



A Roadmap Toward the Definition of Actionable Tumor-Specific Antigens

Robin Minati^{1,2}, Claude Perreault^{2,3*} and Pierre Thibault^{2,4*}

¹ École Normale Supérieure de Lyon, Université Claude Bernard Lyon I, Université de Lyon, Lyon, France, ² Institute for Research in Immunology and Cancer, Université de Montréal, Montréal, QC, Canada, ³ Department of Medicine, Université de Montréal, Montréal, QC, Canada, ⁴ Department of Chemistry, Université de Montréal, Montréal, QC, Canada

OPEN ACCESS

Edited by:

Jennie Lill,
Genentech, Inc., United States

Reviewed by:

Christopher Michael Rose,
Genentech, Inc., United States

Arie Admon,
Technion Israel Institute of
Technology, Israel

Steven Carr,
Broad Institute, United States

Jenn Abelin,
Broad Institute, United States,
in collaboration with reviewer SC

*Correspondence:

Claude Perreault
claude.perreault@umontreal.ca
Pierre Thibault
pierre.thibault@umontreal.ca

Specialty section:

This article was submitted to
Cancer Immunity
and Immunotherapy,
a section of the journal
Frontiers in Immunology

Received: 14 July 2020

Accepted: 30 October 2020

Published: 03 December 2020

Citation:

Minati R, Perreault C and Thibault P
(2020) A Roadmap Toward
the Definition of Actionable
Tumor-Specific Antigens.
Front. Immunol. 11:583287.
doi: 10.3389/fimmu.2020.583287

The search for tumor-specific antigens (TSAs) has considerably accelerated during the past decade due to the improvement of proteogenomic detection methods. This provides new opportunities for the development of novel antitumoral immunotherapies to mount an efficient T cell response against one or multiple types of tumors. While the identification of mutated antigens originating from coding exons has provided relatively few TSA candidates, the possibility of enlarging the repertoire of targetable TSAs by looking at antigens arising from non-canonical open reading frames opens up interesting avenues for cancer immunotherapy. In this review, we outline the potential sources of TSAs and the mechanisms responsible for their expression strictly in cancer cells. In line with the heterogeneity of cancer, we propose that discrete families of TSAs may be enriched in specific cancer types.

Keywords: tumor-specific antigens, neoantigens, immunopeptidome, proteogenomics, alternative antigens, cancer immunotherapy, pan-cancer antigen research

INTRODUCTION

The selection of “Cancer immunotherapy” by the journal *Science* as the breakthrough of the year in 2013 (1) properly illustrates how promoting the patients’ immune response against cancer cells has revolutionized the field of anticancer therapies. More recently, the Nobel prize award to Allison and Honjo, highlights the significance of immunotherapies, and how it changed the way we treat several types of cancers over the past decades. In contrast to classical treatment (i.e., surgery, chemotherapy, and radiation therapy) which target directly and aspecifically the tumoral cells, immunotherapies target the host’s immune system to initiate, augment and/or reestablish an efficient antitumoral immune response (2). Several types of immunotherapies including vaccines (3), antibodies (4), oncolytic viruses (5), immune checkpoint inhibitor (ICI) therapies (6, 7), and T cell-based immunotherapies are currently used in the clinic.

Independently from their mechanisms of action, all of these therapies rely on the ability of the patient’s adaptive immune system to discriminate between healthy (i.e., normal or stressed) cells and cancerous ones. At the molecular level, this distinction is possible because tumoral cells undergo a series of genetic and epigenetic changes leading to the generation of new self-derived antigens which are generally termed tumor-specific antigens (TSAs) or neoantigens. While neoantigens are defined as the subset of TSAs generated by genetic variations only found in the genome of a tumor, TSAs refers to all the antigen types which are specific to cancer cells (8). Because they are not expressed by the medullary thymic epithelial cells (mTECs), which are responsible for the establishment of the central

tolerance (9), TSAs represent a source of potentially immunogenic neoepitopes able to be recognized and targeted by the host's T cells (10). However, before being recognized by T cells through their T cell receptor (TCR), these antigens need to be sequentially processed and presented at the surface of the tumoral cell *via* major histocompatibility complex class I (MHC I) molecules. For MHC I, the antigen processing starts in the cytosol where intrinsic proteins—originating from the self in normal cells or altered-self in tumoral cells—are cleaved into peptides by the proteasome and some aminopeptidases (11). Then, the generated peptides are translocated in the endoplasmic reticulum (ER) *via* transporter associated with antigen processing (TAP) and further processed by the ER aminopeptidase 1 and 2 to reach a size ranging between 8 and 10 residues (11). Peptides are then loaded into the peptide cleft of a MHC I molecule and if the MHC I–peptide complex is stable enough, it is exported at the cell surface and referred to as MHC I-associated peptide (MAP) (11).

MAPs have a central role in T cell activation and more generally in anti-tumoral immunity. However, the question of how we can efficiently identify cancer-specific MAPs and, more generally TSAs, is now rising. As mentioned above, tremendous progress has been made in the way we treat many tumors. Unfortunately, the development of new antitumoral immunotherapies is now partially limited by the difficulty to identify targetable TSA and more precisely cancer-specific MAPs that could be used to initiate an efficient antitumoral immune response. Indeed, immunotherapeutic strategies which are used in the clinic either (i) bypass MHC I presentation (e.g., chimeric antigen receptor, CAR-T cell therapy) or (ii) skip the step of cancer-specific MAPs identification because they rely on a pre-existing antitumoral T cell response (e.g., ICI therapies).

