AI-augmented design of effective therapeutic cancer vaccines and adoptive cell therapies

Giovanni Mazzocco[1], Oleksandr Myronov[1,2], Iga Niemiec[1], Katarzyna Gruba[1,2], Piotr Skoczylas[1], Anna Sanecka-Duin[1], Michał Drwal[1], Jan Kaczmarczyk[1], and Piotr Stepniak[1]

1. Ardigen, Krakow, Poland, 2. Faculty of Mathematics and Information Science, Warsaw University of Technology, Warsaw, Poland

INTRODUCTION

Neoantigens are rapidly gaining interest as central components of personalized cancer therapies, as shown by the increasing number of clinical trials for personalized cancer vaccines and adoptive cell therapies.

One of the main challenges in developing such therapies is that only a small fraction of the detected somatic mutations are really immunogenic and can be used as targets of the treatment. Several biological aspects make the prediction of effective neoantigens a particularly challenging task. These aspects include: (i) highly individual antigenic landscape, (ii) restricted number of targetable mutation-induced neoantigens per tumor, (iii) complex tumor subclonal structure, (iv) private neoantigen specific T-cell repertoire, (v) epitope-induced immune tolerance and immunotoxicity, (vi) variability in neoepitope-associated gene-expression, etc.

Recent peptide-HLA (pHLA) elution and mass spectroscopy data made it possible to develop models and dedicated pipelines accounting for the natural neoantigen presentation [2, 3]. These technologies represent a step forward with respect to previous solutions solely based on pHLA binding affinity prediction.

Predicting neoepitopes' immunogenicity, accounting for safety issues and prioritizing peptides to be used for therapy in a fully personalized manner is the end goal of such technologies. Here we compare the performance of **ArdImmune Vax** with respect to approaches used in literature and industry [1, 2]. The performance of the selected models was evaluated on experimentally validated data for patients' CD8+ T-cell responses.

ARDIMMUNE VAX

Prioritization of cancer neoepitopes is at the heart of personalized therapy process

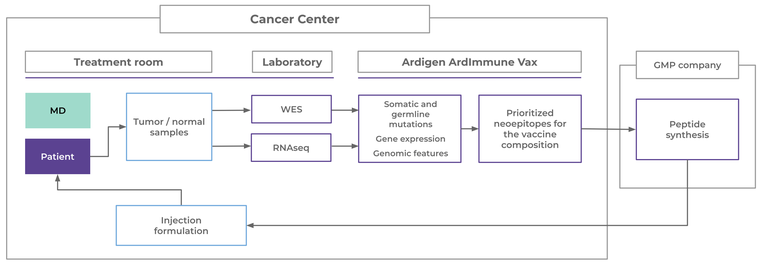

Fig. 1. Personalized cancer vaccine therapy process using peptide delivery platform:

1. DNA and RNA from patient's tumor (FFPE) and normal (PBMC) samples are extracted and sequenced.
2. Raw WES and RNAseq data are processed by Ardigen's computational system.
3. A peptide-based vaccine composition is proposed by ArdImmune Vax.
4. Peptides are produced under GMP requirements.
5. At the Cancer Center the final formulation of injection is prepared including adding any adjuvants.
6. The vaccine is administered to the patient and clinical parameters of response are monitored.

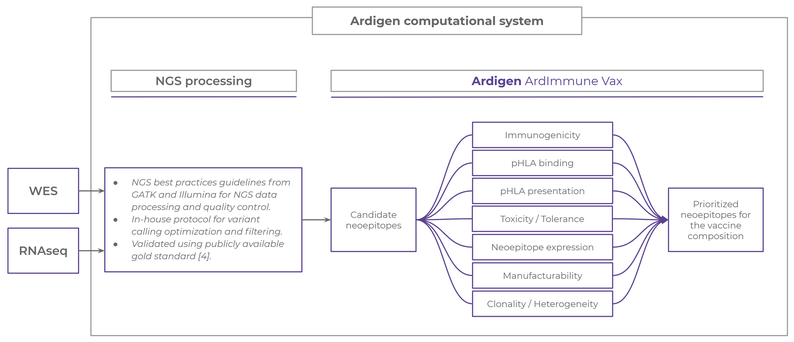

Fig. 2. Close up at the Ardigen ArdImmune Vax - an AI neoepitope vaccine design workflow.

1. Protein effects of the genomic variants identified by the NGS processing phase constitute candidate neoepitopes to be analyzed.
2. Multiple biological aspects contribute to the final call of which neoepitopes are the most likely to effectively elicit immune response of the patient while avoiding potential side effects.
3. In this case, ArdImmune Vax suggests a peptide-based vaccine composition as a result of the analysis.

VALIDATED CD8+ T-CELL RESPONSES

Dataset	No. patients	No. CD8+ responses	Source	Starting input
Dataset A	28	38	Subset of patients from [1]	raw data
Dataset B	10	21	Patients from [2] (Dataset 3A therein)	list of somatic mutations

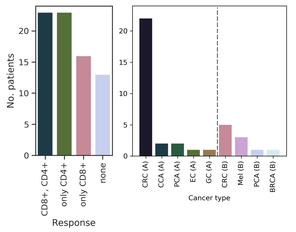

Fig. 3. Composition of datasets. Left: details of datasets included in the presented study. We consider patients with CD8+ response from [1] and [2]. The dataset from [2] (Dataset B) is a curated collection from other publications [5-7]. Dataset A is a subset of patients from [1] from which we dropped 10 patients with no available raw data (WES and RNA-seq). In both datasets we consider the subset of patients with observed CD8+ T cell response. **Right**: Number of patients per response type (for the entire set from [1]) and per cancer type (for Datasets A and B).

