement
L4	<i>In vivo</i> /clinical support	Translation	<i>In vivo</i> endpoints; cohort outcome links; biomarkers	Robust endpoints; cohort replication; feasible assays	Animal; PDX/PDO; cohort analyses; IHC/blood/signatures

This table presents an evidence-organizing framework for virtual knockout studies rather than a formal field-wide standard. The four levels are stratified according to validation source and independence from the original computational model. L1 indicates model-internal computational prediction supported by robustness checks; L2 indicates external computational concordance with independent datasets or curated knowledge bases; L3 indicates controlled *in vitro* perturbation validation using genetic or pharmacologic perturbation with functional and mechanistic readouts; and L4 indicates *in vivo*, patient-derived, or clinically relevant translational support. The framework is intended to guide interpretation, reporting, and validation planning across heterogeneous virtual knockout paradigms. Evidence levels should be interpreted in a context-dependent manner and do not by themselves establish universal biological or clinical validity. **Abbreviations:** QC, quality control; FDR, false discovery rate; LINCS, library of integrated network-based cellular signatures; KB, knowledge base; STRING, search tool for the retrieval of interacting genes/proteins; KEGG, kyoto encyclopedia of genes and genomes; CRISPRi, CRISPR interference; PDX, patient-derived xenograft; PDO, patient-derived organoid; IHC, immunohistochemistry.

EVIDENCE STRATIFICATION AND VALIDATION FRAMEWORK FOR VIRTUAL KNOCKOUT

At present, no dedicated consensus framework exists for evaluating the evidential strength of virtual knockout studies. We therefore propose an evidence-organizing framework rather than a formal field-wide standard. The rationale for this framework was derived from three related methodological traditions. First, clinical prediction-model and medical AI evaluation frameworks emphasize the distinction between internal validation, external validation, risk of bias, and applicability, highlighting that model performance assessed only within the development setting is insufficient to establish generalizability.³²⁻³⁶ Second, systems biology and computational biomedicine have developed model-credibility concepts in which verification, validation, uncertainty quantification, model testability, and context of use are central to determining whether computational predictions can be trusted for a particular scientific or translational purpose.^{37,38} Third, in computational biology and bioinformatics, computationally derived hypotheses generally gain evidential strength when they are corroborated by independent datasets, orthogonal assays, controlled experimental perturbations, and ultimately *in vivo* or patient-relevant evidence.^{39,40}

Based on these principles, we stratify virtual knockout evidence according to two organizing principles: the source of validation evidence and the degree of independence from the original computational model.^{37,41} Thus, the L1-L4

framework is intended to clarify the strength and limitations of evidence supporting virtual knockout claims and to guide transparent reporting and validation planning across heterogeneous virtual knockout paradigms (Table 3).⁴²

L1: In-model computational evidence without independent external validation. L1 represents the lowest level of evidence in this framework because the prediction is evaluated only within the original computational environment and is therefore supported by internal robustness checks rather than independent biological validation.⁴³⁻⁴⁶ The intended scope of L1 evidence is therefore exploratory hypothesis generation and computational prioritization within the development model. It should not be used to support biological causality, external generalizability, or translational validity. At a minimum, L1 evidence should include bootstrap or resampling-based stability evaluation,⁴³ sensitivity analysis,⁴⁴ permutation or randomized negative controls,⁴⁵ and strict avoidance of training-test information leakage that may inflate performance estimates.⁴⁶ Therefore, L1 evidence is useful for hypothesis generation and prioritization, but it should not be interpreted as evidence of biological causality or translational validity.²⁰

L2: External computational concordance with independent datasets or knowledge bases. L2 represents a higher level of support because predictions are compared with evidence outside the original model, such as independent public datasets, perturbation screens, or curated pathway and inter-

action resources. However, L2 should still be interpreted as external computational concordance or biological plausibility, not as controlled wet-lab validation.^{38,39} More specifically, L2 concordance should be assessed using datasets or curated resources not used in model construction, at a prespecified level such as target overlap, effect-direction agreement, pathway-level consistency, perturbation-response signature similarity, dependency concordance, or drug-response concordance.⁴⁷ Curated knowledge-based overlap alone should be interpreted as weak L2 unless supported by independent perturbation or phenotype data. However, L2 evidence should be interpreted as computational-data concordance rather than causal wet-lab validation. Therefore, L2 can strengthen confidence in predicted perturbation effects and support candidate prioritization, but should not be considered sufficient evidence of biological causality or translational validity. Perturbation-specific causal interpretation should require at least L3 evidence, in which selected predictions are tested through controlled biological perturbation.

L3: *In vitro* experimental validation. L3 represents a further increase in evidential strength because key virtual knockout predictions are tested through controlled biological perturbations rather than only through computational concordance. This level is consistent with the principle that computationally derived hypotheses gain credibility when supported by orthogonal experimental evidence and reproducible functional readouts.^{40,48} At this stage, selected predictions are subjected to controlled perturbations, such as clustered regularly interspaced short palindromic repeats (CRISPR)-based knockout, CRISPR interference (CRISPRi), RNA interference, or pharmacological inhibition, combined with functional and mechanistic readouts.⁴⁸ The intended scope of L3 evidence is to support perturbation-specific validity in controlled *in vitro* systems. It can indicate that a selected virtual knockout prediction has experimentally testable functional relevance in a defined cellular context, but it should not by itself be interpreted as evidence of *in vivo* efficacy, patient-level benefit, or clinical applicability.

As a boundary rule, L3 evidence should require at least one functional readout.⁴⁹ Molecular marker changes alone, such as qPCR, immunoblotting, or protein-expression changes without a functional phenotype, should be considered incomplete L3 evidence unless supported by additional perturbation-specific functional validation. To reduce false positives, genetic perturbations should ideally employ at least two independent sgRNAs or constructs and, where feasible, include rescue experiments. Pharmacological interventions should report dose-response relationships and replicate *in vitro* systems.^{49,50} In addition, CRISPR-Cas-mediated knockout should not be assumed to produce complete loss of function in all cases, because knockout escaping can generate residual functional proteins and lead to misinterpretation of knockout phenotypes.⁵¹ Therefore, L3 evidence supports perturbation-specific *in vitro* validity for the predicted effect.

L4: Evidence derived from *in vivo*, patient-derived, or clinically relevant translational systems is categorized as Level 4 (L4) evidence. Validating platforms include animal models, patient-derived xenografts, patient-derived organoids, and genetically engineered mouse models.^{7,52} This level is consistent with closed-loop validation frameworks for AI-driven virtual cell models, in which computational predictions are progressively assessed through CRISPR-based perturbation assays, organoid or organ-on-chip systems, animal models, and translational datasets.¹⁶ At this stage, key predictions or perturbation-prioritized targets should be tested with readouts such as tumor burden, histopathological progression, proliferation, apoptosis, therapeutic response, survival, or clinically feasible biomarker measurement. Furthermore, if a target or pathway shows reproducible risk stratification capability in independent patient cohorts, correlates with therapeutic response or survival outcomes, and can be measured through feasible assays such as IHC, blood biomarkers, or transcriptomic signatures, its translational potential is further strengthened.^{29,53}

As a boundary rule, L4 evidence should require validation in *in vivo*, patient-derived, or clinically relevant systems with translational endpoints, such as tumor burden, survival, therapeutic response, toxicity, biomarker feasibility, or reproducible patient-outcome associations.⁵⁴ Patient-cohort association alone should be interpreted as clinical concordance rather than definitive L4 evidence unless supported by functional, mechanistic, patient-derived, or independent translational validation. The intended scope of L4 evidence is to support translational relevance in *in vivo*, patient-derived, or clinically anno-

tated contexts. However, L4 evidence should still be interpreted within its biological and clinical context of use; it does not imply universal clinical validity unless replicated across independent models, patient-derived systems, or clinical cohorts.

To facilitate the conversion of gene-level virtual knockout predictions into experimentally testable perturbation hypotheses, we suggest a minimal closed-loop validation workflow consisting of virtual prioritization, controlled loss-of-function validation, orthogonal functional and mechanistic assessment, and feedback to define model boundaries. For gene-level targets, validation may involve CRISPR knockout, RNA interference, or pharmacological inhibition, with appropriate controls including untreated or wild-type cells, non-targeting controls, at least two independent sgRNAs or perturbation modalities, and rescue experiments where feasible. Functional readouts, such as proliferation, viability, migration, invasion, drug sensitivity, or immune-related phenotypes, should be paired with mechanistic readouts, such as pathway markers, protein changes, or transcriptomic modules. Reproducibility across batches, models, and external datasets should be assessed whenever possible, and key parameters, effect sizes, and replicate information should be transparently reported. When CRISPR knockout is used, target loss should be confirmed at both molecular and functional levels, and knockout escape should be considered a potential source of misleading phenotypes. For other virtual knockout paradigms, including metabolic reaction knockouts, communication-axis perturbations, or higher-order regulatory abstractions, validation strategies should be adapted according to the biological target, perturbation modality, and model context.

Given that the implementation strategies and validation depth of virtual knockout vary substantially across different tumor contexts and data modalities, we compiled representative application cases cited in this review into Table 4, organized according to “research setting-virtual knockout target-data-model-key findings-evidence level,” to facilitate cross-comparison and methodological reuse. To avoid presenting the L1-L4 framework as a success-only hierarchy, Table 4 also includes a challenged-prediction example. Recent benchmarking work has shown that some genetic perturbation-response models may capture systematic variation or confounding between perturbed and control cells rather than perturbation-specific effects, underscoring the need to distinguish model-internal prediction from externally or experimentally validated virtual knockout evidence.

