A novel computational pipeline supported with in vivo vaccination models for identification and validation of neoantigens

6389

K. Hiranuka[1*], K. Takahashi[1], D. Tang[1], N. Watanabe[1], T. Yamada[1], Y. Mishima[1], N. Fujinami[2], M. Shimomura[2], T. Suzuki[2,3], T. Nakatsura[2], N. Nakamura[1*]

1 BrightPath Biotherapeutics Co., Ltd., Kawasaki, Japan, 2 Exploratory Oncology Research & Clinical Trial Center, National Cancer Center, Kashiwa, Japan, 3 Division of Pharmacology, School of Medicine, Teikyo University, Tokyo, Japan

*Email contact: hiranuka_k@brightpathbio.com, nakamura_n@brightpathbio.com

Background

Personalized neoantigen-based therapies are being actively developed. However, clinical reports have thus far showed that the number of actionable neoantigens was not as high as originally expected. One possible reason is due to a lack of appropriate prediction models, especially for immunogenicity, which is largely hindered by a scarcity of training data based on *in vivo* vaccination models.

Methods

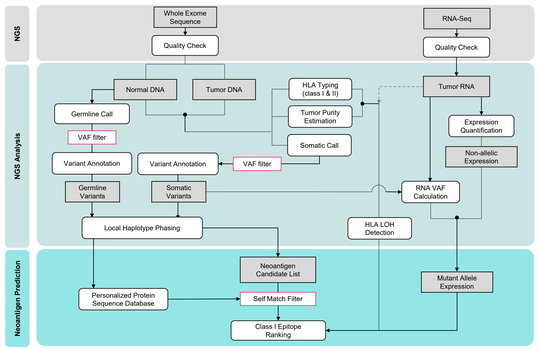
We have developed a computational pipeline to identify neoantigens from patients' NGS data and rank candidates of HLA class I neoepitope (Fig.1). To prioritize neoantigen peptides expected to be presented on patient's HLA class I molecules, we constructed a logistic regression model that takes the results of NetMHCpan-4.0, MHCflurry-1.4, NetChop-3.1 of each peptide, optimizing the model's parameter based on the immunopeptidome data in SysteMHC Atlas database and random peptides generated from human reference protein sequences. A linear predictor was converted using a softplus function to define *SCORE* between 0 and a positive value as well as *SCOREadj* as followed to reflect the effect of mRNA expression on presentation prediction:

$$SCORE = log(1 + e^{z})$$

$$SCORE_{adj} = arctan(TPM\_SUMvar) \cdot SCORE$$

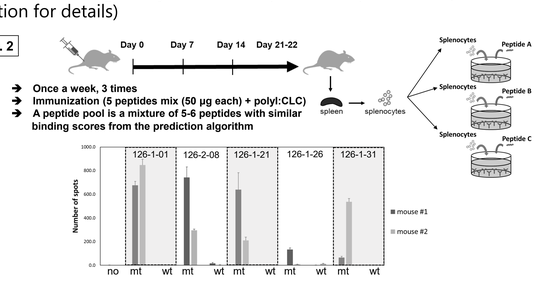
where z is a linear predictor and TPM_SUM_{var} is a sum of TPM(Transcription Per Million) of transcript isoforms in a mutant allele.

Fig. 1



To interrogate *in vivo* immune responses, we immunized our top-ranked neoepitopes into HLA-transgenic mice and performed ELISPOT (Fig.2). Based on the results, we further constructed a neural network model that defines immunogenicity score to improve our prediction (see "Results" section for details)

Fig. 2



Results

Variant call, tumor purity estimation and HLA LOH detection

We generated NGS data(WES & RNA-Seq) from normal/tumor samples of 36 HCC, 26 mCRC, and 11 LC patient and analyzed the data using our computational pipeline. We observed a wide range in the number of somatic variants and estimated tumor purity (Fig.3). We performed HLA typing on DNA and RNA and estimated the allele frequencies (AFs) of HLA alleles (Fig.4). The independently estimated AFs using DNA and RNA had a strong correlation (Spearman's rho = 0.8636) (Fig.5). We used the AFs and tumor purity estimations to call putative loss of heterozygosity (LOH) events in HLA genes (Fig.6).

Fig. 3

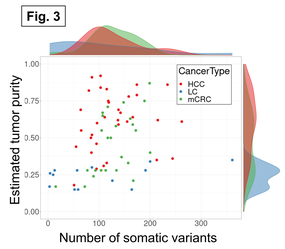


Fig. 4

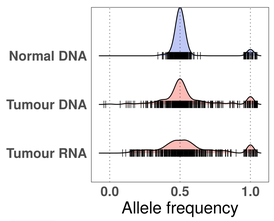


Fig. 5

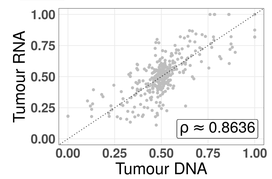
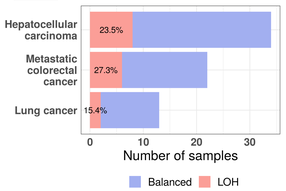


Fig. 6



Presentation prediction mode.

Our presentation model was evaluated on an independent immunopeptidome data (Abelin *et al.*, 2017) and random peptides at 1:100 ratio. It was found highly discriminating between naturally presented ligands and random peptides (Fig.7). *SCOREadj* showed a better performance over *SCORE* (Fig.8), indicating the importance of mRNA expression for presentation prediction.

