Performance of neoantigen prediction for the design of TG4050, a patient specific neoantigen cancer vaccine

AACR 2020 Abstract 20-A-2250-AACR

B. Mallone[1], C. Tosch[2], B. Grellier[2], K.Onoue[3], T. Stylzler[1], K. Rittner[2], Y. Yamashita[3], E. Quemeneur[2], K. Bendjama[2]✉

[1]NEC Laboratory Europe GmbH, Heidelberg, Germany, [2]Transgene SA, Illkirch-Graffenstaden, France (✉: kab@transgene.fr), [3]NEC Corporation, Tokyo, Japan;

BACKGROUND

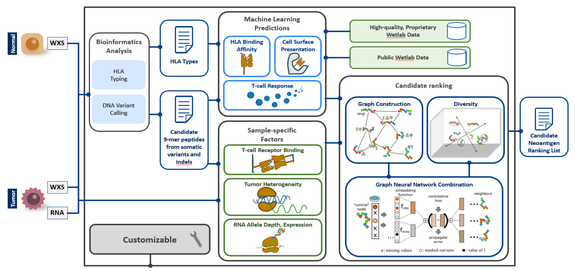
The development of therapeutic cancer vaccines to immunize against tumor antigens constitutes a promising modality. Mutation-associated antigens are considered major targets given their specificity to tumor cells. These mutations are specific to the patients and require tailor-made vaccines targeting the corresponding tumor-specific epitopes. Many mutations are identified in the tumoral genome in most patients, but only a small fraction (around 1%) is suitable as vaccine target. Herein, we report data documenting the prediction performance of the algorithm used for the design of TG4050, a clinical stage patient specific viral-based neoantigen vaccine.

DESIGN AND TRAINING OF PREDICTION SYSTEM

Bioinformatics – Variant calling, MHC typing, and candidate neoantigen identification
We developed a bioinformatics pipeline to identify somatic variants (mutations and indels) from matching healthy and tumor WES samples following best practices. We then extracted all candidate 9-mer peptides which overlap all variants. The healthy WES was further used to identify the Class-I MHC type of the sample.

Machine learning – MHC binding, processing, and immunogenicity
We trained a set of independent machine learning algorithms to score peptides for several steps of the MHC antigen presentation pathway, including MHC binding, intracellular processing, and likelihood to elicit an immune response. These models are then used to make predictions for each candidate neoantigen accounting for the identified MHC alleles of the sample.

Ranking candidate neoantigens - Graph neural network and diversity
In order to rank the candidate neoantigens and determine the vaccine contents, we trained a graph neural network to combine the predictions with sample-specific factors, including expression and conservation of the candidate across clones based on tumor RNA-seq. A final module combined the score with a diversity criterion to create a final ranking of the candidate neoantigens.



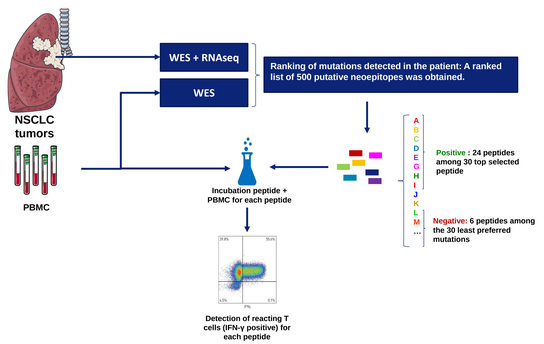
METHODS

Study design
We collected tumoral and peripheral blood samples from patients diagnosed with Non-Small Cell Lung Cancer (NSCLC) who were eligible for surgical resection. Blood samples were processed by centrifugation on Ficoll density gradient to isolate PBMC prior to cryostorage. Tumor samples were rapidly snap frozen on liquid nitrogen upon collection.

Sequencing
Germlines sequences were obtained by WES of PBMC. Tumoral sequences were obtained by WES of tumor DNA. RNA sequencing of tumor samples was also performed for confirmation of expression of tumor genes and evaluation of abundance of mutated transcripts.

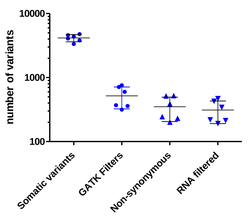
Peptides
Peptides corresponding to targets mutation were synthesized and used for stimulation of autologous PBMC. We first tested pools of 6 peptides batched based on their ranking by the immunogenicity prediction system and then deconvoluted immunogenicity of individual peptides.

Assessment of immunogenicity
Immunogenicity was assessed by counting of IFN-γ secreting cells in patient PBMC after restimulation with peptides encoded by the mutated sequence. Briefly, patient PBMC sample were thawed, exposed to 1 µg/ml of peptide or media (negative control) for 6 hours. After exposure to peptides, cells were washed and incubated with anti-CD3, anti-CD8 and anti-IFN-γ antibodies conjugated with fluorescent probes. Assessment of frequency of antigen specific IFN-γ secreting $CD8^+$ T-cells was performed by flow cytometry (see gating strategy).