ESTABLISHED AND ACTIVELY VALIDATED APPLICATIONS OF VIRTUAL KNOCKOUT IN CANCER RESEARCH

Established or actively validated applications refer to areas in which virtual knockout predictions can be connected to relatively concrete model-based, pathway-level, perturbation-linked, or experimentally supported examples. Currently, this technology is primarily employed in tumor research to explore the functions of driver genes, screen for tumor-specific therapeutic targets, and identify immune-related regulatory targets, thereby offering support for mechanism research and experimental verification. These applications are summarized in Figure 2.

Virtual knockout introduces *in silico* perturbations into computational models of cancer systems, enabling systematic interrogation of gene functions and network responses. (A) Data and model input; (B) virtual knockout in cancer research; (C) driver gene identification and target prioritization; (D) discovery of context-specific regulatory dependencies across tumor subtypes, cellular states, and microenvironmental niches; and (E) dissection of compensatory signaling and network rewiring associated with therapeutic resistance.

Driver gene prioritization and oncogenic mechanism inference

Among current oncologic applications, driver gene prioritization represents one of the most established uses of virtual knockout, because candidate regulators can be evaluated within gene regulatory networks, pathway modules, and experimentally testable perturbation frameworks. Driver genes regulate core cancer phenotypes, including proliferation, apoptosis, cell-fate control, and genome maintenance. Their abnormal activation, loss-of-function, or regulatory imbalance can drive tumor initiation and progression. Therefore, identifying functionally important driver genes remains a central task in cancer research.^{11,14,60} The perturbation effects derived from virtual

Table 4. Representative application paradigms of virtual knockout in cancer research and evidence levels (L1-L4).

Cancer context/study setting	Perturbation target/scope	Data	Computational framework	Key prediction/implication	Evidence level (L1-L4)	Representative literature
Cancer-cell perturbation prediction	Unseen single-gene knockout	Control single-cell profiles	LLM-embedding perturbation model	Predicted post-knockout transcriptional states for L1 prioritization		20
Ovarian cancer metabolism and drug repurposing	Metabolic genes and reactions	Expression constraints + GSMM + DrugBank	GSMM/FBA + DrugBank mapping	Metabolic target prioritization with DrugBank-linked drug nomination	L2	30,31
Multiple myeloma metabolic synthetic lethality	Synthetic-lethal metabolic gene Dknockout pairs	GEM + RNA-seq	gmCS (gmctool)	RNA-seq filtering yielded sample-specific actionable Dknockout candidates	L2	24
HCC metabolism with normal-tissue context	Tumor metabolic genes/reactions	Tumor vs. normal omics constraints	Tumor/normal metabolic modeling	Distinguished shared and patient-specific metabolic liabilities	L2	55
Breast cancer (BRCA) risk-state modulation	CTLA4; CTLA4+IL6 perturbation	scRNA-seq + bulk outcome data	SIDISH	Quantified risk-state shifts following CTLA4 or CTLA4+IL6 perturbation	L2	56
Drug-resistant breast cancer	Single-gene metabolic knockout	Drug-resistant breast cancer metabolic/transcriptomic profiles	Genome-scale knockout simulation	Drug-sensitization targets experimentally validated	L3	57
ccRCC (<i>in vivo</i> + combination therapy)	TRIM28 (genetic perturbation ± drug)	scRNA-seq + <i>in vivo</i> models	Geneformer virtual knockout screening	Nominated TRIM28 as a targetable dependency, with <i>in vivo</i> support for anti-PD-1 combination efficacy	L4	17,58
Benchmarking/negative calibration	Genetic perturbation effects	Multiple perturbation datasets	Systema	Some predicted effects reflected systematic variation rather than perturbation-specific biology	Challenged prediction	59

L1 indicates model-internal prediction without independent validation; L2 indicates external computational or dataset-level concordance; L3 indicates controlled *in vitro* experimental validation; L4 indicates *in vivo*, patient-derived, or clinically relevant translational support. The challenged-prediction row is included to illustrate how benchmarking studies can reveal limitations of model-derived perturbation predictions and help calibrate the evidence framework. **Abbreviations:** GSMM, genome-scale metabolic model; gmCS, genetic minimal cut set; FBA, flux balance analysis; GEM, genome-scale metabolic network; DKnockout, double knockout; LLM, large language model; CTLA4, cytotoxic t-lymphocyte-associated protein 4; IL6, interleukin 6; ccRCC, clear cell renal cell carcinoma; TRIM28, tripartite motif containing 28; PD-1, programmed cell death protein 1.

knockout can be integrated with structural priors such as gene regulatory networks, signaling pathway networks, or protein-protein interaction networks. Through signal propagation or module mapping, affected key pathway modules and potential mediator nodes can be identified, offering clearer validation targets and readout strategies for subsequent mechanistic experiments.^{11,60}

scTenifoldKnk used wild-type single-cell transcriptomic data to simulate Nkx2-1 loss of function. The model identified 171 significantly perturbed genes among 8,647 candidates. The predicted changes in AT1/AT2 marker expression and pathway enrichment were highly consistent with real Nkx2-1 knockout datasets. This example illustrates how virtual knockout can prioritize candidate downstream targets for mechanistic and functional validation.^{22,61} A related application is illustrated by genome-wide virtual knockout screening with Geneformer in ccRCC, where malignant epithelial cells were benchmarked against normal epithelial references to evaluate whether perturbations could shift tumor transcriptional states toward normal-like phenotypes. In this setting, virtual inactivation of TRIM28 was predicted to drive malignant epithelial cells toward a more normal epithelial-like state, nominating TRIM28 as a potential dependency involved in maintaining tumor identity and providing a candidate target for downstream validation.^{17,58} Together, these examples show that virtual knockout can help move driver gene research beyond descriptive association. Its major value is not only to rank candidate genes, but also to link those candidates to pathway-level and cell-state-level consequences that can be experimentally tested.

Tumor context-specific vulnerabilities and subtype/state-dependent effects

Clarifying whether driver genes exert consistent effects across different subtypes and cellular states is a prerequisite for precise tumor classification and personalized therapy.²⁶ Virtual knockout enables stratified perturbation simulations within the context of cell types, cellular states, or molecular subpopulations defined by single-cell sequencing and related datasets. By comparing the silencing effects of the same driver gene across distinct subtypes and examining downstream module responses, subtype-specific regulatory differences and potential vulnerabilities can be identified, thereby providing a more actionable hypothesis framework for stratified validation and therapeutic decision-making.^{11,22,26,60} SIDISH (Single-cell and Bulk Transcriptomics Integration for High-risk Cell Detection and Precision Therapeutics through *in silico* Perturbation) treats clinically defined high-risk cell populations in breast cancer high-risk cell populations as phenotypic readouts and explicitly incorporates a virtual perturbation module into the model. Single-cell and spatial studies of triple-negative breast cancer (TNBC) demonstrate that key malignant phenotypes and immune responses exhibit strong state- and niche-dependent characteristics.^{56,62} Similarly, CellOracle constructs state-resolved gene regulatory networks (GRNs) and demonstrates that transcription factor effects vary across lineage contexts and differentiation stages, thereby guiding the prioritization of specific cell types or temporal windows for CRISPR-based interventions.^{11,63,64}