In this review, we highlight the different sources of TSAs and the mechanisms responsible for their production in cancer cells with the objective to facilitate the identification of multiple targetable cancer-specific MAPs within tumors (**Figure 1**). Although tumor-associated antigens such as cancer-testis antigens can represent valuable source of antigens, we do not discuss them here and focus our topic to antigens absent from healthy tissue. In line with the heterogeneity of cancer, we also propose that discrete families of TSAs may be enriched in specific cancer types.

SINGLE NUCLEOTIDE VARIANT NEOANTIGENS: A RELATIVELY MEAGER SOURCE OF TUMOR-SPECIFIC ANTIGENS FOR IMMUNOTHERAPIES

Single nucleotide variants (SNVs)—and by extension double nucleotide variants—are commonly referred to as non-synonymous point mutations and can result, like most DNA damages, from (i) DNA replication errors, (ii) DNA enzymatic modification, (iii) exogenous or endogenous mutagen exposures, or (iv) defective DNA repair.

SNVs are the most common genetic variation (12) and they have been considered for a long time as the most promising

source of TSAs driving antitumoral responses. This mostly explains why the vast majority of the studies aiming to identify “tumoral neoantigens” have concentrated their efforts in detecting these non-synonymous point mutations located in known exons. Based on the original hypothesis that the TSA number of a given tumor is proportional to its SNV burden (commonly referred to as mutation burden), SNV-derived MAPs appeared as particularly good immunotherapeutic candidates for the most mutated tumor types—i.e., metastatic melanoma and lung cancers (13). However, all studies based on whole-exome or ribonucleic acid sequencing (WES and RNAseq respectively), combined or not with mass spectrometry (MS) analysis, could only identify a very limited number of SNV-derived MAPs. For example, in native human melanoma, which represents the cancer type having the highest mutation burden (13), Bassani-Sternberg et al. could only identify 11 cancer-specific immunopeptides (14). While this might be a limitation of MS sensitivity, it is puzzling that in these ICI-responding tumors, further investigations have shown that most of the identified cancer-specific SNVs were not immunogenic (14). The fact that most predicted TSAs are not validated by MS can be explained by two factors. Some of these TSAs are probably false negatives caused by the sensitivity of shotgun MS analyses (15). However, in-depth genomic analyses suggest that most false negatives are true negatives. Indeed, no evidence was found supporting the negative selection (*via* immunoeediting) of SNV-containing predicted TSAs (16, 17). Furthermore, response to ICI in patients with lung cancer did not correlate more with SNV-containing predicted TSAs than with the global mutation burden (18). These data suggest that the number of genuine SNV-containing TSAs has been overestimated in many studies. Tran et al. estimated that the *in vivo* immunogenicity—i.e., the capacity for antigens to be recognized by a tumor infiltrating lymphocyte (TIL)—of the whole exonic non-synonymous mutations of human gastrointestinal cancers was ranging between 0.009% to 1.25% (19). Although in some cases few neoantigens are sufficient to control the tumor or reach a therapeutic effect (20, 21), these results suggest that most cancer-specific MAPs able to trigger an antitumoral T cell response are likely not originating from exonic point mutations, and that other sources of neoantigens should be explored to drive future antitumoral immunotherapies (22).

MUTATIONAL-FRAMESHIFT NEOANTIGENS: A PROMISING SOURCE OF TUMOR-SPECIFIC ANTIGENS FOR RENAL CELL CARCINOMA, HOMOLOGOUS RECOMBINATION-DEFICIENT TUMORS, AND HIGH MICROSATELLITE INSTABILITY (MSI-H) TUMORS

After SNVs, nucleotide insertions or deletions (indels) represent the second most abundant type of mutation in

