**Introduction**

**THIS IS A TESTkjhgfdxz**

HORMAD1 (HO1) is a protein which is usually expressed in the male testis and is involved in meiosis. However, for most BRCA1/2 and TNBC patients, it is aberrantly expressed in tumour cells, where it is linked to genomic instability and poor prognosis.

The discovery of this protein presents an opportunity for targeting with immunotherapy (Gjerstorff et al 2015). For each protein produced by a cell, a small peptide fragment from it should be displayed on the surface of the cell membrane on an MHC. These antigen-MHC complexes can be recognised by T cells which are constantly being produced recombinantly to have specificity to neoantigens so that they can recognise them as normal or virally infected/cancerous. These T cells or their receptors (TCRs) could be used to develop immunotherapies.

Compared to many other cancer antigens, HO1 is very specific to tumour cells meaning it has little potential for off-target effects (even for male patients), and it is linked to its pathological traits directly meaning that the cells cannot simply stop expressing it to cheat the therapy. In addition to being an attractive antigen to target, it is largely expressed in a group of breast cancer patients which comparatively have worse survival rates, fewer treatment options. For TNBC, options are mainly chemotherapy and mastectomy.

Applications:

Cancer Therapy With TCR-Engineered T Cells (Shafer P et al 2022)

One attractive application in immunotherapy would be to administer a vaccine to healthy gBRCA1/2 carriers to improve their initial immune response, helping to kill the HO1+ tumour cells before they adapt to evade immune response.

**Aims**

The first steps of towards this end goal is to identify TCR sequences which can recognise HO1. There are hypothetically many possible TCR sequences that could recognise HO1, but these will have differing affinities and abilities and they will differ depending on the HLA type of the individual.

For example, one HO1-TCR has been identified and patented already, but when tested by Ernesto he found that although it had great killing ability it was poor at activating other T cells likely making it non-viable for immunotherapy, and additionally would only be valuable as a therapy for one patient HLA type. So the aim of the TCR project is to produce a repertoire of candidate TCR sequences that may be tested to identify the best TCR for all HLA types.

We know these very likely exist as T memory cells in HO1+ patients. In early stages of the HO1 tumour, T cells will have recognised HO1 and likely started to produce many T cells to target the HO1+ cells. From then, the HO1+ portion of the tumour will have either been killed or become immunoresistant, but either way memory T cells would exist in the blood.

**Methods**

With this assumption, blood samples were taken from BRCA1/2 cancer patients at KCL and treated with massive doses of HO1 peptides to trigger these memory T cells to produce large population of T cells.

One complication is that we don’t know which peptide fragment from HO1 is expressed in the cancer patients and therefore which peptide will expand the T cells.

For any given protein, an APC only present a fragment as a peptide around 10aa in length on its MHC class I on its membrane. HO1 has almost 400 possible 10aa peptide fragments it could express. The experiment accounts for half of these, by separating fragments by an amino acid step count of 2, making it 200, and then these were divided into 10 pools containing 20 amino acids which represent a region of the HO1 peptide with pretty good coverage. Blood was also treated for three kinds of control: CEF, DMSO, Neg.

All patients were treated for all peptide groups in the lab with the expectation that a TCR from each HLA type will be specific to at least one of these. Meanwhile, it was known that cells preferentially express a particular peptide fraction for each protein depending on their HLA type, and that DL models exist that had been trained to predict this, including NetMHCpan 4.0 (Andreatta et al 2016). Greg implemented this to help select which samples to send off for sequencing which had the most chance of triggering the T cells, with the aim of getting the best results for the budget.

After treatment and 2 weeks of incubation, only 38 samples from 6 patients yielded enough RNA to send off for sequencing, which also had good prediction from the presentation prediction model.

From the sequencing data, the frequency of distinct CDR3 sequences was counted, which is a region of 10aa confers its antigen-specificity and only the Beta chain was analysed as their RNAseq in bulkseq are separated.

**Data Analysis**

This schematic represents the T cell populations of each patient sample after being exposed to a different pool of HO1 peptides. Each bubble represents a unique TCR sequence with size reflecting the proportion that sequence represents of the whole TCR sequence repertoire.

To find a HO1-specific TCR here, we are looking for a clonotype which has expanded in one of these treatments but not any of the controls, particularly not the CEF.

For each sample, I’ve coloured an expansion index – which is the proportion of the dominant clonotype including all sequences which are below a threshold of distinctness to assume they have similar specificity measured by a python package called TCRdist – which is roughly equivalent to 2 amino acids difference. For our samples, all dominant clonotypes were significantly distinct from other expanded clonotypes. So effectively this figure just colours samples by how expanded the most dominant clonotype is.

When comparing this way, you can see there’s a surprising amount of background expansion, which is unfortunate because it reduces our confidence that any dominant clonotype is triggered by treatment. Despite this, the expansion is significantly greater in some of the pools than treatment, and when the experiments were redesigned to reduce background expansion (protocol B), we still saw more expansion in treatment than control.