Beyond subtype- and state-dependent effects, context dependence can

Analytical Landscape of Virtual Knockout in Cancer Research

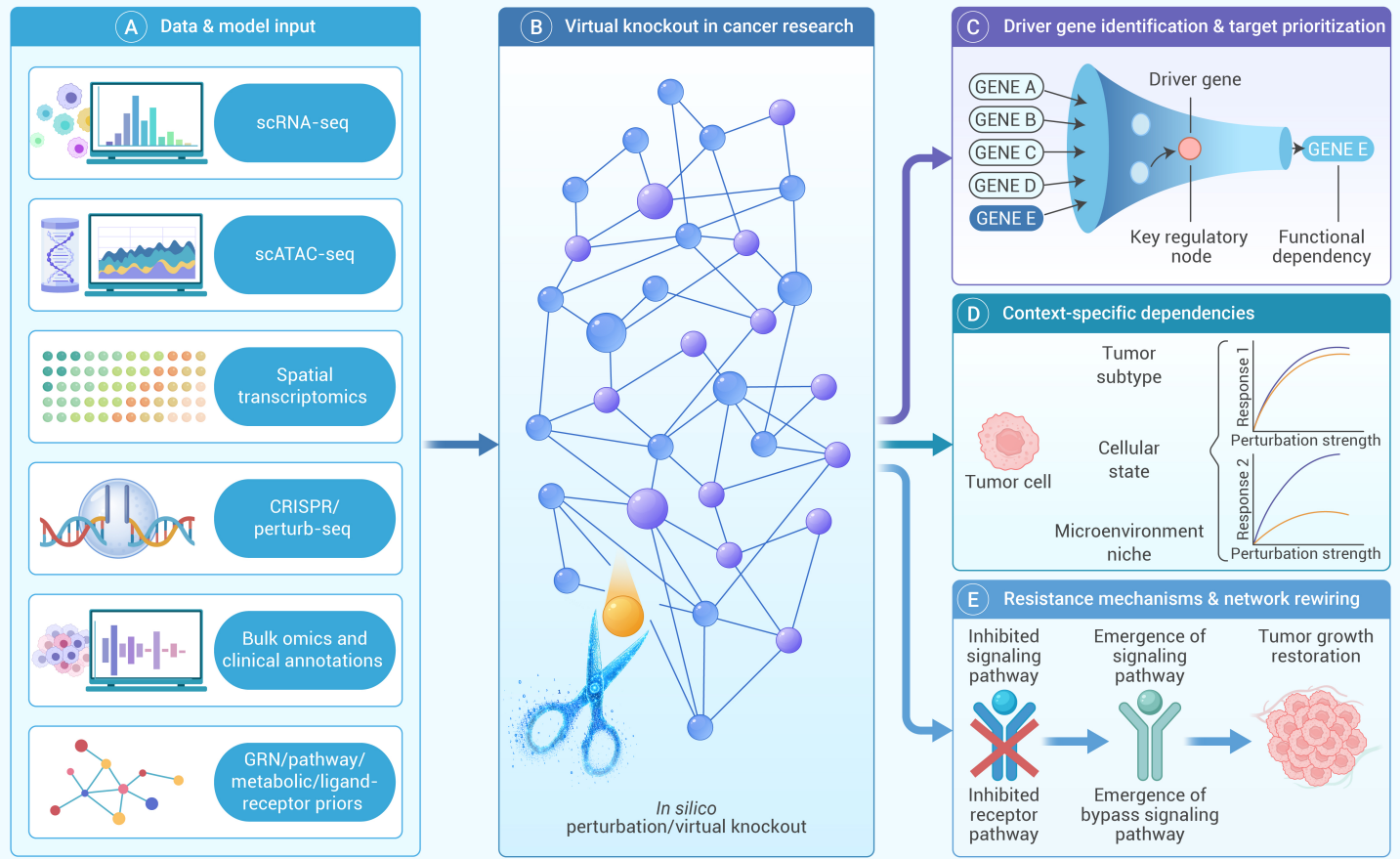


Figure 2. Analytical landscape of virtual knockout in cancer research.

also arise from extracellular constraints.²⁸ Agren and colleagues constructed personalized genome-scale metabolic models (GEMs) based on hepatocellular carcinoma (HCC) patient data. Within these models, virtual perturbations were implemented by setting specific reaction fluxes to zero or imposing inhibitory constraints, thereby simulating target loss-of-function or pharmacological inhibition and assessing their effects on tumor metabolism and growth-associated capabilities. This strategy not only identified candidate antimetabolites shared across patients but also revealed patient-specific metabolic vulnerabilities. Furthermore, potential toxicity was evaluated by integrating normal tissue metabolic models, and several predictions were subsequently validated through cellular experiments.⁵⁵ Likewise, ngMCS treats nutrient deprivation as exchange-reaction knockout and integrates it with gene knockout to infer condition-dependent lethality; case studies such as DHFR illustrate how target sensitivity varies across nutrient contexts, generating experimentally testable, environment-specific hypotheses.²⁷

Tumor immune microenvironment regulation and cancer immune evasion

Tumor immune microenvironment and immune-evasion analysis can be considered an established-to-emerging application area, because many candidate perturbations can be anchored to well-defined immune pathways, antigen-presentation modules, ligand-receptor axes, and increasingly available single-cell, spatial, and perturbation datasets. To help readers grasp the overall logic of the “TIME data input → virtual knockout inference → target prioritization → functional validation” workflow before delving into specific mechanisms such as antigen-presentation loss and immunosuppressive axes, we provide in Figure 3 an overview of a representative closed-loop framework in which virtual knockout is applied to dissect immune evasion, together with the corresponding downstream validation strategies.

Immune evasion via antigen-presentation loss: virtual knockout for perturbation-based hypothesis generation.

A central immune escape route is loss of MHC-I antigen presentation, which prevents CD8⁺ T cells from recognizing tumor antigens.^{65,66} Virtual knockout can be used here in two complementary ways: (1) Tumor-intrinsic virtual knockout (e.g., B2M, NLRC5, TAP1/2) in malignant cell clusters to predict downstream transcriptional consequences (including IFN-response programs, chemokine secretion, and stress responses) and how these changes might reduce immune recruitment/recognition.^{67,68} GRN-based tools (scTenifoldKnk/CellOracle-like) are particularly useful for revealing network-level consequences rather than assuming a simple linear effect.⁶⁹ (2) Communication-aware virtual knockout, where one evaluates how antigen-presentation loss alters tumor-to-immune signaling.⁷⁰ For example, by integrating scRNA-seq with NicheNet, one can ask whether knockout-driven reductions in IFN-stimulated target genes or chemokines (e.g., CXCL9/10/11 programs, when present) weaken recruitment/activation signals for cytotoxic lymphocytes, offering mechanistic hypotheses for “immune-cold” phenotypes.⁶⁹

Immunosuppressive axes: probing checkpoints and suppressive metabolites/cytokines.

Immune suppression in the tumor immune microenvironment (TIME) is driven by overlapping pathways that act across tumor, immune, and stromal compartments.⁷¹⁻⁷³ Virtual knockout can help test axis-level perturbations in a cell type-resolved manner. The TGF- β axis is a major mediator of immune evasion and therapy resistance.⁷⁴ Virtual knockout of TGF- β ligands, receptors, or upstream regulators in tumor cells can predict whether suppressive ligand production decreases, antigen-presenting cell signatures recover, and T-cell infiltration or activation programs are derepressed *in silico*.⁷⁴ CD73, encoded by NT5E, generates extracellular adenosine and suppresses antitumor T-cell immunity through adenosine receptor signaling. It has therefore been described as an emerging immune

Dissecting Tumor Immune Evasion via Virtual Knockout Technologies

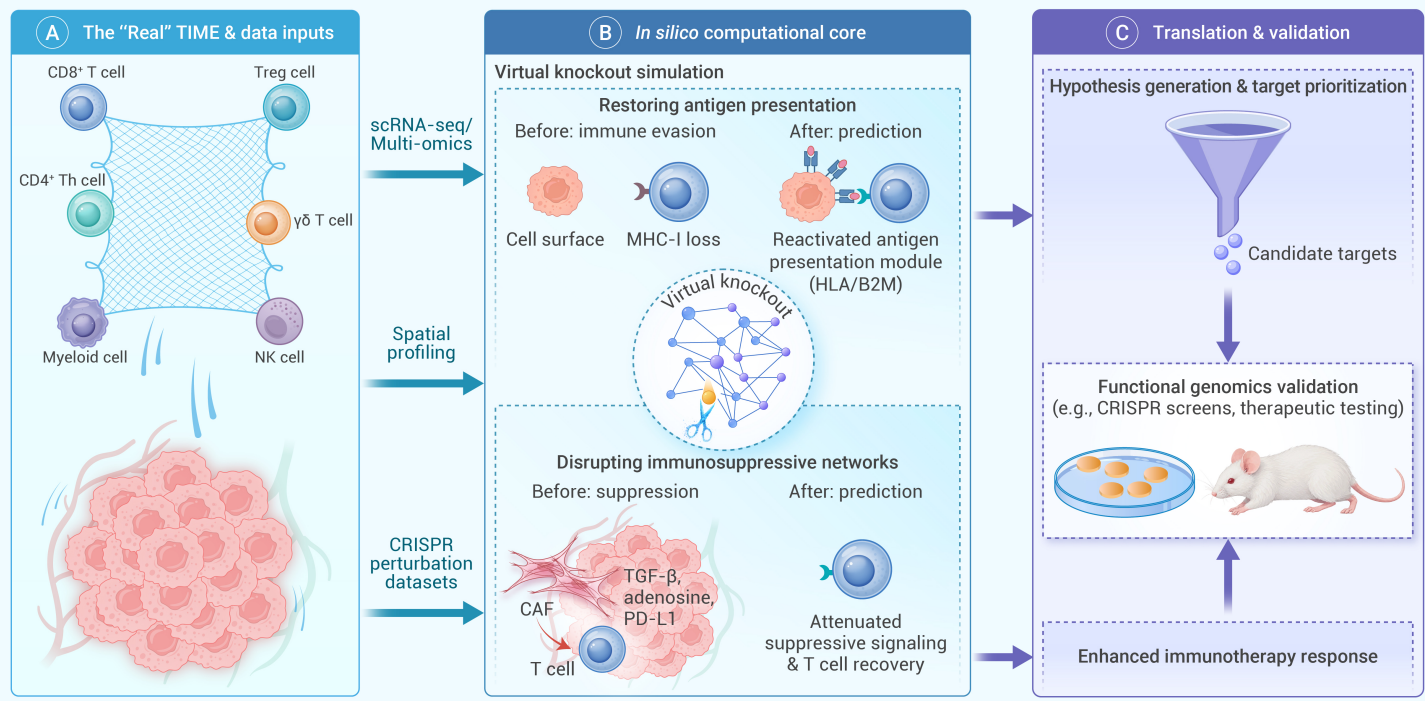


Figure 3. Schematic overview of a virtual knockout-based closed-loop framework for dissecting immune evasion in the tumor immune microenvironment and guiding target prioritization and validation (A) Multi-omics data from the tumor immune microenvironment (TIME) are integrated to characterize immune cell states, tumor-immune interactions, and candidate regulatory nodes. (B) These data are incorporated into computational models to perform virtual knockout simulations of candidate genes or regulatory nodes. (C) The resulting model-derived predictions reveal potential immune system-level responses, thereby supporting target prioritization, functional validation, and immunotherapy-related hypothesis generation.

checkpoint.⁷⁵ Virtual knockout of NT5E can estimate whether adenosine-related suppression is reduced and whether effector T-cell programs recover.^{76,77}

Overall, virtual knockout is useful for prioritizing immunosuppressive pathways and generating testable hypotheses for combination immunotherapy.

