nature communications

Review article

https://doi.org/10.1038/s41467-026-71364-5

Decoding immunotherapy response through computational modeling

Received: 7 February 2025

Accepted: 21 March 2026

Published online: 15 April 2026

Check for updates

Bingrui Li[1,2,6], Ruihan Luo[2,3,4,6], Kexin Huang[2,3,4,6], Jiajia Liu[2], Weiling Zhao[2] & Xiaobo Zhou[2,5]

Immunotherapy has seen success in treating patients with cancer, but variable responses underscore the need for effective patient stratification and therapy planning. Computational tools integrating multi-omics, imaging and machine learning have advanced, yet reliable personalized predictions remain challenging. This review analyzes the field through four converging paradigms: classical machine learning, deep learning, graph and network modeling, and mechanistic systems biology. We examine the evolution from correlational features to representation learning, relational inference, and causal simulation of tumor-immune dynamics, highlighting the shift towards multi-modal fusion and interpretable, clinically deployable models. By providing an integrated review of these computational tools, we hope to bring the community closer to achieving precision immuno-oncology for personalized cancer treatments.

The clinical imperative for predictive biomarkers

Cancer immunotherapy, particularly immune checkpoint inhibitors (ICI), has fundamentally changed the landscape of oncology, offering the prospect of durable, long-term remission for patients with advanced malignancies[1]. Therapies targeting the programmed cell death protein 1 (PD-1)/programmed death-ligand 1 (PD-L1) and cytotoxic T-lymphocyte antigen-4 (CTLA-4) have demonstrated unprecedented success in various cancers, including melanoma[2,3], non-small-cell lung cancer (NSCLC)[4] and renal cell carcinoma[5]. However, the transformative potential of these agents is tempered by the profound heterogeneity in patient response[2,6]. Even within responsive cancer types, a substantial proportion of patients fail to derive clinical benefit, while others may experience severe, and occasionally life-threatening, immune-related adverse events (irAE)[7]. This variability not only exposes non-responding patients to toxicity and substantial financial cost but also underscores an urgent, unmet need for robust predictive biomarkers[8].

The development of such biomarkers is a vital clinical imperative for advancing precision immuno-oncology. An ideal biomarker would accurately stratify patients before treatment, identifying likely responders, candidates for combination strategies, and those better served by alternatives. Achieving these goals often requires computational models that integrate tumor-, host-, and treatment-level factors to help translate biomarker signals into actionable decisions, but the core bottleneck remains in capturing the complex interactions among cancer, immune system and existing therapy, an inherently high-dimensional task well suited to computational tools.

An overview of biological and data-related bottlenecks

Developing reliable predictive models is obstructed by biological complexity and data limitations, which help explain why early biomarkers often proved insufficient (Table 1) and why more sophisticated computational approaches may be necessary.

The primary biological bottleneck is the dynamic and heterogeneous nature of the tumor microenvironment (TME). Intratumor heterogeneity (ITH) creates spatial and genetic diversity, so a single biopsy may not represent the full tumor and can confound spot sampling-based prediction[9]. In addition, the TME is dynamic; a baseline biopsy may miss the subsequent tumor-immune trajectory[10]. Host-

[1]Department of Cancer Biology, University of Texas MD Anderson Cancer Center, Houston, TX, USA. [2]Center for Computational Systems Medicine, School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, TX, USA. [3]West China Biomedical Big Data Centre, West China Hospital, Sichuan University, Chengdu, Sichuan, PR China. [4]Med-X Center for Informatics, Sichuan University, Chengdu, Sichuan, PR China. [5]McGovern Medical School, The University of Texas Health Science Center at Houston, Houston, TX, USA. [6]These authors contributed equally: Bingrui Li, Ruihan Luo, Kexin Huang. e-mail: Xiaobo.Zhou@uth.tmc.edu

Table 1 | Typical performance of predictive models across modalities

Modality/Approach	Performance (AUC)	Cancer Types	Notes
Tumor mutation burden/ neoantigen load	0.58–0.87	Melanoma, non-small cell lung cancer	Performance is inconsistent and context dependent[179].
Transcriptomic signatures (e.g., TIDE)	0.70–0.80	Melanoma, urothelial cancer, non-small cell lung cancer, clear cell renal cell carcinoma	Generalizability can be poor across different cancers[16].
Radiomics (CT/MRI)	0.76–0.90	Multiple cancer types	Performance varies in different cohorts and can be high in specific cohorts[180–182].
Digital pathology (WSI)	0.65–0.88	Pan-cancer	Can predict molecular features (tumor mutation burden, microsatellite instability) from H&E slides[183–185].
Microbiome signatures	0.62–0.75	Melanoma, renal cell carcinoma	Highly experimental and prone to poor generalization due to geographic and dietary factors[186–188].
Multi-modal fusion	>0.80	Pan-cancer	Often shows superior performance but requires complex data integration and large datasets[18,189].

level factors beyond the tumor, such as host germline genetics and gut microbiome[11,12], further modulate immune readiness. Lastly, certain cancers, such as pancreatic and microsatellite-stable colorectal cancer, remain 'cold' and resistant to ICIs, posing a formidable biological barrier that predictive models must account for[13,14].

From a computational perspective, progress in correctly predicting immunotherapy response is hindered by several practical obstacles. Multi-omics cohorts remain scarce and often small, especially for rarer cancer types or novel therapies, thereby increasing the risk of model overfitting. Batch effects and domain shift, in which models trained on data from one institution or sequencing platform, often prevent the generalization of the prediction to new data due to subtle technical variations in data generation and processing. Finally, the field has lacked comprehensive benchmarking datasets and reporting guidelines, despite ongoing efforts[15–17], making it difficult to perform fair, head-to-head comparisons of different predictive models and hindering reproducible research.

A paradigm-based approach to computational modeling and multimodal data integration

Most predictive studies in immuno-oncology remain siloed by data type/drivers: genomics for mutations, transcriptomics for gene signatures, and imaging for radiomic features, so that each data type captures only part of this multi-modal system. Because these drivers are interconnected, single-modality models often miss key dependencies.

Multi-modal integration can improve robustness by learning complementary signals across data types, including genomics, transcriptomics, microbiome profiles, and medical imaging (computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET))[18–21]. Common strategies for multi-modal integration include early fusion (which concatenates feature vectors from each modality and trains a standard classifier, such as XGBoost[22] or an MLP[23]), late fusion (which combines outputs from modality-specific models via averaging, voting, or a meta-learner), and joint representation learning (which uses deep architectures, such as autoencoders, variational autoencoders (VAE)[24] and graph-based models etc., to embed modalities into a shared low-dimensional space that supports prediction and generalization) (Fig. 1A). The pre-trained representation is then coupled to a classifier and optimized on labeled cohorts to predict immunotherapy response classes by minimizing a classification loss (Fig. 1B).

To provide a clear framework, we structure this review into four computational paradigms: Classical machine learning for interpretable models built from structured, feature-engineered data; Deep learning for representation learning from high-dimensional or unstructured data such as images and sequences; Graph and network modeling for capturing relational structure and interactions across scales; and mechanistic and systems biology models for simulating biological processes to generate causal insight and test hypotheses in silico.

The genomic and immunological foundation

Before delving into the computational paradigms, it is essential to understand the biomarkers and features that serve as the inputs for predictive models. These features are derived from three major layers of evidence: tumor genomics and antigen presentation (Fig. 2A), tumor immune microenvironment (Fig. 2B), and microbiome (Fig. 2C). For each layer, we map common feature types to the underlying data modalities and representative computational tools used to generate clinically relevant variables, as detailed in the figure legend. Their accurate quantification relies on computational methods that transform raw sequencing data into biologically meaningful variables (Fig. 3) (Table 2), as explained further below.

Quantifying genomic features: from TMB to mutational signatures

Tumor mutational burden (TMB), defined as the total number of somatic nonsynonymous mutations per megabase of the genome[25], is one of the earliest and most widely adopted biomarkers for ICI response[26]. The biological rationale is that more mutations increase the likelihood of generating novel protein fragments as neoantigens that can be recognized by the immune system as foreign[27,28]. The computational pipeline to estimate TMB (Table 2) from whole-exome sequencing (WES) data involves read alignment, somatic variant calling using tools like Mutect2[29], Strelka2[30], or GATK[31] to distinguish true mutations from sequencing artifacts, filtering of known germline variants present in population databases, and counting mutations normalized by the coding region size. Recognizing the logistical and financial barriers of WES in routine clinical practice (e.g., higher costs, longer turnaround times, greater computational demands, and increased complexity in data interpretation), targeted gene panel-based approaches, such as FoundationOne and MSK-Impact[32] have been adopted clinically. However, extrapolating panel-derived TMB to an exome-wide estimate requires scaling and calibration and introduces harmonization challenges across panels.

Despite its utility, TMB is a crude, single-dimensional metric[33,34] because it does not account for the quality, clonality, or HLA presentation potential of the resulting neoantigens[35], leading to inconsistent performance across cancer types and limited value in some prospective trials[35]. To add mechanistic resolution, mutational signature analysis deconvolves the mutation spectrum of the tumor into underlying processes, such as exposure to ultraviolet light or tobacco smoke, or defects in DNA repair pathways[36]. These signatures (Supplementary Data 1) provide more informative predictors of immune visibility than mutation counts alone[37,38].

https://doi.org/10.1038/s41467-026-71364-5

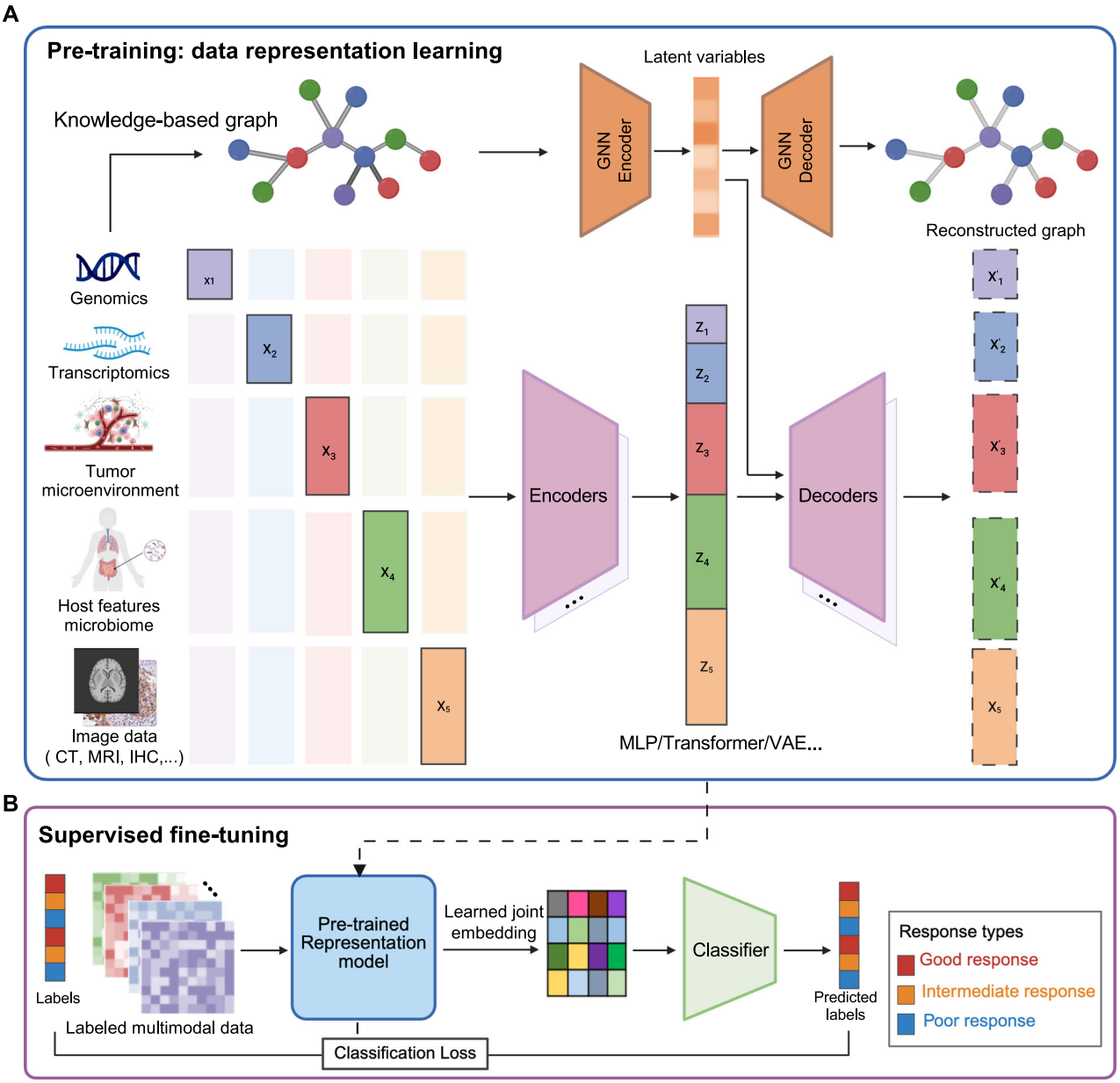


Fig. 1 | Multimodal integration for immunotherapy response prediction via pre-training and supervised fine-tuning. A Pre-training: Patient data from multiple modalities (genomics, transcriptomics, tumor microenvironment, host or microbiome features and imaging) are encoded to produce modality-specific representations that are aligned into a shared latent embedding. In parallel, the same inputs are used to construct a patient knowledge graph, which is learned by a graph neural network (GNN) encoder-decoder through graph reconstruction. Joint training of the modality encoders-decoders and the graph module by methods such as MLP, Transformer and VAE then yields an integrated representation that captures both within-modality signals and cross-patient relationships. **B** Supervised fine-tuning: The pre-trained representation model is combined with a classifier and optimized on labeled cohorts to predict immunotherapy response classes (for example, good, intermediate, and poor response) by minimizing a classification loss. Abbreviations were summarized in Supplementary Data 5. Created in BioRender. Liu, J. (2026) https://BioRender.com/xtyadck.

