Beyond SNVs and indels for neoantigen prediction in cold tumours

Cancer immunotherapy, in particular checkpoint inhibitor (CPI) therapy, has drastically changed the oncology field which has evoked a complete shift of standard care regimens and overall improvement of treatment efficacy of cancer patients. Despite these promising results, many malignancies still do not respond at all to checkpoint immunotherapy exposing patients to unnecessary side effects and society to high costs.

Low TMB malignancies

This low response rate can be explained by the fact that **some tumours are barely different from healthy cells** due to a low mutation frequency, i.e. low mutational burden (TMB) malignancies, and often these tumours are immune deserted or cold tumours[1]. These patients mostly have **no or very limited T-cell activity** since the low amount of mutations cause a limited number of patient-specific antigens (neoantigens) to be present on the tumour cells that can trigger a tumour-directed response. Thus, treatment with CPIs is often useless as there are no T-cells to be reactivated.

Treatment of cold tumours should, therefore, be approached differently by triggering the release of antigens in the tumour vicinity so that an immune response can be naturally evoked or by direct activation of T-cells against neoantigens presented by that tumour. This can be achieved through **low dose chemo/radiotherapy and virotherapy or personalised vaccination** respectively[2]. Here we will focus on personalised vaccination and its challenges in finding the correct targets in low TMB malignancies.

The holy grail of personalised vaccination is centralised around finding correct neoantigens. These are peptides different from the normal peptidome (set of peptides expressed by a normal cell), so they can be used as an identifier for uniquely targeting tumour cells. This will lead to limited side effects since cell death will only appear on cells with these neoantigens on their surface.

However, **finding immunogenic neoantigens** for personalised cancer vaccination in cold tumours **is challenging** due to the low mutation frequency where the number of possible mutational peptides expressed in the tumour cells is limited. It is thus of great importance to initially identify as many disparate peptides as possible. Indeed, Laumont *et al.* postulated that although there are clear indications that the immune system can drive successful tumour eliminations in tumours characterised by a low TMB, different and more thorough neoantigen discovery methods are required[3].

Standard neoantigen prediction – SNVs and indels

The standard industry wireframe for neoantigen prediction is illustrated in **Figure I**. It involves the identification of all somatic non-synonymous mutations (NSMs)[4,5]. To accommodate this discovery, genome sequencing data from matched tumour and normal DNA is required for each patient. Following alignment of these reads to the human reference genome, somatic genetic alterations in the tumour genome are detected using multiple variant-calling algorithms. To avoid the incorrect classification of germline variants as neoantigens and to guarantee the expression of the identified variants, both RNA and DNA from tumour tissue are used to compare with DNA from the matched normal tissue[6,7].

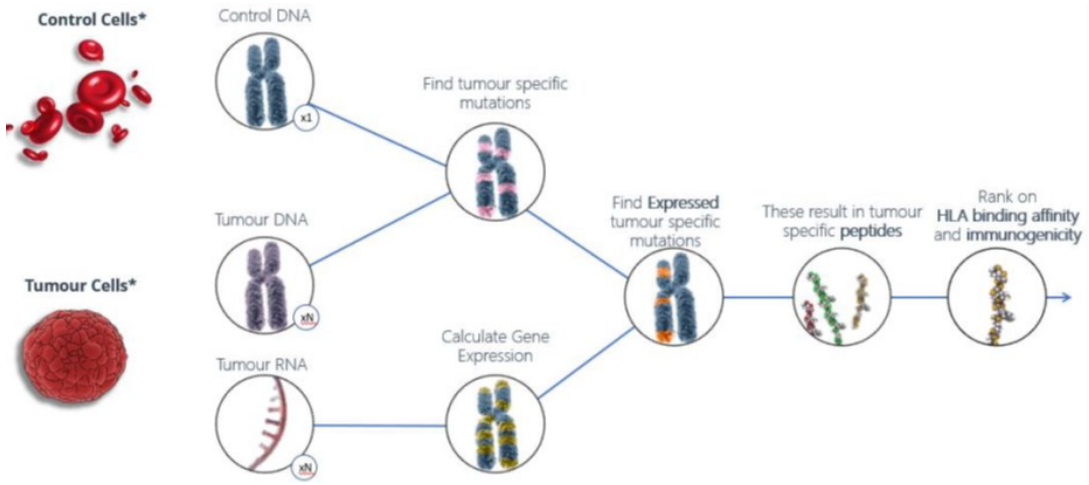

Figure I. Graphical representation of the basis of the myNEO ImmunoEngine wireframe

Most variant calling algorithms have already proven to reliably call **single nucleotide variants** (SNVs) since they are the most abundant type of tumour mutation[8] and because of the relative simplicity and reliability of identifying sequence changes of one base pair[9]. Besides SNVs, **insertion, or deletions** (indels) can lead to highly immunogenic frameshifts[10,11]. These variants are a highly immunogenic mutational class as they may result in highly divergent amino acid sequences and hence produce strong neoantigens. The same variant calling algorithms can be used to detect indels, however, additional realignment tools can be employed to improve the detection power.