EMERGING APPLICATIONS OF VIRTUAL KNOCKOUT IN CANCER BIOLOGY AND ANTITUMOR THERAPY

Emerging applications refer to areas in which virtual knockout has shown clear conceptual and methodological promise, but where evidence remains more heterogeneous across tumor types, datasets, and validation systems.

Therapy resistance-associated network rewiring and antitumor combination prioritization

Because resistance often arises from the interplay of genetic and non-genetic mechanisms, accompanied by the remodeling of molecular pathways and cellular states, relying solely on single-gene functional validation or cross-sectional measurements at a single time point makes it difficult to identify resistance targets within feasible cost and time constraints.⁷⁸ In EGFR-mutant non-small cell lung cancer (NSCLC), mutations within the driver axis itself (e.g., EGFR T790M) or through bypass alterations such as MET amplification both can lead to acquired resistance.^{79,80} Virtual knockout can introduce loss-of-function perturbations to EGFR, MET, and their downstream nodes within stratified models representing pre- and post-resistance states.^{78–80} By comparing changes in pathway modules and drug-response-related readouts, this approach enables prioritization of newly acquired driver dependencies in resistant tumors and identification of rational combination therapy opportunities. GEARS uses single-cell perturbation data to extrapolate transcriptional responses of untested gene combinations and identify non-additive gene-gene interactions.²³ In parallel, gMCS (gmctool) operates at the metabolic network level to enumerate potentially lethal double-knockout or synthetic-lethal combinations and integrates RNA-seq data for sample-

specific filtering, yielding candidate combinations supported by experimental evidence and external datasets in contexts such as multiple myeloma.²⁴

Metastasis-associated regulator prioritization

Compared with driver-gene prioritization and immune-evasion analysis, metastasis-associated virtual knockout remains a more nascent and hypothesis-generating application area. A key aspect of predicting metastasis is to establish a computable link between virtual knockout-induced transcriptional changes and phenotypes such as invasion, colonization, and organ tropism. Figure 4 summarizes an emerging workflow. Using virtual knockout for metastasis-related prioritization requires linking predicted perturbation-induced transcriptional shifts to metastasis-associated phenotypes, intravasation, survival in circulation, colonization, and organotropism.⁸¹ For example, scTenifoldKnk has been used to predict functional roles of genes across multiple tissues, demonstrating that virtual knockout can recover known gene functions and disease phenotypes, which conceptually extends to metastasis-associated regulators in cancer cells or stromal compartments.²² In single-cell studies that define high-metastatic potential gene groups or clonal programs, simulated knockout of transcription factors or chromatin regulators upstream of those programs can indicate whether they are plausible drivers whose inhibition might collapse metastatic traits.^{81,82} This virtual sensitivity analysis may help prioritize putative functional regulators over likely passengers in metastasis-linked networks.

Antitumor drug development and translational filtering

Virtual knockout has increasing relevance for antitumor drug development because it can connect perturbation-derived biological dependencies with measurable pharmacological and translational endpoints.³⁰ At the target discovery stage, essential genes, key enzymes, or metabolic bottleneck reactions identified through virtual knockout can be systematically linked to drug-target resources to generate druggable candidates and further reduce the experimental search space. For example, in an ovarian cancer genome-scale metabolic model (GSMM), virtual knockout prioritized metabolic targets that

Virtual Knockout Framework for Metastasis Prediction & Validation

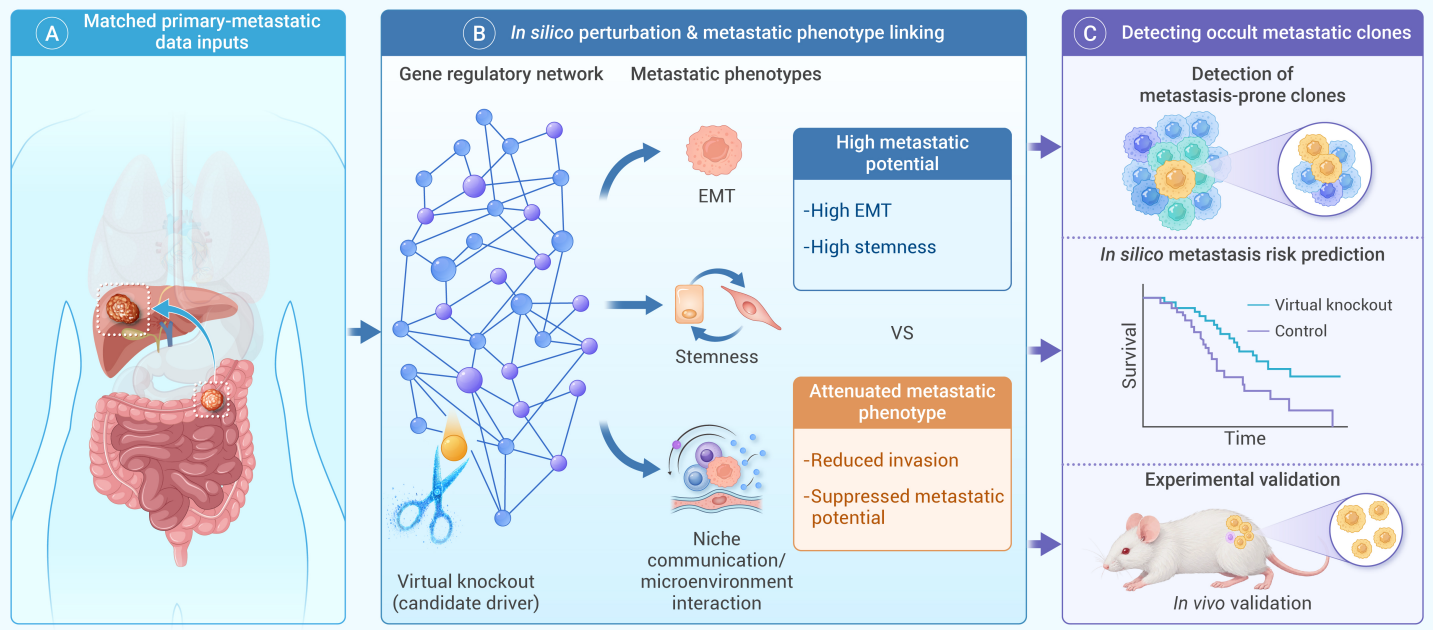


Figure 4. Schematic overview of a virtual knockout-based framework for linking *in silico* perturbation to metastatic phenotypes and downstream validation (A) Matched primary-metastatic data inputs. (B) *In silico* perturbation and metastatic phenotype linking. (C) Detecting occult metastatic clones.

were subsequently mapped to DrugBank to identify candidate compounds, and experimental validation demonstrated that oseltamivir could suppress tumor spheroid formation, illustrating a translational workflow from virtual knockout to drug repurposing and functional validation.³¹ Recent work in drug-resistant breast cancer further extends this logic from drug repurposing to therapeutic resensitization. Lim et al. performed genome-scale single-gene metabolic knockout simulation in doxorubicin- and paclitaxel-resistant breast cancer cells and used clustering analysis to identify metabolic perturbations that could shift resistant cells toward a drug-sensitive state. This framework nominated drug-sensitization targets, and selected predictions were experimentally validated, supporting an L3-level example of metabolic virtual knockout for resistance-oriented drug development.⁵⁷

FUTURE OPPORTUNITIES FOR VIRTUAL KNOCKOUT IN PRECISION ONCOLOGY

Virtual knockout may contribute to therapeutic prioritization in stratified or patient-specific settings. Interpatient and intratumoral heterogeneity can lead to substantial differences in treatment sensitivity across molecular subtypes, cellular states, and ecological niches, making it difficult for averaged cohort-level signals to support truly precise therapeutic decisions.^{83,84} By integrating patient-derived transcriptomic, metabolic, or single-cell profiles with perturbation-informed computational models, virtual knockout can be used to compare candidate vulnerabilities across individual tumors, molecular subtypes, or defined cellular populations, thereby providing a computational basis for prioritizing personalized therapeutic options.⁸⁵ A representative example is personalized genome-scale metabolic modeling in hepatocellular carcinoma, where patient-specific tumor models were used to identify shared and individualized metabolic vulnerabilities, while normal-tissue models were incorporated to evaluate potential toxicity. This illustrates how virtual knockout-related modeling can support individualized vulnerability mapping and therapeutic prioritization before downstream experimental testing.

Individualized drug-response prioritization can be viewed as an extension of patient-specific vulnerability mapping. In this context, virtual knockout has potential value for individualized drug-response prediction and response stratification. By simulating loss-of-function perturbations within patient-specific molecular backgrounds, these models may help estimate the likely sensitivity of a given tumor to targeted agents, metabolic interventions, or

rational combination strategies, while narrowing the range of candidates that require downstream experimental evaluation. When coupled with biologically relevant validation systems, such as patient-derived organoids (PDOs), patient-derived xenografts (PDXs), or other patient-derived models, virtual knockout predictions may be linked to measurable readouts including proliferation, pathway activity, biomarker changes, and drug response, thereby supporting a more translationally meaningful workflow from computational inference to experimental prioritization.⁸⁶

Despite these opportunities, patient-specific virtual knockout faces important translational barriers. Real-time single-cell or spatial profiling from clinical samples is not yet routinely available, and technical constraints such as sample processing, spatial resolution, sequencing depth, and data standardization may limit immediate clinical implementation. In addition, batch effects and biological discrepancies between preclinical models, patient-derived systems, and primary tumors may reduce prediction transferability.⁸⁷⁻⁸⁹ Besides, the regulatory status of computational prediction tools for therapeutic decision-making remains incompletely defined. Its future use will depend on the availability of longitudinal tumor sampling, standardized single-cell or spatial profiling pipelines, rapid model updating, and prospective validation in patient-derived or clinically annotated treatment cohorts.