Identifying tumor-specific targets: neoantigen prediction and immunogenicity scoring

Neoantigens are key targets of anti-tumor T cells, making neoantigen identification central to computational immuno-oncology[39]. The standard computational pipeline for neoantigen prediction begins with somatic mutation and generates candidate mutant peptides, typically 8–11 amino acids for MHC class I (Figs. 3 and 4A). In parallel, the germline HLA genotype of the patient is determined from sequencing data using tools like integer linear programming-based OptiType[40] or Bayesian-based HLA-VBSeq[41](Fig. 3, Supplementary Data 2). Then, candidate neoantigens are filtered by predicted antigen processing, including proteasomal cleavage and TAP transport[42], and prioritized by the binding affinity between peptide and HLA (Fig. 4A). For HLA genotyping from whole-genome sequencing (WGS), WES, and RNA-seq data, two main approaches are used: alignment-based and de novo assembly (Supplementary Data 2). Benchmarking studies have shown that alignment-based HLA typing methods like OptiType and arcasHLA[43] perform comparably well, whereas HLA-VBseq shows reduced accuracy, which may be attributable to the inherent advantages of integer linear programming-based optimization (used in OptiType) over Bayesian inference methods (used in HLA-VBSeq) (Fig. 3). Despite its high accuracy for MHC class I genotyping, OptiType

https://doi.org/10.1038/s41467-026-71364-5

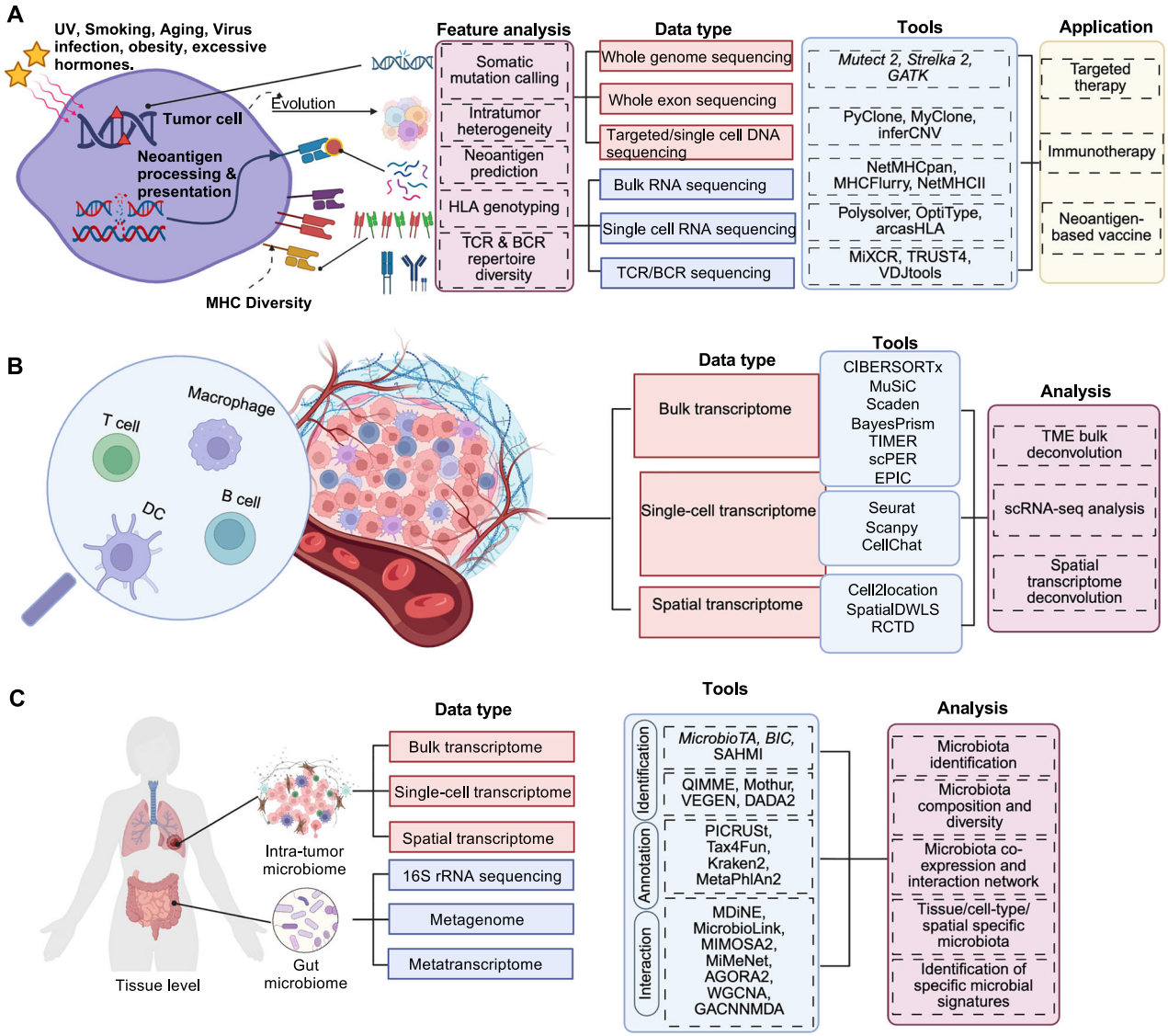


Fig. 2 | Biomarker feature extraction across tumor genomics, the tumor microenvironment, and the microbiome for immunotherapy response modeling. A Tumor-intrinsic and antigen presentation features. Common model inputs are derived from whole-genome, whole-exome, targeted or single-cell DNA, bulk RNA, and TCR/BCR sequencing. Representative analyses include somatic mutation calling, intratumor heterogeneity and clonal reconstruction, neoantigen prediction, HLA genotyping, and immune repertoire diversity. Example tools shown include Mutect2, Strelka2, and GATK for variant calling; PyClone and MyClone for clonal structure; NetMHCpan, MHCflurry, and NetMHCII for peptide-HLA prediction; Polysolver, OptiType, and arcasHLA for HLA inference; and MiXCR, TRUST4, and VDJtools for repertoire reconstruction and analysis. These features support downstream applications including immunotherapy response prediction, targeted therapy context, and neoantigen vaccine design. **B** Tumor microenvironment profiling from transcriptomics and spatial data. Bulk transcriptomes enable immune deconvolution to estimate cell-type fractions, for example, with CIBERSORTx, MuSiC, Scaden, BayesPrism, TIMER, scPER, and EPIC[199], supporting retrospective analysis of large cohorts. Single-cell transcriptomes enable cell-state and interaction analyses using workflows such as Seurat, Scanpy, and CellChat. Spatial transcriptomics, particularly spot-based platforms, benefits from deconvolution methods such as cell2location, SpatialDWLS, and RCTD to infer cell-type composition per location and reveal spatial immune architecture relevant to ICI response. **C** Microbiome-derived features and host-microbe associations. Microbiota can be profiled using 16S rRNA sequencing, metagenomics, and metatranscriptomics, and linked to tissue context using bulk, single-cell, or spatial transcriptomes. Representative pipelines include microbial identification, composition and diversity profiling, functional annotation, and microbe-host interaction or network analyses. Tool lists shown are representative. Abbreviations are provided in Supplementary Data 5. Created in BioRender. Li, B. (2026) https://BioRender.com/rzel3bo.

and similar tools underperform in MHC class II genotyping compared to traditional methods such as PCR-based sequence-specific oligonucleotide (SSO), sequence-specific primer (SSP), and Sanger sequencing. Assembly-based approaches, such as HLAMiner, can better capture novel or rare alleles and structural variants but require higher sequencing depth and are computationally intensive. Additionally, the above-mentioned HLA genotyping tools exhibit higher accuracy in European samples than African samples[44], likely due to the bias of existing testing/validation datasets, thereby underscoring the need for further refinement to improve performance across diverse populations[45].

This neoantigen prediction task has been revolutionized by deep learning, which enables the modeling of nonlinear relationships between peptide sequence, HLA binding groove structure, and antigen processing pathways. Tools such as NetMHCpan-4.1[46], MHCflurry 2.0[47], NetMHCIIpan-4.0[46,48] and MARIA[48] employ modern machine-learning models, such as fully connected neural networks, convolutional neural networks, and

https://doi.org/10.1038/s41467-026-71364-5

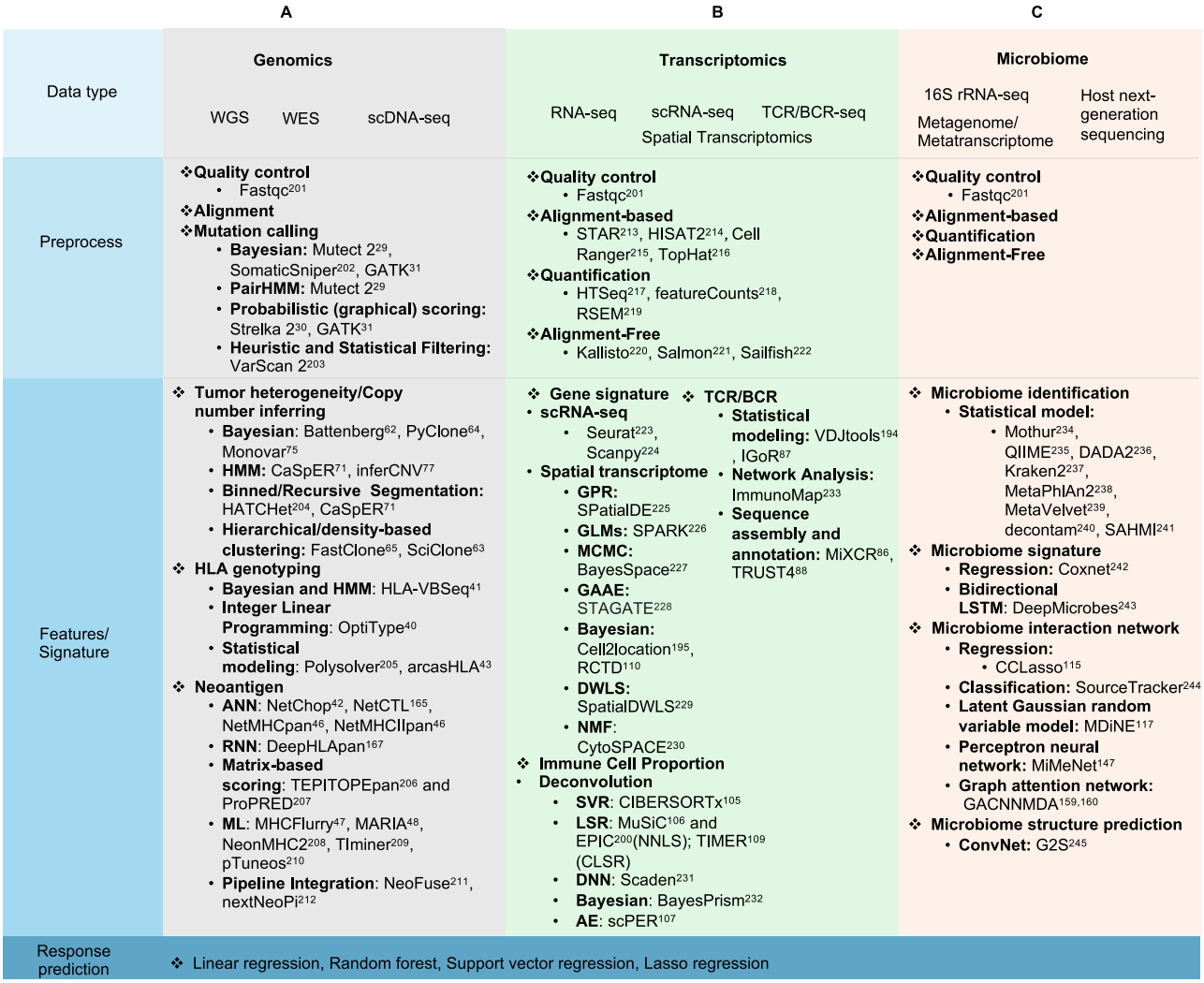


Fig. 3 | Representative computational tools and algorithmic steps for processing multi-omics data to predict immunotherapy response. This figure summarizes commonly used workflows across genomics, transcriptomics, and microbiome modalities, organized from preprocessing to feature extraction and finally response prediction using classical machine learning models. For (**A**) genomics (WGS, WES, scDNA-seq), preprocessing includes quality control with FastQC[200] and somatic variant calling using Bayesian or probabilistic frameworks such as MuTect2, SomaticSniper[201], GATK, Strelka2, and heuristic filtering with VarScan2[202]. Downstream genomic features include tumor heterogeneity and copy-number inference using Battenberg, PyClone, Monovar, CaSpER, inferCNV, HATCHet[203], CaSpER, FastClone, and SciClone; HLA genotyping using HLA-VBSeq, OptiType, Polysolver[204], and arcasHLA; and neoantigen-related prediction modules including NetChop, NetCTL, NetMHCpan, NetMHCIIpan, DeepHLApan, TEPITOPEpan[205], ProPRED[206], MHCflurry, MARIA, NeonMHC2[207], TIminer[208], pTuneos[209], and integrated pipelines such as NeoFuse[210], and nextNeoPi[211]. For (**B**) transcriptomics, alignment-based processing uses STAR[212], HISAT2[213], Cell Ranger[214], and TopHat[215], with quantification by HTSeq[216], featureCounts[217], and RSEM[218], while alignment-free quantification uses Kallisto[219], Salmon[220], and Sailfish[221]. Feature extraction includes scRNA-seq analysis using Seurat[222] and Scanpy[223]; spatial methods including SpatialDE[224], SPARK[225], BayesSpace[226], STAGATE[227], cell2location, RCTD, SpatialDWLS[228], and CytoSPACE[229]; immune cell proportion deconvolution using CIBERSORTx, MuSiC, EPIC, TIMER[109], Scaden[230], BayesPrism[231], and scPER; and immune repertoire analysis including statistical modeling with VDJtools and Igor, network-based analysis with ImmunoMap[232], and sequence assembly and annotation with MiXCR and TRUST4. For (**C**) microbiome data, representative tools include taxonomic identification and profiling using mothur[233], QIIME[234], DADA2[235], Kraken2[236], MetaPhlAn2[237], MetaVelvet[238], decontam[239], and SAHMI[240]; microbiome signature modeling using Coxnet[241], DeepMicrobes[242], and interaction or source tracking using CCLasso, SourceTracker[243], and MDiNE; cross-modal and network learning using MiMeNet and GACNNMDA; and structure prediction using G2S[244]. Across modalities, response prediction is typically performed with linear regression, random forests, support vector regression, and LASSO regression. PairHMM Pair hidden Markov model, ANN artificial neural network, RNN recurrent neural network, ML machine learning, GPR Gaussian process regression, GLMs generalized linear models, MCMC Markov chain Monte Carlo, GAAE graph attention autoencoder, DWLS dampened weighted least squares, NMF non-negative matrix factorization, SVR support vector regression, LSR least squares regression, *DNN* deep neural network, LSTM long short-term memory, AE autoencoder. Abbreviations were summarized in Supplementary Data 5. Detailed explanations of the algorithms were summarized in Supplementary Data 6.

transformer-based encoders (see Fig. 1), to learn complex sequence-binding relationships between peptides and HLA molecules (Supplementary Data 3). These prediction tools incorporate complementary modeling strategies to improve peptide-HLA binding and presentation prediction; they also help yield superior accuracy on benchmark datasets such as Immune Epitope Database[49], PRIDE[50], and SysteMHC[51], particularly for underrepresented HLA alleles. Despite substantial progress, performance across neoantigen prediction models remains context-dependent. Differences in training data curation, peptide length handling, and allele imbalance complicate direct comparison between models and often limit their generalizability across cohorts. Importantly, many benchmarking studies evaluate models using datasets that partially overlap with their original

https://doi.org/10.1038/s41467-026-71364-5

Table 2 | Key genomic and immunological biomarkers and computational tools

Biomarker	Biological rationale	Primary data source	Key computational tools	Clinical significance/limitations
Tumor mutational burden	Proxy for neoantigen load; higher TMB increases the probability of immunogenic mutations.	WES, Targeted Gene Panels	Mutect2[29], Strelka2[30], GATK[31]	FDA-approved biomarker for pembrolizumab, but predictive power is inconsistent and context-dependent.
Mutational signatures	Reflect underlying mutagenic processes (e.g., UV, smoking), some of which are more likely to generate immunogenic neoantigens.	WES, WGS	deconstructSigs[190], SigProfiler[191,192]	Provides mechanistic insight beyond a simple mutation count; helps stratify TMB-high patients.
Neoantigen burden/ quality	Direct targets of the anti-tumor T-cell response; their presence and quality are thought to be necessary for ICI efficacy.	WES/WGS+RNA-seq	OptiType[40], NetMHCpan[46], MHCflurry 2.0[47]	Central to personalized vaccines and T-cell therapies; prediction is computationally intensive and has a high false-positive rate.
Intratumor heterogeneity	Genetic diversity within a tumor drives immune escape and therapeutic resistance.	Bulk WES/WGS, scDNA-seq	PyClone[64], MyClone[66], InferCNV[77]	High ITH is associated with poor outcomes; clonal neoantigens are better targets than subclonal ones.
HLA genotype	Determines which peptides can be presented to T cells; heterozygosity is associated with a broader presentation repertoire.	WES/WGS, RNA-seq	OptiType[40], arcasHLA[43], HLA-VBSeq[41]	HLA-I homozygosity is linked to poorer survival with ICIs; critical for accurate neoantigen prediction.
TCR/BCR repertoire diversity	Measures the breadth and focus of the adaptive immune response; high diversity reflects a robust response.	Bulk RNA-seq, scTCR/BCR-seq	MiXCR[86], TRUST4[88], VDJtools[193]	High intratumoral TCR diversity is a strong positive prognostic marker for ICI response across multiple cancers.

training data, a practice that can inflate reported performance and obscure true out-of-sample accuracy[47]. Moreover, despite some existing efforts have begun to incorporate TCR recognition using clustering and machine-learning approaches (e.g., GLIPH[52], pMTnet[53]) and other deep learning frameworks(e.g., NetTCR[54] and DeepTCR[55]), most models focus on binding affinity rather than true immunogenicity, neglecting TCR recognition and peptide competition within the antigen-processing pathway. Together, these limitations highlight two key drawbacks: context-dependent performance and the predominant focus on peptide-MHC binding rather than true immunogenicity.