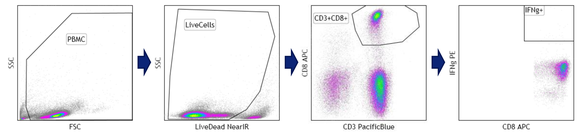
RESULTS

SEQUENCING AND MUTATIONAL PROFILES



Accordingly to data reported in the literature, mutational burden of NSCLC patients was relatively high with a median of 4217 (3339-4782) somatic mutations identified. After filtering using GATK-Broad Institute Best practice recommendations, a median of 484 (316-756) mutations remained. These translated into a median of 316 (201-521) non synonymous mutations of which a median of 281 (192-471) was confirmed at RNA level.

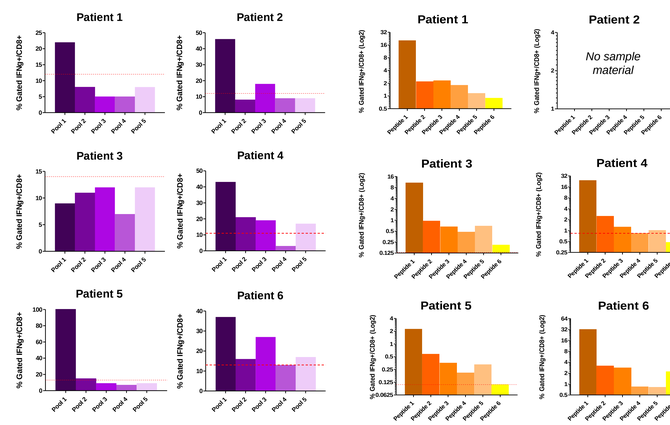
GATING STRATEGY



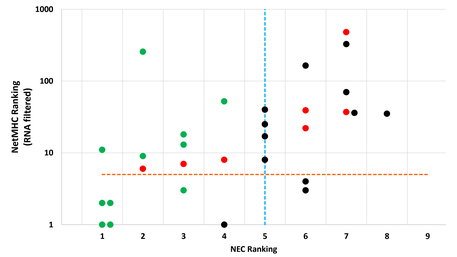
Gating for the identification of $CD3^+CD8^+IFN\text{-}\gamma^+$ cells in patient PBMC.

IMMUNOGENICITY OF PREDICTED PEPTIDES

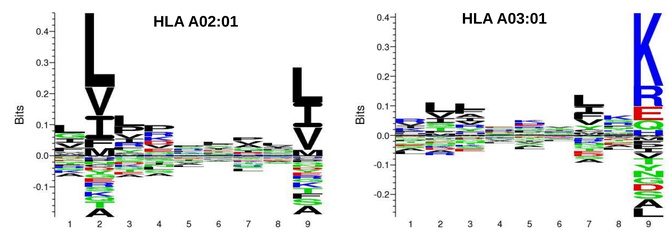
Antigen specific induction of IFN-γ in patient PBMC: on left panel: responses in individual patients for peptide pools, pool composition was based on the peptide ranking with our prediction system (Pool 1: top 6 peptides, pool 2: #7 to #12, pool 3: #13 to #18, pool 4: #19 to #24 and pool 5: pool of low ranked peptides as control for non selected peptide); On right panel, deconvolution of responses for individual peptides in pool 1. Red dotted line: assay background noise.



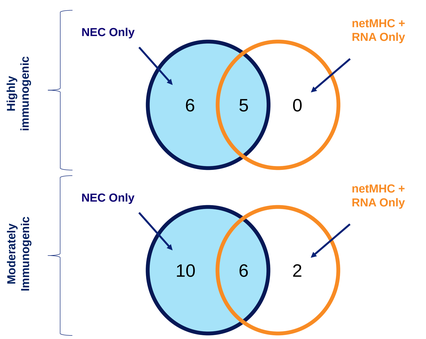
Comparison of ranking using our NEC ranking system and MHC binding prediction + RNA filtering for peptides selected in pool #1 for all patients. In green, highly immunogenic peptides as assessed by frequency of $CD8^+IFN\text{-}\gamma^+$ cells (>2% of positive cells) and in red, peptides with low immunogenicity (<1% of positive cells). Dotted blue line : top 5 peptides NEC cut-off; Dotted orange line: top 5 peptides cut-off netMHC.



Logo plot of peptides selected by the NEC system from tested patients for two common HLA genotypes. Preferred positions are consistent with known HLA-restricted epitopes.



Venn Diagram of immunogenic peptides ranked within the top 5 for NEC prediction system and netMHC + RNA filter.



REFERENCES

- Niepert M et al., Proceedings of the 33 rd International Conference on Machine Learning, 2016
- Vita R et al., Nucleic Acids Res. 2015
- Jurtz V et al., The Journal of Immunology, 2015

ACKNOWLEDGEMENTS

The authors thanks patients consenting to participate to this study and their respective family. Special acknowledgements are addressed to administrative and support function staff at NEC and Transgene who have made this study possible.

CONCLUSIONS

- **The prediction system presented herein is able to identify immunogenic epitopes among a large number of candidate neoepitopes identified in patients with high accuracy.**
- **More than 86% of top ranked epitopes are immunogenic.**
- **Immunogenicity of predicted epitopes is correlated with their ranking by the prediction system**
- **The prediction system was able to select highly immunogenic peptides that were not highly ranked by state-of-the-art methods (netMHCpan 4.0).**