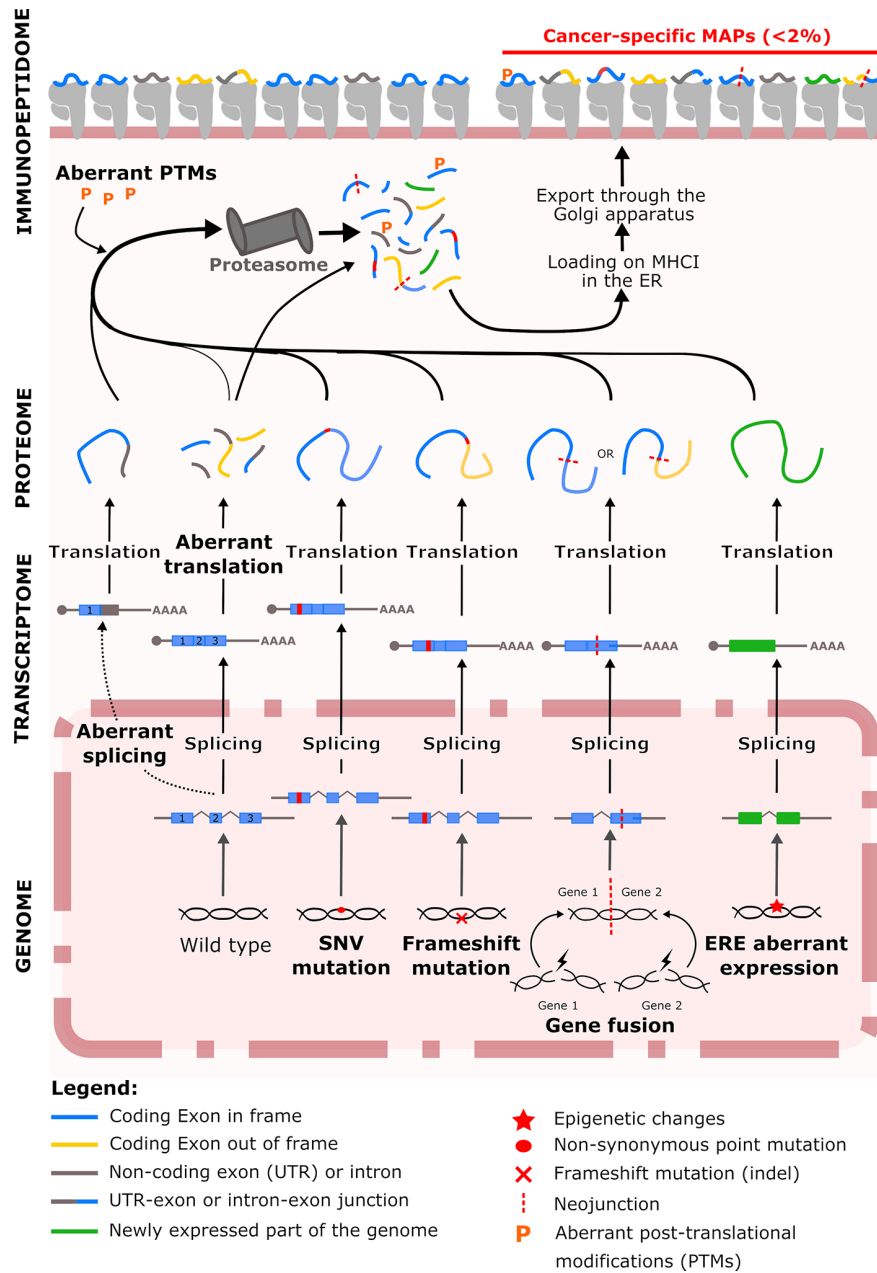


FIGURE 1 | Overview of the tumor-specific antigen production in tumor cells. At the genomic level, cancer cells accumulate tumor-specific genetic and epigenetic changes. Within genomic alteration, single-nucleotide variants (SNVs) represent an historic source of immunogenic neoantigens. Insertions/deletions (indels) or gene fusion events increase the tumor immunogenicity by generating peptide deriving from the out-of-frame translation of coding exons. Epigenetic alterations induce the aberrant expression of endogenous retroelements (EREs) which generated non-mutated cancer-specific peptides with a high immunogenic potential. In addition to genomic alterations, Post-translational modifications (PTMs) cancer-specific events such as aberrant splicing events, ribosomal translation and PTMs also contribute to the generation of cancer-specific major histocompatibility complex class I (MHC I)-associated peptides (MAPs). ER: endoplasmic reticulum.

cancer (23). With a size that ranges between one and more than 60 base pairs (bp), single and trinucleotides indels are the most common events and represent 68% and 13% of all indel events, respectively (23). Biologically and by extension immunologically, all indels are not equals and we can

distinguish two types: (i) a minority of in-frame indels that lead to the production of lowly immunogenic shorter or longer variants and (ii) a majority of frameshift indels that give rise to truncated protein variant containing new (potentially highly immunogenic) fragments derived from the out-of-frame

translation of a coding exon (23). Since frameshift indels (i) have the potential to generate more neoantigens than SNVs (24) and (ii) differ greatly from the germline sequences used for the establishment of the central tolerance, earlier reports suggested that they might serve as better immunotherapeutic targets (24, 25). In accordance with this proposal, several studies have shown a positive correlation between the indel burden and the presence of TILs inside the tumor or the response of the patients to ICI therapies (23, 24, 26–29). More importantly, in a context of non-relevant nonsense-mediated mRNA decay (NMD), truncated mutant proteins resulting from frameshift mutations have been reported to be extensively degraded by the proteasome system (30, 31). While this mechanism protects tumor cells from the potentially harmful effects of truncated proteins, it also promotes the generation of frameshift-derived peptides and their presentation at the cell surface (30).

Interestingly, indel burden also varies significantly across malignancies, though differences were noted regarding their exact proportions in various cancer types. Niavarani et al. reported that the indel proportion across cancers globally ranges between 1.3% and 29.1% (23) while this proportion was estimated between 1% and 12% by Turajlic et al. (24). It is noteworthy that in both studies, renal cell carcinomas, RCCs (i.e., chromophobe renal cell carcinoma, renal papillary cell carcinoma and renal clear cell carcinoma, KIRC) are classified among the cancers with the highest proportion of indels. At the therapeutic level, this is particularly promising because RCCs contain relatively few SNVs and the high proportion of indel opens new perspectives for neoantigen discovery. The presence of frameshift indel-derived antigens could explain the infiltration of RCCs by TILs and their good response to ICI therapies (32). In support of this, Hansen et al. recently reported that TILs from six patients with RCC could recognize both SNV- and frameshift-derived neoantigens (33). While frameshift-derived neoantigens represented only about 16% of the predicted TSAs, they corresponded to 21% of the immunogenic MAPs identified in the study.

In addition to being particularly abundant in RCC, indels have also been reported to accumulate importantly in both homologous-recombination (HR)- and DNA mismatch repair (MMR)-impaired tumors (13). Impaired HR repair pathway has been observed in subpopulations of breast, ovarian and pancreatic cancer where it is associated with an accumulation of numerous large deletions (up to 50 bp) along the genome (13). HR is normally used by dividing cells to guide the error-free repair of double-strand breaks (34) but when it is not available, other error-prone mechanisms ensure the breakpoint junction (13, 35, 36). In ovarian cancer, HR-deficiency is associated with a favorable clinical prognosis (36, 37). This is most likely due to an increase of both the tumoral neoantigen load and immunogenicity resulting from frameshifting indels accumulation. In support of this hypothesis, Strickland et al. showed that for high grade serous ovarian cancer, HR-deficient tumors presented more TILs, higher expression levels of programmed death 1 (PD-1) and programmed cell death

ligand 1 (PD-L1) and more putative neoantigen than HR-proficient ones (36). Similar results were reported for triple-negative breast cancer (38) but, to our knowledge, such studies were not yet conducted in pancreatic cancer. Although the exact contribution of the indel frameshift-derived neoantigens to the immunopeptidome of HR-deficient tumor cells still needs to be determined, this class of antigen seems to play an important role in the antitumoral T cell response and represent a promising immunotherapeutic target for subsets of ovarian, breast and pancreatic cancer.