DISCUSSION

Distinguishing causal-estimation frameworks from virtual knockout prediction

Although most virtual knockout approaches reviewed here generate model-derived post-perturbation predictions, a smaller group of methods explicitly aim to estimate causal effects or interventional relationships. These methods should be distinguished from conventional virtual knockout models. Virtual knockout typically asks how a biological system is predicted to change after a specified loss-of-function perturbation within a model,^{11,22} whereas causal-estimation frameworks attempt to identify intervention effects under explicit assumptions about confounding, transport, intervention definition, or distributional invariance.^{10,90}

Several recent tools illustrate this distinction. CINEMA-OT uses optimal transport to infer single-cell perturbation effects by matching perturbed and control cell states, thereby aiming to estimate causal effects of experimental perturbations rather than simply predict post-perturbation expression profiles.¹² scCausalVI incorporates causal representation learning into single-

cell perturbation analysis to disentangle perturbation-associated responses from nuisance variation¹⁰. More recently, KOCMI introduced a knockoff operation as a virtual intervention and used distributional invariance before and after this operation to infer interventional direct causality.⁹⁰

Therefore, causal-estimation-oriented tools should be evaluated separately from ordinary virtual knockout predictions: model-internal causal estimates without independent validation should be considered low-level evidence, whereas stronger claims require independent perturbation data, controlled experimental validation, or patient-relevant validation consistent with the L1-L4 framework.

Current methodological and translational challenges

A primary limitation lies in data sparsity and noise inherent to tumor datasets. Single-cell measurements often exhibit dropout effects, limited sample sizes, and batch variability, allowing multiple models to explain observational data while yielding divergent knockout predictions.⁹¹⁻⁹³ Imputation, denoising, dimensionality reduction, and probabilistic latent-variable models, such as MAGIC, scImpute, SAVER, DCA, or scVI-like approaches, may partially reduce dropout and technical noise.⁹⁴⁻⁹⁷ However, excessive imputation or oversmoothing may introduce artificial correlations and obscure true low-expression states.^{98,99} Therefore, future virtual knockout tools should incorporate uncertainty quantification rather than reporting only single-point perturbation effects. Potential strategies include ensemble modeling, Bayesian neural networks, variational Bayesian approaches, and conformal prediction, which can provide calibrated uncertainty estimates or prediction intervals around model-derived perturbation effects.^{90,100}

Tumor heterogeneity presents a second challenge that simultaneously defines the value and risk of virtual knockout.²⁵ Stratified modeling across cell types, molecular subtypes, or patient-specific states can reveal context-dependent vulnerabilities that are invisible in bulk analyses.¹⁰¹ However, rare subpopulations may introduce statistical instability because limited cell numbers reduce statistical power and increase sensitivity to sampling variation,¹⁰² whereas modeling heterogeneous tumors as a single mixed cohort may average out cell-state- or subtype-specific effects, mask opposing biological signals, and distort target prioritization.¹⁰³ To mitigate this issue, virtual knockout studies may benefit from subpopulation-specific modeling, virtual patient cohorts, probabilistic or ensemble simulations, network-level robustness analysis, and iterative experimental validation in heterogeneous cell mixtures, organoids, or patient-derived systems.^{97,104} Future frameworks will likely require hierarchical modeling strategies that explicitly encode lineage, differentiation state, clonal structure, and ecological niche context to balance resolution with robustness.^{105,106}

A third limitation arises from cross-platform comparability. Virtual knockout predictions are frequently validated through external resources such as CRISPR screening datasets, Perturb-seq profiles, drug-response signatures, or pathway knowledge bases. Differences in preprocessing pipelines, annotation schemes, and output metrics can generate apparent inconsistencies that reflect technical variability rather than biological disagreement.¹⁰⁷ Establishing standardized benchmarking datasets, harmonized annotation frameworks, and comparable reporting metrics will therefore be critical for enabling reproducible cross-study evaluation.

Common failure modes of virtual knockout approaches

While virtual knockout methods have enabled valuable insights in single-cell and cancer research, their application is accompanied by several fundamental limitations that affect reliability, generalizability, and biological interpretability. Below we summarize key categories of such limitations, each supported by empirical or review evidence:

1) Limited generalizability and out-of-distribution performance

Virtual knockout predictions often fail to generalize across distinct cellular contexts or perturbation regimes, particularly when models are trained on one dataset and applied to another with different biology or technical noise. A recent benchmark of 27 single-cell perturbation prediction methods revealed consistent performance degradation under out-of-distribution (OOD) conditions, indicating that many models capture dataset-specific patterns rather than robust biological responses.¹⁰⁸ More recently, Systema further showed that apparent performance in genetic perturbation-response prediction may

be driven by systematic variation between perturbed and control cells rather than true perturbation-specific biology, highlighting the risk of confounding-driven false prioritization.⁵⁹ This limitation reflects the broader challenge of model transferability in high-dimensional single-cell settings.

2) Sensitivity to data sparsity, noise, and variability

Single-cell omics data are inherently sparse and noisy, with frequent dropout events and cell-to-cell variability. If not properly modeled, such noise can lead to unstable virtual knockout outputs or false prioritizations. Systematic analysis of single-cell differential expression has shown that improper variance modeling leads to extensive false positives, suggesting that downstream perturbation predictions are similarly at risk when analytical assumptions are violated.¹⁰⁹ This data-driven limitation is compounded in rare cell populations with limited representation.

3) Model structural assumptions and interpretability constraints

Many virtual knockout models depend on intermediate representations (e.g., inferred gene regulatory networks, latent spaces, or metabolic objective functions) that are themselves uncertain. Errors or bias in these intermediate structures can propagate, yielding misleading perturbation effects. Reviews of perturbation modelling emphasize challenges in benchmarking, interpretability, and interoperability: model outputs often lack clear mechanistic linkage to biology, making it difficult to distinguish true failures from artifacts of model design.¹¹⁰ For example, generative latent models can produce plausible apparent knockout effects that reflect model inductive bias rather than genuine biological causation.

4) Incomplete biological priors and knowledge bases

Methods that rely on external priors—such as pathway maps, ligand–receptor pairs, or curated interaction databases—are limited by the completeness and accuracy of those resources. Communication-informed perturbation models, for instance, can misprioritize interactions if the underlying LR database lacks key species or contexts. This form of incomplete knowledge bias has been highlighted in reviews of signaling inference tools, which caution that gaps in prior networks can generate false predictions of pathway impact, particularly in less-studied tissues or disease states.¹¹¹

Collectively, these limitations indicate that virtual knockout should be applied with careful consideration of data quality, model assumptions, and validation strategy. Table 1 consolidates common failure modes across model classes and suggests safeguards such as cross-validation, bootstrap/resampling, multiple sgRNAs virtual knockout perturbations, orthogonal experimental validation, and careful dataset selection to mitigate these risks.

Emerging directions toward robust and clinically relevant virtual knockout

Despite these challenges, the future trajectory of virtual knockout in cancer research remains highly promising. Recent advances in AI-driven digital cell modeling further suggest that computational simulation of cellular states, gene perturbations, and disease-relevant responses may provide a broader technical foundation for next-generation virtual knockout frameworks.^{112,113} Advances in perturbation sequencing, combinatorial perturbation design, and experimental systems that more closely recapitulate tumor microenvironmental complexity will provide richer training data and external benchmarks, thereby improving cross-cohort generalization.^{108,114} Methodologically, the field is evolving from static model-derived counterfactual prediction toward dynamic state-transition modeling, transfer learning, and uncertainty-aware inference. Such developments may enable models not only to rank candidate targets but also to provide confidence intervals and mechanistically interpretable predictions that are experimentally falsifiable. Integration of multi-omics and spatial omics data further expands the scope of virtual knockout by constraining regulatory directionality and capturing neighborhood-level interactions, niche remodeling, and immune-stromal crosstalk.¹¹⁵⁻¹¹⁷ These advances are particularly relevant for studying immune escape, therapy resistance, and metastatic evolution, where cellular context and microenvironmental signals shape phenotypic outcomes.^{71,84,118} In parallel, patient-specific modeling and stratified virtual knockout represent a key direction for clinical translation. Predicting therapeutic targets or combination strategies within patient-derived cellular contexts—and validating them through organoids, xenograft models, or stratified clinical cohorts—may bridge the gap between computational inference and precision treatment.⁵²

Outlook: from exploratory inference to a reproducible translational framework

Ultimately, sustainable progress in this field will depend on the development of community-wide benchmarks, minimal reporting standards, and reproducible computational toolchains. Establishing unified evaluation criteria will allow findings from different studies to be compared, reproduced, and iteratively refined, facilitating the transition of virtual knockout from an exploratory computational strategy to a reliable translational framework in cancer research.

