Supplementary Materials

Supplemental Information for Methods

Generation of neoantigen peptides

Tumor neoantigens were predicted and prioritized by in-house bioinformatics pipeline iNeo-SUITE, which consists modules of sequencing read filtering, genome alignment, mutation calling, HLA typing, MHC affinity prediction, gene expression profiling, vaccine peptide sequence design and prioritization based on therapeutic potency. FastQC (v0.11.4) was used for sequencing data quality control. Reads with quality score below 15 or more than 4 N bases were discarded [\[1\]](#page-5-0) . The qualified reads were then mapped to the human reference HG38 (Human Genome version 38) by using Burrows-Wheeler Aligner software (BWA, v0.7.12). Next, comparing to normal sample, tumor somatic mutations were identified by integrating the mutation calling results from Mutect (v 2.0), Varscan 2 (v .3.5.19) [\[2-5\]](#page-5-1), Strelka (1.0.11) and somatic-sniper (v1.0.5.0). Then, somatic mutation candidates were ranked based on their reliability, and were manually inspected by Integrative Genomics Viewer (IGV v.1.0.6) according to their alignment profile. Meanwhile, germline mutations in both normal and tumor samples were identified by GATK HaplotypeCaller. The database of Single Nucleotide Polymorphism (dbSNP) (https://www.ncbi.nlm.nih.gov/snp/) and 1000 Genome datasets were used to filter out high population frequency (PF) mutations (PF > 1%) from somatic mutation candidates. Next, the mutations were further annotated by Variation Effect Predictor (VEP, ensemble v89) [\[5-7\]](#page-5-2). Using the reference sequences from IMGT datasets [\[6\]](#page-5-3), HLA typing and quantification was done by OptiType (v1.3.1), Polysolver (V4), PHLAT (release 1.1) and in-house software iNeo-HLA. Subsequently, the flanking sequence of peptide or the upstream sequence of peptide were extracted from human protein database for single nucleotide mutations or frameshift and stoploss mutations. To predict the neo-epitopes within those peptides, all possible segments that contain mutation-induced amino acid(s) were further extracted with length ranging from 8 to 16 amino acids (8-11 mer for HLA class I and 12-16 mer for HLA class II). The HLA class I neoepitopes were predicted by in-house software iNeo-Pred, a deep-learning and machine-learning integrated predictor trained on datasets from IEDB and mass spectrometry (MS) profiling of HLA ligands. An epitope was considered to have binding affinity if the result from iNeo-Pred was below 500 nM. The HLA class II neo-epitopes were predicted by NetMHCIIPan (v4.0) according to the manual from official website [\[8\]](#page-5-4).

After identifying all the neo-epitope candidates, in-house software iNeo-PRIOR was used to rank mutations based on their therapeutic potency, whereas mutation prevalence, gene expression, affinity change, epitope number, and heterologous level of mutant peptide, etc. were taken into consideration. Since all these factors contribute to the final therapeutic effect, a mathematical formula was designed to integrate all these factors into a single score for prioritization:

$$
iNeo_Score = f_1(Ag) \times f_2(E) \times f_3(M_i) \times f_4(H) + f_5(M_{ii})
$$

In this formula, *iNeo_Score* refers to the score for prioritization, Ag stands for mutation prevalence, E stands for the average gene expression obtained from TCGA database, and H stands for the heterologous level of mutant peptide. M_i and M_i stand for the quality index which take affinity change and epitope number into account for epitopes presented by MHC I molecules and MHC II molecules respectively. Mutations with top ranking scores were subjected for choosing, while other factors such as the reliability of the mutation (assessed by manual examination and Sanger sequencing) and gene function (whether the mutation was in an oncogenic or cancer-driver gene) were also considered.

In-house software iNeo-DESIGN was applied to automatically designed vaccine peptide sequences (length ranging from 15 to 30 amino acids) containing neo-epitopes of both HLA class I and II. Safety issues such as potential peptide toxicity and bioactivity, as well as the difficulties in peptide synthesis were evaluated and optimized accordingly. Finally, the

2

customized long peptides were manufactured by chemical synthesis at GMP-like standard and clinical-grade (bacteria-free, >95.0% purity with less than 10 EU/mg bacterial endotoxin).