Normally involved in the correction of indel loops and bp mismatches occurring during DNA replication, the MMR pathway is crucial to maintaining the microsatellite stability across the genome (39). As a result, its impairment in tumors leads to what is referred to as a microsatellite instability (MSI). Both MMR-deficiency and high MSI (MSI-H) have been primarily documented in both familial (Lynch syndrome) and sporadic subsets of colorectal cancers with a quite high prevalence (40–42). They are now reported in a wide range of malignancies including ovarian (43), endometrial (44), gastric (45), and prostate (46) cancers. On a pan-cancer scale, the endometrial, colon and gastric cancers are the cancer types displaying the three highest proportions of MSI-H cases (47). On an immunologic level, the high accumulation of somatic mutations by MSI-H tumors suggests that they should display several neoantigens (48). In support to this hypothesis, Le et al. demonstrated that pembrolizumab—an anti-PD-1 antibody—was effective in a wide range of MMR-deficient solid tumors (49, 50), opening the way for the first FDA tissue-agnostic approval of an ICI therapy. At the same time, they also showed that the response to pembrolizumab was associated with the *in vivo* expansion of T cell clones specific for tumoral indel-derived neoantigens providing a proof-of-concept of the relevance of targeting frameshift-derived neoantigens in MSI-H malignancies.

GENE FUSION NEOANTIGENS ARE RATHER RARE BUT RECURRENT ACROSS MALIGNANCIES

Gene fusion events are less frequent than SNVs and indel mutations, and consist in the juxtaposition of two previously independent coding sequences by (i) structural rearrangements at the genomic level (i.e., chromosomal translocation, inversion or deletion), (ii) transcription read-through of adjacent genes (51, 52), or (iii) *trans*- and *cis*-splicing of pre-mRNAs (53–55). In these three cases, the result is the production of a fusion transcript that can be translated into what we refer to as a fusion or chimeric protein. Although fusion events also occur in non-tumoral cells (56, 57), many cancer-specific fusion proteins have already been associated with a different malignancies including leukemia (58), sarcoma (59), breast (60), bladder (61), colon (62), and lung (63) cancers where they can be used as diagnosis and prognostic markers.

So far, most of the studies conducted on oncogenic fusion proteins have been focusing on either leukemia—i.e., acute myeloid leukemia, acute lymphocytic leukemia, or chronic myeloid leukemia—or sarcoma (64) where SNV burdens is relatively low (13, 65). In these malignancies, the hope of developing a vaccine was stimulated by the observation that some gene fusion events, such as the translocations t(11;22)(p13;q12) and t(12;22)(q13;q12), are particularly recurrent in given specific cancer subtypes (66). In a pan-cancer analysis of fusion events, Vellichirammal et al. have recently shown that the 40 most recurrent fusion events cover a wide spectrum of malignancies (67), and as such represent a promising source of multivalent neoantigens that could be used to drive cross-cancer immunotherapies.

Moreover, several gene fusions have been reported to act as driver mutations favoring tumorigenesis (68). Mechanistically, this oncogenic influence can be exerted either by altering the expression or activity of tumor suppressor or proto-oncogenes or by forming a fusion product with oncogenic properties (e.g., a constitutively activated tyrosine kinase domain). Because, in this case, fusion products are a source of oncogenesis and they are functionally linked with tumor fitness, targeting them may be more beneficial clinically than targeting other types of passenger mutations. Several pharmaceutical inhibition-based therapies targeting oncogenic fusion products have already shown promising results in the clinics (69–71). Regarding immunotherapies, several attempts were made to design fusion neoantigen-based vaccines but could only demonstrate moderate clinical efficacy (72–75). In these studies, even though the fusion peptides used for vaccination were able to activate the patients' T cells, all immunizations were performed with a single gene fusion epitope that may have favored the emergence of vaccine-resistant sub-clones. Although driver mutations have been shown to be highly clonal during the early stages of cancer, they tend to become highly heterogeneous and sub-clonal at later stages of the disease (76). Therefore, a driving fusion protein that may be essential for the survival of a transformed cell during cancer initiation can be completely absent from part of its progeny once the tumor is well established. In addition to this loss of clonality, cancer cells also develop different mechanisms enabling them to escape the immune surveillance. These mechanisms include the expression of immune checkpoints (7), a complete (77), or partial (78) loss of MHC I expression and the epigenetic silencing of neoantigens recognized by the immune system (79). These observations suggest that immunotherapies against fusion products-derived neoantigens would be more effective against early-stage cancers rather than later ones. On a pan-cancer level, patients with malignancies characterized by relatively low SNV and indel burden and a minimal immune-infiltration, such as leukemia (8), sarcoma (8), adenoid cystic carcinomas (80), or head and neck tumors (80), will most likely benefit the most from fusion product targeting. In the context of vaccine development, fusion proteins are a meaningful source of neoantigens, and their therapeutic value could be enhanced by combining several “driver” and “passenger” neoepitopes originating from different fusion proteins, and by

including ICI in the vaccination protocol to minimize the risk of immune evasion (80).