REFERENCES

- Luo Q. and Smith D.P. (2025). Global cancer burden: Progress, projections, and challenges. *Lancet* **406**:1536–1537. DOI:10.1016/S0140-6736(25)01570-3
- Hanahan D. and Weinberg R.A. (2011). Hallmarks of cancer: The next generation. *Cell* **144**:646–674. DOI:10.1016/j.cell.2011.02.013
- Gaillard H., Garcia-Muse T. and Aguilera A. (2015). Replication stress and cancer. *Nat. Rev. Cancer* **15**:276–289. DOI:10.1038/nrc3916
- Hanahan D. (2022). Hallmarks of cancer: New dimensions. *Cancer Discov.* **12**:31–46. DOI:10.1158/2159-8290.CD-21-1059
- Togar T., Desai S., Mishra R., et al. (2020). Identifying cancer driver genes from functional genomics screens. *Swiss Med. Wkly.* **150**:w20195. DOI:10.4414/smw.2020.20195
- Pinto B., Henriques A.C., Silva P.M.A., et al. (2020). Three-Dimensional spheroids as in vitro preclinical models for cancer research. *Pharmaceutics* **12**. DOI:10.3390/pharmaceutics12121186.
- Liu Y., Wu W., Cai C., et al. (2023). Patient-derived xenograft models in cancer therapy: Technologies and applications. *Signal Transduct. Target Ther.* **8**:160. DOI:10.1038/s41392-023-01419-2
- DiMasi J.A., Feldman L., Seckler A., et al. (2010). Trends in risks associated with new drug development: Success rates for investigational drugs. *Clin. Pharmacol. Ther.* **87**:272–277. DOI:10.1038/clpt.2009.295
- The ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium. (2020). Pan-cancer analysis of whole genomes. *Nature* **578**:82–93. DOI:10.1038/s41586-020-1969-6
- An S., Cho J.W., Cao K., et al. (2025). scCausalVI disentangles single-cell perturbation responses with causality-aware generative model. *Cell Syst.* **16**:101443. DOI:10.1016/j.cels.2025.101443
- Kamimoto K., Stringa B., Hoffmann C.M., et al. (2023). Dissecting cell identity via network inference and in silico gene perturbation. *Nature* **614**:742–751. DOI:10.1038/s41586-022-05688-9
- Dong M., Wang B., Wei J., et al. (2023). Causal identification of single-cell experimental perturbation effects with CINEMA-OT. *Nat. Methods* **20**:1769–1779. DOI:10.1038/s41592-023-02040-5
- Weng J., Ju F., Lyu Z., et al. (2025). Single-cell insights into tumor microenvironment heterogeneity and plasticity: Transforming precision therapy in gastrointestinal cancers. *J. Exp. Clin. Cancer Res.* **44**:314. DOI:10.1186/s13046-025-03567-5
- Lotfollahi M., Wolf F.A. and Theis F.J. (2019). scGen predicts single-cell perturbation responses. *Nat. Methods* **16**:715–721. DOI:10.1038/s41592-019-0494-8
- Lotfollahi M., Klimovskaia Susmelj A., De Donno C., et al. (2023). Predicting cellular responses to complex perturbations in high-throughput screens. *Mol. Syst. Biol.* **19**:e11517. DOI:10.15252/msb.202211517
- Ma C., Zhang H., Rao Y., et al. (2025). AI-driven virtual cell models in preclinical research: Technical pathways, validation mechanisms, and clinical translation potential. *npj Digit. Med.* **9**:25. DOI:10.1038/s41746-025-02198-6
- Theodoris C.V., Xiao L., Chopra A., et al. (2023). Transfer learning enables predictions in network biology. *Nature* **618**:616–624. DOI:10.1038/s41586-023-06139-9
- Cui H., Wang C., Maan H., et al. (2024). scGPT: Toward building a foundation model for single-cell multi-omics using generative AI. *Nat. Methods* **21**:1470–1480. DOI:10.1038/s41592-024-02201-0
- Hao M., Gong J., Zeng X., et al. (2024). Large-scale foundation model on single-cell transcriptomics. *Nat. Methods* **21**:1481–1491. DOI:10.1038/s41592-024-02305-7
- Zhu O. and Li J. (2026). Scouter predicts transcriptional responses to genetic perturbations with large language model embeddings. *Nat. Comput. Sci.* **6**:21–28. DOI:10.1038/s43588-025-00912-8
- Szappanos B., Kovacs K., Szamecz B., et al. (2011). An integrated approach to characterize genetic interaction networks in yeast metabolism. *Nat. Genet.* **43**:656–662. DOI:10.1038/ng.846
- Osorio D., Zhong Y., Li G., et al. (2022). scTenifoldKnk: An efficient virtual knockout tool for gene function predictions via single-cell gene regulatory network perturbation. *Patterns (N Y)* **3**:100434. DOI:10.1016/j.patter.2022.100434
- Roohani Y., Huang K. and Leskovec J. (2024). Predicting transcriptional outcomes of novel multigene perturbations with GEARS. *Nat. Biotechnol.* **42**:927–935. DOI:10.1038/s41587-023-01905-6
- Valcarcel L.V., San Jose-Eneriz E., Ordonez R., et al. (2024). An automated network-based tool to search for metabolic vulnerabilities in cancer. *Nat. Commun.* **15**:8685. DOI:10.1038/s41467-024-52725-4
- Marusyk A., Almendro V. and Polyak K. (2012). Intra-tumour heterogeneity: A looking glass for cancer. *Nat. Rev. Cancer* **12**:323–334. DOI:10.1038/nrc3261
- Dagogo-Jack I. and Shaw A.T. (2018). Tumour heterogeneity and resistance to cancer therapies. *Nat. Rev. Clin. Oncol.* **15**:81–94. DOI:10.1038/nrclinonc.2017.166
- Apaolaza I., San Jose-Eneriz E., Valcarcel L.V., et al. (2022). A network-based approach to integrate nutrient microenvironment in the prediction of synthetic lethality in cancer metabolism. *PLoS Comput. Biol.* **18**:e1009395. DOI:10.1371/journal.pcbi.1009395
- O'Brien Edward J., Monk J.M. and Palsson B.O. (2015). Using genome-scale models to predict biological capabilities. *Cell* **161**:971–987. DOI:10.1016/j.cell.2015.05.019
- Abbasian M.H., Sobhani N., Sisakht M.M., et al. (2025). Patient-derived organoids: A game-changer in personalized cancer medicine. *Stem Cell Rev. Rep.* **21**:211–225. DOI:10.1007/s12015-024-10805-4
- Arora G., Banerjee M., Langthasa J., et al. (2023). Targeting metabolic fluxes reverts metastatic transitions in ovarian cancer. *iScience* **26**:108081. DOI:10.1016/j.isci.2023.108081
- Heirendt L., Arreckx S., Pfau T., et al. (2019). Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. *Nat. Protoc.* **14**:639–702. DOI:10.1038/s41596-018-0098-2
- TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. (2024). *BMJ* **385**:q902. DOI:10.1136/bmj.q902.
- Vogelstein B., Papadopoulos N., Velculescu V.E., et al. (2013). Cancer genome landscapes. *Science* **339**:1546–1558. DOI:10.1126/science.1235122
- Collins G.S., Dhiman P., Ma J., et al. (2024). Evaluation of clinical prediction models (part 1): From development to external validation. *BMJ* **384**:e074819. DOI:10.1136/bmj-2023-074819
- Riley R.D., Archer L., Snell K.I.E., et al. (2024). Evaluation of clinical prediction models (part 2): How to undertake an external validation study. *BMJ* **384**:e074820. DOI:10.1136/bmj-2023-074820
- Riley R.D., Snell K.I.E., Archer L., et al. (2024). Evaluation of clinical prediction models (part 3): Calculating the sample size required for an external validation study. *BMJ* **384**:e074821. DOI:10.1136/bmj-2023-074821
- Patterson E.A. and Whelan M.P. (2017). A framework to establish credibility of computational models in biology. *Prog. Biophys. Mol. Biol.* **129**:13–19. DOI:10.1016/j.pbiomolbio.2016.08.007
- Tatka L.T., Smith L.P., Hellerstein J.L., et al. (2023). Adapting modeling and simulation credibility standards to computational systems biology. *J. Transl. Med.* **21**:501. DOI:10.1186/s12967-023-04290-5
- Jafari M., Guan Y., Wedge D.C., et al. (2021). Re-evaluating experimental validation in the Big Data Era: A conceptual argument. *Genome Biol.* **22**:71. DOI:10.1186/s13059-021-02292-4
- Dancic G.M. and Vlahopoulos S.A. (2025). Editorial for the special issue: Bioinformatics and computational biology for cancer prediction and prognosis. *Genes (Basel)* **16**. DOI:10.3390/genes16020167
- Moons K.G.M., Wolff R.F., Riley R.D., et al. (2019). PROBAST: A tool to assess risk of bias and applicability of prediction model studies: Explanation and elaboration. *Ann. Intern. Med.* **170**:W1–W33. DOI:10.7326/M18-1377
- Gross F. and MacLeod M. (2017). Prospects and problems for standardizing model validation in systems biology. *Prog. Biophys. Mol. Biol.* **129**:3–12. DOI:10.1016/j.pbiomolbio.2017.01.003
- Steyerberg E.W., Harrell F.E.Jr., Borsboom G.J.J.M., et al. (2001). Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. *J. Clin. Epidemiol.* **54**:774–781. DOI:10.1016/s0895-4356(01)00341-9
- Meinshausen N. and Bühlmann P. (2010). Stability selection. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **72**:417–473. DOI:10.1111/j.1467-9868.2010.00740.x
- Golland P. and Fischl B. (2003). Permutation tests for classification: Towards statistical significance in image-based studies. *Inf. Process. Med. Imaging* **18**:330–341. DOI:10.1007/978-3-540-45087-0_28
- Rosenblatt M., Tejaviyula L., Jiang R., et al. (2024). Data leakage inflates prediction performance in connectome-based machine learning models. *Nat. Commun.* **15**:1829. DOI:10.1038/s41467-024-46150-w
- Liu Z., Xue Z., Song H., et al. (2026). Molecular mechanisms linking cadmium chloride exposure to ankylosing spondylitis: An integrative network-based study. *Naunyn-Schmiedeberg's Arch. Pharmacol.* DOI:10.1007/s00210-026-05394-7
- Qi L.S., Larson M.H., Gilbert L.A., et al. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**:1173–1183. DOI:10.1016/j.cell.2013.02.022
- Bock C., Datlinger P., Chardon F., et al. (2022). High-content CRISPR screening. *Nat. Rev. Methods Primers* **2**:8. DOI:10.1038/s43586-022-00098-7
- Joung J., Konermann S., Gootenberg J.S., et al. (2017). Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. *Nat. Protoc.* **12**:828–863. DOI:10.1038/nprot.2017.016
- Wang Y., Zhai Y., Zhang M., et al. (2024). Escaping from CRISPR-Cas-mediated knockout: The facts, mechanisms, and applications. *Cell Mol. Biol. Lett.* **29**:48. DOI:10.1186/s11658-024-00565-x
- Tong L., Cui W., Zhang B., et al. (2024). Patient-derived organoids in precision cancer medicine. *Med* **5**:1351–1377. DOI:10.1016/j.medj.2024.08.010