In parallel, recent methods are expanding the neoantigen search space to include peptides derived from alternative splicing (NeoSplice[56], SNAF[57]), transposable elements (TEProf2[58]), and structural variants (NeoSV[59]), though these specialized pipelines face additional challenges in data scarcity and model interpretability. LENS is also used for the identification of non-synonymous mutations, insertions, deletions, and gene fusions[60]. Overall, while deep learning has substantially improved peptide-MHC binding prediction, systematic and transparent validation across biological contexts remains a critical next step for clinical translation.

Characterizing tumor evolution: intratumor heterogeneity and clonal reconstruction

Intratumor heterogeneity is a major driver of therapeutic resistance and immune evasion, as genetically distinct subclones allow the tumor to adapt under selective pressure[61]. Computational methods infer subclonal architecture from bulk sequencing data. While Bayesian clustering tools like Battenberg[62], SciClone[63], PyClone[64], and FastClone[65] offer deep probabilistic inference of clonal structure, their requirement for substantial computing time and high-performance hardware has led to the development of faster alternatives like MyClone[66], which uses a variational expectation-maximization algorithm, trading some model complexity for significant gains in speed suitable for large-scale clinical cohorts. These tools analyze the variant allele frequencies (VAF) of somatic mutations in conjunction with local copy number information to group mutations into clusters that likely co-occur in the same cancer cell populations. Each cluster is assigned a cellular prevalence, representing the fraction of cancer cells harboring that set of mutations. Benchmark studies[67,68] suggest that clonality reconstructing pipelines combining Mutect2, FACETS, and PyClone achieve optimal accuracy, though scalability issues persist for larger datasets[67–70].

Tumor clonal reconstruction analysis allows for distinction between clonal and subclonal neoantigens[71], with T cell responses targeting clonal neoantigens often associated with durable benefit from ICBs[72–74]. However, bulk sequencing struggles to detect rare tumor subclones, which may harbor resistance mutations. Single-cell DNA and RNA sequencing (scDNA-seq, scRNA-seq) improves resolution, with tools like Monovar[75] and LiRA[76] (for somatic SNV detection) and InferCNV[77] (for CNV estimation) helping bypass sequencing errors and amplification biases. Meanwhile, SCcaller[78] employs a likelihood ratio test to differentiate true nsSNVs from artifacts, aiding discrimination based on aging and smoking status. Strategies for identifying CNV in single cells include Alleloscope[79] for allele-specific CNVs in scDNA-seq, while InferCNV[77] and CONICS[80] use scRNA-seq data to estimate copy number states, offering insights into ITH across cancer types. These technologies allow for the direct measurement of genetic variation at the cellular level, enabling high-resolution reconstruction of clonal phylogenies. However, specialized computational tools are required to handle the technical noise inherent in single-cell data. For instance, methods like InferCNV and Numbat[81] infer copy number variation from scRNA-seq and spatial transcriptomics data, respectively. Additionally, integrative methods like B-SCITE[82] and SCARLET[83] integrate bulk and single-cell data to reconstruct clonal lineages and

https://doi.org/10.1038/s41467-026-71364-5

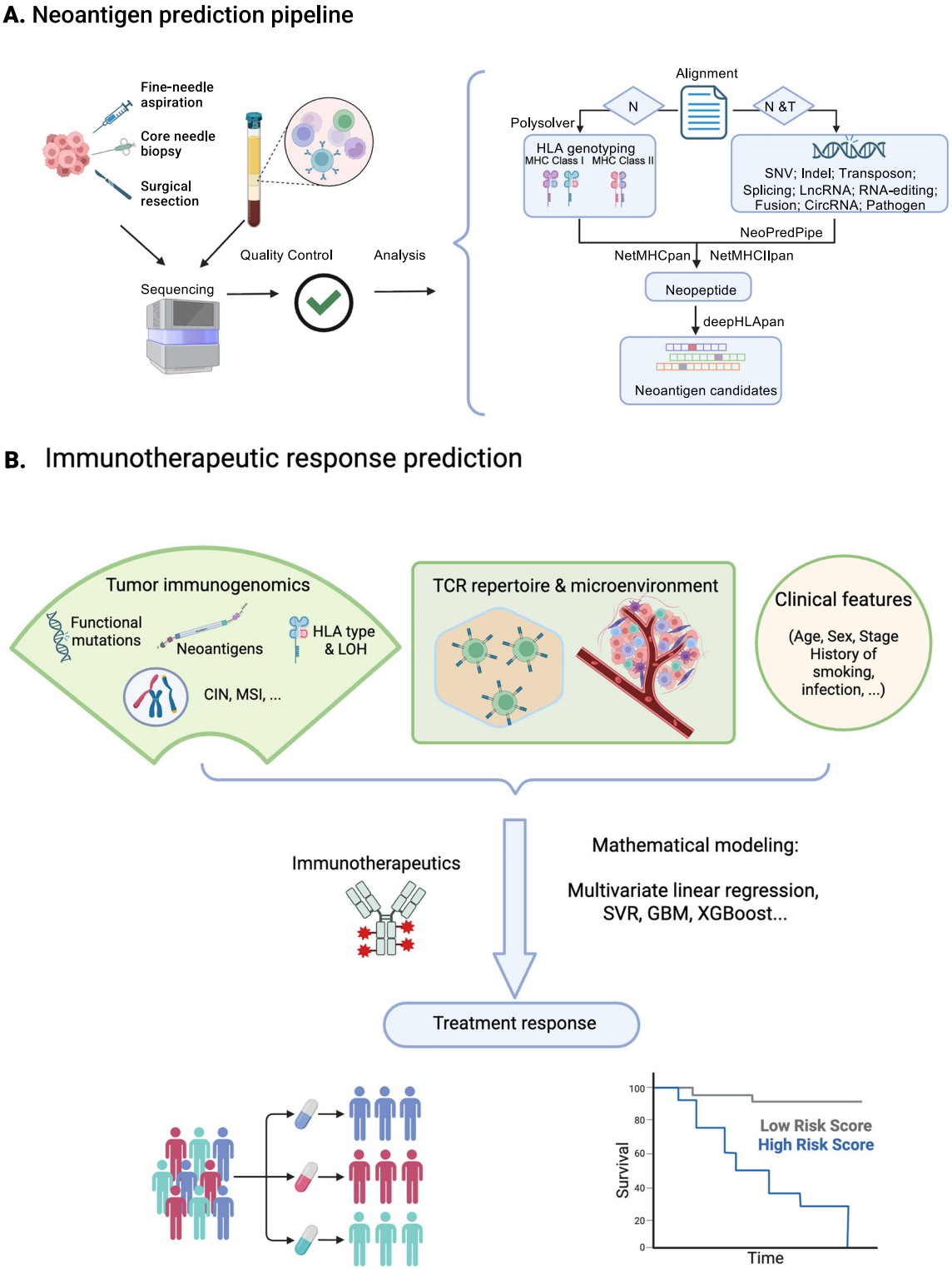


Fig. 4 | The computational workflow for neoantigen prediction and immunotherapy response evaluation. A Neoantigen prediction pipeline. Tumor samples obtained via fine-needle aspiration or core-needle biopsy are subjected to NGS, including WES and RNA-seq. Somatic mutations are identified and integrated with HLA typing to predict candidate neoantigens. **B** Prediction of ICIs treatment response using mathematical models. Additional genomic features, such as CIN and MSI, are computed to characterize tumor genomic context. Parallel immune profiling, including TCR repertoire analysis, provides information on clonal expansion and immune dynamics. These multi-dimensional features are then incorporated into predictive mathematical models to estimate ICI response probability. Ultimately, this integrated framework supports individualized immunotherapy stratification and longitudinal treatment monitoring. Created in BioRender. Li, B. (2026) https://BioRender.com/qx6ihx2.

tumor evolution. Together, these methods advance our understanding of ITH, offering opportunities to improve cancer diagnostics and treatment strategies.

Modeling the immune repertoire: TCR and BCR diversity and clonality

The diversity and composition of the T cell receptor (TCR) and B cell receptor (BCR) repertoires provide a direct readout of the adaptive anti-tumor immunity[84,85]. Broad diversity reflects the capacity to recognize many different antigens, while clonal expansion indicates antigen-driven responses. High-throughput sequencing of the hypervariable CDR3 regions of TCR and BCR genes enables quantitative profiling, and computational tools support TCR/BCR sequence recovery, clustering and annotation from bulk RNA-seq, scTCR-seq, and scBCR-seq data[86–88].

Computational tools such as MiXCR and TRUST4 can reconstruct TCR and BCR sequences from bulk RNA-seq data, enabling retrospective analysis of large clinical cohorts[2]. Single-cell TCR and BCR sequencing links each T or B cell receptor to cell state and supports metrics such as diversity, clonality and V(D)J gene usage. Additionally, methods such as Elastic net models with Monte Carlo cross-validation have been developed to integrate TCR-modeled entropy into biomarker development, improving the discrimination of responders in cancer immunotherapy datasets[89]. More recently, predicTCR[90] introduced the first machine learning classifier for tumor-reactive TCRs, enabling the identification of reactive TCR clonotypes in TILs across cancer types. As single-cell technologies continue to advance, the methods and tools discussed are expected to evolve for deeper analyses.

These biomarkers define the input feature space for immunotherapy response modeling. We next organize computational approaches into four paradigms with distinct workflows and strengths (Fig. 5), beginning with classical machine learning.

Paradigm I: classical machine learning for pattern discovery

Classical machine learning (ML) was among the first widely used computational approaches in immuno-oncology. Models such as logistic regression, support vector machines, and tree-based ensembles identify statistical patterns in structured, feature-engineered data[91,92]. They are commonly used for biomarker discovery, patient stratification, and interpretable clinical scores, though their performance is often constrained by the feature quality and cohort-specific biases (Table 3).

Supervised learning for response classification and survival prediction

Supervised ML is most often used to classify responders versus non-responders and to model endpoints such as overall survival (OS) or progression-free survival (PFS)[93]. These models are trained on labeled datasets where patient outcomes are known, using a matrix of input features derived from genomic, transcriptomic, or clinical data. Logistic regression and Cox models remain popular for their simplicity and interpretability[94]. The coefficients assigned to each feature provide a clear indication of its contribution to the prediction, making these models ideal for developing transparent clinical scores.

For high-dimensional data such as transcriptomes or microbiome profiles, ensemble methods including random forests (RF) and gradient boosting methods like XGBoost and LightGBM can capture non-linear interactions and handle many features[95]. These methods operate by constructing a multitude of decision trees during training and combining their outputs, which makes them more robust for noisy or missing data and often leads to superior predictive performance. Reported applications include OS prediction in metastatic renal cell carcinoma patients using an ensemble machine-learning model, achieving an AUC around 0.78–0.80[96], and XGBoost-based prediction with 97% accuracy[97]. Feature selection using univariate filters, or embedded methods such as LASSO[98,99] and Elastic Net regression, is critical to limit overfitting. Numerous predictive gene signatures, such as macrophage-related or lncRNA-based panels, have been developed using LASSO models, often reporting high AUCs in their discovery cohorts[100–102].

Unsupervised ML for identifying latent immunophenotypes

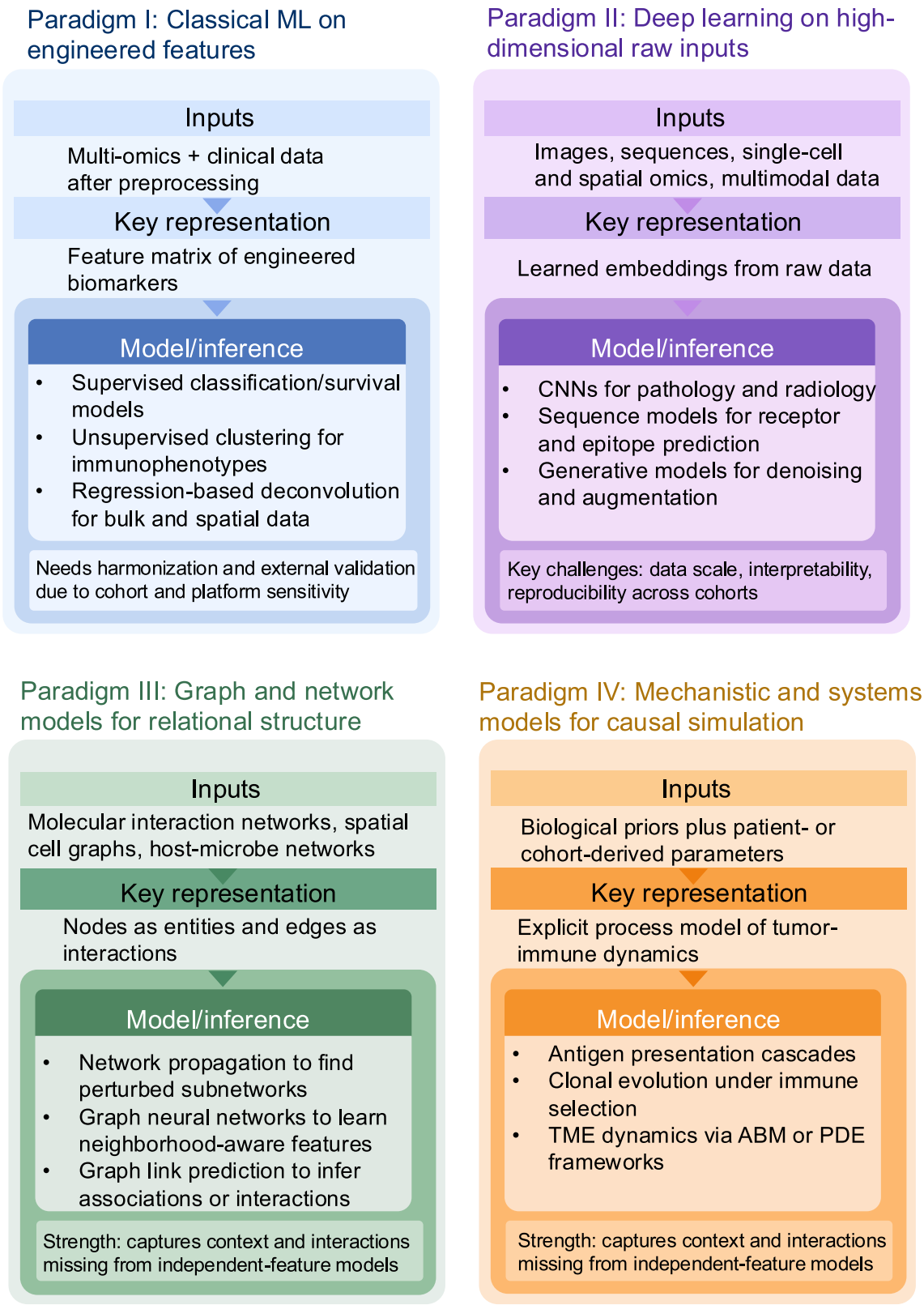
In contrast to supervised ML, unsupervised methods are used for exploratory analysis when outcome labels are unavailable or when the goal is to discover novel, underlying structures in the data. In immuno-oncology, the primary application of unsupervised ML is for patient stratification by identifying latent immunophenotypes[2]. Clustering approaches, such as k-means, hierarchical clustering, and non-negative matrix factorization (NMF), are applied to high-dimensional omics data to group patients or samples with similar molecular profiles. One example is consensus clustering of bulk RNA-seq data to identify pan-cancer immune subtypes[103,104], which has consistently revealed the existence of 'hot' (immune-inflamed) tumors as characterized by high T-cell infiltration and interferon signaling and typically responsive to ICIs. By contrast, 'cold' (immune-excluded) tumors lack immune infiltration and are generally resistant. These unsupervised classifications provide a fundamental framework for understanding the immune context of a tumor and also strong predictive value.