ENDOGENOUS RETROELEMENTS-DERIVED TUMOR-SPECIFIC ANTIGENS: A PREDOMINANT SOURCE OF NON-MUTATED ANTIGENS FOR A VACCINE AGAINST CANCER

Endogenous retroelements (EREs) represent about 42% of the human genome (81) and result from the integration of transposable elements into our genome millions of years ago. They comprise both long terminal repeat (LTR) elements (i.e., human endogenous retroviruses, HERVs, and mammalian apparent LTR-retrotransposons)—and non-LTR elements—which include long and short interspersed nuclear elements (LINEs and SINEs, respectively). Following their long co-evolution with human, the vast majority of EREs are now truncated and/or mutated and have lost their capacity to transpose in the genome (82–84). For those still able to “replicate” a strict epigenetic repression is maintained on their open reading frames (ORFs) to prevent the insertional mutagenesis and chromosomal rearrangements associated with their expression (85). Considered for a long time as “junk” DNA, the remnants of retroelements still contains functional promoters, enhancers, ORFs, splice donor/acceptor sites and polyadenylation sites able to impact cell physiology (85) and can contribute to several key processes of our development and adulthood (86–88). In line with this, Larouche et al. recently reported that various levels of ERE transcripts can be found in all human somatic tissues and that their expression is particularly predominant in mTECs which are responsible for T cell negative selection (89). These findings suggest that some antigens derived from these “domesticated” EREs are tolerated by the immune system.

In the context of cancer, the alteration of the epigenetic landscape or the use of demethylating therapies can result in the loss of repression marks along the genome and dysregulate ERE expression leading to the transcription and translation of aberrantly expressed EREs (aeEREs) (90–94). These aeERE have been reported to affect cancer progression through both pro- and antitumoral mechanisms (95). Previous reports indicated that aeERE could generate viral-like neoantigens able to increase both the antigenicity and immunogenicity of tumor cells (89, 94). Unlike the ERE-derived antigens expressed in normal tissues, those restricted to cancer cells (i.e., aeEREs) can be recognized by the immune system although they originate from non-mutated sequences. Indeed, several aeERE-derived MAPs were shown to activate CD8⁺ T cells in both B-lymphoblastoid cell line and KIRC (89, 96). Because aeEREs can produce non-mutated immunogenic neoantigens, they are now considered as a particularly attractive source of TSAs for the development of cancer vaccines. Unlike mutated neoantigens which are “private”, non-mutated TSA, such as aeEREs, are very likely to be shared across tumors and malignancies.

On a pan-cancer level, Attig et al. compared the expression levels of “cancer-specific LTR element-overlapping transcripts” (CLTs) across 31 cancer types and showed that the three malignancies with the highest number of CLTs were respectively the testicular germ cell tumors (TGCTs), the esophageal carcinoma and the ovarian serous cystadenocarcinoma (82). Although most of the overlap in CLT expression was observed in related tissues such as KIRC and renal papillary cell carcinoma, the study highlighted that 44 CLTs were shared by ten or more cancer types (82). Although a pan-cancer study including LINES and SINES is still needed, this LTR analysis supports the notion that aeEREs represent meaningful targets for the generation of shared TSAs.

THE POST-TRANSCRIPTIONAL ANTIGENS AND THEIR RISING INTEREST FOR IMMUNOTHERAPIES

Although most studies have focused on TSA classes arising from genomic alterations (e.g., SNV-, indels-, gene fusion-, and ERE-derived antigens), other classes of TSA exist and can still contribute to the development of antitumoral immunotherapies. Broadly referred to as post-transcriptional TSAs, this wide class of antigen regroups antigens derived from aberrant (i) alternative splicing, (ii) ribosomal events, and (iii) post-translational modifications (PTMs).

Aberrant Splicing-Derived Tumor-Specific Antigens

Alternative splicing of premature messenger RNAs (pre-mRNAs) is responsible for the diversification of both the transcriptome and the proteome of eukaryotic cells. This cellular process explains how one protein-coding gene can generate multiple alternative transcripts, also called variants, and give rise to different protein isoforms which are structurally and sometimes functionally different (97). Tightly regulated in time and space in normal cells, alternative splicing is carried out by the spliceosome machinery, and plays a key role in both cellular differentiation and identity (98). On a mechanistic level, alternative splicing events traditionally include intron retention, exon skipping, the use of alternative 5'- or 3'-splice site which lead to the retention of exon fragments, and exon mutual exclusion. However, since alternative promoters and alternative polyadenylation sites can generate transcripts with alternative 5'- and 3'-ends, they are sometimes considered as alternative splicing events although they are not directly carried out by the spliceosome machinery (99).

In cancer, it is now well established that both aberrant alternative splicing events (i.e., novel transcripts absent in normal cell) and alterations in the ratio of alternatively spliced transcripts occurs in a wide range of malignancies including breast (100, 101), brain (102), colon (103), prostate (103, 104), lung (105), and ovarian (101) cancers. Although both are cancer landmarks, only aberrant splicing events can generate cancer-specific transcript that can be translated into new protein

isoforms and produce immunogenic TSAs. Arising either from *cis*-acting splice junction mutations (106) or *trans*-acting spliceosome dysregulation (107, 108), aberrantly spliced transcripts lead to the formation and translation of cancer-specific junctions termed neojunctions. Based on the position and the nature of the neojunction—i.e., (i) in-frame exon-exon junctions (ii) out-of-frame exon-exon junctions, (iii) exon-intron junctions, or (iv) exon-untranslated region (UTR) junctions, the impact on protein's function and immunogenicity can be significant. At the functional level, all aspects of tumor development, progression, and response to treatments can be affected by aberrant alternative splicing and several known aberrantly splice variants have been shown to affect key processes such as metabolism, apoptosis, cell cycle control, angiogenesis, invasiveness, metastatic potential, and resistance to treatments of tumors (99). Similar to gene fusions, cancer-specific variants deriving from aberrant alternative splicing events can be functionally linked to the fitness of the tumor. However, unlike gene fusion products that can only be turned on or off, these splicing variants offer additional levels of plasticity which seem to be used by tumors to evade both immune surveillance (109) and therapies. In line with that, several aberrant spliced variants appear to be implicated with the resistance to several antitumoral treatments such as imatinib (110, 111), poly adenosine triphosphate (ADP)-ribose polymerase (PARP) inhibitor (112), cisplatin (112, 113), and tamoxifen (114). Although this link between aberrant protein isoform and drug resistance still needs to be confirmed, recent reports suggest that targeting aberrant splicing could resensitize cancer cells to existing therapies (99).