Present in the control – promising. However, our candidates from Pool7 are found in the control. This doesn’t necessarily mean that their expansion wasn’t triggered by the treatment, as I don’t think it’s too unlikely that if the blood contains these T cells that they would be present at low levels without treatment. However, I want to adapt these figures to reflect which control (positive or negative) and at what level of expression to help with this.

Over the next weeks, I want to add apply a machine learning model to each of these TCR sequences to predict their binding specificity to the peptides in the pool their were found, using pMTnet (Lu et al 2021). I think it would be very useful to colour each clonotype in this schematic by this models prediction – we would hopefully see a much higher binding affinity in the treatments than control, and hopefully for our expanded clonotypes. It might also help to discard any candidates if they have poor predictions.

**Transduction of candidates**

From the preliminary experiments and this analysis, candidate sequences were chosen for validation experiments. Beta chain sequences were paired with alpha chains which were most probably from the same T cell population based on similar expression levels, and T cells were transduced with TCRs with the same CDR3 sequence. These transduced T cells were treated with the peptides from the pools they were found, this time a separate peptide per pool, and tested to see if they had expanded. Results: KCL763-9 massive, 7 small. KCL725-9 medium

Pool9 is a promising candidate and further experiments will be conducted for each TCR on both its ability to kill and activate more T cells, but the more candidates the better.

**In the future**

Next month we will be receiving back single cell sequencing data. This will help us to pair alpha and beta sequences rather than having to infer from large proportions and try different combinations. It will also increase our sample size which will be helpful. We might see the same sequences again in the control, increasing our confidence in them, or see new sequences to add to our repertoire to test.

We also plan to try and locate these candidate TCR sequences in tissue samples. Bulk RNA sequencing *had* been done on tissue samples already in the hopes of finding these, however there was no significant expansion, nor presence of any candidate sequences from the blood samples. This is surprising, but it’s entirely likely that this is due to the region of tissue sample from the tumour. TNBC tumour we expect to have high TILs but the distribution of these is likely not uniform. So with spatial transcriptomics, we have more of a chance of finding these HO1 TCR sequences, but also will be able to categorise the distribution of these in the tissue.

**Data analysis pipeline**

With all this new data coming in, I am working to make a robust pipeline for analysing both bulk and single cell sequencing data. Analysis scripts pull latest data and runs analyses and produces figures which are pushed to another repository. All the generated content is displayed on a flask application which is being hosted on an Apache server on an Ubuntu VM on KCL’s CREATE cloud. So, project members can view analyses as they come out in one place.

**References**

Gjerstorff MF, Andersen MH, Ditzel HJ. Oncogenic cancer/testis antigens: prime candidates for immunotherapy. Oncotarget. 2015 Jun 30;6(18):15772-87. doi: 10.18632/oncotarget.4694. PMID: 26158218; PMCID: PMC4599236.

Lu, T., Zhang, Z., Zhu, J. et al. Deep learning-based prediction of the T cell receptor–antigen binding specificity. Nat Mach Intell 3, 864–875 (2021). <https://doi.org/10.1038/s42256-021-00383-2>

Ilyas S, Yang JC. Landscape of Tumor Antigens in T Cell Immunotherapy. J Immunol. 2015 Dec 1;195(11):5117-22. doi: 10.4049/jimmunol.1501657. PMID: 26589749; PMCID: PMC4656134.

Shafer P, Kelly LM, Hoyos V. Cancer Therapy With TCR-Engineered T Cells: Current Strategies, Challenges, and Prospects. Front Immunol. 2022 Mar 3;13:835762. doi: 10.3389/fimmu.2022.835762. PMID: 35309357; PMCID: PMC8928448.

WO2021092223 - HLA RESTRICTED HORMAD1 T CELL RECEPTORS AND USES THEREOF <https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2021092223>

Jurtz V, Paul S, Andreatta M, Marcatili P, Peters B, Nielsen M. NetMHCpan-4.0: Improved Peptide-MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. J Immunol. 2017 Nov 1;199(9):3360-3368. doi: 10.4049/jimmunol.1700893. Epub 2017 Oct 4. PMID: 28978689; PMCID: PMC5679736.

Massimo Andreatta, Morten Nielsen, Gapped sequence alignment using artificial neural networks: application to the MHC class I system, Bioinformatics, Volume 32, Issue 4, February 2016, Pages 511–517, <https://doi.org/10.1093/bioinformatics/btv639>

O'Donnell TJ, Rubinsteyn A, Laserson U. MHCflurry 2.0: Improved Pan-Allele Prediction of MHC Class I-Presented Peptides by Incorporating Antigen Processing. Cell Syst. 2020 Jul 22;11(1):42-48.e7. doi: 10.1016/j.cels.2020.06.010. Epub 2020 Jul 14. Erratum in: Cell Syst. 2020 Oct 21;11(4):418-419. PMID: 32711842.