



Published in final edited form as:

*Immunoinformatics (Amst)*. 2021 October ; 1-2: . doi:10.1016/j.immuno.2021.100002.

## Integrating single cell sequencing with a spatial quantitative systems pharmacology model spQSP for personalized prediction of triple-negative breast cancer immunotherapy response

Shuming Zhang<sup>a,\*</sup>, Chang Gong<sup>a</sup>, Alvaro Ruiz-Martinez<sup>a</sup>, Hanwen Wang<sup>a</sup>, Emily Davis-Marcisak<sup>b,c</sup>, Atul Deshpande<sup>b</sup>, Aleksander S. Popel<sup>a,b,1</sup>, Elana J. Fertig<sup>a,b,d,1</sup>

<sup>a</sup>Department of Biomedical Engineering, Johns Hopkins University School of Medicine, Baltimore, MD, United States

<sup>b</sup>Department of Oncology and Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University School of Medicine, Baltimore, MD, United States

<sup>c</sup>Department of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, United States

<sup>d</sup>Department of Applied Mathematics and Statistics, Johns Hopkins University, Baltimore, MD, United States

### Abstract

Response to cancer immunotherapies depends on the complex and dynamic interactions between T cell recognition and killing of cancer cells that are counteracted through immunosuppressive pathways in the tumor microenvironment. Therefore, while measurements such as tumor mutational burden provide biomarkers to select patients for immunotherapy, they neither universally predict patient response nor implicate the mechanisms that underlie immunotherapy resistance. Recent advances in single-cell RNA sequencing technology measure cellular heterogeneity within cells of an individual tumor but have yet to realize the promise of predictive oncology. In addition to data, mechanistic multiscale computational models are developed to predict treatment response. Incorporating single-cell data from tumors to parameterize these computational models provides deeper insights into prediction of clinical outcome in individual patients. Here, we integrate whole-exome sequencing and scRNA-seq data from Triple-Negative Breast Cancer patients to model neoantigen burden in tumor cells as input to a spatial Quantitative

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

\*Corresponding author. szhan121@jhu.edu (S. Zhang).

<sup>1</sup>These authors jointly supervised this work.

Authors' contributions

SZ, ASP, and EJJ conceptually designed the study. SZ, CG, AR-M, and HW, built the model. ED-M and AD took part in the omics analysis. SZ performed all simulations, analyzed the simulation data, and prepared a draft of the manuscript. All authors and especially EJJ and ASP critically revised the manuscript. All authors have read and approved the final manuscript.

Declaration of Competing Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.immuno.2021.100002.

System Pharmacology model. The model comprises a four-compartmental Quantitative System Pharmacology sub-model to represent a whole patient and a spatial agent-based sub-model to represent tumor volumes at the cellular scale. We use the high-throughput single-cell data to model the role of antigen burden and heterogeneity relative to the tumor microenvironment composition on predicted immunotherapy response. We demonstrate how this integrated modeling and single-cell analysis framework can be used to relate neoantigen heterogeneity to immunotherapy treatment outcomes. Our results demonstrate feasibility of merging single-cell data to initialize cell states in multiscale computational models such as the spQSP for personalized prediction of clinical outcomes to immunotherapy.

## Keywords

Systems biology; Computational model; Mathematical model; Agent-based model; Bioinformatics; Immuno-oncology; Immune checkpoint inhibitor

---

## Introduction

Breast cancer is the most frequently diagnosed and leading cause of cancer death among female population in the world [1]. Triple-Negative Breast Cancer (TNBC), which is defined as the type of breast cancer characterized by absence of expression of estrogen receptor (ER), progesterone receptor (PR) and Erb-B2 receptor tyrosine-protein kinase (HER2), has the poorest treatment outcome among all breast cancer subtypes due to lack of therapeutic molecular targets. Whereas hormone receptor-positive subtypes of breast cancer benefit from targeted therapies, the absence of receptor expression does not allow for such precision therapeutic selection in this cancer type. Immunotherapies that reinvigorate the host's immune system to eradicate cancer cells have emerged as promising alternatives to chemotherapy and targeted therapy [2]. Immune-checkpoint inhibitor (ICI) therapy or immunotherapy, including nivolumab, pembrolizumab (anti-PD1), atezolizumab, durvalumab (anti-PDL1), ipilimumab, and tremelimumab (anti-CTLA4), were studied in multiple clinical trials among patients with TNBC to investigate efficacy of immunotherapy in either monotherapy or combination therapy [3]. The Objective Response Rate (ORR) ranges from 4.8% to 62% for all recent TNBC clinical trials involving different types of ICIs including combinations with chemotherapy [4-8]. The wide range of ORR suggests the need for effective, mechanistic biomarkers to predict treatment outcome for individual patients. (All abbreviations are presented in Appendix 1.)

The interplay between the immune cells and malignant cells within the tumor ultimately drives successful responses to ICI. Within the tumor, tumor mutational burden (TMB), defined as the total number of somatic mutations per megabase in tumor genome, has been recognized as a biomarker to predict effectiveness of immunotherapy in multiple clinical trials [9,10]. The greater number of mutations reflected in higher TMB correlates with a higher probability of displaying neoantigens, which can be recognized by T cells to elicit immune response [11]. However, other studies have shown high TMB cannot guarantee patients' responses to immunotherapy due to either the insufficient immune cell infiltration or therapeutic resistance resulting from tumor cell heterogeneity [12,13].

Tumor heterogeneity has three major sources: genetic, phenotypic, and microenvironmental. Intratumoral heterogeneity leads to the acquired resistance during treatments, and intertumoral heterogeneity leads to different patients' responses to the same treatment. The spatial heterogeneity, specifically the structural arrangement of epithelial cells, stromal cells, and vasculature intrinsically impact the immune cell infiltration, which significantly influences the immunotherapy efficacy [14]. Experimental studies have further demonstrated that high expression of diverse neoantigens can lead to a reduced immune attack over systems with similar expression of a homogeneous population of neoantigens [15]. Thus, predicting individual outcomes to ICI could be enhanced by extending from population-level biomarkers to modeling the impact of this tumor- and immune-cell heterogeneity on individual tumors.

Advanced omics technologies spanning DNA, RNA, and proteomic scales have enabled researchers to gain deeper insights into tumor heterogeneity at the individual patient level. Many groups characterized immune cell landscapes through single-cell RNA sequencing data (scRNA-seq) for various types of cancer [16-20]. These technologies can be used to determine immune cell compositions in the tumor microenvironment (TME) along with the molecular states of tumor cells [21]. These technologies have been widely applied to study the state of breast tumors [16]. Likewise, resolved molecular profiling technologies provide further opportunities to characterize the tumor microenvironment [22,23], but are only just emerging at a high-dimensional molecular resolution to characterize the pathways in both tumor and immune cells [24,25]. Characterizing the heterogeneous molecular and cellular states with greater resolution helps not only to identify novel biomarkers, but also to understand the significance of these pathways and intercellular interactions in ICI response.

At the same time as high-throughput tumor atlases are emerging, mechanistic computational models are also developing as powerful tools to predict patients' responses to cancer immunotherapy [26]. For instance, Quantitative System Pharmacology (QSP) models that simulate biological processes, pharmacokinetics (PK), and pharmacodynamics (PD) of selected drugs, have become an indispensable tool for drug discovery and designing dosing regimens [27-29]. QSP models are often validated by results of clinical trials to reflect their predictive power [30-33]. Despite QSP models' ability to reasonably reproduce clinical outcomes at population level, the models are unable to characterize cellular or spatial heterogeneity for individual patients due to their compartmental nature and their lumped representation of tumors. Several authors developed multiscale agent-based models (ABM) to simulate spatiotemporal tumor progression and simulation results can be visualized with single cell resolution [34-38]. The addition of spatial dimension enables direct comparison between patient-specific digital pathology data and model predictions [39-41]. With finer granularity and discrete cell states, the spatially resolved model provides a deeper understanding of intratumoral heterogeneity [42]. Further integrating omics data into QSP models can parameterize these models for individual patients, providing the prospect to simulate a virtual patients' longitudinal response to various therapeutic regimens. Lazarou et al. proposed integrating omics data into QSP models at multiscale levels (tissue, cellular, and molecular) [43]. Johnson et al. integrated single cell RNA sequencing data with mechanistic models to improve the predictive accuracy of chemotherapy responses [44]. The robust

characterization of cellular heterogeneity by single cell technologies makes them ideally suited for integration with QSP models predicting ICI response.

In this study, we extend our spatial Quantitative System Pharmacology (spQSP) model to integrate single-cell RNA-sequencing data. We incorporate single-cell sequencing data and whole-exome sequencing data of TNBC tumors from Chung et al. into a spQSP model to enable patient-specific simulations [45]. We leverage this model to utilize the single cell data to simulate the impact of tumor neoantigen profiles and neoantigen heterogeneity on response to immunotherapy. Altogether, this spQSP model provides a new platform to combine omics data with the computational QSP model to simulate response to ICI administration.

## Methods

### spQSP model for TNBC

The spQSP model consists of two sub-models: a whole-patient QSP model and a tumor compartment-specific agent-based model (ABM). The QSP model is comprised of four major compartments: tumor, tumor-draining lymph nodes (TDLN), central, and peripheral. The tumor compartment models cancer cell proliferation and anti-tumoral activities. TDLN compartment simulates naïve CD8+ T cell priming initiated by tumor neoantigen on antigen-presenting cells (APCs) followed by T cell expansion. Central compartment represents blood vessels in human body, transporting endogenous molecules, cells, and drugs to different parts of the body. Peripheral compartment represents other organs in the body [28,30,31]. In total, the QSP model contains 120 variables, 230 parameters, and 154 reactions. The ABM, representing a  $1\text{ mm} \times 1\text{ mm} \times 1\text{ mm}$  region of interest in the tumor, simulates spatio-temporal molecular and cellular interactions in a three-dimensional space [34]. Specifically, we modified the original framework to further include the effect of myeloid derived suppressor cells (MDSCs) and antigen recognition to better tailor this model to breast cancer. A schematic of the spQSP model comprising a whole-patient compartmental ordinary differential equation-based QSP model and a spatial agent-based model (ABM) representing a region-of-interest tumor volume is shown in Fig. 1A. The control flow of spQSP simulation at each time step is shown in Fig. 1B. With each time point  $\tau$  and simulation time interval  $t$ , the ABM sub-model is updated by QSP variables at  $t = \tau$ . Next, both ABM and QSP sub-models are solved for  $t = \tau + t$ . Then, the number of recruited cells and tumor antigen production from the ABM are updated back to the QSP, so that both sub-models are synchronized at  $t = \tau + t$ .

