Enhanced Neo-antigen prioritization for CAR-T Therapy by fusion of Protein Language Model embeddings with PSSM Scores

Anasuya Chatterjee, Tehemton Khairabadi, Vishal Pagidipally

Background

- Personalised immunomics enhances precision oncology by identifying robust, clinically actionable therapeutic targets and minimising off-target toxicities of conventional therapies.
- Recent advances in highly immunogenic, tumour-specific neoepitope-directed targeted therapies like CAR-T demonstrate promising efficacy for early diagnosis, immune priming, and treatment of solid tumours.

Limitation of existing methods

Lack of a rapid,cost-effective robust pipeline

Current methods are CPU-intensive, fragmented and lack robustness in feasibility, validity and speed

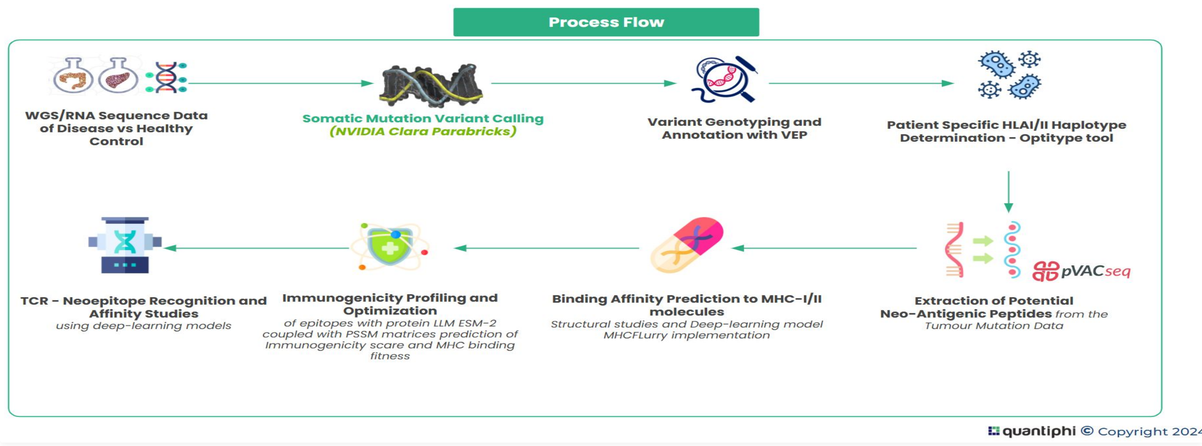
Lack of Functional significance

Deep learning methods are not fully biologically grounded or immunologically interpretable

Main contributions

- An integrated, multimodal, high-throughput end-to-end platform for for Neoantigen Discovery and Prioritization for precision therapies in oncology from HCC1395 cell line TNM Breast cancer WGS Data.
- Hybrid ESM2-PSSM embeddings with an attention-based supervised head for a synergistic knowledge-guided consensus score for immunogenicity and MHC fitness prediction