While mutations responsible for the occurrence of an aberrant splicing event were initially presumed to occur only in the consensus intronic dinucleotide splice donor (GT) or acceptor (AG) sites, it is now clear that other mutations can also affect RNA maturation (115–118). As a result, many mutations originally misannotated as silent, missense, insertion/deletion or nonsense mutations are now being considered as *cis*- and *trans*-acting splicing mutations (106). Unlike *cis*-acting splicing mutations which only affect the local splicing of the genes carrying them, *trans*-acting splicing mutations affect directly or indirectly the regulation of the spliceosome machinery and therefore the splicing process of many genes. By looking for splicing quantitative trait loci (sQTLs), Kahles et al. recently identify seven *trans*-sQTLs, two of which were associated with mutations in genes encoding the core spliceosome factors splicing factor 3b subunit 1 (SF3B1) and U2 small nuclear RNA auxiliary factor 1 (U2AF1) (107). Unexpectedly, other identified mutated genes (i.e., transcriptional adaptor 1, TADA1, protein phosphatase 2 scaffold subunit A alpha, PPP2R1A, epidermal growth factor receptor, EGFR, and isocitrate dehydrogenase 1, IDH1) were not previously known to impact the splicing of other genes. While the mechanistic basis behind this association still needs to be clarified, the landscape of *trans*-acting splicing variant continues to evolve. In another TCGA study, Seiler et al. identified far more genetic alterations able to impact other genes' alternative splicing in various ways by focusing on mutation

affecting 119 known splicing factor genes (108). While it may still be too early to determine how much *trans*-acting splicing mutations contribute to the immunopeptidome of a tumor, this class of mutation has the potential to generate much more neojunctions than its *cis*-acting homologue.

At the scale of different cancer types, like for SNVs or indels, the “neojunction/aberrant splicing burden” of the different malignancies varies considerably (107). On average, ovarian serous cystadenocarcinoma, liver hepatocellular carcinoma, esophageal carcinoma, and stomach adenocarcinoma are the four cancer types displaying the highest numbers of neojunctions. However, when looking at the median values, the global ranking changes drastically suggesting that important internal variations do exist within cancer types. In that context, in addition to this cancer-based analysis, it seems that a cancer-subtype approach could be used to identify more patient subgroups likely to benefit from the development of aberrant splicing targeting therapies.

Despite the fact that it is still unknown if the increase of aberrant splicing events in tumors is positively selected or not, splicing-derived cancer-specific proteins can potentially be used to inform cancer immunotherapies. In their studies, both Khaled et al. and Jayasinghe et al. predicted that aberrantly splicing events were much more likely to generate TSAs able to bind MHC I than SNV mutations (106, 107). Although the translation of several of the alternative splicing-derived putative neopeptides could be validated using publicly available MS data from other studies, most of their prediction relied on both RNAseq data and prediction algorithms. Since aberrant splicing is predominantly associated with low-abundance isoforms (119), the presentation of the corresponding peptides and their immunogenic potentials require experimental validation before they can be of immunotherapeutic value.

Tumor-Specific Antigen Derived From Aberrant Translation

Aberrant ribosomal translation events (ARTEs), sometimes called non-canonical translation events, correspond to the translation of either allegedly non-coding sequences or coding sequences in a non-canonical reading frame. Such events generally include non-canonical initiation, elongation and termination events. Briefly, a non-canonical initiation event occurs when the ribosome does not start the translation at the primary AUG codon—but at a non-primary AUG codon (120) or at a near-cognate start codon (CUG, UUG, or GUG) (121)—as a result of a start codon scan-through (122), a translation reinitiation (123) or the presence of an internal ribosome entry site (IRES) on the messenger RNA (124). Non-canonical elongation events happen when a frameshift occurs spontaneously during elongation and lead to the translation of a part of the protein in a non-canonical reading frame. Some slippage-prone sequences present within transcripts have already been reported to promote what is called a programmed ribosomal frameshift (125). Non-canonical termination events, although rare, are possible and consist of either a stop-codon read-through (126)—some stop codons such as UGA and UAG

appears to be leakier than UAA—or a ribosomal frameshift at the stop codon. When such ARTEs occur, they lead to the production of non-canonical proteins and cryptic MAPs.

While cryptic MAPs were initially observed as marginal and irrelevant, this view is now changing as we get a better understanding of their immunotherapeutic potentials. Recent studies indicated that at least 10% of the MAP repertoire is represented by cryptic peptides that are common among individuals carrying the same MHC I alleles (123). Interestingly, cryptic MAPs are involved in the establishment of the central tolerance and the priming of CD8⁺ T cells in mice (127). They are also recognized and targeted by both TILs (128–131) and auto-reactive T cells (120) in human. As their relevance to adaptive immunity becomes clearer, TSA research is virtually expanded from the ~2% of protein-coding genes to the ~75% of the transcribed genome (132). As such, ARTEs are redefining translation events at the whole transcriptome level while aberrant splicing events discussed in the previous section are delineating the boundaries of exons and introns.