The T cell recognition is modified from our previously published model of TNBC [31] to pair specific TCRs to neoantigens. We leveraged this new model of T cell recognition to study the impact of tumor cell heterogeneity on patient-specific immunotherapy response. The spQSP model further requires that T cells can recognize each antigen to yield effective T cell killing. Specifically, alignment between the hypervariable loop (CDR3 $\alpha$ , CDR3 $\beta$ ) and epitopes is required for immune response and killing of cancer cells, thus it is crucial to describe TCR-epitope specificity in our mathematical model [46,47]. We assume that neoantigen-specific TCR must be present in the adjacent voxels of the target cancer cell to initiate cancer apoptosis, and the graphical illustration is shown in Supplemental Figure 3.

The comprehensive formulation of the spQSP model is described in Appendix 2, including all equations, parameters, mechanisms, and cases simulated, and all model-related parameters are presented in the Supplemental Material.

### **Genomic data availability and neoantigen identification**

To identify tumor neoantigens and their expression in each cancer cell, we downloaded TNBC scRNA-seq data and WES from Chung et al. [45], available in the NCBI Gene Expression Omnibus database under the accession code GSE75688. Briefly, the WES data were used to define neoantigens following the steps outlined below. Then the scRNA-seq data from the tumor cells were used to quantify the heterogeneity of expression of identified neoantigens. The schematic workflow of neoantigen identification is presented in Fig. 1C. All genes identified as neoantigen genes from this pipeline were filtered for inclusion in TSNAdb, a database that stores all immunogenic mutations from The Cancer Genome Atlas (TCGA), and the immune response was confirmed by Immune Epitope Database (IEDB) [48]. This analysis yields patient-specific estimates of neoantigen expression, able to model the heterogeneity of neoantigens tumor cells for that patient.

### **MHC (HLA) selection**

As most immunotherapies restore cytotoxic activity of CD8+ T lymphocytes, our QSP model focuses on the effect of CD8+ T cells. Therefore, we primarily focus on MHC-I binding with epitopes. The raw sequencing data for the WES data are not available, challenging direct MHC estimation. Therefore, to best ensure MHC-I alleles that we selected are expressed across the population, we chose 16 MHC-I alleles that are expressed in more than 5% of the overall population (shown in Supplemental Table 1) from 1000 Genome HLA frequency Data [49]. Since Chung et al. study was conducted in Korea, we only rank the frequency within the Asian population assuming the majority patients in the study were Asian.

### **MHC-I binding prediction**

WES data from Chung et al. contains only single nucleotide polymorphism (SNP), so indels and frameshifts were excluded from forming neoantigens [45]. We used netMHCpan 4.0, a web tool predicting peptide-MHC affinity based on neural network approach to select MHC-epitopes [50]. The inputs are 16 HLA alleles selected from the previous step and 21-mer-peptide with the mutational site at the center, and the output is the MHC-epitope complex predicted by netMHCpan 4.0.

### **MHC-epitope immunogenicity prediction**

To ensure predicted MHC-epitopes complexes are immunogenic, we used IEDB Class I Immunogenicity web tool to predict if selected MHC-epitopes complexes can elicit immune responses. The prediction model was trained based on the experimental data in the database. The model uses “immunogenicity score” to quantify the strength of immune response elicited by epitopes [51]. Based on the experimental data, immunogenic epitopes have an average immunogenicity score 0.097 vs. non-immunogenic epitopes score 0.01. Therefore,

epitopes with immunogenicity scores higher than 0.1 are considered immunogenic, and the genes that express the peptide are identified as tumor neoantigens.

### Tumor neoantigen expression in cancer cell and T cell in ABM module

We further use the single-cell RNA-seq data from Chung et al. to model neoantigen heterogeneity within the tumor cells [45]. We apply SAVER to the single cell expression profiles to estimate expression profiles in spite of the missing data from the technical dropout in single cell RNA-seq and log transform estimated transcript per million (TPM) values [52]. We use kernel density estimation to approximate the expression distribution of one antigen across all cancer cells within a single patient, excluding cells with expression values in the bottom 10% in this estimate to account for dropout. This binary expression value is used as input to the spQSP model, using this simplification of binary tumor neoantigen expression (expressed vs. not expressed) in each cancer cell as input to the model. We use these data to calculate the Coefficient of Variation (CoV) of neoantigen frequency to quantify tumor neoantigen heterogeneity, expressed as  $\frac{\sigma_{antigen/cell}}{\mu_{antigen/cell}}$ , where

$\sigma_{antigen/cell}$  is the standard deviation of neoantigen frequency, and  $\mu_{antigen/cell}$  is the mean of neoantigen frequency. This metric enables us to represent the difference between diverse neoantigen expression within a small sub-population of tumor cells from uniform neoantigen expression across distinct subclones of cancer cells within the tumor.

### Simulated digital pathology data

To resemble digital pathology data, the spatial result from ABM sub-model is sliced every 0.05 mm in the y direction; therefore, twenty immunofluorescence (IF) panels are generated at each time point. Our simulated IF panels are compared with the multiplexed data in Keren et al. We applied the same metric from Keren et al. - mixing score, a method quantifying the separation between the immune cells and cancer cells. The mixing score is calculated based on the fraction of immune cells adjacent to cancer cells. We define cell *A* is in contact with the target cell if cell *A* is in the 2D von Neumann neighborhoods (range = 1 pixel) of the target cell [53]. All patient TME are distinguished between two categories: mixed (mixing score > 0.22) and compartmentalized (mixing score < 0.22).

### Model initialization and simulation

Since the number of cancer cells sampled in the scRNA-seq data was low (13 to 28 cancer cells per sample), we use bootstrapping to increase the number of cancer cells to 1000 as the initial condition for all simulations. The virtual patient cohort is generated to resemble realistic TNBC patient population [54,55]. A fraction of model parameters, including initial tumor volume and antigen binding affinity are varied to represent inter-patient variability (26 parameters are varied). The distributions of varied parameters are approximated based on either normal physiological data (e.g., T cell death rate, antigen binding affinity) or published clinical data on TNBC (e.g., initial tumor volume, PK/PD parameters of nivolumab). Values of varied parameters are sampled using Latin Hypercube Sampling (LHS) based on estimated distributions [31]. A set of complete model parameters is defined as a virtual patient, and each virtual patient cohort contains 100 patients in this study. Thus, the cohorts referred to as Patient 7, 8, 10, 11 below each represent 100 virtual patients.

All model parameters are provided in the Supplemental Material. In the model, anti-PD-1 immune checkpoint inhibitor nivolumab at 3mg/kg is administered to every virtual patient via bolus injection every two weeks (Supplemental Figure 4), and each simulation lasts for 200 days. The spQSP model is built with C++ language, and all simulations are run on a Linux computer cluster.

## Results

### Tumor neoantigen abundance impact on tumor progression

We adapted our previous parameterization of the spQSP model for TNBC [31] to model patient-specific responses to anti-PD-1 (nivolumab) monotherapy using combined WES data and scRNA-seq data of tumor cells from Chung et al. [45]. This dataset contains samples from 11 patients with breast cancer, of which 4 patients were diagnosed with TNBC (patient 7, 8, 10, and 11, Supplemental Figure 1). Samples from these 4 patients were used for this study, which reflect a range of antigen burdens (Patient 7: 15 neoantigens; Patient 8: 17 neoantigens; Patient 10: 6 neoantigens, and Patient 11: 4 neoantigens in total) and CoV of neoantigen frequency reflective of intratumoral heterogeneity of neoantigen expression (Patient 7: 0.140, Patient 8: 0.193, Patient 10: 0.396, Patient 11: 0.142, Supplemental Figure 2). A statistically significant difference in the CoV of neoantigen was observed between Patient 10 and Patient 11 ( $p=0.005$  by the asymptotic test from Feltz et al.) [56]. Therefore, incorporating this dataset into the tumor compartment of the spQSP model enables us to study how tumor neoantigen abundance and heterogeneity influence our simulations of immunotherapy efficacy.

Treatments were simulated for 200 days, with tumor progression snapshots being taken at Day 0, 30, and the end of treatment for all patient samples (Fig. 2, Movie 1.1 and Movie 1.2 in Supplemental Materials). Only patients 7 and 8 who had the highest neoantigen burden responded to the nivolumab therapy. Our results qualitatively agree with clinical data suggesting that immunotherapy is more effective in patients with more tumor neoantigens [57-59].

In clinical practices, TMB is used as a surrogate to tumor neoantigen as immunogenic neoantigens are more complicated to identify, and cancer genes with higher mutation rates can potentially produce more tumor neoantigens [60]. To further verify our findings, four virtual patient cohorts are generated based on each patient's neoantigen profile (400 virtual patients in total). Then, we simulate tumor progressions for all four virtual patient cohorts receiving nivolumab monotherapy with the dosing regimen of 3mg/kg every 2 weeks starting at Day 0 (Fig. 3, Supplemental Figure 5). We define a patient as a "responder" if cancer cell count is reduced by 50% compared with cancer cell counts on Day 10. Among 400 virtual patients who received nivolumab monotherapy 31, (8 %) responded. Simulated cancer cell counts were reduced significantly only in the two patients with highest neoantigen burden (Patient 7: 12 responders, Patient 8: 16 responders, Patient 10: 1 responder, and Patient 11: 2 responders). This further demonstrates our model prediction that patients with more tumor neoantigens potentially have better clinical outcomes, in qualitative agreement with clinical data [61].

Next, we use these simulations to further analyze immune biomarkers associated with individual patients. We use simulated cell type abundances to investigate whether clinically established immune biomarkers also have predictive power in our model. All values are taken at day 30 of the simulation to mimic biopsy at early-stage treatments. The density of CD8<sup>+</sup> T cells and the ratio of CD8<sup>+</sup> T to FoxP3<sup>+</sup> T cells, which are regarded biomarkers in response to immunotherapy, are higher in responders (Fig. 4A and D;  $p = 3.9 \times 10^{-16}$  and  $p = 1.7 \times 10^{-7}$  respectively, Student's t-test) [62]. Significantly higher MDSC densities are found in non-responders than responders reflecting their immune-suppressive nature (Fig. 4B,  $p = 1.3 \times 10^{-21}$ , Student's t-test). Somewhat counterintuitively, responders have higher FoxP3<sup>+</sup> density in the tumor (Fig. 4C), but this result is also consistent with the clinical data [63]. Altogether, integration of scRNA-seq data of tumor cells into the spQSP model qualitatively reproduced immune profiles that resemble clinical results and showed that tumor progression could be predicted by conventional biomarkers.