CD8+ responses were verified experimentally as follows: autologous APCs were either transfected with minigenes encoding the mutation flanked by nucleotides from the wild-type gene or pulsed with long peptides (~25AA) including the mutation of interest. Reactivity of patients' tumor TILs or PBMC was tested using IFNγ ELISPOT assay. Additionally, for dataset A flow cytometric analyses for 4-1BB up-regulation on CD8+ T cells were performed [1].

RESULTS

Dataset A

Fig. 4. Results on dataset A. Top: two ranked lists of mutations for each patient with the immunogenic mutations marked in purple and potentially toxic/tolerated peptides marked in red. The green bar denotes top-20 peptides. **Top-left:** ArdImmune Vax, **Top-right:** netMHCpan 4.0 [8]. These panels involve 28 patients with 37 immunogenic mutations. **Bottom-left:** mean number of immunogenic peptides detected per patient using selected tools: netMHCpan 4.0, netMHCpan 4.0 with filtering by expression (TPM > 0), and MHCflurry [9]. **Bottom-right:** mean number of immunogenic peptides detected per patient for two indication groups (CRC and CCA, PCA, EC, GC) using ArdImmune Vax. Confidence intervals in the plots represent 90% confidence intervals of the mean value. The maximum of the y-axis corresponds to the maximum possible value of the metric (equal to 1.3).

Dataset B

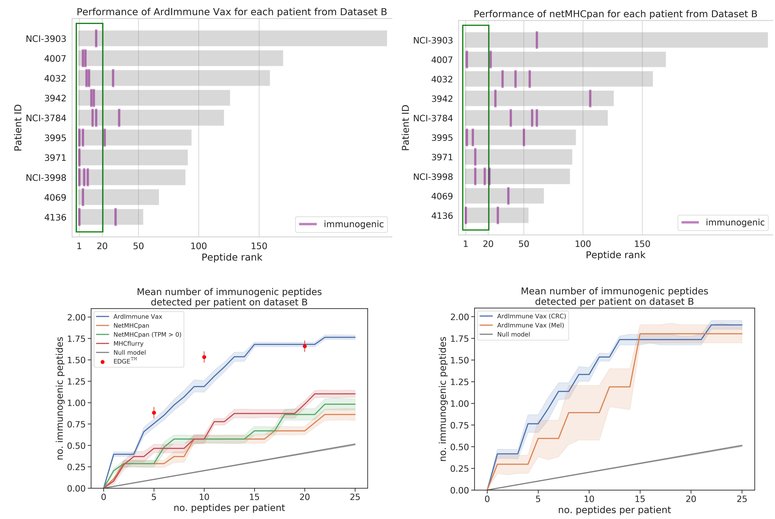

Fig. 5. Results on dataset B. Top: two ranked lists of mutations for each patient with the immunogenic mutations marked in purple. The green bar denotes top-20 peptides. **Top-left:** ArdImmune Vax, **top-right:** netMHCpan 4.0 [8]. These panels involve 10 patients with 21 immunogenic mutations. **Bottom-left**: mean number of immunogenic peptides detected per patient using selected tools: netMHCpan 4.0, netMHCpan 4.0 with filtering by expression (TPM > 0), and MHCflurry [9]. We also include the results of the EDGE™ algorithm [2] (using the scores reported in [2] for the considered dataset). **Bottom-right**: mean number of immunogenic peptides detected per patient for two indications (CRC and Mel) using ArdImmune Vax. Confidence intervals in the plots represent 90% confidence intervals of the mean value. The maximum of the y-axis corresponds to the maximum possible value of the metric (equal to 2.1).

DISCUSSION

- We benchmarked vaccine design methods on two datasets. Dataset A contains immunologically validated mutations not filtered by affinity prediction tools. As a result this dataset is relatively large and unbiased.
- Vaccine design workflows considering solely pHLA affinity prediction (even when filtered by expression) allow the inclusion of potentially toxic/tolerated peptides within the proposed vaccine formulation.
- Neoepitope relative clonal abundance and manufacturability should be also considered in personalized cancer vaccine design.
- Dataset A [1] contains patients with missing or incomplete existing immune responses (only CD4+, only CD8+) which indicates that stimulating novel immune responses is a needed and valid therapeutic strategy.
- Current work on ArdImmune Vax is focused on extending the method to HLA-class II epitopes in order to enable the design of personalized cancer vaccine formulations accounting for both CD8+ and CD4+ T cells activation.

CONCLUSIONS

- For both datasets ArdImmune Vax significantly outperforms methods based on pHLA binding prediction.
- ArdImmune Vax performs on a par with the EDGE™ algorithm [2] for the dataset B.
- The high performance of our model is consistent across cancer types and across cohorts.

REFERENCES

1. Parkhurst M, *et al. Cancer Discov. 2019*
2. Bulik-Sullivan B, *et al. Nat Biotechnol. 2018*
3. Abelin JG, *et al. Cell Immunity 2017*
4. Fang LT, *et al. BioRxiv 2019 (preprint)*
5. Gros A, *et al. Nat. Med.* 2016
6. Tran E, *et al. Science* 2015
7. Zacharakis N, *et al. Nat. Med.* 2018
8. Jurtz V, *et al. J Immunol. 2017*
9. O'Donnell TJ, *et al. Cell Systems* 2018

contact: piotr.stepniak@ardigen.com