53. Viceconti M., Pappalardo F., Rodriguez B., et al. (2021). In silico trials: Verification, validation and uncertainty quantification of predictive models used in the regulatory evaluation of biomedical products. *Methods* **185**:120–127. DOI:10.1016/j.ymeth.2020.01.011
54. Wang Y., Li J., Yan Y., et al. (2026). A metabolic-immune subtype of breast cancer defined by G6PD and SHMT2: From single-cell dissection to dual-targeted therapy. *J. Steroid Biochem. Mol. Biol.* **260**:106983. DOI:10.1016/j.jsbmb.2026.106983
55. Agren R., Mardinoglu A., Asplund A., et al. (2014). Identification of anticancer drugs for hepatocellular carcinoma through personalized genome-scale metabolic modeling. *Mol. Syst. Biol.* **10**:721. DOI:10.1002/msb.145122
56. Jolasun Y., Song K., Zheng Y., et al. (2025). SIDISH integrates single-cell and bulk transcriptomics to identify high-risk cells and guide precision therapeutics through in silico perturbation. *Nat. Commun.* **16**:11271. DOI:10.1038/s41467-025-66162-4
57. Lim J., Jung H. D., Park S. Y., et al. (2025). Genome-scale knockout simulation and clustering analysis of drug-resistant breast cancer cells reveal drug sensitization targets. *Proc. Natl. Acad. Sci. USA* **122**:e2425384122. DOI:10.1073/pnas.2425384122
58. Zhan X., Hu H., Liu Y., et al. (2025). TRIM28 drives immune evasion via PARP1 SUMOylation and NAD(+) depletion in clear cell renal cell carcinoma. *J. Immunother. Cancer* **13**. DOI:10.1136/jitc-2025-013025
59. Vinas T.R., Wiatrak M., Piran Z., et al. (2025). Systema: A framework for evaluating genetic perturbation response prediction beyond systematic variation. *Nat. Biotechnol.* DOI:10.1038/s41587-025-02777-8
60. Yang Y., Li G., Zhong Y., et al. (2023). Gene knockout inference with variational graph autoencoder learning single-cell gene regulatory networks. *Nucleic Acids Res.* **51**:6578–6592. DOI:10.1093/nar/gkad450
61. Little D.R., Gerner-Mauro K.N., Flodby P., et al. (2019). Transcriptional control of lung alveolar type 1 cell development and maintenance by NK homeobox 2-1. *Proc. Natl. Acad. Sci. USA* **116**:20545–20555. DOI:10.1073/pnas.1906663116
62. Shiao S. L., Gouin K.H.III, Ing N., et al. (2024). Single-cell and spatial profiling identify three response trajectories to pembrolizumab and radiation therapy in triple negative breast cancer. *Cancer Cell* **42**:70–84.e78. DOI:10.1016/j.ccell.2023.12.012
63. Paul F., Arkin Y., Giladi A., et al. (2016). Transcriptional heterogeneity and lineage commitment in myeloid progenitors. *Cell* **164**:325. DOI:10.1016/j.cell.2015.12.046
64. Orkin S.H. and Zon L.I. (2008). Hematopoiesis: An evolving paradigm for stem cell biology. *Cell* **132**:631–644. DOI:10.1016/j.cell.2008.01.025
65. Yang K., Halima A. and Chan T.A. (2023). Antigen presentation in cancer – mechanisms and clinical implications for immunotherapy. *Nat. Rev. Clin. Oncol.* **20**:604–623. DOI:10.1038/s41571-023-00789-4
66. Garrido F., Ruiz-Cabello F. and Aptsiauri N. (2017). Rejection versus escape: The tumor MHC dilemma. *Cancer Immunol. Immunother.* **66**:259–271. DOI:10.1007/s00262-016-1947-x
67. Maggs L., Sadagopan A., Moghaddam A.S., et al. (2021). HLA class I antigen processing machinery defects in antitumor immunity and immunotherapy. *Trends in Cancer* **7**:1089–1101. DOI:10.1016/j.trecan.2021.07.006
68. Kobayashi K.S. and van den Elsen P.J. (2012). NLRC5: A key regulator of MHC class I-dependent immune responses. *Nat. Rev. Immunol.* **12**:813–820. DOI:10.1038/nri3339
69. Kok V.C., Wang C.C.N., Liao S.H., et al. (2022). Cross-platform in-silico analyses exploring tumor immune microenvironment with prognostic value in triple-negative breast cancer. *Breast Cancer (Dove Med Press)* **14**:85–99. DOI:10.2147/BCTT.S359346
70. Jin S., Guerrero-Juarez C.F., Zhang L., et al. (2021). Inference and analysis of cell-cell communication using CellChat. *Nat. Commun.* **12**:1088. DOI:10.1038/s41467-021-21246-9
71. Aliasis K., Christofides A., Shah R., et al. (2025). The tumor microenvironment's role in the response to immune checkpoint blockade. *Nat. Cancer* **6**:924–937. DOI:10.1038/s43018-025-00986-3
72. Allard B., Allard D., Buisseret L., et al. (2020). Publisher correction: The adenosine pathway in immuno-oncology. *Nat. Rev. Clin. Oncol.* **17**:650. DOI:10.1038/s41571-020-0415-x
73. Mariathasan S., Turley S.J., Nickles D., et al. (2018). TGFbeta attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* **554**:544–548. DOI:10.1038/nature25501
74. Ghahremanifard P., Chanda A., Bonni S., et al. (2020). TGF-beta mediated immune evasion in cancer-spotlight on cancer-associated fibroblasts. *Cancers (Basel)* **12**:3650. DOI:10.3390/cancers12123650
75. Liu X., Ding Q., Zhang H., et al. (2025). The CD39-CD73-adenosine axis: Master regulator of immune evasion and therapeutic target in pancreatic ductal adenocarcinoma. *Biochim. Biophys. Acta Rev. Cancer* **1880**:189443. DOI:10.1016/j.bbcan.2025.189443
76. Zhang B. (2010). CD73: A novel target for cancer immunotherapy. *Cancer Res.* **70**:6407–6411. DOI:10.1158/0008-5472.CAN-10-1544
77. Chen S., Wainwright D. A., Wu J. D., et al. (2019). CD73: An emerging checkpoint for cancer immunotherapy. *Immunotherapy* **11**:983–997. DOI:10.2217/imt-2018-0200
78. Salgia R. and Kulkarni P. (2018). The genetic/non-genetic duality of drug 'Resistance' in cancer. *Trends Cancer* **4**:110–118. DOI:10.1016/j.trecan.2018.01.001
79. Kobayashi S., Boggon T.J., Dayaram T., et al. (2005). EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. *N. Engl. J. Med.* **352**:786–792. DOI:10.1056/NEJMoa044238
80. Engelman J.A., Zejnullahu K., Mitsudomi T., et al. (2007). MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science* **316**:1039–1043. DOI:10.1126/science.1141478
81. Asghari M., Abazari M.F., Bokharaei H., et al. (2018). Key genes and regulatory networks involved in the initiation, progression and invasion of colorectal cancer. *Future Sci. OA* **4**:FSO278. DOI:10.4155/fsoa-2017-0108
82. Bravo Gonzalez-Blas C., De Winter S., Hulselmans G., et al. (2023). SCENIC+: single-cell multimic inference of enhancers and gene regulatory networks. *Nat. Methods* **20**:1355–1367. DOI:10.1038/s41592-023-01938-4
83. Tomasik B., Garbicz F., Braun M., et al. (2024). Heterogeneity in precision oncology. *Camb. Prism. Precis. Med.* **2**:e2. DOI:10.1017/pcm.2023.23
84. Roerden M. and Spranger S. (2025). Cancer immune evasion, immunoediting and intratumour heterogeneity. *Nat. Rev. Immunol.* **25**:353–369. DOI:10.1038/s41577-024-01111-8
85. Jie J., Wang Q., Chen Z., et al. (2026). Multi-omics analysis and preliminary experimental validation of acetyl tributyl citrate (ATBC) promoting bladder cancer progression via the AKR1B1/EMT axis. *Environ. Pollut.* **397**:127970. DOI:10.1016/j.envpol.2026.127970
86. Li R., Liu X., Huo C., et al. (2026). An endothelial-centered regulatory framework reveals context-dependent roles of MYLK in lung adenocarcinoma. *Front. Immunol.* **17**:1719296. DOI:10.3389/fimmu.2026.1719296
87. Frangieh C.J., Fan J.L., Melms J.C., et al. (2026). Single-cell and spatial profiling in cancer biology and clinical oncology. *Nat. Cancer* **7**:597–607. DOI:10.1038/s43018-026-01142-1
88. Wensink G.E., Elias S.G., Mullenders J., et al. (2021). Patient-derived organoids as a predictive biomarker for treatment response in cancer patients. *npj Precis. Oncol.* **5**:30. DOI:10.1038/s41698-021-00168-1
89. Thorel L., Perreard M., Florent R., et al. (2024). Patient-derived tumor organoids: A new avenue for preclinical research and precision medicine in oncology. *Exp. Mol. Med.* **56**:1531–1551. DOI:10.1038/s12276-024-01272-5
90. Zhang X. and Chen L. (2025). Quantifying interventional causality by knockoff operation. *Sci. Adv.* **11**:eadu6464. DOI:10.1126/sciadv.adu6464
91. Song F., Chan G.M.A. and Wei Y. (2020). Flexible experimental designs for valid single-cell RNA-sequencing experiments allowing batch effects correction. *Nat. Commun.* **11**:3274. DOI:10.1038/s41467-020-16905-2
92. Haghighi L., Lun A.T.L., Morgan M.D., et al. (2018). Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat. Biotechnol.* **36**:421–427. DOI:10.1038/nbt.4091
93. Bouland G.A., Mahfouz A. and Reinders M.J.T. (2023). Consequences and opportunities arising due to sparser single-cell RNA-seq datasets. *Genome Biol.* **24**:86. DOI:10.1186/s13059-023-02933-w
94. van Dijk D., Sharma R., Nainys J., et al. (2018). Recovering gene interactions from single-cell data using data diffusion. *Cell* **174**:716–729.e727. DOI:10.1016/j.cell.2018.05.061
95. Li W.V. and Li J.J. (2018). An accurate and robust imputation method scImpute for single-cell RNA-seq data. *Nat. Commun.* **9**:997. DOI:10.1038/s41467-018-03405-7
96. Huang M., Wang J., Torre E., et al. (2018). SAVER: Gene expression recovery for single-cell RNA sequencing. *Nat. Methods* **15**:539–542. DOI:10.1038/s41592-018-0033-z
97. Lopez R., Regier J., Cole M.B., et al. (2018). Deep generative modeling for single-cell transcriptomics. *Nat. Methods* **15**:1053–1058. DOI:10.1038/s41592-018-0229-2
98. Hou W., Ji Z., Ji H., et al. (2020). A systematic evaluation of single-cell RNA-sequencing imputation methods. *Genome Biol.* **21**:218. DOI:10.1186/s13059-020-02132-x
99. Zhang R., Atwal G.S. and Lim W.K. (2021). Noise regularization removes correlation artifacts in single-cell RNA-seq data preprocessing. *Patterns (N Y)* **2**:100211. DOI:10.1016/j.patter.2021.100211
100. Ahlmann-Eltze C., Huber W. and Anders S. (2025). Deep-learning-based gene perturbation effect prediction does not yet outperform simple linear baselines. *Nat. Methods* **22**:1657–1661. DOI:10.1038/s41592-025-02772-6
101. Xue Y., Friedl V., Ding H., et al. (2024). Single-cell signatures identify microenvironment factors in tumors associated with patient outcomes. *Cell Rep. Methods* **4**. DOI:10.1016/j.crmeth.2024.100799
102. Schmid K.T., Höllbacher B., Cruceanu C., et al. (2021). scPower accelerates and optimizes the design of multi-sample single cell transcriptomic studies. *Nat. Commun.* **12**:6625. DOI:10.1038/s41467-021-26779-7
103. Hegenbarth J.C., Lezocche G., De Windt L.J., et al. (2022). Perspectives on bulk-tissue RNA sequencing and single-cell RNA sequencing for cardiac transcriptomics. *Front. Mol. Med.* **2**:839338. DOI:10.3389/fmmed.2022.839338
104. Vlachogiannis G., Hedayat S., Vatsiou A., et al. (2018). Patient-derived organoids model treatment response of metastatic gastrointestinal cancers. *Science* **359**:920–926. DOI:10.1126/science.aao2774
105. Wagner D.E. and Klein A.M. (2020). Lineage tracing meets single-cell omics: Opportunities and challenges. *Nat. Rev. Genet.* **21**:410–427. DOI:10.1038/s41576-020-0223-2
106. Binnewies M., Roberts E.W., Kersten K., et al. (2018). Understanding the tumor