### Tumor neoantigen heterogeneity influence treatment outcome

Our simulations predict that Patient 10 would have a worse treatment outcome (1 responder) than Patient 11 (2 responders), despite the fact that Patient 10 has more tumor neoantigens than Patient 11 (6 and 4, respectively). This observation seems contradictory to the results presented above and the utility of TMB as an immunotherapy biomarker. Given that Patient 10 has the highest CoV of tumor neoantigen, we hypothesize that patients with higher tumor neoantigen heterogeneity have worse prognosis due to a lower probability of T cell recognition of an individual antigen. We test this hypothesis by increasing the number of T cell clones in the simulations of Patient 10 and 11 to levels that enable these patients to respond to nivolumab therapy. As we did in the previous section, we sampled two additional virtual patient cohorts under both Patient 10's and 11's neoantigen profiles (200 virtual patients total) and simulate tumor progression under nivolumab monotherapy for patients from both cohorts. The results show that tumors in Patient 10 acquired immunotherapy resistance at a later period of the treatment (Fig. 5A; Movie 2), whereas the resistance was not found in Patient 11's simulation (Fig. 5B; Movie 2). On average, tumors in Patient 11 have 25.7% fewer cancer cells than those in Patient 10 by the end of the treatment ( $p = 6.3 \times 10^{-3}$ , Student's t-test) (Fig. 5H, I). Although the immunotherapy response rate of patients with TNBC remains low according to clinical data, the result indicates that patients with lower neoantigen heterogeneity tend to have better responses to the nivolumab monotherapy [7,8].

To further investigate how higher tumor neoantigen heterogeneity leads to immunotherapy resistance, we focused on the neoantigen composition of cancer cells in the tumor. We found that cancer cells with fewer tumor neoantigens tend to survive throughout the treatment because they are less likely to be recognized by CD8<sup>+</sup> T cells (Fig. 5C - F, Movie 3). Since the remaining cancer cells contain fewer tumor neoantigens, T cell diversity also reduces throughout the treatment (Fig. 5E-G). To summarize, high tumor neoantigen heterogeneity leads to the reduction of the abundance of CD8<sup>+</sup> T cell in the tumor which in turn causes immunotherapy resistance. This process is better reflected in case of Patient 10. In contrast, since at least 3 neoantigens are expressed in cancer cells in Patient 11, the diversity of CD8<sup>+</sup> T cell stays stable. Therefore, the efficacy of nivolumab is preserved. This is in agreement

with clinical studies that demonstrated that higher CD8+ T cell diversity demonstrate superior responses to immunotherapy in NSCLC and mesothelioma [64,65].

We explore the simulated post-treatment cancer cell composition of both virtual patient cohorts to reinforce our findings (Supplemental Figure 6). The average antigen expressed on cancer cell is significantly reduced for simulations of Patient 10 (pre-treatment: 4.02, post-treatment: 2.47,  $p = 1.45 \times 10^{-12}$ , Student's t-test), whereas the result is not significant for simulations of Patient 11 (pre-treatment: 3.45, post-treatment: 3.35,  $p = 0.49$ , Student's t-test). Our simulations are consistent with the findings of Gejman et al. that diverse neoantigens expression leads to failure of immune-mediated cancer cell elimination [15]. Clinical results for NSCLC patients also indicate that high neoantigen heterogeneity negatively impacts treatment outcomes [66]. Our results indicate that tumor neoantigen heterogeneity, which is quantified as CoV, can be regarded as a potential biomarker to predict post-treatment outcomes.

### Results of simulations resemble spatial clinical data

While the single-cell RNA-sequencing data used to initialize our model is from dissociated cells, an advantage of the spQSP model is its ability to capture the spatial characteristics of tumors. Comparing spatial cellular data from digital pathology with the spatial distribution of cells from the spQSP simulations helps to validate the model. Thus, we qualitatively compare the simulated cell distribution from the ABM sub-model with multiplexed pathology proteomics imaging data from patients with TNBC collected by Keren et al. [23]. The data use multiplexed ion beam imaging to quantify 36 protein expressions in 41 patients with TNBC *in situ*. The study shows that patients with compartmentalized TME quantified through a mixing ratio between tumor and immune cells exhibited better overall survival rates. To compare our simulations to the structure of response reported by Keren et al., under the same metric, we calculate tumor mixing score on the simulated spatial distribution of cells on Day 60 and observed consistent results: both Patient 10's samples form mixed tumors (Sample 1 mean mixing score: 0.71, Sample 2 mean mixing score: 0.75), and they are predicted to have worse treatment outcomes (Fig. 6, Movie 4.1, 4.2) [53]. In contrast, both samples from Patient 11 have compartmentalized TME (Sample 1 mean mixing score: 0.17, Sample 2 mean mixing score: 0.12) (Fig. 6, Movie 4.3, 4.4). We hypothesize that, in compartmentalized tumors, cancer cells' growth is impeded by the immune system. Even though the results are from small sample size and are compared qualitatively to the distributions in this larger cohort, these simulations demonstrate that our spQSP model can simulate spatial distributions that qualitatively resemble clinical multiplexed data.

### Sensitivity analysis

To test the uncertainty of the spQSP model predictions, we performed a global sensitivity analysis for 18 input parameters, including Cancer cell growth rate, T cell clone per antigen, PD-L1(2) expression level, etc. Output parameters include counts of post-treatment cancer cells, CD8+ T cells, Treg, and MDSC in both QSP and ABM sub-models. We used the Partial Rank Correlation Coefficient (PRCC) to quantify the uncertainty of input parameters (Fig. 7A). In general, the output values in the QSP module have very similar sensitivity as output values in the ABM module for the same input parameter. The results are consistent

with the notion that the QSP and ABM modules are generally coupled. However, we saw discrepancies for some input parameters that will be discussed in the next section. We found that cancer cell growth rate, T cell clones per antigen, T cell killing rate, PD-L1 expression level, and maximum MDSCs are significantly correlated with post-treatment cancer cell counts (both positively and negatively). Then, we qualitatively analyzed the impacts of both neoantigen abundance and heterogeneity on the predicted post-treatment results (Fig. 7B). Three values of each parameter were selected (Neoantigen mean: 5, 11, 17; Neoantigen heterogeneity: 0.1, 0.2, 0.35). The number of simulated responders gradually increases as neoantigen burden increases (5 neoantigens: 10 responders, 11 neoantigens: 16 responders, 17 neoantigens: 22 responders,  $p$ -value = 0.005, Wilcoxon test). Our results corroborate that higher number of neoantigens correlates with better predicted prognosis. Then, we compared post-treatment outcome for low-heterogeneity group (CoV = 0.1) vs. high-heterogeneity group (CoV = 0.35) at different neoantigen levels. The impact of neoantigen heterogeneity on simulated prognosis is dependent on the total number of neoantigens per patient (5 neoantigens,  $p$  = 0.16, 11 neoantigens,  $p$  = 0.07, 17 neoantigens,  $p$  = 0.03, Wilcoxon test). The results suggest that neoantigen may play a more important role when neoantigen burden is high [61,67].

## Discussion

In this study, we investigate how tumor neoantigen burden and heterogeneity impact the efficacy of immunotherapy by incorporating high-throughput sequencing data into a computational spQSP model of immunotherapy response. This model does not attempt to reproduce clinical outcomes of any specific clinical trial; rather, conceptually, it provides a new methodology of combining high-throughput data with computational models for more personalized post-treatment outcome prediction. To enable the integration of high-throughput data, our spQSP model blends multiple mathematical modeling frameworks. First, the QSP module simulates tumor progression at organ and whole-patient levels with parameters that are specifically calibrated for TNBC. However, due to the limitation of the compartmental QSP model, spatial heterogeneity is not represented. The addition of the ABM module of the tumor and its microenvironment allows us to overcome this limitation. The discretized agents with finer granularity allow us to further differentiate phenotypic characteristics. The ABM module recapitulates the heterogeneity and spatio-temporal phenomena in the tumors. In our current spQSP platform, cancer cells have three distinctive states (stem-like, progenitor, and senescent) and express distinctive neoantigens based on available scRNA-seq data. In addition, CD8+ T-cells can be further categorized by their TCR. Therefore, the ABM module captures more realistic immune/tumor interaction by recapitulating T-cell specificity and enables direct integration of single cells from high-throughput data to initialize the model. Leveraging the available public domain data, we focused this study on the integration of single-cell data from tumor cells. Simulations based on this patient-specific distribution of tumor cells confirm that high antigen expression, as well as homogeneity in antigen expression are associated with immunotherapy response [61]. Although the initial condition of the ABM module relied on non-spatially resolved scRNA-seq data, we can still compare our simulated data with available multiplexed spatial data qualitatively. Our results demonstrate that the spQSP model reflects the spatial

distribution of tumor microenvironments that are sensitive and resistant to immunotherapy, suggesting that this model is a suitable platform for incorporating omics data with high spatiotemporal resolution.

Although integrating highly personalized data into comprehensive spatial computational models will facilitate designing optimal treatment regimens, we recognize that the current spQSP model has limitations. We observed that cancer cells grow slower between days 15 and 30, which might not reflect the real biological processes. This indicates the QSP and ABM modules might require tighter coupling in subsequent versions of the model. The need for a stricter coupling between the two modules is also reflected in the sensitivity analysis. The maximum Treg number in the tumor has a smaller impact on Treg abundance in the ABM sub-model compared to the QSP sub-model (Fig. 7A). This is due to the fact that, in the QSP module, the number of Tregs in the tumor is capped to a maximum Treg density, which was removed in the ABM sub-model since the tumor volume is not clearly defined (as in the ABM sub-model we only consider one or several regions of interest (ROI) volumes and scale the results to the whole tumor). The detailed immune cell recruitment mechanisms in both QSP and ABM modules are described in Appendix 2. In addition, the boundary between tumor and normal tissue needs to be defined in the ABM sub-model as Mi et al. had shown that the invasive front (IF) of TNBC tumors plays a significant role in forming tumor immune landscape, and the immune architecture is spatially heterogeneous especially between the IF and core regions of the tumor [68]. Including normal tissue into our current spQSP platform will help tracking the dynamics of tumor IF and better defining tumor volume in the ABM sub-model. In this project, we used the “mixing score” from Keren et al. to make a qualitative comparison between the digital pathology data and the spatial distribution from spQSP simulation [23]. Including additional biological details and modules in the spQSP model, such as tumor-associated macrophages, cancer-associated fibroblasts, NK cells, and surface protein markers, should allow quantitative comparison of the simulation results with various types of spatial data, such as multiplex digital pathology and spatial transcriptomics.