ARTEs are found in both normal cells and tumoral cells, though their products differ depending on the genetic and epigenetic instability associated with cancer cells. This difference leads to the generation of cancer-specific cryptic MAPs that are relevant targets for vaccine development (93, 133, 134). Indeed, MAPs deriving from aberrant expression of non-mutated non-coding regions of the genome are much more likely to be shared by multiple tumors than randomly mutated sequences. Moreover, contrary to canonical MAPs, the generation of cryptic MAPs can be enhanced by inflammatory stimuli (135) such as type I interferon or tumor necrosis factor alpha (TNF α) and by drugs (136) such as aminoglycoside that might be used to increase the global immunogenicity of cold tumors.

Since cryptic MAPs cannot be identified using canonical protein databases, Laumont et al. have recently developed a proteogenomic-based approach to identify both of canonical and cryptic MAPs specific to tumor cells (e.g., mutated and aberrantly expressed TSA) (93). In parallel of this, two proof-of-principle studies established that MAPs can also be identified using reference databases built from ribosome profiling (Ribo-seq) (137, 138). Ribo-Seq is based on the sequencing of mRNA fragments protected for ribonuclease digestion by their location within the ribosome decoding site. It provides quantitative information on the nature of translated mRNAs including their reading frame and start and termination codons. While both of these approaches open new avenues for identification of cryptic MAPs which are potentially shared between patients, one unanswered question is the identification of TSAs among cryptic MAPs identified using databases built on Ribo-Seq. Indeed, this would require Ribo-Seq data for all types of normal cells.

Tumor-Specific Antigen Derived From Post-Translation Modifications

There are approximately 300 PTMs that have been described to modify proteins in normal condition (139). Among them, we find very diverse modifications—such as acetylation,

ubiquitination, glycosylation, SUMOylation, etc.—which are important to control the stability, localization, and conformation of proteins within the cell. Particularly important for the cell signaling, many PTMs have been shown to be dysregulated in the context of cancer (140–143).

Among the most studied PTMs, phosphorylation is the one associated with the largest number of diseases (144). In cancer, Reimand et al. showed, that SNV mutations affecting the phosphorylation sites could be found in nearly 90% of the tumors where they were associated with gains or losses of signaling contributing to what they called the “oncogenic rewriting of the kinase network” (145). On an antigenic level, phosphorylated proteins were shown to be processed normally by the antigen presentation pathway of both normal and tumoral cells (146–152). In line with the idea that they could be used for immunotherapies, several studies reported that phosphorylated peptides, but not their dephosphorylated counterparts, could activate T cells in the mice model (150, 153). Although their immunogenic potential has not been demonstrated in human, attempts are currently being made to integrate peptide phosphorylation into MHC I binding prediction tools (154). If they succeed, prediction tools should facilitate the detection of phosphorylated peptide and clarify their potential as a source of TSAs.

In addition to phosphorylation, other PTMs such as citrullination, ubiquitination and O-glycosylation might also contribute to both the antigenicity and immunogenicity of cancer cells in different ways. Citrullination results from the deimination of arginine residues into citrulline by a peptidyl-arginine deiminase (PAD). Despite being involved in several physiological processes, citrullination is predominantly known for its involvement in several autoimmune disorders including rheumatoid arthritis, multiple sclerosis, and type I diabetes where it was shown to be immunogenic (155–159). It is important to note that this PTM has also been identified in cancer (160). Citrullination levels seems to be higher in ovary, uterus, colon, bladder, breast, liver, lung, esophagus, kidney, and prostate tumors than in their corresponding normal tissues due to the overexpression of either PAD4 or PAD2 (161–163). Although the presentation of citrullinated MAPs on MHC I molecules has never been demonstrated, citrullination was shown to increase peptides binding affinity for HLA-DRB1 (a MHC class II allele) (164, 165) which could then be recognized by both mice and human repertoires of “cytotoxic” CD4⁺ T cells (166–168).

The β O-linked N-acetylglucosamine (O-GlcNAc) is a ubiquitous PTM modifying both serine and threonine residues and is involved in cell signaling of in all eukaryotic cells (169). This modification is reversibly attached and removed from its substrates in the cytosol or the nucleus of the cell by the O-linked N-acetylglucosaminyltransferase (OGT) and the β -N-acetylglucosaminidase (OGA), respectively (169). In normal cells, O-GlcNAcylation modulates several important biological functions such as the enzymatic and transcription activities, protein turnover, protein-protein interactions, and subcellular localization of several proteins (170, 171). Dysregulations of the

O-GlcNAcylation as well as aberrant expression of OGT and/or OGA have been observed in cancer where they are associated with increased cancer cell proliferation and survival, invasiveness, and metastasis (171). Because O-GlcNAcylated proteins are present at the level of the cytosol, O-GlcNAcylated MAPs were shown to be displayed at the cell surface and activate T cells (172, 173). The resolution of two different MHC I-glycopeptide structures by X-ray crystallography highlighted that the accessibility of the O-GlcNAc group to the TCR was key for the T cell reactivity (174). More recently, using MS, Malaker et al. could identified 36 unique O-GlcNAcylated MAPs from primary human leukemic and Epstein-Barr virus-transformed B cell (175). While the MAPs they identified presented various levels of glycosylation and methylation, five out of the seven tested could activate T cells from healthy donors. Although these antigens have not been proven tumor-specific, the authors reported that 92% of the identified O-GlcNAcs MAPs could not be detected in their healthy tissue samples (175).