Deconvolution of bulk and spatial omics data

Deconvolution estimates cell-type proportions from bulk omics data, enabling inference of the immune context from widely available bulk RNA-seq data (Table 4). Reference-based deconvolution methods leverage a signature matrix, typically derived from a curated scRNA-seq dataset, which defines the characteristic gene expression profiles of each cell type. Approach like CIBERSORTx[105] applies v-support vector regression (v-SVR) to estimate the cell fractions in bulk samples, while MuSiC[106] uses weighted non-negative least squares to improve robustness when using references from multiple subjects. Recently, scPER further improved the robustness to batch effects in reference panels by employing adversarial autoencoders[107]. These tools have been instrumental in retrospectively analyzing large cohorts like The Cancer Genome Atlas (TCGA), linking the inferred abundance of cytotoxic T cells or immunosuppressive myeloid cells to clinical outcomes[108,109].

This reference-based deconvolution strategy extends to the deconvolution of spots in spatial transcriptomics, particularly for spot-based platforms where each capture location aggregates multiple cells. While newer spatial technologies can approach single-cell resolution, deconvolution remains valuable when measurements are mixed or sparsely sampled. Approaches like RCTD[110] and SPOTlight[111] use regression and NMF-based approaches, respectively, to estimate the cell-type composition within each spatial spot using a single-cell reference, producing high-resolution maps that can reveal spatial patterns such as immune exclusion at the tumor boundary for predicting ICI response[110,111].

Classical ML performance is often unstable when features are sensitive to technical variations, making models vulnerable to the domain shift described above. A model trained on data from one clinical center or sequencing platform may perform poorly on data from another, not because the underlying biology is different, but because the input features themselves have shifted. Accordingly, classical ML models require careful preprocessing and harmonization, and should be evaluated with rigorous external validation to assess real-world robustness.



ML-based prediction models of microbiome data

Microbiome features can also be critical biomarkers for ICI efficacy or toxicity. Increasing reports support within-cohort discrimination, but variable cross-cohort generalization remains. Multi-study and leave-one-study-out (LOSO) validations consistently show that the microbiome features are robust yet cohort-dependent: a large meta-analysis integrating 147 metagenomes confirmed microbiome-ICB links but found limited reproducibility of signatures across cohorts when models were transferred, reiterating domain shift as the principal failure mode rather than pure overfitting within a study[112]. For adverse events, ML pipelines pairing microbiome features with clinical data have identified microbial and metabolite markers for irAE risk stratification, again with study- and cancer-type dependence that necessitate external validation as a reporting prerequisite[113].

Fig. 5 | High-level workflows of four computational paradigms for predicting immunotherapy response. (Paradigm I) Classical machine learning operates on engineered biomarker features derived from multi-omics and clinical data to perform supervised response or survival prediction, unsupervised discovery of latent immunophenotypes, and regression-based deconvolution of bulk and spatial measurements. Its utility depends on preprocessing, harmonization, and rigorous external validation to mitigate cohort and platform sensitivity. (Paradigm II) Deep learning learns representations directly from high-dimensional or unstructured inputs, including pathology and radiology images, biological sequences, and single-cell or spatial omics, enabling multimodal integration. Typical applications include image-based prediction, receptor-epitope modeling, and generative modeling for denoising or data augmentation, with key challenges in interpretability and cross-cohort reproducibility. (Paradigm III) Graph and network modeling represents biological systems as nodes and edges to capture interactions across genes, cells, and microbes. Network propagation and graph neural networks support pathway-level inference, spatial neighborhood analysis, and relational biomarker discovery by leveraging molecular interaction networks, spatial graphs, and host-microbe networks. (Paradigm IV) Mechanistic and systems biology models use knowledge-driven simulations to model tumor-immune dynamics, including antigen processing and presentation, clonal evolution under immune pressure, and tumor microenvironment dynamics using agent-based or equation-based frameworks. These models provide causal insight and enable counterfactual testing of interventions to inform therapy design.

Table 3 | Comparison of machine learning and deep learning paradigms

Attribute	Classical machine learning (e.g., XGBoost, RF)	Deep learning (e.g., CNN, Transformer)
Feature engineering	Manual; requires domain expertise to create structured feature tables.	Automatic; learns hierarchical features directly from raw data.
Data requirement	Effective on small to medium-sized datasets (hundreds to thousands of samples).	Requires large, labeled datasets (thousands to millions of samples) to perform well.
Performance on unstructured data	Poor; not designed for raw images, sequences, or text.	State-of-the-art performance on unstructured data like images (CNNs) and sequences (Transformers).
Interpretability	Generally high (e.g., feature importance in RF, coefficients in logistic regression).	Generally low ("black box"); requires specialized methods (e.g., SHAP, saliency maps) for explanation.
Computational cost	Low to moderate; can often be trained on a standard CPU.	High; typically requires specialized hardware like GPUs or TPUs for training.
Key application in IO	Building predictive scores from clinical/genomic features; cell type deconvolution.	Pathology/radiology image analysis; neoantigen prediction; TCR-epitope binding modeling.

Table 4 | Leading computational tools for spatial and single-cell analysis

Tool name	Primary task	Computational approach	Key feature/advantage
Cell2location	Spatial deconvolution	Bayesian regression (variational inference)	Probabilistic framework that models cell-type proportions and states, improving accuracy for heterogeneous cell states[194].
RCTD	Spatial deconvolution	Robust regression	Assigns cell type proportions to spatial spots while down-weighting outliers, making it robust to noisy data[110].
SpaGCN	Spatial domain identification	Graph convolutional network (GCN)	Integrates gene expression and spatial location to identify spatially coherent tissue domains without requiring single-cell reference data[195].
CellChat	Cell-cell interaction mapping	Mass action-based modeling	Infers and visualizes intercellular communication networks from scRNA-seq data by considering ligand-receptor interactions and signaling pathways[196].
NicheNet	Cell-cell interaction mapping	Network propagation & ML	Predicts ligand-receptor pairs driving gene expression changes in receiver cells by integrating expression data with prior knowledge networks[197].
Tangram	Spatial mapping/integration	Optimization-based alignment	Aligns scRNA-seq data onto spatial transcriptomics data by optimizing the mapping of cells to spatial locations to match gene expression profiles[198].
CellLENS	Integrated subtype discovery	Hybrid CNN and GNN	Fuses microscopy images, spatial location, and molecular data to identify functionally distinct cell subtypes based on their neighborhood context[155].

At the systems level, microbial co-occurrence networks reveal ecological relationships within microbial communities. Tools such as SparCC[114] and CCLasso[115] infer microbial correlations while correcting for compositional bias, thereby identifying taxa that co-vary across environmental or disease conditions (Supplementary Data 4). SparCC approximates correlations among latent absolute abundances via log-ratios to recover sparse co-occurrence graphs with typical edge-recovery AUROC ~ 0.70–0.85 in simulations; meanwhile CCLasso estimates a sparse precision matrix on CLR-transformed data to target conditional dependencies with higher precision, which often has AUROC ~ 0.75–0.90[116]. Dynamic extensions such as MDiNE[117] incorporate temporal information from longitudinal datasets, capturing how microbial associations evolve with tumor progression or treatment exposure. Complementing these correlation networks, genome-scale metabolic reconstructions, exemplified by AGORA[118] and AGORA2[119], enable mechanistic modeling of host-microbiome metabolic flux. These frameworks simulate the exchange of metabolites and energy between human and microbial pathways, revealing how microbial metabolites modulate host immunity, inflammation and tumor metabolism.

Paradigm II: deep learning for representation learning from high-dimensional data

Deep learning (DL) shifts prediction from hand-crafted features to learned representations. Unlike classical ML, deep neural networks learn directly from high-dimensional or unstructured input such as pathology images, DNA or protein sequences, and spatial omics[120–122] (Table 3), enabling improved performance and multimodal integration in immune-oncology.

Convolutional neural networks (CNN) in pathological and radiological imaging

CNNs are widely used for image analysis because they learn relevant visual patterns directly from pixels[120]. In digital pathology, CNNs applied to gigapixel whole-slide images (WSI) of H&E-stained tumor sections can predict molecular and immune features from tissue

https://doi.org/10.1038/s41467-026-71364-5

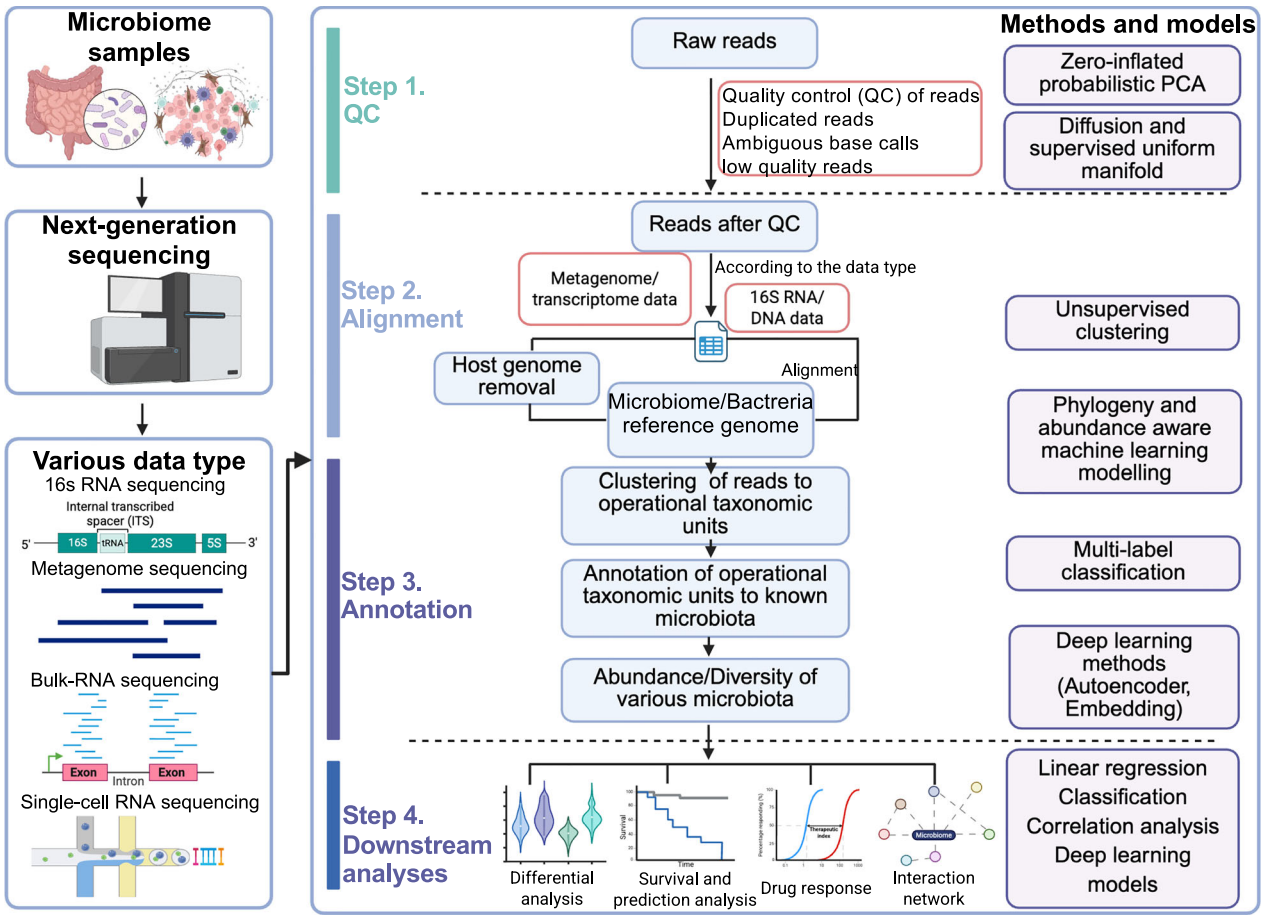


Fig. 6 | Workflow for microbiome identification from NGS data and representative computational methods. Microbiome samples are profiled by NGS and may generate multiple data modalities, including 16S rRNA gene (or 16S RNA/DNA) sequencing, shotgun metagenome sequencing, bulk RNA-seq, and single-cell RNA-seq (left). The core identification and downstream analysis can be summarized into four steps (center): (Step 1) read-level QC to remove duplicated reads, ambiguous base calls, and low-quality reads, yielding high-confidence reads; (Step 2) alignment and host filtering performed according to data type, typically including host genome removal followed by alignment to microbiome reference genomes, and clustering reads into operational taxonomic units (OTUs) or equivalent operational units; (Step 3) taxonomic/functional annotation by mapping OTUs to known microbiota, followed by quantification of microbial abundance and diversity; and (Step 4) downstream analyses such as differential analysis, survival/prognostic modeling, drug-response association, and interaction/network inference (bottom). Representative statistical and machine-learning models frequently used across these steps are summarized on the right panel, including zero-inflated and probabilistic dimension-reduction models such as probabilistic PCA, manifold learning such as UMAP, unsupervised clustering, phylogeny- and abundance-aware machine-learning modeling, multi-label classification, deep-learning approaches, as well as classical regression and classification and correlation-based analyses. Abbreviations are listed in Supplementary Data 5. Created in BioRender. Li, B. (2026) https://BioRender.com/rzel3bo.

morphology, including driver mutations, microsatellite instability, PD-L1 expression, and TMB with remarkable accuracy, using only the standard H&E slide as input[123–126]. For example, TiDo[127] predicted *TP53* mutation-linked phenotypes in prostate cancer by learning subtle histological patterns in the tumor and surrounding stroma, supporting low-cost molecular inference from routine slides.

In radiology, CNNs can learn quantitative imaging signatures from CT, MRI, and PET scans that reflect tumor shape, texture, and intensity, and have been used to predict immune infiltration, TME phenotypes and ICI response, particularly when combined with clinical or genomic data[128–130]. A key challenge with image-based DL is interpretability: accurate predictions may not be easily explained or may rely on spurious cues. Saliency maps such as Grad-CAM highlight image regions most influential for a prediction, supporting clinical review[131,132].

Sequence-based models for TCR and BCR binding specificity

Deep learning is increasingly used to predict which T cell or B cell receptors recognize specific epitopes by learning patterns directly from amino acid sequences[133,134]. Early frameworks such as pMTnet[53], combining Long Short-Term Memory (LSTM) for pMHC sequence encoding, autoencoders for TCR features, and a deep neural network (DNN) fusion layer, enables integrated binding prediction across antigen and receptor modalities. Similarly, ImRex[135] used attention, a neural network mechanism, to model TCR-peptide binding, capturing sequence complementarity.