The modeling of cellular agents of the tumor in the spQSP model provides the framework that allows the integration of scRNA-seq profiles into the model. As mentioned previously, our simulations have demonstrated the feasibility of leveraging high-throughput data to simulate a patient-specific tumor microenvironment. Here, we have incorporated tumor neoantigen expression from scRNA-seq into the spQSP model, and built an extensive computational model using omics data. While this current study focused on tumor cells, future studies with comprehensive single cell characterization of both tumor and immune cells can be used to initialize the immune cell composition in the computational model directly, expanding to the spatial organization for emerging spatial molecular platforms. Beyond cell types, single-cell RNA-seq also captures additional biological features including cellular proliferation, cell state transitions, and signaling pathways not directly modeled in the agents in the tumor compartment of our spQSP model. To integrate newly discovered pathways, appropriate computational models are required to describe cell-to-cell interactions and protein expression dynamics. The conventional computational models include, but not limited to, mechanistic models [69-71], statistical models [72,73], and data-driven models [74].

All model structures require a certain extent of prior knowledge. Due to the complexity and the noise of the biological system, the prior knowledge obtained even from single-cell data can be insufficient, and pretreatment data can fail to account for evolutionary processes in an individual tumor; therefore, they are not modeled in our spQSP model. These challenges lead to inaccurate treatment outcome predictions for individual patients. A potential solution is using machine learning to train the optimal parameter sets on a defined mechanistic model [75]. Importantly, spatially-resolved computational models driven by high-throughput data should provide insights into intra- and intertumoral heterogeneity and interactions between treatments like immunotherapy and the tumor microenvironment.

## Availability of data and material

The authors confirm that the data supporting the findings of this study are available within the article and the Supplementary Material. C++ code for model generation and in silico clinical trials can be found at <https://doi.org/10.5281/zenodo.5152703> and Github (<https://github.com/popellab/spQSP-omics-2021>).

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Funding

Supported by NIH grants R01CA138264 and U01CA212007.

## References

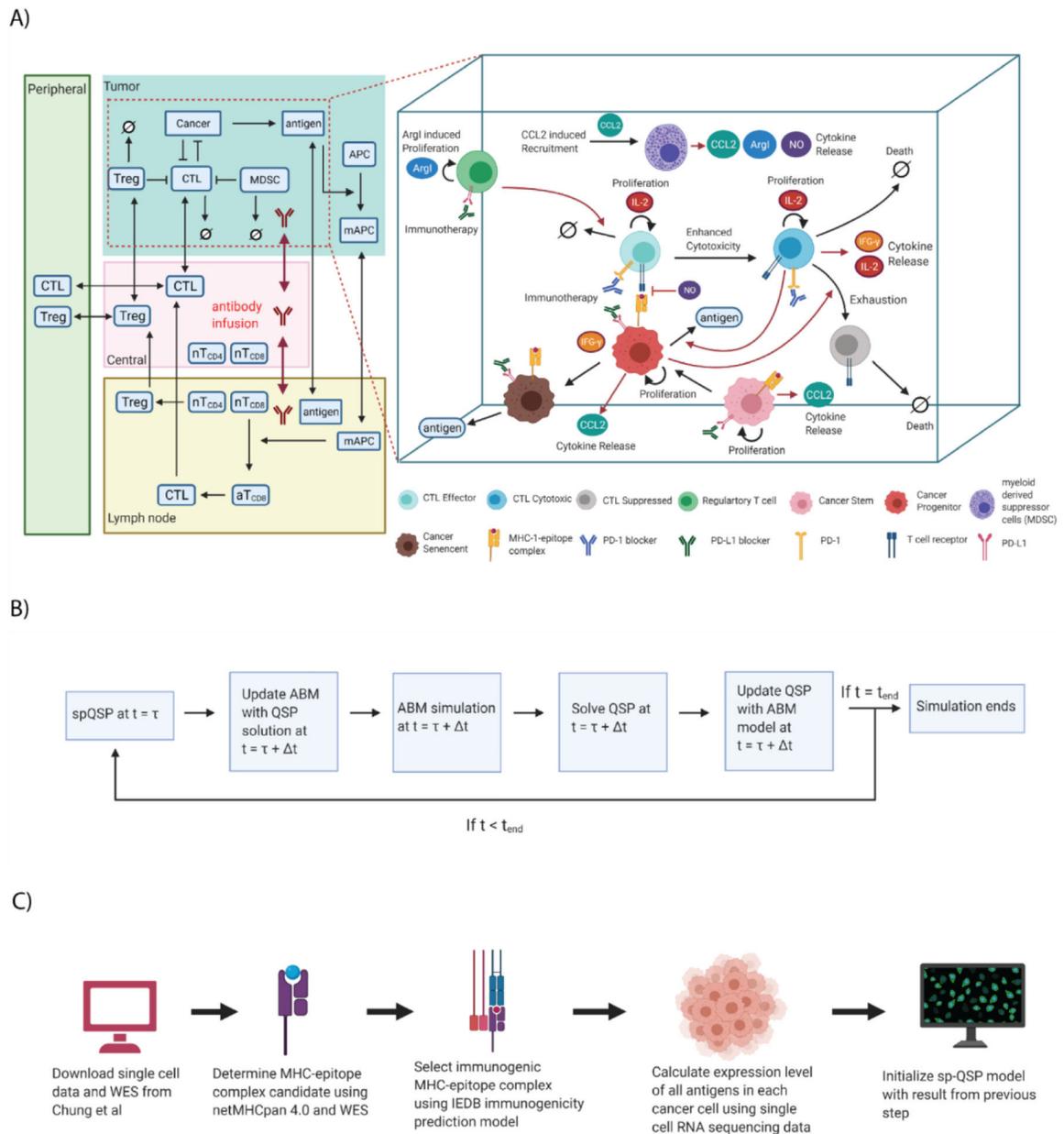
- [1]. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68(6):394–424. doi:10.3322/caac.21492. [PubMed: 30207593]
- [2]. Benvenuto M, Focaccetti C, Izzi V, Masuelli L, Modesti A, Bei R. Tumor antigens heterogeneity and immune response-targeting neoantigens in breast cancer. *Semin Cancer Biol* 2019:1–12. doi:10.1016/j.semcancer.2019.10.023.
- [3]. Bianchini G, Balko JM, Mayer IA, Sanders ME, Gianni L. Triple-negative breast cancer: Challenges and opportunities of a heterogeneous disease. *Nat Rev Clin Oncol* 2016;13(11):674–90. doi:10.1038/nrclinonc.2016.66. [PubMed: 27184417]
- [4]. Emens LA. Immunotherapy in Triple-Negative Breast Cancer. *Cancer J* 2021;27(1):59–66. doi:10.1097/PPO.0000000000000497. [PubMed: 33475294]
- [5]. Planes-Laine G, Rochigneux P, Bertucci F, Chrétien AS, Viens P, Sabatier R, et al. PD-1/PD-11 targeting in breast cancer: the first clinical evidences are emerging. a literature review. *Cancers* 2019;11(7):1–25. doi:10.3390/cancers11071033.
- [6]. Malhotra MK, Emens LA. The evolving management of metastatic triple negative breast cancer. *Semin Oncol* 2020;47(4):229–37. doi:10.1053/j.seminoncol.2020.05.005. [PubMed: 32563561]
- [7]. Schmid P, Adams S, Rugo HS, Schneeweiss A, Barrios CH, Iwata H, et al. Atezolizumab and nab-paclitaxel in advanced triple-negative breast cancer. *N Engl J Med* 2018;379(22):2108–21. doi:10.1056/nejmoa1809615. [PubMed: 30345906]
- [8]. Emens LA, Cruz C, Eder JP, Braiteh F, Chung C, Tolaney SM, et al. Long-term clinical outcomes and biomarker analyses of atezolizumab therapy for patients with metastatic triple-negative breast cancer: a phase 1 study. *JAMA Oncol* 2019;5(1):74–82. doi:10.1001/jamaoncol.2018.4224. [PubMed: 30242306]

- [9]. Fumet JD, Truntzer C, Yarchoan M, Ghiringhelli F. Tumour mutational burden as a biomarker for immunotherapy: current data and emerging concepts. *Eur J Cancer* 2020;131:40–50. doi:10.1016/j.ejca.2020.02.038. [PubMed: 32278982]
- [10]. Brown SD, Warren RL, Gibb EA, Martin SD, Spinelli JJ, Nelson BH, et al. Neoantigens predicted by tumor genome meta-analysis correlate with increased patient survival. *Genome Res* 2014;24(5):743–50. doi:10.1101/gr.165985.113. [PubMed: 24782321]
- [11]. Jiang T, Shi T, Zhang H, Hu J, Song Y, Wei J, et al. Tumor neoantigens: from basic research to clinical applications. *J. Hematol. Oncol* 2019;12(1):1–13. doi:10.1186/s13045-019-0787-5. [PubMed: 30606227]
- [12]. Maleki Vareki S High and low mutational burden tumors versus immunologically hot and cold tumors and response to immune checkpoint inhibitors. *J Immunother Cancer* 2018;6(1):4–8. doi:10.1186/s40425-018-0479-7. [PubMed: 29368638]
- [13]. Kazdal D, Endris V, Allgäuer M, Kriegsmann M, Leichsenring J, Volckmar AL, et al. Spatial and temporal heterogeneity of panel-based tumor mutational burden in pulmonary adenocarcinoma: separating biology from technical artifacts. *J Thorac Oncol* 2019;14(11):1935–47. doi:10.1016/j.jtho.2019.07.006. [PubMed: 31349062]
- [14]. Marusyk A, Janiszewska M, Polyak K. Intratumor heterogeneity: the Rosetta stone of therapy resistance. *Cancer Cell* 2020;37(4):471–84. doi:10.1016/j.ccell.2020.03.007. [PubMed: 32289271]
- [15]. Gejman RS, Chang AY, Jones HF, Dikun K, Hakimi AA, Schietinger A, Scheinberg DA. Rejection of immunogenic tumor clones is limited by clonal fraction. *eLife* 2018;7:1–22. doi:10.7554/eLife.41090.
- [16]. Azizi E, Carr AJ, Plitas G, Cornish AE, Konopacki C, Prabhakaran S, et al. Single-cell map of diverse immune phenotypes in the breast tumor microenvironment. *Cell* 2018;174(5):1293–1308.e36. doi:10.1016/j.cell.2018.05.060. [PubMed: 29961579]
- [17]. Zheng C, Zheng L, Yoo JK, Guo H, Zhang Y, Guo X, et al. Landscape of infiltrating t cells in liver cancer revealed by single-cell sequencing. *Cell* 2017;169(7):1342–1356.e16. doi:10.1016/j.cell.2017.05.035. [PubMed: 28622514]
- [18]. Zhang J, Cai J, Bello A, Roy A, Sheng J. Model-based population pharmacokinetic analysis of Nivolumab in Chinese patients with previously treated advanced solid tumors, including non-small cell lung cancer. *J Clin Pharmacol* 2019;59(10):1415–24. doi:10.1002/jcph.1432. [PubMed: 31115908]
- [19]. Ma L, Hernandez MO, Zhao Y, Mehta M, Tran B, Kelly M, et al. Tumor cell biodiversity drives microenvironmental reprogramming in liver cancer. *Cancer Cell* 2019;36(4):418–430.e6. doi:10.1016/j.ccell.2019.08.007. [PubMed: 31588021]
- [20]. Guo X, Zhang Y, Zheng L, Zheng C, Song J, Zhang Q, et al. Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. *Nat Med* 2018;24(7):978–85. doi:10.1038/s41591-018-0045-3. [PubMed: 29942094]
- [21]. Lim B, Lin Y, Navin N. Advancing cancer research and medicine with single-cell genomics. *Cancer Cell* 2020;37(4):456–70. doi:10.1016/j.ccell.2020.03.008. [PubMed: 32289270]
- [22]. Jackson HW, Fischer JR, Zanotelli VRT, Ali HR, Mechera R, Soysal SD, et al. The single-cell pathology landscape of breast cancer. *Nature* 2020;578(7796):615–20. doi:10.1038/s41586-019-1876-x. [PubMed: 31959985]
- [23]. Keren L, Bosse M, Marquez D, Angoshtari R, Jain S, Varma S, et al. A structured tumor-immune microenvironment in triple negative breast cancer revealed by multiplexed ion beam imaging. *Cell* 2018;174(6):1373–1387.e19. doi:10.1016/j.cell.2018.08.039. [PubMed: 30193111]
- [24]. Andersson A, Larsson L, Stenbeck L, Salmén F, Ehinger A, Wu S, Lundeberg J. Spatial deconvolution of HER2-positive breast tumors reveals novel intercellular relationships. *bioRxiv* 2020:1–64. doi:10.1101/2020.07.14.200600.
- [25]. He B, Bergenstråhle L, Stenbeck L, Abid A, Andersson A, Borg Å, et al. Integrating spatial gene expression and breast tumour morphology via deep learning. *Nat Biomed Eng* 2020;4(8):827–34. doi:10.1038/s41551-020-0578-x. [PubMed: 32572199]