Fig. 7

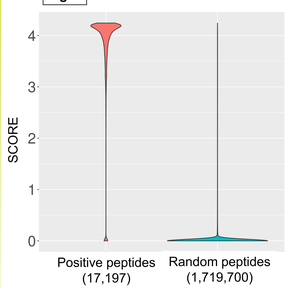
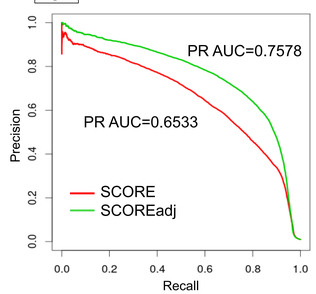


Fig. 8

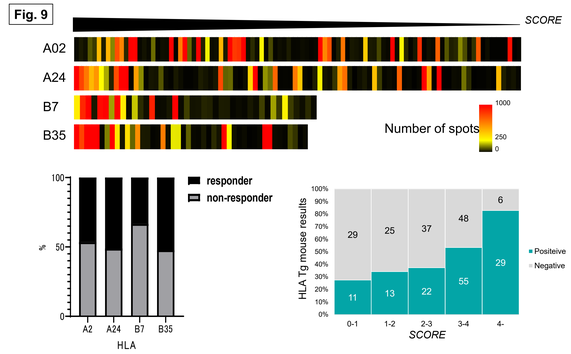


Note that, however, in the following Tg-mouse experiment where the source of epitopes was the synthesized peptides, we ranked the neoepitopes based on *SCORE* instead of *SCOREadj* due to irrelevance of mRNA expression in patients.

Testing predicted neoantigens using HLA-Tg mice

From 27 patient cases, 275 of HLA-A02:01-, A24:02-, B07:02-, or B35:01-predicted neoepitopes were tested in the HLA-transgenic mice. The rate of positively reacted epitopes increased in a correlated manner to *SCORE*, achieving 47% (ROC AUC=0.6871) overall and 82% if focusing on those with *SCORE*>4.0 (Fig.9). The number of neoepitopes whose *SCORE* exceeds 4.0, however, is usually very low or zero per patient depending on the number of somatic variants. This brought up the question whether an additional prediction model trained with these immunogenicity data could improve the prediction performance.

Fig. 9



Immunogenicity prediction model based on Tg-mouse data

The 199 of the validated HLA–A peptides were used to construct an immunogenicity prediction model (Fig.10) where a peptide sequence was converted into a numeric vector of physicochemical properties which was passed through neural network layers and the final score was obtained as *IMMUNE_SCORE*. The model training was performed as followed: 1) the peptide dataset was split into Train and Test dataset at 8:2 ratio, 2) a model with initially chosen hyper-parameters was trained by the Training dataset in a 5-fold cross validation, 3) the average F1 score of 5 cv models (Dev. F1 score) was calculated, and 4) 2)&3) were repeated by progressively updating hyper-parameters.

Fig. 10

After generating a large number of such models, each model was evaluated on the Test dataset to calculate Test F1 score. Although the correlation with Dev. F1 score was weak (Fig.11), top10 models whose average Dev. F1 score was 82.07 showed comparably high Test F1 score (73.31 on average). Afterall, we selected the one with the best Dev. F1 score as the first release model (V1 model).

Fig. 11

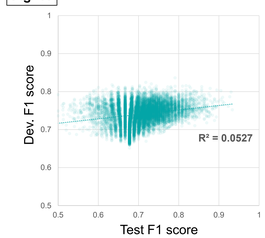
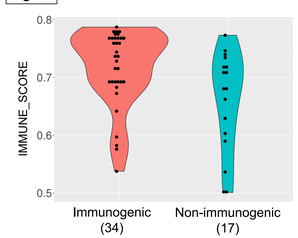


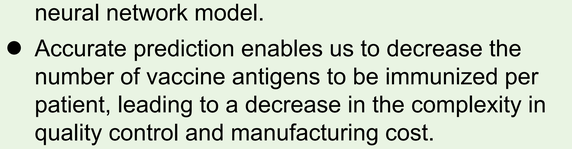
Fig. 12



To further evaluate our immunogenicity prediction model, we analyzed 15 additional LC patients to prioritize neoepitopes by *IMMUNE_SCORE* after pre-filtered by *SCORE* >=2.5. The immunogenicity of 51 newly selected neoepitopes were validated with HLA transgenic mice, resulting in 66.7% overall positive rate(Fig. 12). Notably, those with *IMMUNE_ SCORE* >=0.75 showed 93.75% positive rate (15/16) although all but only one had *SCORE* < 4.0, indicating high efficiency of the model on selecting immunogenic neoepitopes. We augmented the training dataset with the rest of the previously validated peptides and the newly validated peptides to update the model. Robust prediction improvement was observed, and the second release model(V2 model) was selected (Fig.13). To repeat this learning cycle, over 100 additional peptides are currently being tested in Tg-mice.

Fig. 13

Leaning cycle	Input data (Pos. + Neg.)	Ave. Test F1 score over top 10 models
1st	97 + 102	73.31
2nd	169 + 169	79.15



Conclusion

- We developed neoantigen prediction pipeline and analyzed clinical samples from over 100 patients.
- HLA LOH frequently occurs in tumor, and thus a robust detection algorithm was implemented.
- We validated predicted neoepitopes with HLA transgenic mice to see *in vivo* immune response.
- We showed the feasibility of improving prediction algorithms using HLA-transgenic mice data and a neural network model.
- Accurate prediction enables us to decrease the number of vaccine antigens to be immunized per patient, leading to a decrease in the complexity in quality control and manufacturing cost.

Acknowledgement

The super-computing resource was provided by Human Genome Center (the Univ. of Tokyo).

Disclosure
Hiranuka K.: Employee of BrightPath Biotherapeutics

BrightPath Biotherapeutics