IFN-γ enzyme-linked immunospot (ELISpot) assay

Peripheral blood (10-30 mL) was collected from each patient, followed by the isolation of peripheral blood mononuclear cells (PBMCs) by Ficoll/Hypaque density-gradient centrifugation (GE Healthcare). IFN-γ ELISpot assays were performed with Human IFN-γ precoated ELISpot kit (DAKEWEI). Briefly, after adding 200μL serum-free medium into each well, the plate was incubated at room temperature for 5-10 minutes before discarding the solution. 100 μL cell suspension was added to each well at a density of 2×10^5 cells per well, followed by the addition of 5-10 μg/mL neoantigen peptide into the same well as sample or 2 μg/mL of CEF peptide as positive control. Then the mixtures were incubated at 37 °C for 16-24 hours. 200 μL pre-cooled deionized water was added into each well to lyse at 4 °C for 10 minutes. The plates were washed 6 times before the addition of 100 μL biotin-labeled antibody and then incubated at 37 °C for 1 hour. After washing the plates, 100 μL enzyme-labeled avidin working solution was added into each well and incubated at 37 °C for 1 hour. AEC solution mix was then added into each well after washing the plates, and the plates were kept in the dark for 25 minutes at room temperature before adding deionized water to stop the reaction. ELISpot plate was then placed in an automatic plate reader set with appropriate parameters, spot count and statistical analysis. The samples with more than 100 spots after noise subtraction (based on negative control group) were considered to show strong positive results, while samples with 8 to 20 spots were considered to showweak positive results.

Cytometric analysis of T-lymphocyte activity through surface biomarker

Antibodies were purchased from Biolegend, as shown in Table A. PBMCs were isolated, and T cells were labeled following manual instruction. In brief, the corresponding antibodies were added into an empty flow tube, mixed with 100 μL T cell sample thoroughly, and then incubated in the dark for 15 minutes. 2 mL of erythrocyte lysate (Zhejiang Bozhen Biotechnology Co., Ltd.) was added into the sample, mixed entirely, and then incubated in the dark for 10 minutes. The sample was centrifuged at 500×g for 5 minutes, and 1620 μL of supernatant (440 μL remained) was removed. Next, 10 μL of absolute count microspheres was added into the tube and mixed well. Cytometric analysis was conducted after sample preparation.

Table A: Antibodies for flow cytometry

Cytometric Bead Array (CBA) Analysis of Cytokines

The concentrations of serum cytokines were measured by CBA, according to the manufacture's protocol (Hangzhou Saiji Biotechnology Co., Ltd). Th1/Th2 cytokine kit was applied. In brief, 25 μL solution of captured microspheres was added into a blank flow tube, followed by the addition of 25 μL buffer solution of microspheres. The mixture was incubated in the dark for 30 minutes. 25 μL fluorescence detection reagent and 25 μL serum were added successively. The solution was vortex-mixed and then incubated in the dark for 2.5 hours. After the addition of 1mL of PBS solution, the sample was centrifuged at 200×g for 5 minutes. Following the removal of supernatant, 100 μL PBS solution was added to resuspend the sample. The samples were tested by a flow cytometer, and the acquired data were analyzed using FlowJo V10 software.