- [26]. Rozenblatt-Rosen O, Regev A, Oberdoerffer P, Nawy T, Hupalowska A, Rood JE, et al. The human tumor atlas network: charting tumor transitions across space and time at single-cell resolution. *Cell* 2020;181(2):236–49. doi:10.1016/j.cell.2020.03.053. [PubMed: 32302568]
- [27]. Helmlinger G, Sokolov V, Peskov K, Hallow KM, Kosinsky Y, Voronova V, et al. Quantitative systems pharmacology: an exemplar model-building workflow with applications in cardiovascular, metabolic, and oncology drug development. *CPT* 2019;8(6):380–95. doi:10.1002/psp4.12426.
- [28]. Sové RJ, Jafarnejad M, Zhao C, Wang H, Ma H, Popel AS. QSP-IO: a quantitative systems pharmacology toolbox for mechanistic multiscale modeling for immuno-oncology applications. *CPT* 2020;9(9):484–97. doi:10.1002/psp4.12546.
- [29]. Bai JPF, Earp JC, Pillai VC. Translational quantitative systems pharmacology in drug development: from current landscape to good practices. *AAPS J* 2019;21(4):1–13. doi:10.1208/s12248-019-0339-5.
- [30]. Wang H, Ma H, Sové RJ, Emens LA, Popel AS. Quantitative systems pharmacology model predictions for efficacy of atezolizumab and nab-paclitaxel in triple-negative breast cancer. *J Immunother Cancer* 2021;9(2):e002100. doi:10.1136/jitc-2020-002100. [PubMed: 33579739]
- [31]. Wang H, Sové RJ, Jafarnejad M, Rahmeh S, Jaffee EM, Stearns V, et al. Conducting a virtual clinical trial in HER2-negative breast cancer using a quantitative systems pharmacology model with an epigenetic modulator and immune checkpoint inhibitors. *Front Bioeng Biotechnol* 2020;8(2):1–16. doi:10.3389/fbioe.2020.00141. [PubMed: 32039188]
- [32]. Ma H, Wang H, Sové RJ, Wang J, Giragossian C, Popel AS. Combination therapy with T cell engager and PD-L1 blockade enhances the antitumor potency of T cells as predicted by a QSP model. *J Immunother Cancer* 2020;8(2):1–11. doi:10.1136/jitc-2020-001141.
- [33]. Jafarnejad M, Gong C, Gabrielson E, Bartelink IH, Vicini P, Wang B, et al. A computational model of neoadjuvant PD-1 inhibition in non-small cell lung cancer. *AAPS J* 2019(5):21. doi:10.1208/s12248-019-0350-x. [PubMed: 30690680]
- [34]. Gong C, Milberg O, Wang B, Vicini P, Narwal R, Roskos L, et al. A computational multiscale agent-based model for simulating spatio-temporal tumour immune response to PD1 and PDL1 inhibition. *J R Soc, Interface* 2017(134):14. doi:10.1098/rsif.2017.0320.
- [35]. Norton KA, Wallace T, Pandey NB, Popel AS. An agent-based model of triple-negative breast cancer: the interplay between chemokine receptor CCR5 expression, cancer stem cells, and hypoxia. *BMC Syst Biol* 2017;11(1):1–15. doi:10.1186/s12918-017-0445-x. [PubMed: 28061857]
- [36]. Wang Z, Butner JD, Kerketta R, Cristini V, Deisboeck TS. Simulating cancer growth with multiscale agent-based modeling. *Semin Cancer Biol* 2015;30:70–8. doi:10.1016/j.semcancer.2014.04.001. [PubMed: 24793698]
- [37]. Ji Z, Zhao W, Lin HK, Zhou X. Systematically understanding the immunity leading to CRPC progression. *PLoS Comput Biol* 2019;15(9):1–29. doi:10.1371/journal.pcbi.1007344.
- [38]. Bull JA, Mech F, Quaiser T, Waters SL, Byrne HM. Mathematical modelling reveals cellular dynamics within tumour spheroids. *PLoS Comput Biol* 2020;16(8):1–25. doi:10.1371/journal.pcbi.1007961.
- [39]. Norton KA, Jin K, Popel AS. Modeling triple-negative breast cancer heterogeneity: effects of stromal macrophages, fibroblasts and tumor vasculature. *J Theor Biol* 2018;452:56–68. doi:10.1016/j.jtbi.2018.05.003. [PubMed: 29750999]
- [40]. Almendro V, Cheng YK, Randles A, Itzkovitz S, Marusyk A, Ametller E, et al. Inference of tumor evolution during chemotherapy by computational modeling and in situ analysis of genetic and phenotypic cellular diversity. *Cell Rep* 2014;6(3):514–27. doi:10.1016/j.celrep.2013.12.041. [PubMed: 24462293]
- [41]. Wang Y, Brodin E, Nishii K, Frieboes HB, Mumenthaler SM, Sparks JL, et al. Impact of tumor-parenchyma biomechanics on liver metastatic progression: a multi-model approach. *Sci Rep* 2021;11(1):1–20. doi:10.1038/s41598-020-78780-7. [PubMed: 33414495]
- [42]. Norton KA, Gong C, Jamalian S, Popel AS. Multiscale agent-based and hybrid modeling of the tumor immune microenvironment. *Processes* 2019;7(1):1–23. doi:10.3390/pr7010037.