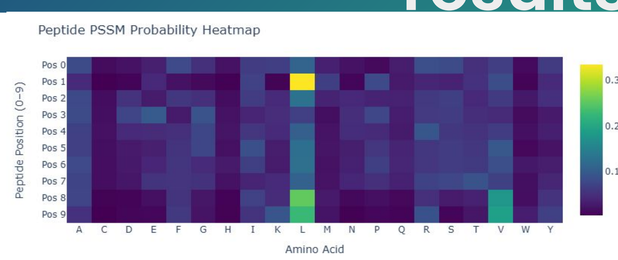
FusionNeoRanker Architecture



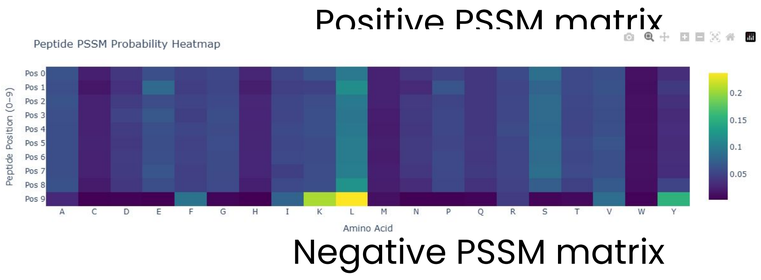
The integrated model of FusionNeoRanker

- **GPU-accelerated scalable optimised variant calling:** Nvidia Clara Parabricks has been leveraged for GPU-accelerated faster optimised somatic mutation variant calling
- **Grounded and validated in neoantigen optimised features** : The putative features of HLA-I binding, presentation potential and surface TCR reactivity and binding affinity have been considered and validated in this pipeline.
- **Integration of a pioneering pLLM–Transformer architecture with optimised loss-function and attention head:** Our novel ESM2-PSSM synergistic model has been incorporated which combines the statistical power of established binding matrices with the deep contextual learning of modern large language models to make an accurate and informed prediction on peptide immunogenicity

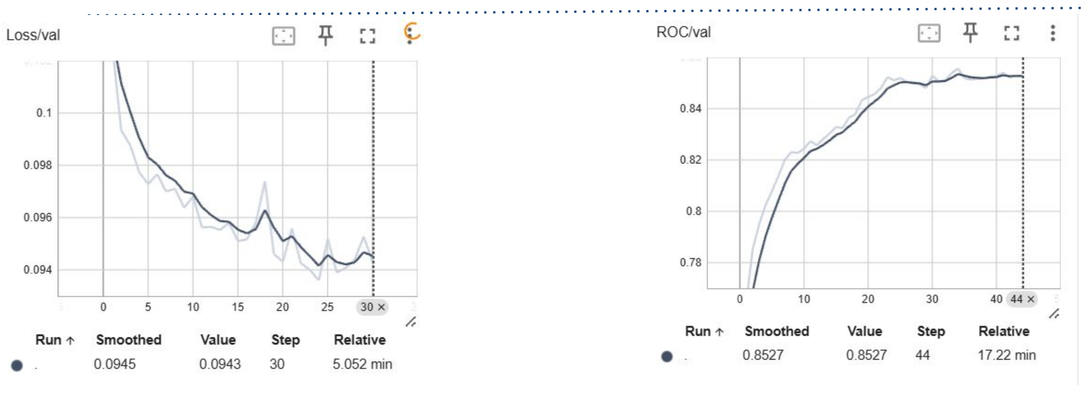
Model Accuracy and Optimization results



Positive PSSM matrix

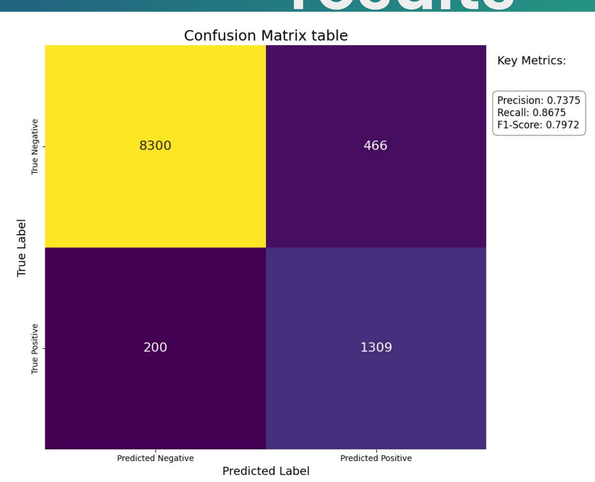


Negative PSSM matrix



ESM2-PSSM Contrastive Immunogenicity model with ESM-2 attention-based sequence representation and PSSM matrix based prior positional embeddings

Model Accuracy and Optimization results



Conclusion and Future perspectives

- We leveraged a novel in-silico pipeline to uncover 23 high-priority neoantigens based on the model and TCR binding, they are the clinically actionable targets for next-generation CAR-T therapies. .
- Future efforts will focus on the expedited in-vitro validation and clinical trials of the candidates, to bridge the gap between computational discovery and transformative patient care in oncology.