References

- 1. Hundal, J., et al., *Cancer Immunogenomics: Computational Neoantigen Identification and Vaccine Design.* Cold Spring Harb Symp Quant Biol, 2016. **81**: p. 105-111.
- 2. Chen, F., et al., *Neoantigen identification strategies enable personalized immunotherapy in refractory solid tumors.* J Clin Invest, 2019. **130**.
- 3. Hundal, J., et al., *pVAC-Seq: A genome-guided in silico approach to identifying tumor neoantigens.* Genome Med, 2016. **8**(1): p. 11.
- 4. Ott, P.A., et al., *An immunogenic personal neoantigen vaccine for patients with melanoma.* Nature, 2017. **547**(7662): p. 217-221.
- 5. Ng, A.W.R., et al., *In silico-guided sequence modifications of K-ras epitopes improve immunological outcome against G12V and G13D mutant KRAS antigens.* PeerJ, 2018. **6**: p. e5056.
- 6. Hilf, N., et al., *Actively personalized vaccination trial for newly diagnosed glioblastoma.* Nature, 2019. **565**(7738): p. 240-245.
- 7. Keskin, D.B., et al., *Neoantigen vaccine generates intratumoral T cell responses in phase Ib glioblastoma trial.* Nature, 2019. **565**(7738): p. 234-239.
- 8. Andreatta, M., et al., *Accurate pan-specific prediction of peptide-MHC class II binding affinity with improved binding core identification.* Immunogenetics, 2015. **67**(11-12): p. 641-50.

Supplementary Tables

Supplementary Table 1. The treatment process for patients before and after neoantigen vaccine therapy

Supplementary Table 2. QC metrics of whole-exome sequencing for each patient

Supplementary Table 3. Summary of the number of identified somatic mutations, predicted neoantigens and synthesized vaccine peptides

Supplementary Table 4. HLA allotypes in both normal and tumor cells for each patient

Supplementary Table 5. Summary of designed and synthesized peptides for each patient

Supplementary Table 6. The best response of each peptide/peptide pool for all 7 patients

Supplementary Table 7. Cytokine titers in peripheral blood for each patient

Supplementary Table 8. T-cell subsets before and after vaccination

Supplementary Table 9. Mutation of KRAS for each patient

Supplementary Table S2. QC metrics of whole-exome sequencing for each patient

Patient ID		P01	P ₀ 2	P ₀ 3	P ₀₄	P ₀₅	P ₀₆	P ₀ 7
Total mutations	Total number	82340	169	3330	11815	8688	5394	114
	Missense Mutation	33663	104	2045	7404	5613	3335	96
	In Frame InDel	474		17	249	51	40	14
	Frame Shift InDel	566	4	53	360	55	111	3
	Nonstop Mutation	104	θ	33	51	20	41	
	Other	47533	54	1182	3751	2949	1867	θ
Mutations with more than one neoantigen	Total number	1808	111	2131	4957	5711	3493	104
	Missense Mutation	1751	103	2077	4615	5598	3316	96
	In Frame InDel	22	5	14	135	41	34	8
	Frame Shift InDel	31	3	40	179	52	105	θ
	Nonstop Mutation	4	Ω	Ω	28	20	38	θ
Predicted	Class I	24800	1082	38792	66616	69884	54049	760
epitopes	Class II	71175	3957	152515	189842	233565	229590	766
Number of peptides included in vaccine		12	11	12	13	4	13	5

Supplementary Table S3. Summary of the number of identified somatic mutations, predicted neoantigens and synthesized vaccine peptides

Supplementary Table S4. HLA allotypes in both normal and tumor cells for each patient

Supplementary Table 5. Summary of designed and synthesized peptides for each patient

Supplementary Table 7. Cytokine titers in peripheral blood for each patient

Supplementary Table 8. T-cell subsets before and after vaccination

Patient ID	KRAS mutation	Chr	Position	Ref bp	Alt bp	AA Change	Mutation type	Included in iNeo- Vac-P01
P01	No		$\overline{}$					
P ₀ 2	Yes	chr12	25380275	T	A	O61H	SNV	Yes
P ₀ 3	Yes	chr12	25245350	C		G12D	SNV	Yes
P ₀₄	Yes	chr12	25398284	C	T	G12D	SNV	Yes
P ₀₅	No		$\overline{}$		$\overline{}$		$\overline{}$	$\overline{}$
P ₀₆	Yes	chr12	25398284	C	A	G12V	SNV	Yes
P ₀ 7	Yes	chr12	25245350	C	A	G12V	SNV	Yes

Supplementary Table 9. Mutation of KRAS for each patient