- immune microenvironment (TIME) for effective therapy. *Nat. Med.* **24**:541–550. DOI:10.1038/s41591-018-0014-x
107. Hu X., Li H., Chen M., et al. (2025). Reference-informed evaluation of batch correction for single-cell omics data with overcorrection awareness. *Commun. Biol.* **8**:521. DOI:10.1038/s42003-025-07947-7
108. Wei Z., Wang Y., Gao Y., et al. (2025). Benchmarking algorithms for generalizable single-cell perturbation response prediction. *Nat. Methods* **23**:451–464. DOI:10.1038/s41592-025-02980-0
109. Squair J.W., Gautier M., Kathe C., et al. (2021). Confronting false discoveries in single-cell differential expression. *Nat. Commun.* **12**:5692. DOI:10.1038/s41467-021-25960-2
110. Li W., Zhang Z., Xie B., et al. (2024). HiOmics: A cloud-based one-stop platform for the comprehensive analysis of large-scale omics data. *Comput. Struct. Biotechnol. J.* **23**:659–668. DOI:10.1016/j.csbj.2024.01.002
111. Huang Y., Luo J., Zhang Y., et al. (2023). Identification of MKNK1 and TOP3A as ovarian endometriosis risk-associated genes using integrative genomic analyses and functional experiments. *Comput. Struct. Biotechnol. J.* **21**:1510–1522. DOI:10.1016/j.csbj.2023.02.001
112. Yang T., Ma F., Qian H., et al. (2024). AI-driven construction of digital cell model. *Innov. Life* **2**:100102. DOI:10.59717/j.xinn-life.2024.100102
113. Wei R., Wang B., Yan B., et al. (2026). From equations to agents: The artificial intelligence virtual cell reshaping precision oncology. *Innov. Oncol.* **1**:100002. DOI:10.59717/j.xinn-oncol.2026.100002
114. Polak R., Zhang E.T. and Kuo C.J. (2024). Cancer organoids 2.0: Modelling the complexity of the tumour immune microenvironment. *Nat. Rev. Cancer* **24**:523–539. DOI:10.1038/s41568-024-00706-6
115. Badia I.M.P., Wessels L., Muller-Dott S., et al. (2023). Gene regulatory network inference in the era of single-cell multi-omics. *Nat. Rev. Genet.* **24**:739–754. DOI:10.1038/s41576-023-00618-5
116. Schäfer P.S.L., Dimitrov D., Villablanca E.J., et al. (2024). Integrating single-cell multi-omics and prior biological knowledge for a functional characterization of the immune system. *Nat. Immunol.* **25**:405–417. DOI:10.1038/s41590-024-01768-2
117. Walker C. and Angelo M. (2024). Toward clinical applications of spatial-omics in

- cancer research. *Nat. Cancer* **5**:1771–1773. DOI:10.1038/s43018-024-00868-0
118. Lambert A.W., Zhang Y. and Weinberg R.A. (2024). Cell-intrinsic and microenvironmental determinants of metastatic colonization. *Nat. Cell Biol.* **26**:687–697. DOI:10.1038/s41556-024-01409-8

FUNDING AND ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (82460034), China Foundation for Youth Entrepreneurship and Employment (P24032887713) and Shandong Provincial Natural Science Foundation (ZR2024QH296). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

AUTHOR CONTRIBUTIONS

Xiangyu Wang, Min Liu, and Meng Jia contributed to literature collection, manuscript drafting, and revision. Xiangyu Wang and Wei Fang conceived the topic and designed the overall framework of the review. Xiangyu Wang and Xinrong Lv coordinated literature synthesis, table preparation, figure organization, and manuscript revision. Min Liu, Meng Jia, Quan Wang, Xiangwen Li, Qinbiao Zou, Xiuxiang Huang, and Lina Zhang contributed to literature screening, content organization, methodological interpretation, and revision of specific sections. Jie Wang, Lizeng Peng, and Wei Fang provided supervision, critical revision, and intellectual guidance. Wei Fang, Jie Wang, and Lizeng Peng reviewed and refined the manuscript for important intellectual content. All authors contributed to the manuscript and approved the final version.

DECLARATION OF INTERESTS

The authors declare no competing interests.

DATA AND CODE AVAILABILITY

No new data were created or analyzed in this study.