Going beyond – targeting a broad antigen landscape

As mentioned, personalised vaccination targeting neoantigens is a promising strategy to revert cold tumours into the inflamed immune type. However, the current neoantigen prediction tools (vide supra) used in the industry lack depth. There is a high need for tools that allow detection of a broader antigen landscape, hence tools that identify tumour-specific alterations on distinct levels within a cell.

This means taking mutations into account that go beyond and explore the domain of **gene fusion events, transposable elements activity, neoisoforms, alternative proteasomic splicing**, and other differential processing events. This extends the search domain from DNA level onto RNA and proteomic level.

In addition, alterations in the tumour caused by aberrant expression instead of mutations need to be examined, which are responsible for the formation of aberrantly expressed tumour specific antigens (aeTSAs). Detection requires further extending the variant search domain with **non-canonical variants** such as novel exon-exon junctions, alternative start codons, uncanonical ORFs (in introns, antisense, etc…) and smORFS (in lncRNAs, 5'UTRs, etc…). These variants are often deemed more interesting than intracellular changes caused by mutational processes, as they are not dependent on the random occurrence of a specific mutation in a tumour, and thus require a less individualised therapeutic approach. Also, it has been shown that nonmutated aeTSA events greatly outnumber mutational TSA events[3].

Altogether, this means that the focus of variant calling for cold tumours needs to **shift beyond purely exonic mutational event detection towards non-exonic variant calling**, as depicted in **Figure II**. This shift requires patient-specific mass spectrometry datasets combined with whole-genome sequencing (WGS), ribosome depleted RNA-Sequencing, and even Ribo-Sequencing.

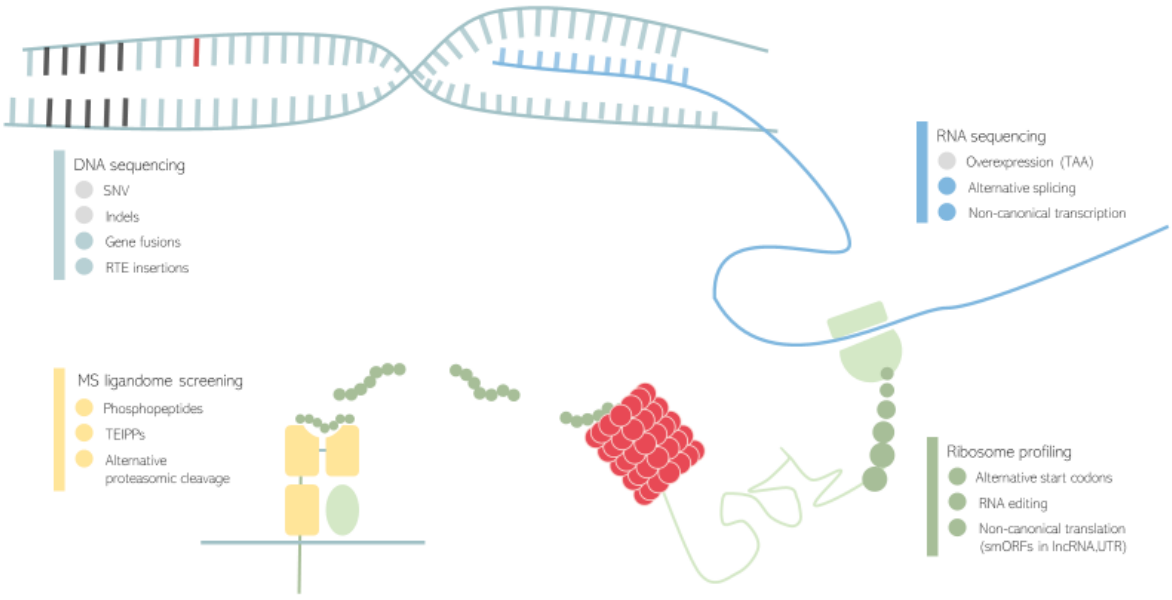

Figure II. Neoantigen search domain expanding from purely exonic towards non-exonic variant calling

The relevance of the in-depth alteration detection

The importance of this approach was again confirmed in 2019 by Löffler[12], when he stressed the relevance of non-exonic tumour antigens in tumours with low tumour mutational burden (e.g. MSS CRC). But several other research groups have already stumbled upon relevant conclusions in the past decades.