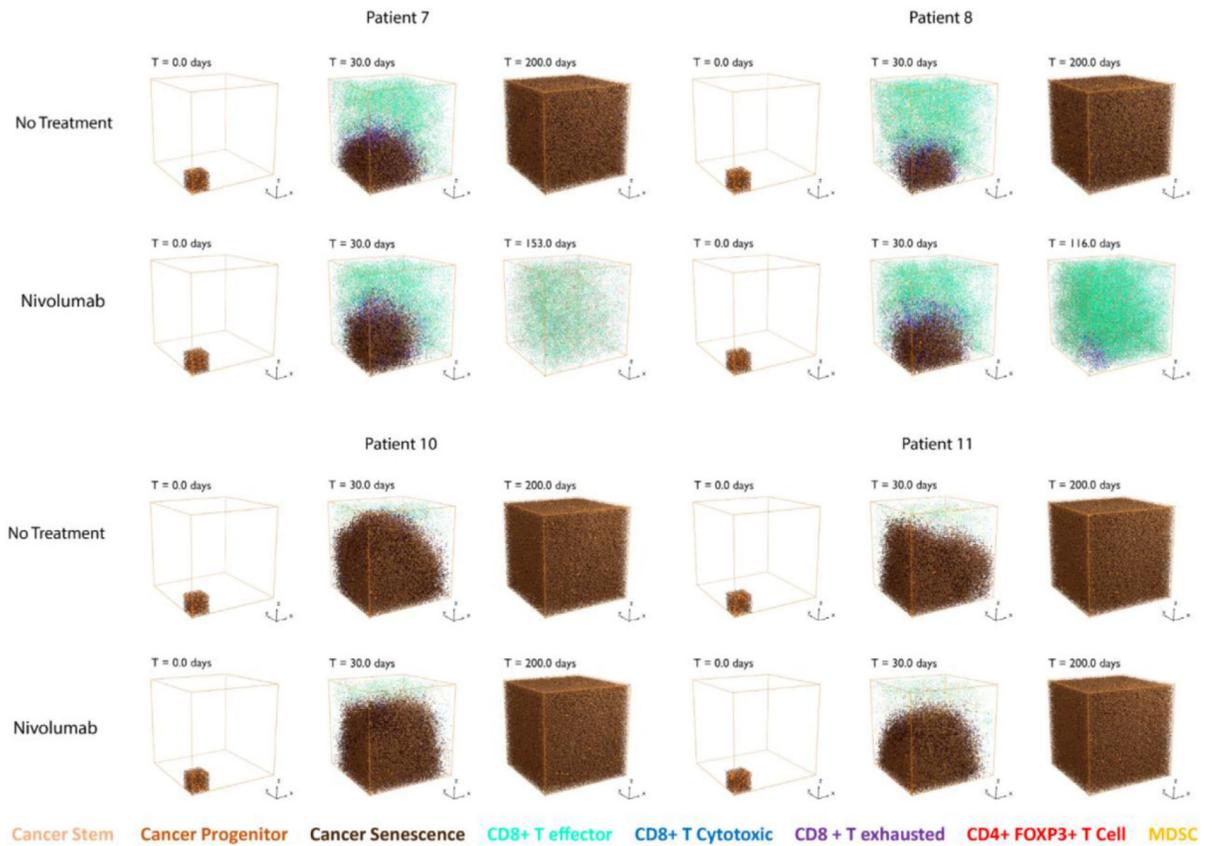
- [43]. Lazarou G, Chelliah V, Small BG, Walker M, Graaf PH, Kierzek AM. Integration of omics data sources to inform mechanistic modeling of immune-oncology therapies: a tutorial for clinical pharmacologists. *Clin Pharmacol Ther* 2020;0(0):1–13. doi:10.1002/cpt.1786.
- [44]. Johnson K, Howard GR, Morgan D, Brenner EA, Gardner AL, Durrett RE, Brock A. Integrating multimodal data sets into a mathematical framework to describe and predict therapeutic resistance in cancer. *bioRxiv* 2020.02.11.943738. doi:10.1101/2020.02.11.943738.
- [45]. Chung W, Eum HH, Lee HO, Lee KM, Lee HB, Kim KT, et al. Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. *Nat Commun* 2017;8(5):1–12. doi:10.1038/ncomms15081. [PubMed: 28232747]
- [46]. Singh NK, Riley TP, Baker SCB, Borrmann T, Weng Z, Baker BM. Emerging concepts in TCR specificity: rationalizing and (maybe) predicting outcomes. *J Immunol* 2017;199(7):2203–13. doi:10.4049/jimmunol.1700744. [PubMed: 28923982]
- [47]. Glanville J, Huang H, Nau A, Hatton O, Wagar LE, Rubelt F, et al. Identifying specificity groups in the T cell receptor repertoire. *Nature* 2017;547(7661):94–8. doi:10.1038/nature22976. [PubMed: 28636589]
- [48]. Wu J, Zhao W, Zhou B, Su Z, Gu X, Zhou Z, et al. TSNADB: A Database For Tumor-Specific Neoantigens From Immunogenomics Data Analysis. *Genom Proteom Bioinform* 2018;16(4):276–82. doi:10.1016/j.gpb.2018.06.003.
- [49]. Gourraud PA, Khankhanian P, Cereb N, Yang SY, Feolo M, Maiers M, et al. HLA diversity in the 1000 genomes dataset. *PLoS One* 2014(7):9. doi:10.1371/journal.pone.0097282.
- [50]. Jurtz V, Paul S, Andreatta M, Marcatili P, Peters B, Nielsen M. NetMHCpan-4.0: improved peptide–MHC class I interaction predictions integrating eluted ligand and peptide binding affinity data. *J Immunol* 2017;199(9):3360–8. doi:10.4049/jimmunol.1700893. [PubMed: 28978689]
- [51]. Calis JJA, Maybeno M, Greenbaum JA, Weiskopf D, De Silva AD, Sette A, et al. Properties of MHC class I presented peptides that enhance immunogenicity. *PLoS Comput Biol* 2013(10):9. doi:10.1371/journal.pcbi.1003266.
- [52]. Huang M, Wang J, Torre E, Dueck H, Shaffer S, Bonasio R, et al. SAVER: Gene expression recovery for single-cell RNA sequencing. *Nat Methods* 2018;15(7):539–42. doi:10.1038/s41592-018-0033-z. [PubMed: 29941873]
- [53]. Das D A Survey on Cellular Automata and Its Applications. *Communications in Computer and Information Science*, Vol. 269; 2011. doi:10.1007/978-3-642-29219-4\_84.
- [54]. Rieger TR, Allen RJ, Bystricky L, Chen Y, Colopy GW, Cui Y, et al. Improving the generation and selection of virtual populations in quantitative systems pharmacology models. *Prog Biophys Mol Biol* 2018;139:15–22. doi:10.1016/j.pbiomolbio.2018.06.002. [PubMed: 29902482]
- [55]. Rieger TR, Allen RJ, Musante CJ. Modeling is data driven: use it for successful virtual patient generation. *CPT* 2021;10(5):393–4. doi:10.1002/psp4.12630.
- [56]. Feltz CJ, Miller GE. An asymptotic test for the equality of coefficients of variation from k populations. *Stat Med* 1996;15(6):647–58. doi:10.1002/(SICI)1097-0258(19960330)15:6<647::AID-SIM184>3.0.CO;2.
- [57]. Goodman AM, Kato S, Bazhenova L, Patel SP, Frampton GM, Miller V, et al. Tumor mutational burden as an independent predictor of response to immunotherapy in diverse cancers. *Mol Cancer Ther* 2017;16(11):2598–608. doi:10.1158/1535-7163.MCT-17-0386. [PubMed: 28835386]
- [58]. Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, et al. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* 2015;348(6230):124–8. doi:10.1126/science.aaa1348. [PubMed: 25765070]
- [59]. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* 2015;160(1–2):48–61. doi:10.1016/j.cell.2014.12.033. [PubMed: 25594174]
- [60]. Fancello L, Gandini S, Pelicci PG, Mazzarella L. Tumor mutational burden quantification from targeted gene panels: Major advancements and challenges. *J Immunother Cancer* 2019;7(1):1–13. doi:10.1186/s40425-019-0647-4. [PubMed: 30612589]
- [61]. Osipov A, Lim SJ, Popovic A, Azad NS, Laheru DA, Zheng L, et al. Tumor mutational burden, toxicity, and response of immune checkpoint inhibitors targeting PD(L)1,

- CTLA-4, and combination: a meta-regression analysis. *Clin Cancer Res* 2020;26(18):4842–51. doi:10.1158/1078-0432.CCR-20-0458. [PubMed: 32586938]
- [62]. Liu F, Lang R, Zhao J, Zhang X, Pringle GA, Fan Y, et al. CD8+ cytotoxic T cell and FOXP3 + regulatory T cell infiltration in relation to breast cancer survival and molecular subtypes. *Breast Cancer Res Treat* 2011;130(2):645–55. doi:10.1007/s10549-011-1647-3. [PubMed: 21717105]
- [63]. Yeong J, Thike AA, Lim JCT, Lee B, Li H, Wong SC, et al. Higher densities of Foxp3 + regulatory T cells are associated with better prognosis in triple-negative breast cancer. *Breast Cancer Res Treat* 2017;163(1):21–35. doi:10.1007/s10549-017-4161-4. [PubMed: 28233108]
- [64]. Han J, Duan J, Bai H, Wang Y, Wan R, Wang X, et al. TCR repertoire diversity of peripheral PD-1<sup>hi</sup>CD8<sup>hi</sup> T cells predicts clinical outcomes after immunotherapy in patients with non-small cell lung cancer. *Cancer Immunol Res* 2020;8(1):146–54. doi:10.1158/2326-6066.CIR-19-0398. [PubMed: 31719056]
- [65]. Vroman H, Balzaretto G, Belderbos RA, Klarenbeek PL, Van Nimwegen M, Bezemer K, et al. T cell receptor repertoire characteristics both before and following immunotherapy correlate with clinical response in mesothelioma. *J Immunother Cancer* 2020;8(1):1–8. doi:10.1136/jitc-2019-000251.
- [66]. McGranahan N, Furness AJS, Rosenthal R, Ramskov S, Lyngaa R, Saini SK, et al. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science* 2016;351(6280):1463–9. doi:10.1126/science.aaf1490. [PubMed: 26940869]
- [67]. Chan TA, Yarchoan M, Jaffee E, Swanton C, Quezada SA, Stenzinger A, et al. Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. *Ann Oncol* 2019;30(1):44–56. doi:10.1093/annonc/mdy495. [PubMed: 30395155]
- [68]. Mi H, Gong C, Sulam J, Fertig EJ, Szalay AS, Jaffee EM, et al. Digital pathology analysis quantifies spatial heterogeneity of CD3, CD4, CD8, CD20, and FoxP3 immune markers in triple-negative breast cancer. *Front Physiol* 2020;11(10):1–22. doi:10.3389/fphys.2020.583333. [PubMed: 32038307]
- [69]. Zhao C, Medeiros TX, Sové RJ, Annex BH, Popel AS. A data-driven computational model enables integrative and mechanistic characterization of dynamic macrophage polarization. *iScience* 2021;102112. doi:10.1016/j.isci.2021.102112. [PubMed: 33659877]
- [70]. Zhao C, Miranda AC, Sové RJ, Medeiros TX, Annex BH, Popel AS. A mechanistic integrative computational model of macrophage polarization: Implications in human pathophysiology. *PLoS Comput Biol* 2019;15(11):1–28. doi:10.1371/journal.pcbi.1007468.
- [71]. Bouhaddou M, Barrette AM, Stern AD, Koch RJ, DiStefano MS, Riesel EA, et al. A mechanistic pan-cancer pathway model informed by multi-omics data interprets stochastic cell fate responses to drugs and mitogens. *PLoS Comput Biol* 2018;Vol. 14. doi:10.1371/journal.pcbi.1005985.
- [72]. Avanzini S, Antal T. Cancer recurrence times from a branching process model. *PLoS Comput Biol* 2019;15(11):1–30. doi:10.1371/journal.pcbi.1007423.
- [73]. Szczurek E, Krüger T, Klink B, Beerenwinkel N. A mathematical model of the metastatic bottleneck predicts patient outcome and response to cancer treatment. *PLoS Comput Biol* 2020;16(10):1–20. doi:10.1371/journal.pcbi.1008056.
- [74]. Aguilar B, Gibbs DL, Reiss DJ, McConnell M, Danziger SA, Dervan A, Shmulevich I. A generalizable data-driven multicellular model of pancreatic ductal adenocarcinoma. *GigaScience* 2020;9(7):1–15. doi:10.1093/gigascience/giaa075.
- [75]. Yuan B, Shen C, Luna A, Korkut A, Marks DS, Ingraham J, et al. Article CellBox : interpretable machine learning for perturbation biology with application to the design of cancer combination therapy CellBox : interpretable machine learning for perturbation biology with application to the design of cancer combination T. *Cell Syst* 2021:1–13. doi:10.1016/j.cels.2020.11.013. [PubMed: 33476552]