More recent models focus on generalization to unseen epitopes and limited-data settings. DeepTCR[55] adopts a CNN backbone with optional supervised or VAE layers, capturing motifs across CDR3α/β and V(D)J gene usage while supporting repertoire-level classification tasks. PanPep[136] applies meta-learning coupled with a neural Turing machine (NTM), allowing rapid adaptation to novel peptide-TCR pairs with limited data. TEIM-Res[137] introduces a CNN-based interaction extractor that outputs both binding probabilities and residue-level affinity maps, improving interpretability by identifying contact residues contributing to epitope recognition. Pep2TCR[138] integrates transfer learning and ensemble modeling, fine-tuned for $CD4^+$ T cell contexts, thereby addressing MHC-II-restricted epitopes that are often underrepresented in training data. EpicPred[139] extends this paradigm through a bidirectional encoder representation from transformers

(BERT)-based encoder integrated with an open-set recognition and attention-driven multiple-instance learning (MIL) module, enabling joint modeling of binding specificity and T cell phenotype and bridging sequence-level and functional prediction. By contrast, TCR-VALID[140] utilizes a VAE to embed TCRs into a latent feature space that preserves biophysical constraints, facilitating unsupervised representation learning and clustering of receptor repertoires. Finally, TPepRet[141] employs a hybrid BiGRU-RetNet framework, combining bidirectional recurrence with self-attention variants to learn contextual representations from amino acid composition, outputting continuous binding probabilities.

Despite accuracy gains, challenges remain: limited paired TCR-epitope data, incomplete structural context, and variable reproducibility across cohorts. Integrating structural priors, larger-scale pretraining, and uncertainty estimates may improve robustness and interpretability.

Generative models for modeling heterogeneity and denoising data

Generative models are a class of unsupervised DL models that learn the data distribution, enabling low-dimensional latent representations, denoising and synthetic data generation.

VAEs are widely used for this purpose[24]. In immuno-oncology, VAEs capture structure in high-dimensional omics data, such as gene expression or mutation frequencies, supporting trajectory analysis in scRNA-seq data (e.g., scVI[142]) or tumor subclones. VAEs also denoise scRNA-seq data to improve the downstream accuracy, and can embed multi-modal patient data like genomics, transcriptomics, and clinical ones into a shared latent space for patient stratification or outcome prediction.

Another generative methods, generative adversarial networks (GAN), consist of two competing neural networks, a generator and a discriminator, which are trained together. GANs have been used to augment small cohorts and improve response prediction[143,144].

Deep learning approaches for microbiome-based prediction

Microbiome modeling is shifting from single-taxon associations to higher-order community structure and host-microbe relationships[145] (Fig. 6), including host or pathogen metabolites as soluble mediators.

For the microbiome-metabolites interaction network, cross-modal mappers like MiMeNet use multi-task neural regression to learn nonlinear mappings from microbial compositions to metabolite levels, yielding microbe-metabolite embeddings and cohort-level predictions[146]. By providing surrogate biochemical readouts when metabolomics is missing and attributing metabolite variance to specific features, MiMeNet supplies mechanism-linked representations that can be integrated into downstream clinical predictors (e.g., ICB response or irAE risk) and be subjected to standard validation workflows[146]. As a representative example, MiMeNet moves beyond lists of taxonomic groups (taxa) toward 'quantified networks' and 'flux-level hypotheses', enabling interpretable modules that can be plugged into downstream predictors of ICB response or irAE risk and calibrated for clinical decision support[147].

In the context of cancer immunology, deep neural networks, convolutional architectures, and autoencoders can automatically learn latent microbial features that encapsulate higher-order co-occurrence structures and complex host-microbe relationships. Frameworks such as DeepGeni[148], an interpretable autoencoder trained on multi-cohort datasets, have demonstrated improved ICI response prediction with approximately 10-15% higher AUROC relative to regularized logistic regression, while maintaining biological interpretability of latent nodes[148]. Similar results have been observed with DeepMicro[149], which compresses sparse, compositional abundance matrices into low-dimensional embeddings that generalize better across studies. These model architectures (including feed-forward DNNs, cross-attention–based models, and autoencoders) are most effective when combined with stringent overfitting control, dropout, early stopping, and leave-one-cohort-out validation, ensuring that latent features represent genuine biological signals rather than dataset-specific noise. A natural next step is to encode these microbiome-metabolite-host systems as heterogeneous graphs (e.g., microbes-metabolites-host features as nodes, with edges derived from co-occurrence, biochemical associations, or learned cross-modal links) and learn representations using graph neural networks or graph autoencoders. Such graph-structured learning can unify network topology with predictive embeddings, enabling more transferable predictors and more localized, subnetwork-level mechanistic hypotheses for ICB response.

Paradigm III: graph and network models for relational insights

Graph and network models represent biological systems as interacting networks rather than independent features. By modeling entities such as genes, cells, or microbes as nodes and their relationships as edges, these methods capture context and influence in complex biological ecosystems[150], supporting a more systems-level understanding of immunotherapy response.

Leveraging molecular interaction networks for pathway-level inference

A common application uses curated molecular interaction networks, such as protein-protein interaction (PPI) and gene regulatory networks to interpret genomic signals at the pathway level.

Methods based on 'network propagation' or 'heat diffusion', such as HotNet2[151] and NetBio[152] overlay gene-level data, such as mutations or differential expression, onto a network and apply propagation or diffusion to identify perturbed subnetworks. This approach can uncover pathway modules and functionally important hub genes even when individual genes are not strongly altered by standard per-gene tests.

Graph neural networks for spatial and single-cell ecosystems

Graph Neural Networks (GNN) extend deep learning to graph-structured data by iteratively aggregating feature information from neighboring nodes, enabling context-aware features.

In spatial transcriptomics data, the tumor microenvironment can be modeled as a spatial graph, with cells or spots as nodes and edges connecting neighbors. GNNs can then detect cellular neighborhoods and spatial patterns linked to immunotherapy response. For example, SpaGCN[153] uses graph convolution to identify spatially coherent tissue domains, while DSTG[154] incorporates a GNN to model cell-cell interactions and resolve the spatial organization of immune cell subtypes with high resolution. CellLENS[155] uses a hybrid CNN-GNN architecture to integrate microscopy images, spatial coordinates, and molecular data to identify functionally distinct cell subtypes based on local environment and neighborhood composition; CellLENS also differentiates T cells actively engaged at the tumor boundary versus those that are not, representing a crucial shift in analysis from cellular composition to cellular architecture.

Modeling the host-microbiome interaction network

Moving beyond correlation and constraint-based reconstructions, recent graph-based deep learning models learn microbiome structure directly from graphs and multi-modal relations. At the genome-reconstruction level, GraphMB[156] uses graph neural networks on contig assembly graphs to improve long-read metagenomic binning, increasing the number of high-quality metagenome-assembled genomes versus non-graph baselines by leveraging connectivity and coverage as inductive bias, thereby delivering cleaner nodes for downstream ecological networks (Fig. 6, Step 4). This assembly-graph binning setting is distinct

from association-based graph learning models (i.e., models that construct heterogeneous knowledge graphs or multi-view relation graphs over microbes, diseases, drugs, metabolites, and hosts) to support phenotype discovery and intervention design. For phenotype discovery across entities, heterogeneous GNNs (i.e., graphs with multiple node and edge types) such as GCATCMDA and related variants encode microbe-disease knowledge graphs and perform link prediction with GCN and GAT layers, typically reporting AUC gains over matrix-factorization and heuristic baselines, often ≥0.85 when rich priors are available[157]. Extending to therapy design, GCNATMDA and GPUDMDA/GACNNMDA integrate microbe-drug-disease tripartite information and 'positive-unlabeled learning' or attention mechanisms to prioritize novel microbe-drug interactions, improving early retrieval (top-k) and overall discrimination in cross-validation on public datasets[158,159]. Generative models like MADGAN[160] impose adversarial constraints on multi-view similarity graphs to enhance the separability of candidate microbe-disease edges, further boosting link prediction when labeled associations are sparse. Bridging composition to mechanism, MMvec[161] learns a neural log-bilinear mapping from microbes to metabolites, producing low-rank microbe-metabolite embeddings that transfer across studies better than raw co-occurrence and should be treated as mechanism-bridging priors rather than stand-alone predictors. Finally, emerging multi-modal GNN frameworks fuse composition, function, metabolites and host-immune features into a single heterogeneous graph with pathway and interaction priors, yielding more stable and transportable patient-level representations than single-view models and enabling clearer mechanistic read-outs via attention, particularly when evaluated with carefully designed negative sampling (i.e., constructing putative negatives in a way that avoids treating all unlabeled edges as true negatives and reduces closed-world evaluation artifacts), external validation, and leave-site-out testing to mitigate closed-world effects and domain shift[162,163].

Paradigm IV: mechanistic and systems biology models for causal simulation

Mechanistic and systems biology models complement data-driven ML by simulating tumor-immune processes using established biological principles, enabling in silico hypothesis testing and counterfactual perturbations, and linking predictions to explicit processes. Although purely data-driven predictive models, particularly high-capacity approaches with limited interpretability, may achieve higher predictive accuracy on some datasets, mechanistic frameworks provide causal insight and can delineate mechanistic drivers of response or resistance to inform rational design of novel therapies.

Simulating the antigen processing and presentation pathway

Mechanistic models of antigen presentation simulate steps from mutated protein to peptide-MHC complex, providing a more biologically faithful estimate of immunogenicity than binding affinity alone. Early integrated tools like NetCTL[164] combined predictions for three key steps: proteasomal cleavage of the source protein using models like NetChop[42], transport of the resulting peptides into the endoplasmic reticulum by the TAP transporter, and subsequent binding to MHC class I molecules. Each step was modeled independently, and the scores were combined to produce a final prediction of antigen presentation. More recent systems-level predictors include LENS[60], which computes a clonal fitness score integrating peptide-MHC binding with sequence similarity to known pathogenic epitopes, based on bias of T cell repertoire toward pathogen-derived recognition. By incorporating this prior biological knowledge, LENS[60] was able to distinguish responders from non-responders in melanoma and NSCLC more effectively than TMB or simple neoantigen counts alone.

Modeling tumor-immune co-evolution and immunoediting pressure

Immune escape is a major cause of immunotherapy failure, driven by the selection and expansion of tumor subclones that evade immune control. Mechanistic models simulate these evolutionary dynamics by treating tumors as evolving clonal populations under immune pressure. For example, the population bottleneck model of fitness simulates the population dynamics of tumor clones, allowing the estimation of the likelihood that a resistant clone will emerge and escape immune control. Network-based models like the Network-based Immune Escape Index (NETIE)[165] quantify the potential for immune escape by analyzing the structure of the neoantigen-HLA-TCR interaction network. NEO2IS/NEO2ITHS[74] applies deep attention-based networks (deepHLApan[166]) to jointly model peptide features and tumor clonality, boosting ICB response prediction.

These simulation-based approaches move beyond static, correlational biomarkers to provide a dynamic and causal framework for understanding how immune pressure sculpts the tumor's genomic landscape, leading to acquired resistance. For example, NEOPS integrates HLA loss of heterozygosity (LOH) via DASH with peptide presentation predictions from XGBoost[167]. In addition, NEOPS is not merely an explanatory exercise by identifying the likely mechanisms of escape, such as loss of HLA expression and downregulation of a specific neoantigen. The above models can generate actionable hypotheses for combination therapies designed to counteract or preempt resistance.

Agent-based and equation-based models of the tumor microenvironment

Mechanistic models also capture spatial and temporal TME dynamics as in silico laboratories. Agent-Based Models (ABMs) are a bottom-up approach where individual cells, such as cancer cells, T cells, and macrophages, are simulated as autonomous agents. Each agent is endowed with a set of rules governing its behavior, such as proliferation, migration, cytokine secretion, and cell-cell killing[168]. When thousands of these agents interact within a simulated 2D or 3D space, complex, emergent phenomena can be observed, such as the formation of immunosuppressive niches or the failure of T cells to penetrate a dense stromal barrier.

In contrast, partial differential equation (PDE) models are a top-down approach that uses systems of equations to describe how the densities of different cell populations and the concentrations of molecules, such as chemokines and cytokines, change over space and time[169]. While they do not track individual cells, PDEs are computationally efficient and can effectively capture macro-level dynamics, such as the wave of T cell infiltration into a tumor following ICI administration or the spatial gradients of oxygen and nutrients that influence tumor growth. By varying the parameters in these ABM or PDE models, researchers can simulate different therapeutic interventions, such as increasing T cell killing efficacy and blocking an immunosuppressive cytokine, and predict their impact on tumor control, guiding the rational design of combination therapies.

Toward digital twins for immunotherapy

Digital twins are patient-specific virtual models that integrate longitudinal, multi-scale data and are updated as new measurements become available[170]. An immunotherapy digital twin would integrate the genomics, proteomics, imaging, spatial omics, microbiome data, and clinical history of the patient into a hybrid model that links data-driven prediction with mechanistic simulations[171]. In principle, clinicians could run simulations on the patient's digital twin to compare therapies, optimize dosing and schedules and anticipate resistance or toxicity, updating recommendations as the patient evolves. While the full realization of comprehensive digital twins remains a grand challenge, current advances in multimodal integration and mechanistic

BOX 1

Deep learning: models that learn predictive patterns directly from raw data.

CNN: model for images that learns visual patterns from pixels.

Transformer/attention: model that learns relationships among sequence elements.

Embedding: a compact numeric representation of an input for prediction.

VAE: model that compresses data and can denoise it.

GAN: model that generates realistic synthetic data.

Saliency map: heatmap showing which input regions most influenced a prediction.

BOX 2

From algorithm to clinic: validation, translation, and ethical imperatives

Despite the rapid pace of algorithmic innovation, the real-world clinical adoption of computational models for immunotherapy prediction remains limited. The journey from a high-performing model in a research paper to a trusted tool in the clinic is fraught with challenges that extend beyond predictive accuracy. Bridging this algorithm-to-action gap requires a rigorous commitment to robust validation, transparent interpretation, and the thoughtful navigation of complex ethical and societal issues.

Key considerations include:

1. Model validation and generalizability

The vast majority of predictive models in immuno-oncology are developed and validated on retrospective, often publicly available datasets (e.g., TCGA) or single-institution cohorts. While this is a necessary first step, models trained in this manner are highly susceptible to several biases that limit their real-world utility. Overfitting is a constant risk, where a model learns the specific noise and idiosyncrasies of a small training set rather than a true, generalizable biological signal. Even more pernicious is domain shift, where a model that performs well on data from one hospital or sequencing platform fails when applied to data from another due to subtle but systematic differences in patient populations, clinical practices, or data processing pipelines.

To build trust and ensure that a model is robust, a stringent hierarchy of validation strategies must be employed, including internal validation, external validation, prospective validation and subgroup analysis. Cross-validation within the training dataset is the minimum standard, but it is insufficient on its own. The model must be tested on at least one, and preferably several, completely independent datasets from different institutions or clinical trials. This is the most critical test of generalizability. The gold standard is to evaluate the performance of the model in a prospective clinical trial, where its predictions are used to guide patient care or are evaluated in real time. Few models have reached this stage. Performance should also be explicitly evaluated across different cancer types, ancestries, and other relevant clinical subgroups to ensure the model is equitable and to define its specific context of use.

2. Implementation challenges and failure cases

Even well-validated models face significant barriers to clinical deployment. Technical hurdles include the lack of user-friendly, regulatory-approved software and the difficulty of integrating predictive tools into existing hospital workflows and Electronic Health Record (EHR) systems. Biologically, the inherent dynamism of the immune response and the challenge of tumor heterogeneity mean that a prediction based on a single, pre-treatment biopsy may be an oversimplification.

3. Regulatory and legal landscape

As algorithmic tools become more sophisticated, they are attracting greater scrutiny from regulatory bodies like the U.S. Food and Drug Administration (FDA) and the European Medicines Agency (EMA). For a model to be used in clinical decision-making, especially as a companion diagnostic to determine eligibility for a specific therapy, it must demonstrate both analytical validity and clinical validity. Proving clinical validity typically requires evidence from well-designed clinical trials. The regulatory pathway for these software as medical device (SaMD) products is still evolving, creating uncertainty for developers and clinicians.