Many of the earliest validated epitopes, discovered via the older non-sequence-based cDNA libraries, are **non-canonical peptides**. As a prime example, the first human neoantigen identified was a point mutation located across the exon-intron junction of the MUM1 gene[13], and is thus not identifiable via the basic exome-based prediction approach.

Since these non-canonical peptides have shown to be recognised by anti-tumour cytotoxic T-lymphocytes, T-cell-mediated surveillance of the integrity of the genome thus extends to some intronic regions. In **Table I** below, several other tumour-specific alterations that appeared in non-canonical genomic regions have been described to lead to immunogenic neoantigens.

More recently, mass spectrometry datasets combined with immune assays have confirmed that non-canonical epitopes derived from annotated non-coding regions are able to elicit immune responses[14]. It has even been demonstrated in murine mouse models that the aberrant expression of allegedly noncoding regions is the main source of targetable tumour-specific antigens, accounting for over 90% of the total antigen landscape[3]. Secondly, focusing on the exome as the only source of tumour alterations is very restrictive, since over 99 % of cancer mutations are located in noncoding regions. Standard exome-based variation identification approaches miss thus not only a large fraction of the tumour-specific alterations but also a great fraction of the targetable immune repertoire.

To accurately identify these non-canonical cancer peptides, an **integrated approach between sequencing datasets and mass-spec ligandome sets** is essential. New technological developments such as Ribo-sequencing, for example, allows to verify the expression of non-canonical genes (lncRNAs and pseudogenes), to evaluate the expression of genomic regions annotated as intronic, and to discover alternative Open Reading Frames within known genes. By closely coupling this sequence-based prediction

approach with mass spectrometry-based detection, non-canonical tumour alterations and their resulting neoantigens are identified and confirmed.

Epitope identified[a]	Type of tumor	Origin of non-canonical peptide	Expression in normal tissue	Gene name	HLA restriction element	Reference
MSLQRQFLR	Melanoma	aORF	Unknown	*TRP1*	HLA-A*31	(69)
VYFFLPDHL	Melanoma	Intronic	Yes	*GP100*	HLA-A*24	(70)
RSDSGQQARY	Melanoma	Intronic	Yes/low[b]	*AIM2*	HLA-A*01	(71)
VLPDVFIRC/VLPDVFIRCV	Melanoma	Intronic	No	*GNTV*	HLA-A*02:01	(72)
EEKLIVVLF	Melanoma	Intronic	No	*MUM1*	HLA-B*44:02	(4)
LPAVVGLSPGEQEY	Renal cell carcinoma	aORF	Yes	*MCSF*	HLA-B*35:01	(73)
SPRWWPTCL	Renal cell carcinoma	aORF	Yes/low[b]	*iCE*	HLA-B*07:02	(74)
EVISCKLIKR	Melanoma	Intronic	No	*TRP2*	HLA-A*68:011/HLA-A*33:01	(75)
LAAQERRVPR	Melanoma and breast cancer	aORF	Unknown	*NYESO1*	HLA-A*31	(76)
MLMAQEALAFL	Melanoma	aORF	Yes	*LAGE1*	HLA-A*02:01	(77)
CQWGRLWQL/MCQWGRLWQL	Melanoma	aORF	Unknown	*BING4*	HLA-A*02	(78)
LPRWPPPQL	Renal cell carcinoma	Intronic	Yes	*RU2*	HLA-B*07	(79)

[a]identified by cDNA library screens; [b]compared to cancer tissue; aORF, alternative open reading frame; HLA, human leukocyte antigen.

Table I. Tumour-rejection antigens derived from non-canonical protein sequences (Adapted from reference [15]).

Case study: smORFs in lncRNA

Ribo-Sequencing has shown the presence of hundreds of putative small open reading frames (smORFs) within long non-coding RNAs (Moumtaz and Couso, 2015), potentially translated by leaky scanning and reinitiation. Although the role of non-coding RNAs in cancer has long been thought mainly one of regulation, these observations indicate that the smORFs may have a functional role as well. The latter has been shown in colorectal cancer for a peptide derived from the HOXB-AS3 lncRNA (Huang et al, 2017). Translatable sORFs are also found within the 5′ leader and 3′ trailer regions of mRNAs, even overlapping onto the main protein-coding sequence of mRNAs.