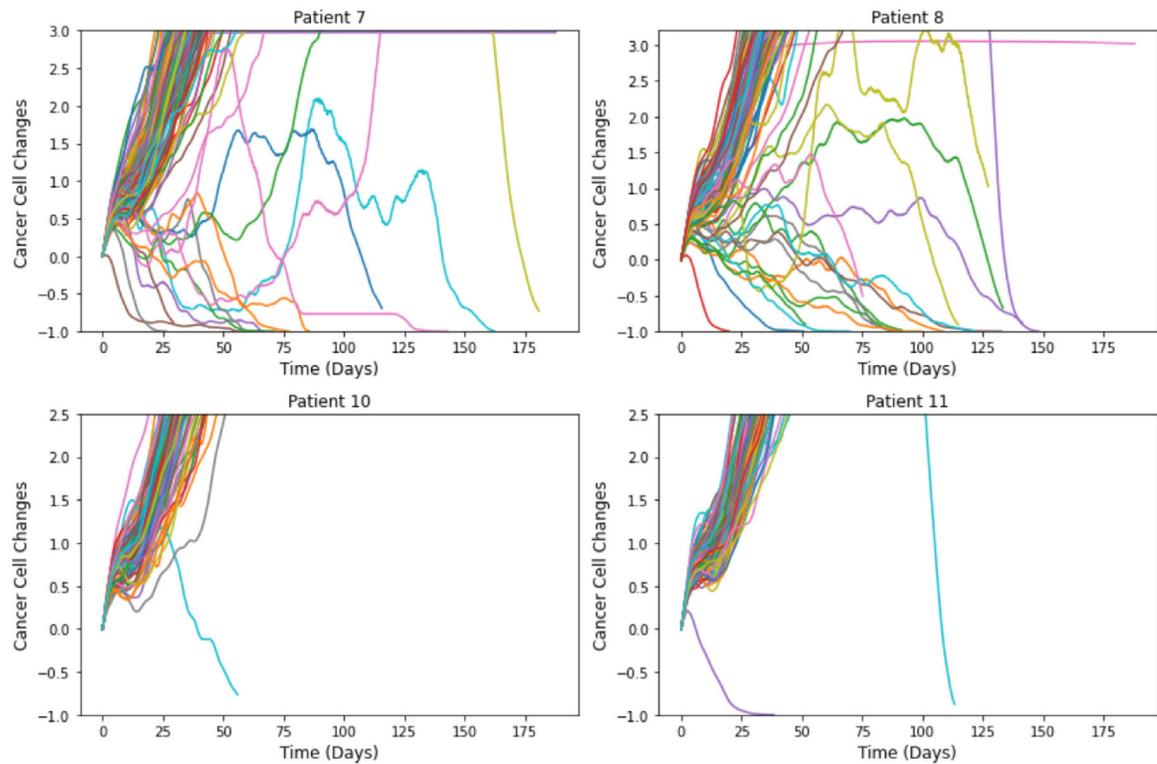
**Fig. 1.**

A) Schematic of the spQSP model comprised of a QSP sub-model and an ABM sub-model. Left: The four-compartment QSP model, including tumor (69 reactions, 19 variables), central blood (32 reactions, 9 variables), TDLNs (38 reactions, 15 variables), and peripheral (15 reactions, 8 variables). After the death of cancer cells, mature antigen-presenting cells (mAPCs) collect and process neoantigens in the tumor. mAPCs are then transported to the TDLN through the lymphatic vessels to facilitate the priming of cytotoxic T lymphocytes (CTL) and Tregs. The primed T cells are circulated in the central compartment and recruited into the tumor compartment. Right: The spatio-temporal ABM sub-model partially represents the tumor compartment, accounting for the anti-tumoral activities, immune cell suppression, and immune-checkpoint inhibitor effects with finer granularity and spatial

resolution. The spatial distribution of cytokines, including ArgI, CCL2, NO, IL-2, and IFN- $\gamma$ , are solved by PDEs. B) Flowchart of the spQSP model: starting with updating ABM model with QSP variables, the ABM is then simulated over time interval  $t$ , followed by QSP model simulation over time interval  $t$ . Finally, relevant QSP variables are updated by ABM simulation results. C) Workflow for identifying immunogenic neoantigen from single cell RNA-seq of TNBC in Chung et al, including data source, MHC-peptide binding prediction, immunogenicity prediction, neoantigen expression on cancer cell, and integration with spQSP model [45].

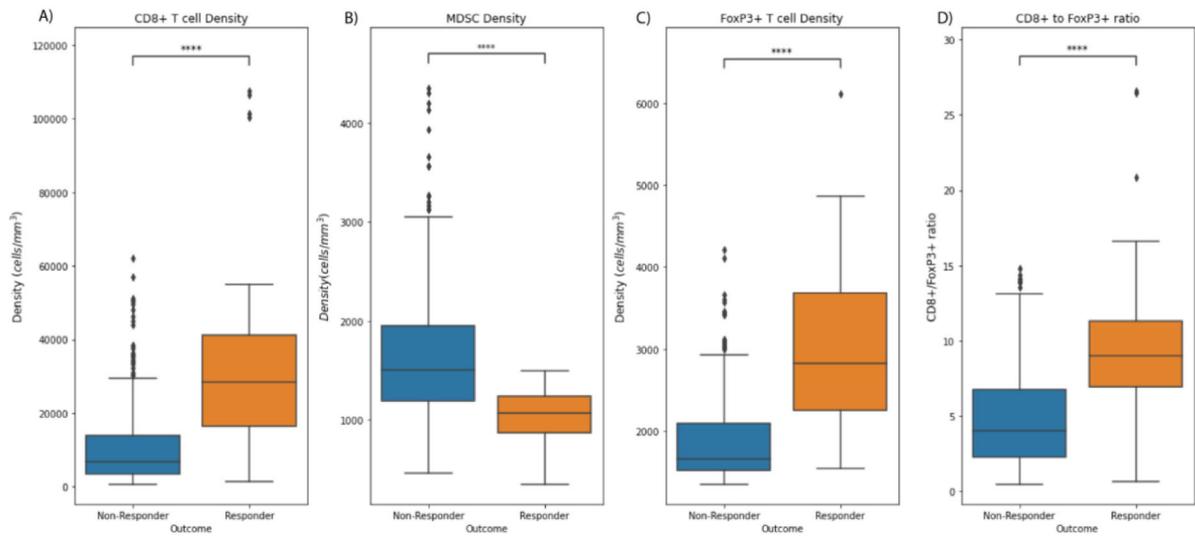


**Fig. 2.** Visualization of the ABM simulation result. Tumor growth is simulated for 200 days total (no treatment vs. nivolumab monotherapy). The dosing regimen for nivolumab is 3mg/kg every 2 weeks starting from day 0. Snapshots are taken at day 0, day 30, and end of simulation (day 200). Light pink: Stem-like cancer cell; Brown: Progenitor cancer cell; Dark Brown: Senescent cancer cell; Cyan: CD8+ effector T cell; Blue: CD8+ cytotoxic T cell; Purple: Exhausted CD8+ T cell; Red: FoxP3+ Regulatory T cell; Yellow: Myeloid Derived Suppressor cell (MDSC).



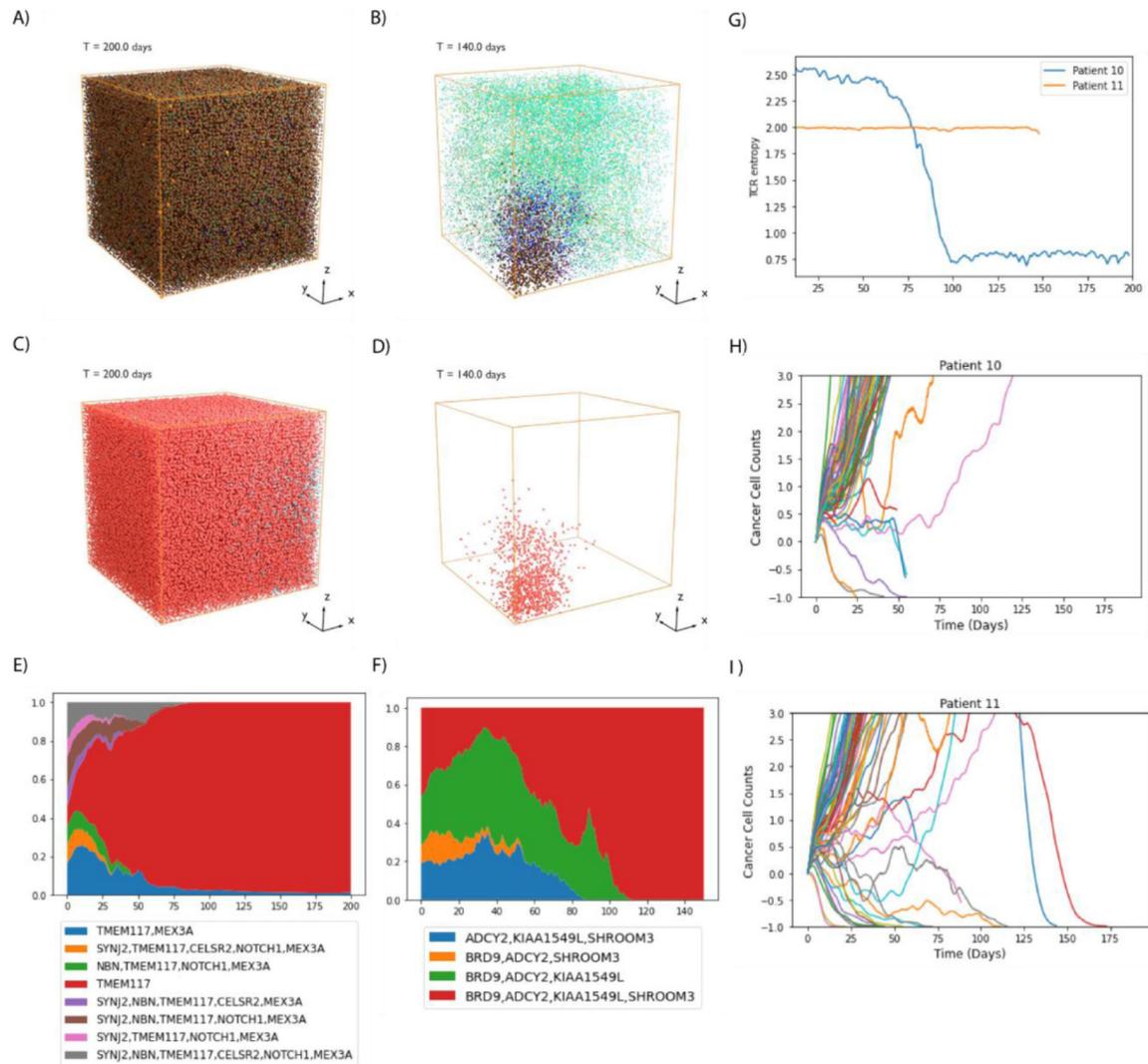
**Fig. 3.**

Spider plot of 400 selected virtual patients receiving 200-day nivolumab monotherapy treatment (3mg/kg, every 2 weeks) under four different patient neoantigen profiles. The y-axis is the relative cancer cell count changes compared to cancer cell counts at Day 10, and the time is reset to 0 on Day 10 (x-axis). The ranges of varied parameters are derived from Wang et al. (2020) to represent inter-patient variabilities. Simulated cancer cell counts were reduced significantly only in the two patients with highest neoantigen burden (compared with no-treatment, Patient 7: 12.0% reduction, p-value of 0.0011, Patient 8: 22.7% reduction, p-value of  $8.1 \times 10^{-7}$ , Patient 10: 0.9% reduction, p-value of 0.99; Patient 11: 2.9% reduction, p-value of 0.72, Student's t-test). The corresponding spider plots with absolute cell counts are shown in Supplemental Figure 5.