4. Ethics and algorithmic equity

Ethical considerations are not an afterthought but a central component of responsible model development and deployment. A primary concern is algorithmic bias and fairness. Most large genomic and clinical datasets are overwhelmingly derived from patients of European ancestry. Models trained on this biased data may perform significantly worse for patients from underrepresented racial and ethnic groups, thereby perpetuating and even exacerbating existing health disparities. Auditing model performance across diverse demographic groups is an ethical imperative.

Together, these considerations outline the practical and ethical requirements for translating immunotherapy prediction models into clinically trusted tools.

The black box nature of many complex DL models raises issues of interpretability and accountability. For a clinician to trust and act upon a recommendation from a model, they must have some understanding of why the model made its prediction. Deploying opaque algorithms in high-stakes clinical decisions without clear, human-understandable explanations is medically and ethically problematic. The goal should be to create human-AI collaboration frameworks that augment, rather than replace, clinical judgment. Finally, the use of vast amounts of sensitive patient data, particularly for the construction of digital twins, raises profound privacy and data security concerns that must be addressed through robust technical safeguards and transparent governance policies.

modeling are assembling the required building blocks and point toward simulation-guided, personalized cancer care.

Applications of large language and foundation models

A new class of Transformer-based models, including Large Language Models (LLM) and Foundation Models, is emerging. Unlike the task-specific DL models above, which is typically trained on a single modality and task with supervised labels, foundation models are first pre-trained at scale using self-supervised objectives to learn general-purpose representations, then adapted to downstream tasks. LLMs are pre-trained primarily on language, and in immune oncology they are applied either directly to clinical text or to biological data encoded as token sequences.

LLMs can extract structured variables, such as cancer staging and biomarker status, from unstructured clinical notes and pathology reports, and multi-modal foundation models can fuse pathology images (H&E slides) with their corresponding text. Models like MUSK[172], which were pre-trained on pathology images and text, have demonstrated strong performance in predicting outcomes like melanoma relapse and immunotherapy response. Frameworks such as HONeYBEE[173] are being developed to generate unified embeddings from clinical notes, pathology slides, radiology scans, and molecular profiles.

Protein language models are also being applied to biological sequences. For TCR-epitope binding, LANTERN[174] reported that LLM-based models achieved a mean ROC-AUC of 0.8165, compared with 0.5846 for other end-to-end deep learning models, while STAG-LLM[175] achieved a median ROC-AUC of 0.815 for TCR-pHLA binding and outperformed several existing baselines. LLMs are also being explored to summarize and assess the immunogenicity of candidate neoantigens.

Despite this promise, limitations remain, including training data bias, incomplete clinical reliability, and the need for rigorous benchmarking; some foundation models do not outperform simpler baselines on specific tasks (Box 1).

Integrate computational tools into clinical frameworks

Several real-world platforms have begun to translate multi-modal integration into clinical research to guide trial stratification and treatment planning. Conceptually, these examples primarily align with Paradigm I and Paradigm II above. Tempus has developed multimodal immuno-oncology assays, including the Immune Profile Score (IPS), a DNA- and RNA-based algorithmic test that stratifies outcomes in patients treated with ICIs, and digital pathology algorithms such as p-MSI to support biomarker assessment[176,177] (Paradigm I–II). SCORPIO Trial Platform[178] uses routine blood tests and clinical variables to predict immunotherapy outcomes, enabling risk stratification and earlier identification of patients at risk of poor benefit or early progression (Paradigm I). Together, these examples underscore how integrating complementary clinical, molecular, and tissue-derived signals can improve patient stratification and prognostication for immunotherapy, ultimately contributing to more effective and tailored clinical decision-making (Box 2).

Conclusions and future directions

Predicting immunotherapy response remains an open, high-impact problem because it requires models that are accurate, generalizable, and clinically actionable. The next phase of the field should prioritize prospective validation in diverse populations, with clear definitions of endpoints, treatment context, and data standards. Community benchmarking resources, transparent reporting, and rigorous external validation will be essential to distinguish true signal from cohort-specific artifacts and to enable fair comparisons across methods.

Methodologically, progress will likely come from hybrid approaches that combine multimodal foundation models with mechanistic constraints. Key needs include robust multimodal fusion across incomplete data, uncertainty quantification to communicate confidence at the patient level, and interpretability tools that support clinical review rather than post hoc explanations. Equally important is moving beyond prediction toward decision support, including causal inference for treatment selection, counterfactual reasoning for combination strategies, and adaptive modeling that updates with longitudinal data.

Finally, translation will depend on practical deployment: reproducible pipelines, privacy-preserving learning across institutions, model monitoring for drift, and alignment with regulatory and ethical requirements. If these technical and clinical gaps are addressed, computational modeling can evolve from retrospective association to prospective, simulation-guided personalization of immunotherapy.

References

1. Abbosh, C. et al. Phylogenetic ctDNA analysis depicts early-stage lung cancer evolution. *Nature* **545**, 446–451 (2017).
2. Hugo, W. et al. Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma. *Cell* **165**, 35–44 (2016).
3. Hodi, F. S. et al. Improved survival with ipilimumab in patients with metastatic melanoma. *N. Engl. J. Med.* **363**, 711–723 (2010).
4. Topalian, S. L. et al. Five-Year Survival and Correlates Among Patients With Advanced Melanoma, Renal Cell Carcinoma, or Non-Small Cell Lung Cancer Treated With Nivolumab. *JAMA Oncol.* **5**, 1411–1420 (2019).
5. Motzer, R. J. et al. Nivolumab versus everolimus in patients with advanced renal cell carcinoma: Updated results with long-term follow-up of the randomized, open-label, phase 3 CheckMate 025 trial. *Cancer* **126**, 4156–4167 (2020).
6. Rizvi, N. A. et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* **348**, 124–128 (2015).
7. Postow, M. A., Sidlow, R. & Hellmann, M. D. Immune-Related Adverse Events Associated with Immune Checkpoint Blockade. *N. Engl. J. Med.* **378**, 158–168 (2018).
8. Tran, G. & Zafar, S. Y. Financial toxicity and implications for cancer care in the era of molecular and immune therapies. *Ann. Transl. Med.* **6**, 166 (2018).
9. Gerlinger, M. et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* **366**, 883–892 (2012).
10. Riaz, N. et al. Tumor and Microenvironment Evolution during Immunotherapy with Nivolumab. *Cell* **171**, 934–949 e916 (2017).
11. Routy, B. et al. Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. *Science* **359**, 91–97 (2018).
12. Groha, S. et al. Germline variants associated with toxicity to immune checkpoint blockade. *Nat. Med.* **28**, 2584–2591 (2022).
13. Ju, Y. et al. Barriers and opportunities in pancreatic cancer immunotherapy. *NPJ Precis. Oncol.* **8**, 199 (2024).
14. Chen, E. & Zhou, W. Immunotherapy in microsatellite-stable colorectal cancer: Strategies to overcome resistance. *Crit. Rev. Oncol. Hematol.* **212**, 104775 (2025).
15. Kang, H. et al. A Comprehensive Benchmark of Transcriptomic Biomarkers for Immune Checkpoint Blockades. *Cancers* **15**, https://doi.org/10.3390/cancers15164094 (2023).
16. Sun, S. et al. Systematic Assessment of Transcriptomic Biomarkers for Immune Checkpoint Blockade Response in Cancer Immunotherapy. *Cancers* **13**, https://doi.org/10.3390/cancers13071639 (2021).
17. Oisakede, E. O. et al. Predictive models for immune checkpoint inhibitor response in cancer: A review of current approaches and future directions. *Crit. Rev. Oncol. Hematol.* **216**, 104980 (2025).
18. Vanguri, R. S. et al. Multimodal integration of radiology, pathology and genomics for prediction of response to PD-(L)1 blockade in patients with non-small cell lung cancer. *Nat. Cancer* **3**, 1151–1164 (2022).
19. Captier, N. et al. Integration of clinical, pathological, radiological, and transcriptomic data improves prediction for first-line immunotherapy outcome in metastatic non-small cell lung cancer. *Nat. Commun.* **16**, 614 (2025).
20. Li, Y., Wu, X., Fang, D. & Luo, Y. Informing immunotherapy with multi-omics driven machine learning. *NPJ Digit Med.* **7**, 67 (2024).
21. Rashid, M. M. & Selvarajoo, K. Advancing drug-response prediction using multi-modal and -omics machine learning integration

(MOMLIN): a case study on breast cancer clinical data. *Brief Bioinform.* **25**, https://doi.org/10.1093/bib/bbae300 (2024).
22. Chen, T., Guestrin, Carlos. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794 (ACM, 2016).
23. Rumelhart, D. E., Hinton, G. E. & Williams, R. J. Learning representations by back-propagating errors. *nature* **323**, 533–536 (1986).
24. Kingma, D. P. & Welling, M. "Auto-encoding variational bayes." arXiv preprint arXiv:1312.6114 (2013).
25. Chalmers, Z. R. et al. Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med.* **9**, 34 (2017).
26. Kruger, S. et al. Advances in cancer immunotherapy 2019 - latest trends. *J. Exp. Clin. Cancer Res.* **38**, 268 (2019).
27. Chan, T. A. et al. Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. *Ann. Oncol.* **30**, 44–56 (2019).
28. Davis, A. A. & Patel, V. G. The role of PD-L1 expression as a predictive biomarker: an analysis of all US Food and Drug Administration (FDA) approvals of immune checkpoint inhibitors. *J. Immunother. Cancer* **7**, 278 (2019).
29. do Valle, I. F. et al. Optimized pipeline of MuTect and GATK tools to improve the detection of somatic single nucleotide polymorphisms in whole-exome sequencing data. *BMC Bioinforma.* **17**, 341 (2016).
30. Kim, S. et al. Strelka2: fast and accurate calling of germline and somatic variants. *Nat. Methods* **15**, 591–594 (2018).
31. McKenna, A. et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
32. Samstein, R. M. et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat. Genet* **51**, 202–206 (2019).
33. Diaz, L. A. Jr. et al. Pembrolizumab versus chemotherapy for microsatellite instability-high or mismatch repair-deficient metastatic colorectal cancer (KEYNOTE-177): final analysis of a randomised, open-label, phase 3 study. *Lancet Oncol.* **23**, 659–670 (2022).
34. Le, D. T. et al. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* **357**, 409–413 (2017).
35. Jardim, D. L., Goodman, A., de Melo Gagliato, D. & Kurzrock, R. The Challenges of Tumor Mutational Burden as an Immunotherapy Biomarker. *Cancer Cell* **39**, 154–173 (2021).
36. Pham, T. V. et al. Role of ultraviolet mutational signature versus tumor mutation burden in predicting response to immunotherapy. *Mol. Oncol.* **14**, 1680–1694 (2020).
37. Mandal, R. et al. Genetic diversity of tumors with mismatch repair deficiency influences anti-PD-1 immunotherapy response. *Science* **364**, 485–491 (2019).
38. Binnewies, M. et al. Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat. Med.* **24**, 541–550 (2018).
39. Garcia-Garijo, A., Fajardo, C. A. & Gros, A. Determinants for Neoantigen Identification. *Front. Immunol.* **10**, 1392 (2019).
40. Yi, J., Chen, L., Xiao, Y., Zhao, Z. & Su, X. Investigations of sequencing data and sample type on HLA class Ia typing with different computational tools. *Brief Bioinform.* **22**, https://doi.org/10.1093/bib/bbaa143 (2021).
41. Wang, Y. Y. et al. HLA-VBSeq v2: improved HLA calling accuracy with full-length Japanese class-I panel. *Hum. Genome Var.* **6**, 29 (2019).
42. Nielsen, M., Lundegaard, C., Lund, O. & Kesmir, C. The role of the proteasome in generating cytotoxic T-cell epitopes: insights obtained from improved predictions of proteasomal cleavage. *Immunogenetics* **57**, 33–41 (2005).
43. Orenbuch, R. et al. arcasHLA: high-resolution HLA typing from RNAseq. *Bioinformatics* **36**, 33–40 (2020).
44. Yu, D. et al. A rigorous benchmarking of alignment-based HLA typing algorithms for RNA-seq data. Preprint at https://doi.org/10.1101/2023.05.22.541750 (2024).
45. Lang, K. et al. Full-Length HLA Class I Genotyping with the MinION Nanopore Sequencer. *Methods Mol. Biol.* **1802**, 155–162 (2018).
46. Reynisson, B., Alvarez, B., Paul, S., Peters, B. & Nielsen, M. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res.* **48**, W449–W454 (2020).
47. O'Donnell, T. J., Rubinsteyn, A. & Laserson, U. MHCflurry 2.0: Improved Pan-Allele Prediction of MHC Class I-Presented Peptides by Incorporating Antigen Processing. *Cell Syst.* **11**, 42–48 e47 (2020).
48. Jensen, K. K. et al. Improved methods for predicting peptide binding affinity to MHC class II molecules. *Immunology* **154**, 394–406 (2018).
49. Vita, R. et al. The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Res.* **47**, D339–D343 (2019).
50. Vizcaino, J. A. et al. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* **44**, 11033 (2016).
51. Shao, W. et al. The SysteMHC Atlas project. *Nucleic Acids Res.* **46**, D1237–D1247 (2018).
52. Huang, H., Wang, C., Rubelt, F., Scriba, T. J. & Davis, M. M. Analyzing the Mycobacterium tuberculosis immune response by T-cell receptor clustering with GLIPH2 and genome-wide antigen screening. *Nat. Biotechnol.* **38**, 1194–1202 (2020).
53. Lu, T. et al. Deep learning-based prediction of the T cell receptor-antigen binding specificity. *Nat. Mach. Intell.* **3**, 864–875 (2021).
54. Montemurro, A. et al. NetTCR-2.0 enables accurate prediction of TCR-peptide binding by using paired TCRalpha and beta sequence data. *Commun. Biol.* **4**, 1060 (2021).
55. Sidhom, J. W., Larman, H. B., Pardoll, D. M. & Baras, A. S. DeepTCR is a deep learning framework for revealing sequence concepts within T-cell repertoires. *Nat. Commun.* **12**, 1605 (2021).
56. Chai, S. et al. NeoSplice: a bioinformatics method for prediction of splice variant neoantigens. *Bioinform Adv.* **2**, vbac032 (2022).
57. Li, G. et al. Splicing neoantigen discovery with SNAF reveals shared targets for cancer immunotherapy. *Sci. Transl. Med.* **16**, eade2886 (2024).
58. Shah, N. M. et al. Pan-cancer analysis identifies tumor-specific antigens derived from transposable elements. *Nat. Genet* **55**, 631–639 (2023).
59. Shi, Y., Jing, B. & Xi, R. Comprehensive analysis of neoantigens derived from structural variation across whole genomes from 2528 tumors. *Genome Biol.* **24**, 169 (2023).
60. Vensko, S. P. et al. LENS: Landscape of Effective Neoantigens Software. *Bioinformatics* **39**, https://doi.org/10.1093/bioinformatics/btad322 (2023).
61. Jamal-Hanjani, M. et al. Tracking the Evolution of Non-Small-Cell Lung Cancer. *N. Engl. J. Med.* **376**, 2109–2121 (2017).
62. Nik-Zainal, S. et al. The life history of 21 breast cancers. *Cell* **149**, 994–1007 (2012).
63. Miller, C. A. et al. SciClone: inferring clonal architecture and tracking the spatial and temporal patterns of tumor evolution. *PLoS Comput. Biol.* **10**, e1003665 (2014).
64. Roth, A. et al. PyClone: statistical inference of clonal population structure in cancer. *Nat. Methods* **11**, 396–398 (2014).
65. Xiao, Y. et al. FastClone is a probabilistic tool for deconvoluting tumor heterogeneity in bulk-sequencing samples. *Nat. Commun.* **11**, 4469 (2020).