A mutation discovery process targeting smORF-producing lncRNAs, therefore, expands on the traditional scope of neoantigen discovery. These smORFs are, however, not yet well characterised as their annotation is lacking and prediction tools that identify them purely based on DNA/RNA-sequencing datasets are scarce. Standard gene prediction methods fall short since they aim to recognise patterns in genomic sequences that denote features (such as canonical initiation codons, termination sites, splice sites, promoter sequences, and polyadenylation signals), or innate DNA characteristics (such as codon usage bias, nucleotide composition and in-frame hexamer frequency that might indicate coding potential), both absent in smORFs. Unadjusted proteomics studies face similar issues since they usually only consider large proteins of several kDAs.

By optimally using both Ribo-sequencing and mass spectrometry datasets, however, novel tumour-specific smORFs can be detected in previously considered untranslated RNA strands. Both aberrant translation of small ORFs within these lncRNAs and mutations within these smORFs need to be identified, and their derived neoantigens should be subsequently identified and scored.

About myNEO

myNEO (Ghent, Belgium) developed a platform enabling genomic-informed drug discovery in the key therapeutic areas of oncology and immunology. The data-driven ImmunoEngine identifies the most efficacious targets (epitopes) for each cancer patient, uniquely presented on the tumour cells and capable of redirecting a patient's immune system, leading to elimination of the cancer cells. The discovery platform enables targets to be identified even in hard-to-treat tumours with a cold/lowly mutated profile. Similarly, the company has applied its technology to identify immunogenic sequences in infectious diseases, capable of protecting populations with strong broad immune responses. myNEO is one of the companies that emerged from the Novalis biotech incubator fund at the end of 2018, founded by two leading entrepreneurs already known for several successes in the biotech industry: Wim Van Criekinge, professor of computational biology at Ghent University, and childhood friend Jan Van den Berghe.

Contact us

Interested in more information about myNEO? Contact us!

Lien Lybaert, PhD
Scientific Alliance Manager
Lien.lybaert@myneotx.com

[1] Bonaventura P. *et al.* Cold tumors: a therapeutic challenge for immunotherapy. *Front Immunol* (2019)
[2] Galon J. *et al.* Approaches to treat immune hot, altered and cold tumours with combination immunotherapies. *Nat Rev* (2019)
[3] Laumont, C. *et al.* (2018) Noncoding regions are the main source of targetable tumour-specific antigens. *Sci Transl Med* (2018)
[4] Ott P. *et al.* An immunogenic personal neoantigen vaccine for patients with melanoma. *Nature* (2017)
[5] Carreno B. *et al.* Cancer immunotherapy. A dendritic cell vaccine increases the breadth and diversity of melanoma neoantigen-specific T cells. *Science* (2015)
[6] Sahin U. *et al.* Personalized vaccines for cancer immunotherapy. *Science* (2018)
[7] Xu C. A review of somatic single nucleotide variant calling algorithms for next-generation sequencing data. *Computational and Structural Biotechnology Journal* (2018)
[8] Vogelstein B. *et al.* Cancer genome landscapes. *Science* (2013)
[9] Turajlic S. *et al.* Insertion-and-deletion-derived tumour-specific neoantigens and the immunogenic phenotype: a pan-cancer analysis. *The Lancet Oncology* (2017)
[10] Türeci *et al.* Targeting the Heterogeneity of Cancer with Individualized Neoepitope Vaccines. *Clinical Cancer Research* (2016)
[11] Yang H. *et al.* Important role of indels in somatic mutations of human cancer genes. *BMC Medical Genetics* (2010)
[12] Löffler *et al.* Multi-omics discovery of exome-derived neoantigens in hepatocellular carcinoma. *Genome Med* (2019)
[13] Coulie P. *et al.* A mutated intron sequence codes for an antigenic peptide recognized by cytolytic T lymphocytes on a human melanoma. *Proc Natl Acad Sci USA* (1995)
[14] Chong C. *et al.* Integrated Proteogenomic Deep Sequencing and Analytics Accurately Identify Non-Canonical Peptides in Tumour Immunopeptidomes. *BioRxiv* (2019)
[15] Garcia-Garijo, A. *et al.* Determinants for neoantigen identification. Frontiers in Immunology. *Frontiers Media* (2019)