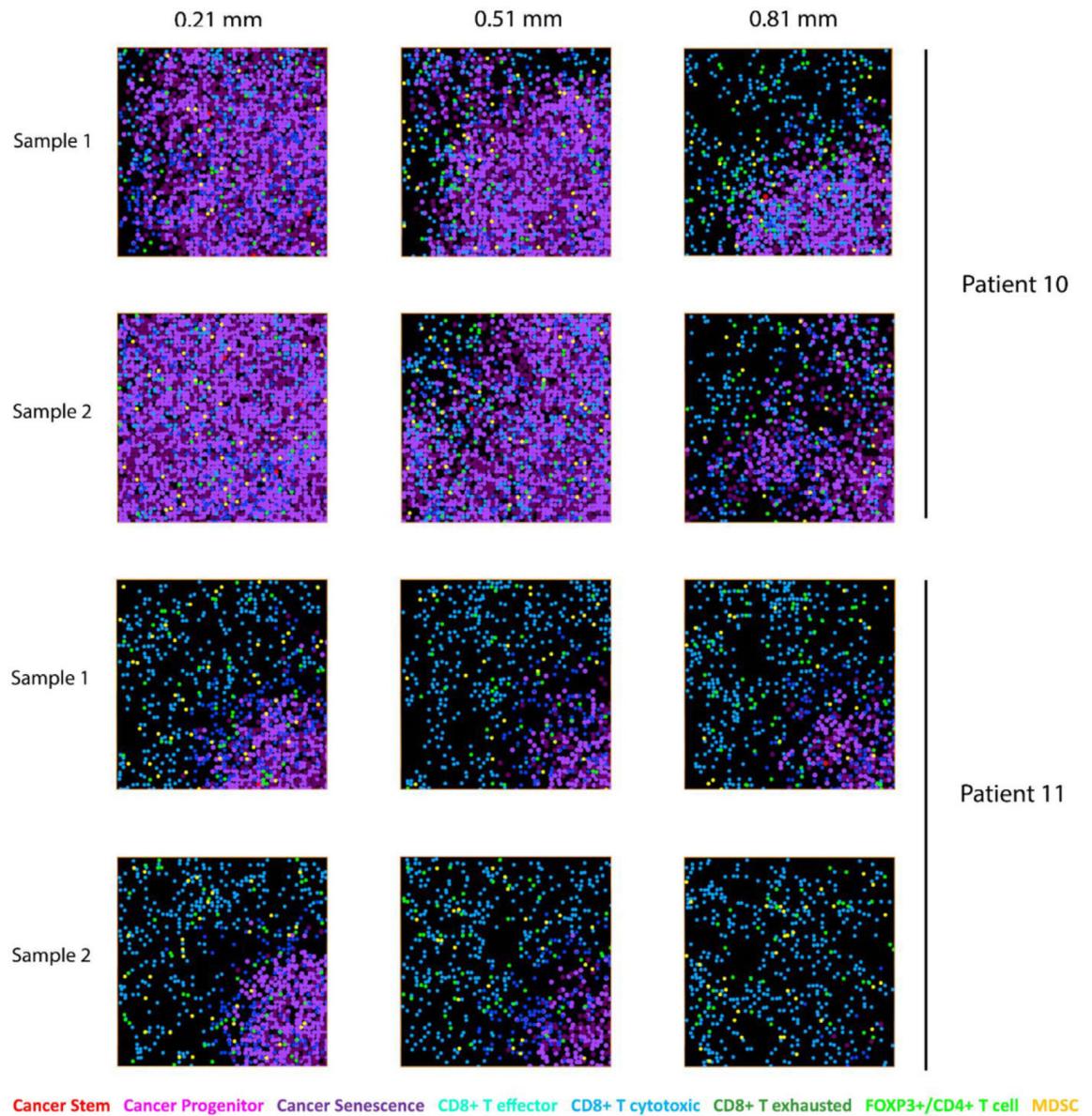
**Fig. 4.**

Immune biomarker comparison between responders and non-responders at Day 30 of the treatment. A) CD8+ T cells; B) MDSCs; C) FoxP3+ regulatory T cells; D) CD8+ T cell and FoxP3+ regulatory T cell ratio. (Student's t-test, ns:  $5.00e-02 < p < 1.00e+00$ , \*:  $1.00e-02 < p < 5.00e-02$ , \*\*:  $1.00e-03 < p < 1.00e-02$ , \*\*\*:  $1.00e-04 < p < 1.00e-03$ , \*\*\*\*:  $p < 1.00e-04$ ).



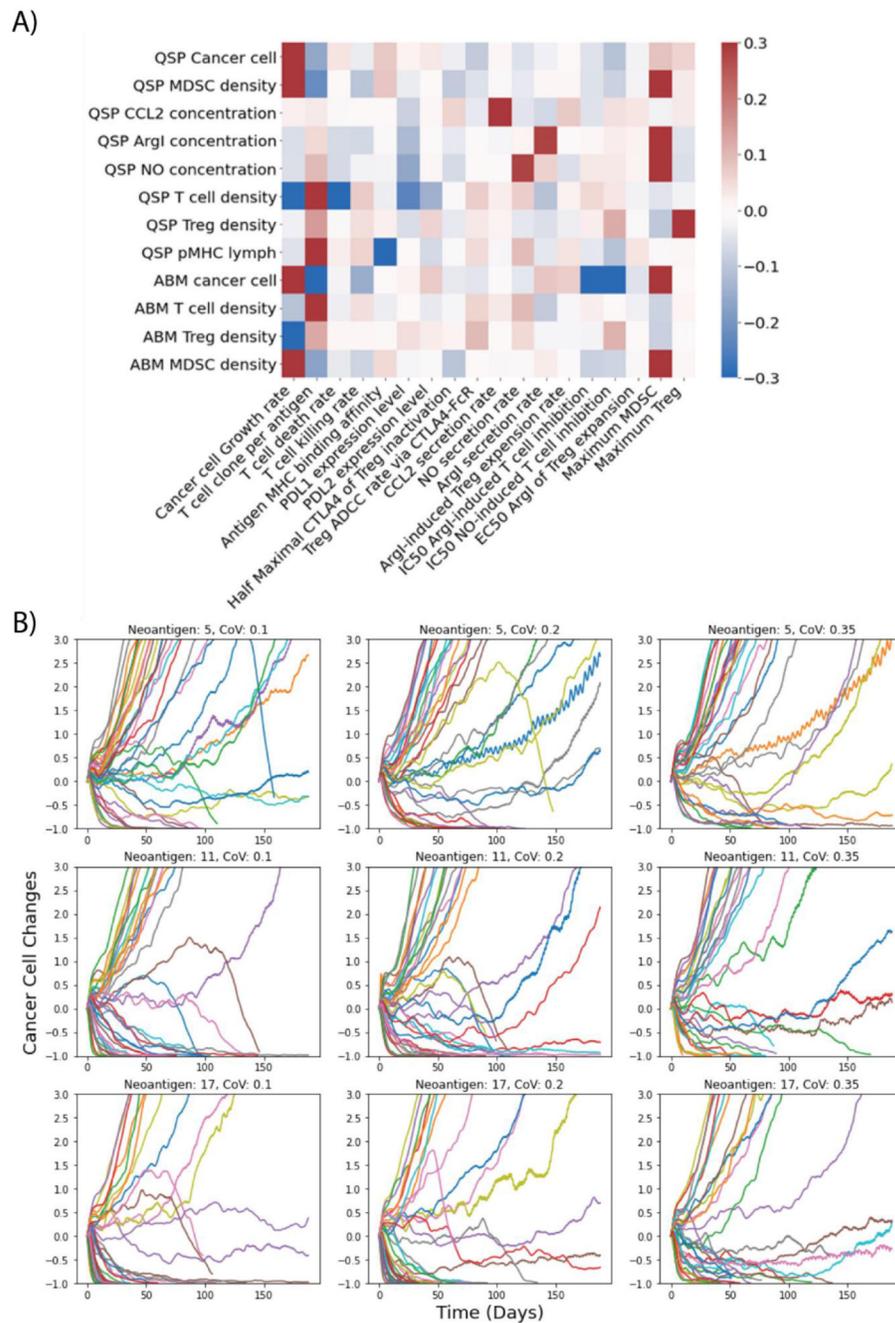
**Fig. 5.**

A, B) 3-D visualization of the whole tumor for Patient 10 and Patient 11 by the end of the treatment, respectively. C, D) Visualization of cancer cells for Patient 10 and Patient 11 by the end of the treatment, respectively. Cancer cells are colored by neoantigen clones (same legend for E and F). E, F) Cancer cell composition in the ABM module over the 200-day treatment for Patient 10 and 11, cancer cells are grouped by the set of tumor neoantigen they contain. G) Time-dependent T cell receptor (TCR) entropy over the course of nivolumab treatment. H, I) Spider plot of 100 selected virtual patients receiving 200-day nivolumab monotherapy treatment (3mg/kg, every 2 weeks) under Patient 10's and Patient 11's neoantigen profiles, respectively. The y-axis is the relative cancer cell count changes compared to cancer cell counts at Day 10, and the time is reset to 0 on Day 10 (x-axis).



**Fig. 6.**

Two samples of simulated immunofluorescence (IF) for patient 10 and 11 on Day 60. The mixing score is calculated based on the method from Keren et al (2018) (Patient 10: Sample 1 mean mixing score: 0.71, Sample 2 mean mixing score: 0.75; Patient 11: Sample 1 mean mixing score: 0.17, Sample 2 mean mixing score: 0.12). The cross section of IF is taken along the y-axis at 0.21 mm, 0.51 mm, and 0.81 mm, with 0.05 mm thickness. Red: Stem-like cancer cell, Magenta: Progenitor cancer cell, Dark Purple: Senescent cancer cell, Cyan: CD8+ Effector T cell, Light Blue: CD8+ Cytotoxic T cell, Dark Green: Exhausted CD8+ T cell, Light Green: FoxP3+ T cell, Yellow: MDSC.



**Fig. 7.** A) Global sensitivity analysis for 18 parameters sampled by Latin Hypercube Sampling (LHS) method. The Partial Rank Correlation Coefficient (PRCC) between simulated post-treatment outcome and input physiological parameters. B) Sensitivity analysis of neoantigen abundance and heterogeneity. We selected 3 values of neoantigen burdens (5, 11, 17) and 3 values of Coefficient of Variation (0.1, 0.2, 0.35). Each set of simulations represents 40 virtual patients. Each virtual patient receives 200-day nivolumab monotherapy treatment (3mg/kg, every 2 weeks) via bolus injection.