66. Guo, Q. et al. MyClone: rapid and precise reconstruction of clonal population structures for tumors. *BMC Bioinforma.* **26**, 242 (2025).
67. Salcedo, A. et al. A community effort to create standards for evaluating tumor subclonal reconstruction. *Nat. Biotechnol.* **38**, 97–107 (2020).
68. Tanner, G., Westhead, D. R., Droop, A. & Stead, L. F. Benchmarking pipelines for subclonal deconvolution of bulk tumour sequencing data. *Nat. Commun.* **12**, 6396 (2021).
69. Dentro, S. C. et al. Characterizing genetic intra-tumor heterogeneity across 2,658 human cancer genomes. *Cell* **184**, 2239–2254 e2239 (2021).
70. Martinez-Ruiz, C. et al. Genomic-transcriptomic evolution in lung cancer and metastasis. *Nature* **616**, 543–552 (2023).
71. Serin Harmanci, A., Harmanci, A. O. & Zhou, X. CaSpER identifies and visualizes CNV events by integrative analysis of single-cell or bulk RNA-sequencing data. *Nat. Commun.* **11**, 89 (2020).
72. Dijkstra, K. K. et al. Subclonal immune evasion in non-small cell lung cancer. *Cancer Cell* **43**, 1833–1849 e1810 (2025).
73. Donia, M. & Svane, I. M. Harnessing neoantigen-specific T cells for precision cancer immunotherapy. *Nat. Rev. Clin. Oncol.* **21**, 253–254 (2024).
74. Luo, R. et al. A novel integrated approach to predicting cancer immunotherapy efficacy. *Oncogene* **42**, 1913–1925 (2023).
75. Zafar, H., Wang, Y., Nakhleh, L., Navin, N. & Chen, K. Monovar: single-nucleotide variant detection in single cells. *Nat. Methods* **13**, 505–507 (2016).
76. Bohrson, C. L. et al. Linked-read analysis identifies mutations in single-cell DNA-sequencing data. *Nat. Genet* **51**, 749–754 (2019).
77. Tickle, T., Christophe Georgescu, I. T., Brown, M. & Haas, B. inferCNV of the Trinity CTAT Project. https://github.com/broadinstitute/inferCNV (2019).
78. Dong, X. et al. Accurate identification of single-nucleotide variants in whole-genome-amplified single cells. *Nat. Methods* **14**, 491–493 (2017).
79. Wu, C. Y. et al. Integrative single-cell analysis of allele-specific copy number alterations and chromatin accessibility in cancer. *Nat. Biotechnol.* **39**, 1259–1269 (2021).
80. Muller, S., Cho, A., Liu, S. J., Lim, D. A. & Diaz, A. CONICS integrates scRNA-seq with DNA sequencing to map gene expression to tumor sub-clones. *Bioinformatics* **34**, 3217–3219 (2018).
81. Gao, T. et al. Haplotype-aware analysis of somatic copy number variations from single-cell transcriptomes. *Nat. Biotechnol.* **41**, 417–426 (2023).
82. Malikic, S., Jahn, K., Kuipers, J., Sahinalp, S. C. & Beerenwinkel, N. Integrative inference of subclonal tumour evolution from single-cell and bulk sequencing data. *Nat. Commun.* **10**, 2750 (2019).
83. Satas, G., Zaccaria, S., Mon, G. & Raphael, B. J. SCARLET: Single-cell tumor phylogeny inference with copy-number constrained mutation losses. *Cell Syst.* **10**, 323–332 e328 (2020).
84. Linnemann, C., Mezzadra, R. & Schumacher, T. N. TCR repertoires of intratumoral T-cell subsets. *Immunol. Rev.* **257**, 72–82 (2014).
85. Yusko, E. et al. Association of Tumor Microenvironment T-cell Repertoire and Mutational Load with Clinical Outcome after Sequential Checkpoint Blockade in Melanoma. *Cancer Immunol. Res.* **7**, 458–465 (2019).
86. Bolotin, D. A. et al. MiXCR: software for comprehensive adaptive immunity profiling. *Nat. Methods* **12**, 380–381 (2015).
87. Marcou, Q., Mora, T. & Walczak, A. M. High-throughput immune repertoire analysis with IGoR. *Nat. Commun.* **9**, 561 (2018).
88. Song, L. et al. TRUST4: immune repertoire reconstruction from bulk and single-cell RNA-seq data. *Nat. Methods* **18**, 627–630 (2021).
89. Bortone, D. S., Woodcock, M. G., Parker, J. S. & Vincent, B. G. Improved T-cell Receptor Diversity Estimates Associate with Survival and Response to Anti-PD-1 Therapy. *Cancer Immunol. Res.* **9**, 103–112 (2021).
90. Tan, C. L. et al. Prediction of tumor-reactive T cell receptors from scRNA-seq data for personalized T cell therapy. *Nat. Biotechnol.* https://doi.org/10.1038/s41587-024-02161-y (2024).
91. Szincsak, S. et al. Predicting Immunotherapy Efficacy with Machine Learning in Gastrointestinal Cancers: A Systematic Review and Meta-Analysis. *Int. J. Mol. Sci.* **26**, 5937 (2025).
92. Li, J., Dan, K. & Ai, J. Machine learning in the prediction of immunotherapy response and prognosis of melanoma: a systematic review and meta-analysis. *Front Immunol.* **15**, 1281940 (2024).
93. Abuhelwa, A. Y. et al. Machine Learning for Prediction of Survival Outcomes with Immune-Checkpoint Inhibitors in Urothelial Cancer. *Cancers* **13**, https://doi.org/10.3390/cancers13092001 (2021).
94. Negrao, M. V. et al. PD-L1 Expression, Tumor Mutational Burden, and Cancer Gene Mutations Are Stronger Predictors of Benefit from Immune Checkpoint Blockade than HLA Class I Genotype in Non-Small Cell Lung Cancer. *J. Thorac. Oncol.* **14**, 1021–1031 (2019).
95. Dong, Y. et al. Prediction of immunotherapy responsiveness in melanoma through single-cell sequencing-based characterization of the tumor immune microenvironment. *Transl. Oncol.* **43**, 101910 (2024).
96. Barkan, E. et al. Artificial intelligence-based prediction of overall survival in metastatic renal cell carcinoma. *Front Oncol.* **13**, 1021684 (2023).
97. Iivanainen, S., Ekstrom, J., Virtanen, H., Kataja, V. V. & Koivunen, J. P. Electronic patient-reported outcomes and machine learning in predicting immune-related adverse events of immune checkpoint inhibitor therapies. *BMC Med Inf. Decis. Mak.* **21**, 205 (2021).
98. Zhang, P. et al. A novel PD-1/PD-L1 pathway-related seven-gene signature for the development and validation of the prognosis prediction model for breast cancer. *Transl. Cancer Res.* **13**, 1554–1566 (2024).
99. Wang, X. et al. Identification of Immune-Related LncRNA Pairs for Predicting Prognosis and Immunotherapeutic Response in Head and Neck Squamous Cell Carcinoma. *Front. Immunol.* **12**, 658631 (2021).
100. Xiang, Y. et al. Macrophage-Related Gene Signatures for Predicting Prognosis and Immunotherapy of Lung Adenocarcinoma by Machine Learning and Bioinformatics. *J. Inflamm. Res.* **17**, 737–754 (2024).
101. Ding, Z. et al. Identification of an Immune-Related LncRNA Signature in Gastric Cancer to Predict Survival and Response to Immune Checkpoint Inhibitors. *Front. Cell Dev. Biol.* **9**, 739583 (2021).
102. Cao, R., Yuan, L., Ma, B., Wang, G. & Tian, Y. Immune-related long non-coding RNA signature identified prognosis and immunotherapeutic efficiency in bladder cancer (BLCA). *Cancer Cell Int.* **20**, 276 (2020).
103. Thorsson, V. et al. The Immune Landscape of Cancer. *Immunity* **48**, 812–830 e814 (2018).
104. Charoentong, P. et al. Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. *Cell Rep.* **18**, 248–262 (2017).
105. Newman, A. M. et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat. Biotechnol.* **37**, 773–782 (2019).
106. Wang, X., Park, J., Susztak, K., Zhang, N. R. & Li, M. Bulk tissue cell type deconvolution with multi-subject single-cell expression reference. *Nat. Commun.* **10**, 380 (2019).

107. Li, B., Zhou, X. & Kalluri, R. scPER: A Rigorous Computational Approach to Determine Cellular Subtypes in Tumors Aligned With Cancer Phenotypes From Total RNA Sequencing. *Adv. Sci.* **13**, e14502 (2025).
108. Gentles, A. J. et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nat. Med.* **21**, 938–945 (2015).
109. Li, T. et al. TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. *Cancer Res.* **77**, e108–e110 (2017).
110. Cable, D. M. et al. Robust decomposition of cell type mixtures in spatial transcriptomics. *Nat. Biotechnol.* **40**, 517–526 (2022).
111. Elosua-Bayes, M., Nieto, P., Mereu, E., Gut, I. & Heyn, H. SPOTlight: seeded NMF regression to deconvolute spatial transcriptomics spots with single-cell transcriptomes. *Nucleic Acids Res.* **49**, e50 (2021).
112. Lee, K. A. et al. Cross-cohort gut microbiome associations with immune checkpoint inhibitor response in advanced melanoma. *Nat. Med.* **28**, 535–544 (2022).
113. Dora, D. et al. Implication of the gut microbiome and microbial-derived metabolites in immune-related adverse events: emergence of novel biomarkers for cancer immunotherapy. *Int. J. Mol. Sci.* **24**, 2769 (2023).
114. Friedman, J. & Alm, E. J. Inferring correlation networks from genomic survey data. *PLoS Comput. Biol.* **8**, e1002687 (2012).
115. Fang, H., Huang, C., Zhao, H. & Deng, M. CCLasso: correlation inference for compositional data through Lasso. *Bioinformatics* **31**, 3172–3180 (2015).
116. Weiss, S. et al. Correlation detection strategies in microbial data sets vary widely in sensitivity and precision. *ISME J.* **10**, 1669–1681 (2016).
117. McGregor, K., Labbe, A. & Greenwood, C. M. MDiNE: a model to estimate differential co-occurrence networks in microbiome studies. *Bioinformatics* **36**, 1840–1847 (2020).
118. Babaei, P., Shoaie, S., Ji, B. & Nielsen, J. Challenges in modeling the human gut microbiome. *Nat. Biotechnol.* **36**, 682–686 (2018).
119. Heinken, A. et al. Genome-scale metabolic reconstruction of 7,302 human microorganisms for personalized medicine. *Nat. Biotechnol.* **41**, 1320–1331 (2023).
120. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
121. Litjens, G. et al. A survey on deep learning in medical image analysis. *Med. Image Anal.* **42**, 60–88 (2017).
122. Campanella, G. et al. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nat. Med.* **25**, 1301–1309 (2019).
123. Coudray, N. et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat. Med.* **24**, 1559–1567 (2018).
124. Kather, J. N. et al. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nat. Med.* **25**, 1054–1056 (2019).
125. Yamashita, R. et al. Deep learning model for the prediction of microsatellite instability in colorectal cancer: a diagnostic study. *Lancet Oncol.* **22**, 132–141 (2021).
126. Dammak, S., Cecchini, M. J., Breadner, D. & Ward, A. D. Using deep learning to predict tumor mutational burden from scans of H&E-stained multicenter slides of lung squamous cell carcinoma. *J. Med Imaging* **10**, 017502 (2023).
127. Pizurica, M. et al. Whole Slide Imaging-Based Prediction of TP53 Mutations Identifies an Aggressive Disease Phenotype in Prostate Cancer. *Cancer Res.* **83**, 2970–2984 (2023).
128. Tong, H. et al. A Machine Learning Model Based on PET/CT Radiomics and Clinical Characteristics Predicts Tumor Immune Profiles in Non-Small Cell Lung Cancer: A Retrospective Multicohort Study. *Front. Immunol.* **13**, 859323 (2022).
129. Jiang, Y. et al. Non-invasive tumor microenvironment evaluation and treatment response prediction in gastric cancer using deep learning radiomics. *Cell Rep. Med.* **4**, 101146 (2023).
130. Lu, J. et al. Predicting PD-L1 status in NSCLC patients using deep learning radiomics based on CT images. *Sci. Rep.* **15**, 12495 (2025).
131. Zhang, H. & Ogasawara, K. Grad-CAM-based explainable artificial intelligence related to medical text processing. *Bioengineering* **10**, https://doi.org/10.3390/bioengineering10091070 (2023).
132. van der Velden, B. H. M., Kuijf, H. J., Gilhuijs, K. G. A. & Viergever, M. A. Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. *Med. Image Anal.* **79**, 102470 (2022).
133. Mhanna, V. et al. Enhancing comparative T cell receptor repertoire analysis in small biological samples through pooling homologous cell samples from multiple mice. *Cell Rep. Methods* **4**, 100753 (2024).
134. Wang, M., Patsenker, J., Li, H., Kluger, Y. & Kleinstein, S. H. Language model-based B cell receptor sequence embeddings can effectively encode receptor specificity. *Nucleic Acids Res.* **52**, 548–557 (2024).
135. Moris, P. et al. Current challenges for unseen-epitope TCR interaction prediction and a new perspective derived from image classification. *Brief. Bioinforma.* **22**, bbaa318 (2021).
136. Gao, Y. et al. Pan-Peptide Meta Learning for T-cell receptor–antigen binding recognition. *Nat. Mach. Intell.* **5**, 236–249 (2023).
137. Peng, X. et al. Characterizing the interaction conformation between T-cell receptors and epitopes with deep learning. *Nat. Mach. Intell.* **5**, 395–407 (2023).
138. Diao, K. et al. Pep2TCR: Accurate prediction of CD4 T cell receptor binding specificity through transfer learning and ensemble approach. *iMetaOmics* **1**, e43 (2024).
139. Jeon, J. et al. EpicPred: predicting phenotypes driven by epitope-binding TCRs using attention-based multiple instance learning. *Bioinformatics* **41**, btaf080 (2025).
140. Leary, A. Y. et al. Designing meaningful continuous representations of T cell receptor sequences with deep generative models. *Nat. Commun.* **15**, 4271 (2024).
141. Wang, M., Fan, W., Wu, T. & Li, M. TPepRet: a deep learning model for characterizing T-cell receptors–antigen binding patterns. *Bioinformatics* **41**, btaf022 (2025).
142. Lopez, R., Regier, J., Cole, M. B., Jordan, M. I. & Yosef, N. Deep generative modeling for single-cell transcriptomics. *Nat. Methods* **15**, 1053–1058 (2018).
143. Waters, M. R. et al. GAiN: An integrative tool utilizing generative adversarial neural networks for augmented gene expression analysis. *Patterns* **5**, 100910 (2024).
144. Lan, H. et al. Integrating deep generative model with active learning for predicting immunotherapy responses in gastric cancer. *Comput. Methods Programs Biomed.* **273**, 109133 (2025).
145. Luo, M. et al. Progress on network modeling and analysis of gut microecology: a review. Appl. Environ. Microbiol. **90**, e00092–00024 (2024).
146. Reiman, D., Layden, B. T. & Dai, Y. MiMeNet: Exploring microbiome-metabolome relationships using neural networks. *PLoS Comput. Biol.* **17**, e1009021 (2021).
147. Kurtz, Z. D. et al. Sparse and compositionally robust inference of microbial ecological networks. *PLoS Comput. Biol.* **11**, e1004226 (2015).
148. Oh, M. & Zhang, L. J. S. R. DeepGeni: deep generalized interpretable autoencoder elucidates gut microbiota for better cancer immunotherapy. *Sci. Rep.* **13**, 4599 (2023).

149. Oh, M. & Zhang, L. J. S. R. DeepMicro: deep representation learning for disease prediction based on microbiome data. *Sc.i Rep.* **10**, 6026 (2020).
150. Ozturk, K., Dow, M., Carlin, D. E., Bejar, R. & Carter, H. The Emerging Potential for Network Analysis to Inform Precision Cancer Medicine. *J. Mol. Biol.* **430**, 2875–2899 (2018).
151. Leiserson, M. D. et al. Pan-cancer network analysis identifies combinations of rare somatic mutations across pathways and protein complexes. *Nat. Genet.* **47**, 106–114 (2015).
152. Kong, J. et al. Network-based machine learning approach to predict immunotherapy response in cancer patients. *Nat. Commun.* **13**, 3703 (2022).
153. Wang, J. et al. scGNN is a novel graph neural network framework for single-cell RNA-Seq analyses. *Nat. Commun.* **12**, 1882 (2021).
154. Song, Q. & Su, J. DSTG: deconvoluting spatial transcriptomics data through graph-based artificial intelligence. *Brief Bioinform.* **22**, https://doi.org/10.1093/bib/bbaa414 (2021).
155. Zhu, B. et al. CellLENS enables cross-domain information fusion for enhanced cell population delineation in single-cell spatial omics data. *Nat. Immunol.* **26**, 963–974 (2025).
156. Lamurias, A., Sereika, M., Albertsen, M., Hose, K. & Nielsen, T. D. Metagenomic binning with assembly graph embeddings. *Bioinformatics* **38**, 4481-4487 (2022).
157. Jiang, C. et al. Predicting microbe-disease associations via graph neural network and contrastive learning. *Front. Microbiol.* **15**, 1483983 (2024).
158. Wang, B. et al. Microbe-drug association prediction model based on graph convolution and attention networks. *Sci. Rep.* **14**, 22327 (2024).
159. Peng, L. et al. Predicting potential microbe-disease associations with graph attention autoencoder, positive-unlabeled learning, and deep neural network. *Front. Microbiol.* **14**, 1244527 (2023).
160. Hu, W., Yang, X., Wang, L. & Zhu, X. J. MADGAN: A microbe-disease association prediction model based on generative adversarial networks. *Front. Microbiol.* **14**, 1159076 (2023).
161. Morton, J. T. et al. Learning representations of microbe–metabolite interactions. *Nat. Methods.* **16**, 1306–1314 (2019).
162. Irwin, C., Mignone, F., Montani, S. & Portinale, L. J. A. P. A. Graph Neural Networks for Gut Microbiome Metaomic data: A preliminary work. Preprint at https://arxiv.org/abs/2407.00142 (2024).
163. Saber, M. M. & Shapiro, B. Benchmarking bacterial genome-wide association study methods using simulated genomes and phenotypes. *Microb. Genom.* **6**, e000337 (2020).
164. Larsen, M. V. et al. Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. *BMC Bioinforma.* **8**, 424 (2007).
165. Lu, T. et al. Netie: inferring the evolution of neoantigen-T cell interactions in tumors. *Nat. Methods* **19**, 1480–1489 (2022).
166. Wu, J. et al. DeepHLApan: A Deep Learning Approach for Neoantigen Prediction Considering Both HLA-Peptide Binding and Immunogenicity. *Front. Immunol.* **10**, 2559 (2019).
167. Abbott, C. W. et al. Prediction of Immunotherapy Response in Melanoma through Combined Modeling of Neoantigen Burden and Immune-Related Resistance Mechanisms. *Clin. Cancer Res.* **27**, 4265–4276 (2021).
168. Norton, K. A., Gong, C., Jamalian, S. & Popel, A. S. Multiscale agent-based and hybrid modeling of the tumor immune microenvironment. *Processes* **7**, https://doi.org/10.3390/pr7010037 (2019).
169. McDonald, T. O. et al. Computational approaches to modelling and optimizing cancer treatment. *Nat. Rev. Bioeng.* **1**, 695–711 (2023).
170. Shen, S. et al. From virtual to reality: innovative practices of digital twins in tumor therapy. *J. Transl. Med.* **23**, 348 (2025).
171. Wang, H., Arulraj, T., Ippolito, A. & Popel, A. S. From virtual patients to digital twins in immuno-oncology: lessons learned from mechanistic quantitative systems pharmacology modeling. *NPJ Digit Med.* **7**, 189 (2024).
172. Xiang, J. et al. A vision-language foundation model for precision oncology. *Nature* **638**, 769–778 (2025).
173. Tripathi, A., Waqas, A., Schabath, M. B., Yilmaz, Y. & Rasool, G. HONeYBEE: enabling scalable multimodal AI in oncology through foundation model-driven embeddings. *NPJ Digit Med.* **8**, 622 (2025).
174. Jang, D. et al. Lantern: Accelerating visual autoregressive models with relaxed speculative decoding. Preprint at https://arxiv.org/abs/2410.03355 (2024).
175. Slone, J. K. et al. STAG-LLM: Predicting TCR-pHLA binding with protein language models and computationally generated 3D structures. *Comput. Struct. Biotechnol. J.* (2025).
176. Beaubier, N. et al. Clinical validation of the tempus xT next-generation targeted oncology sequencing assay. *Oncotarget* **10**, 2384–2396 (2019).
177. Boulos, H. et al. Analytical Validation of Next-Generation Sequencing-Based Comprehensive Liquid Biopsy Assay for Therapy Selection. *J. Mol. Diagn.* **27**, 383–394 (2025).
178. Yoo, S. K. et al. Prediction of checkpoint inhibitor immunotherapy efficacy for cancer using routine blood tests and clinical data. *Nat. Med.* **31**, 869–880 (2025).
179. Lyu, G. Y., Yeh, Y. H., Yeh, Y. C. & Wang, Y. C. Mutation load estimation model as a predictor of the response to cancer immunotherapy. *NPJ Genom. Med.* **3**, 12 (2018).
180. Yu, Y. et al. Radiomics-based prediction of response to immune checkpoint inhibitor treatment for solid cancers using computed tomography: a real-world study of two centers. *BMC Cancer* **22**, 1241 (2022).
181. Peng, J. et al. A novel sub-regional radiomics model to predict immunotherapy response in non-small cell lung carcinoma. *J. Transl. Med.* **22**, 87 (2024).
182. Trebeschi, S. et al. Predicting response to cancer immunotherapy using noninvasive radiomic biomarkers. *Ann. Oncol.* **30**, 998–1004 (2019).
183. Hu, J. et al. Using deep learning to predict anti-PD-1 response in melanoma and lung cancer patients from histopathology images. *Transl. Oncol.* **14**, 100921 (2021).
184. Rakaee, M. et al. Deep Learning Model for Predicting Immunotherapy Response in Advanced Non-Small Cell Lung Cancer. *JAMA Oncol.* **11**, 109–118 (2025).
185. Patkar, S. et al. Predicting the tumor microenvironment composition and immunotherapy response in non-small cell lung cancer from digital histopathology images. *NPJ Precis Oncol.* **8**, 280 (2024).
186. Limeta, A., Ji, B., Levin, M., Gatto, F. & Nielsen, J. Meta-analysis of the gut microbiota in predicting response to cancer immunotherapy in metastatic melanoma. *JCI Insight* **5**, https://doi.org/10.1172/jci.insight.140940 (2020).
187. Huang, X. et al. Multi-kingdom gut microbiota analyses define bacterial-fungal interplay and microbial markers of pan-cancer immunotherapy across cohorts. *Cell Host Microbe* **31**, 1930–1943 e1934 (2023).
188. Gunjur, A. et al. A gut microbial signature for combination immune checkpoint blockade across cancer types. *Nat. Med.* **30**, 797–809 (2024).
189. Wang, X. et al. Cancer immunotherapy response prediction from multi-modal clinical and image data using semi-supervised deep learning. *Radiother. Oncol.* **186**, 109793 (2023).
190. Rosenthal, R., McGranahan, N., Herrero, J., Taylor, B. S. & Swanton, C. DeconstructSigs: delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. *Genome Biol.* **17**, 31 (2016).

191. Nik-Zainal, S. et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* **534**, 47–54 (2016).
192. Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Campbell, P. J. & Stratton, M. R. Deciphering signatures of mutational processes operative in human cancer. *Cell Rep.* **3**, 246–259 (2013).
193. Shugay, M. et al. VDJtools: Unifying Post-analysis of T Cell Receptor Repertoires. *PLoS Comput. Biol.* **11**, e1004503 (2015).
194. Kleshchevnikov, V. et al. Cell2location maps fine-grained cell types in spatial transcriptomics. *Nat. Biotechnol.* **40**, 661–671 (2022).
195. Hu, J. et al. SpaGCN: Integrating gene expression, spatial location and histology to identify spatial domains and spatially variable genes by graph convolutional network. *Nat. Methods* **18**, 1342–1351 (2021).
196. Jin, S., Plikus, M. V. & Nie, Q. CellChat for systematic analysis of cell-cell communication from single-cell transcriptomics. *Nat. Protoc.* **20**, 180–219 (2025).
197. Browaeys, R., Saelens, W. & Saeys, Y. NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat. Methods* **17**, 159–162 (2020).
198. Biancalani, T. et al. Deep learning and alignment of spatially resolved single-cell transcriptomes with Tangram. *Nat. Methods* **18**, 1352–1362 (2021).
199. Racle, J. & Gfeller, D. EPIC: A Tool to Estimate the Proportions of Different Cell Types from Bulk Gene Expression Data. *Methods Mol. Biol.* **2120**, 233–248 (2020).
200. Andrews, S. FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (2010).
201. Larson, D. E. et al. SomaticSniper: identification of somatic point mutations in whole genome sequencing data. *Bioinformatics* **28**, 311–317 (2012).
202. Koboldt, D. C. et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* **22**, 568–576 (2012).
203. Zaccaria, S. & Raphael, B. J. Accurate quantification of copy-number aberrations and whole-genome duplications in multi-sample tumor sequencing data. *Nat. Commun.* **11**, 4301 (2020).
204. Shukla, S. A. et al. Comprehensive analysis of cancer-associated somatic mutations in class I HLA genes. *Nat. Biotechnol.* **33**, 1152–1158 (2015).
205. Zhang, L. et al. TEPITOPEpan: extending TEPITOPE for peptide binding prediction covering over 700 HLA-DR molecules. *PLoS One* **7**, e30483 (2012).
206. Singh, H. & Raghava, G. P. ProPred: prediction of HLA-DR binding sites. *Bioinformatics* **17**, 1236–1237 (2001).
207. Abelin, J. G. et al. Defining HLA-II Ligand Processing and Binding Rules with Mass Spectrometry Enhances Cancer Epitope Prediction. *Immunity* **51**, 766–779 e717 (2019).
208. Kirchmair, A. & Finotello, F. In Silico Prediction of Tumor Neoantigens with TIminer. *Methods Mol. Biol.* **2120**, 129–145 (2020).
209. Zhou, C. et al. pTuneos: prioritizing tumor neoantigens from next-generation sequencing data. *Genome Med.* **11**, 67 (2019).
210. Fotakis, G., Rieder, D., Haider, M., Trajanoski, Z. & Finotello, F. NeoFuse: predicting fusion neoantigens from RNA sequencing data. *Bioinformatics* **36**, 2260–2261 (2020).
211. Rieder, D. et al. nextNEOpi: a comprehensive pipeline for computational neoantigen prediction. *Bioinformatics* **38**, 1131–1132 (2022).
212. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
213. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915 (2019).
214. Zheng, G. X. et al. Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8**, 14049 (2017).
215. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
216. Anders, S., Pyl, P. T. & Huber, W. HTSeq-a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
217. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
218. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinforma.* **12**, 323 (2011).
219. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).
220. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).
221. Patro, R., Mount, S. M. & Kingsford, C. Sailfish enables alignment-free isoform quantification from RNA-seq reads using lightweight algorithms. *Nat. Biotechnol.* **32**, 462–464 (2014).
222. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 e3529 (2021).
223. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
224. Svensson, V., Teichmann, S. A. & Stegle, O. SpatialDE: identification of spatially variable genes. *Nat. Methods* **15**, 343–346 (2018).
225. Sun, S., Zhu, J. & Zhou, X. Statistical analysis of spatial expression patterns for spatially resolved transcriptomic studies. *Nat. Methods* **17**, 193–200 (2020).
226. Zhao, E. et al. Spatial transcriptomics at subspot resolution with BayesSpace. *Nat. Biotechnol.* **39**, 1375–1384 (2021).
227. Dong, K. & Zhang, S. Deciphering spatial domains from spatially resolved transcriptomics with an adaptive graph attention auto-encoder. *Nat. Commun.* **13**, 1739 (2022).
228. Dong, R. & Yuan, G. C. SpatialDWLS: accurate deconvolution of spatial transcriptomic data. *Genome Biol.* **22**, 145 (2021).
229. Vahid, M. R. et al. High-resolution alignment of single-cell and spatial transcriptomes with CytoSPACE. *Nat. Biotechnol.* **41**, 1543–1548 (2023).
230. Menden, K. et al. Deep learning-based cell composition analysis from tissue expression profiles. *Sci. Adv.* **6**, eaba2619 (2020).
231. Chu, T., Wang, Z., Pe'er, D. & Danko, C. G. Cell type and gene expression deconvolution with BayesPrism enables Bayesian integrative analysis across bulk and single-cell RNA sequencing in oncology. *Nat. Cancer* **3**, 505–517 (2022).
232. Sidhom, J. W. et al. ImmunoMap: A Bioinformatics Tool for T-cell Repertoire Analysis. *Cancer Immunol. Res.* **6**, 151–162 (2018).
233. Schloss, P. D. et al. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* **75**, 7537–7541 (2009).
234. Bolyen, E. et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
235. Callahan, B. J. et al. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
236. Wood, D. E., Lu, J. & Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome Biol.* **20**, 257 (2019).
237. Truong, D. T. et al. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nat. Methods* **12**, 902–903 (2015).
238. Namiki, T., Hachiya, T., Tanaka, H. & Sakakibara, Y. MetaVelvet: an extension of Velvet assembler to de novo metagenome assembly from short sequence reads. *Nucleic Acids Res.* **40**, e155 (2012).

239. Davis, N. M., Proctor, D. M., Holmes, S. P., Relman, D. A. & Callahan, B. J. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* **6**, 226 (2018).
240. Ghaddar, B., Blaser, M. J. & De, S. Denoising sparse microbial signals from single-cell sequencing of mammalian host tissues. *Nat. Comput. Sci.* **3**, 741–747 (2023).
241. Simon, N., Friedman, J. H., Hastie, T. & Tibshirani, R. Regularization paths for Cox's proportional hazards model via coordinate descent. *J. Stat. Softw.* **39**, 1–13 (2011).
242. Liang, Q., Bible, P. W., Liu, Y., Zou, B. & Wei, L. DeepMicrobes: taxonomic classification for metagenomics with deep learning. *NAR Genom. Bioinform* **2**, lqaa009 (2020).
243. Knights, D. et al. Bayesian community-wide culture-independent microbial source tracking. *Nat. Methods* **8**, 761–763 (2011).
244. Rampelli, S. et al. G2S: A New Deep Learning Tool for Predicting Stool Microbiome Structure From Oral Microbiome Data. *Front. Genet.* **12**, 644516 (2021).

Acknowledgements

R.L. was partially supported by the Center of Excellence-International Collaboration Initiative Grant, West China Hospital, Sichuan University (139170052). K.H., J.L., and X.Z. were partially supported by the National Institutes of Health (R01LM014156, R01CA241930, R01GM153822 and R01AA032723), CPRIT (RP250043), and the National Science Foundation (NSF2217515 and NSF2326879).

Author contributions

X.Z. conceived the original idea of this review. B.L., R.L., K.H., J.L.,W.Z., X.Z. contributed to the literature review and wrote sections of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-026-71364-5.

Correspondence and requests for materials should be addressed to Xiaobo Zhou.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work.