While phosphorylation, citrullination or O-GlcNAcylation can be explored as potential sources of immunogenic TSAs, other PTMs such as ubiquitination do not directly provide tumor-specific epitopes but have been shown to affect peptide presentation. Ubiquitin is usually used as a degradation signal in the cell: when a protein reaches a threshold of ubiquitination, it is addressed to the proteasome where it is hydrolyzed into peptides (176). These peptides (among others) are a source of endogenous antigens for the MHC I presentation pathway. As a result, due to its key role in providing peptides for the MHC I immunosurveillance, dysregulation of the ubiquitination in cancer could lead to a modification of the immunopeptidome landscape. In case of a decrease in ubiquitination (intrinsic or pharmacologically-induced (177)), more peptides from the ubiquitin-independent presentation pathway will be presented at the cell surface. This includes peptides originating from small defective ribosome products (DRiPs) and aberrant translation products which seems to generate MAPs in a proteasome-independent manner (11). While the potential of PTMs for immunotherapies is still not yet fully accessed, preliminary results are promising and open additional perspective to target cancer.

FROM IDENTIFIED TUMOR-SPECIFIC ANTIGENS TO ACTIONABLE THERAPEUTICS: VALIDATION AND SELECTION

By definition the identification of a candidate requires to determine the correct amino acid sequence as well as the precise nucleic sequence its originates from. When the TSA candidate sequence is closed from similar to the reference sequence like it is the case for SNVs, the identification is relatively easy especially as soon as WES or RNAseq data are available. However for other TSA classes, it can be more challenging. In the case of frameshift mutations, indels are difficult to identify from Sanger and next-generation

sequencing which is why special tools such as pVACseq (178), Neopepsee (179), MuPeXI (180), Epidisco (181) and Antigen garnish (182) have been developed. While pVACtools, Epidisco, and Antigen.garnish also support the prediction of gene fusion-derived peptides, other tools such as INTEGRATE-neo have been specifically designed to predict fusion neoantigens (183). Because no prediction software is currently able to deal with repetitive regions, the prediction of ERE-derived TSA can only be made combining quantifiers such as hervQuant (184) or RepeatMasker (185) with a classic epitope prediction software such as NetMHCpan (186).

To face issues associated with the large search space of unbiased identification, one may want to conduct targeted searches with databases from human mRNA annotated sequences with associated variation information derived from the Single Nucleotide Polymorphism Database (dbSNP) and remove all non-polymorphic information (187). Searches for alternative reading frames, transcripts from non-coding regions (93) or EREs (89) can be achieved separately and results combined subsequently. As mentioned in this review, TSA classes have different abundances in different tumor types. Therefore, knowing which class might be predominant in a tumor of interest, could be used to guide identification by reducing search space.

Once identified, the actionability of the TSA is determined based on five parameters which evaluate/validate different aspects of its therapeutic potential.

First, TSA candidates have to be validated as “truly” tumor-specific. Because targeting antigens which are also expressed in healthy tissues could result in severe side effects or autoimmunity, Laumont et al. developed a stringent validation strategy based on the resources gathered by the GTEx consortium (93). In their study, candidates were only considered as “true” TSA if their corresponding reads were absent from the transcriptomes of a wide range healthy tissue. Although well adapted for most TSA classes, this validation strategy could not be apply to TSAs deriving from aberrant translations or PTMs. For these particular classes, their absence from healthy tissues could only be validated at the proteomic level, but, to our knowledge, it has never been done.

The second key criteria to assess TSA’s actionability is their immunogenic potential. To be targetable, a presented TSA has to be recognized by a TCR and able to trigger T cell activation. This is usually determined *ex vivo* by interferon gamma ELISpot, but a wide range of well-established assays have been described and can quantify other aspects of the T cell-dependent immunogenicity than interferon gamma production (188). We can also mention that many efforts are currently being made to develop machine-learning approaches to predict the immunogenicity of a given peptide from its sequence (189). Although this could both faster and ease the selection of therapeutic epitopes, this approach is still limited by the type of data available to train the programs.

Third, it is also important to estimate the incidence/prevalence of newly discovered TSA on a pancancer, tumoral and subtumoral scale. The more an immunogenic antigen is shared between patients (i.e., frequent within a cancer type/subtype or across malignancies), the higher its therapeutic

interest will be. This evaluation is usually performed by looking for TSA-corresponding transcripts in a large number of cancer transcriptomes from the TCGA database. By using this approach, Zhao et al. recently showed that 78% and 18% of the transcripts encoding for aberrantly expressed TSA in ovarian cancer were respectively expressed by at least 10% and 80% of the ovarian cancer samples (190). While this parameter is key to establishing the wide scale potential of an antigen, here again this strategy could not be applied for TSAs deriving from aberrant translations or PTMs.

Forth, the large-scale therapeutic potential of a candidate peptide is also affected by its MHC restriction (i.e., the number and frequencies of MHC I alleles it can bind to). If an immunogenic TSA is shared by an extensive number of tumors but only presented by a rare MHC I allele, its therapeutic interest for universal therapies will be decreased compared with another antigen which can be presented by a large portion of the population. TSA binding profiles are generally determined using MHC I-binding prediction tools such as NetMHC or NetMHCpan (186, 191) but needs to be experimentally validated either using a proteogenomic approach but ideally this should be done by T2 or RMA-S peptide binding assay (192, 193).

Fifth, the potential of TSAs will finally depend on the type of immunotherapy/strategy which is considered. In the particular context of vaccine design, all classes of TSA are not necessarily suitable for all vaccination strategy (i.e., DNA/mRNA-based vaccination (194), peptide-based vaccination (195), or TSA-loaded antigen presenting cell infusion (196, 197)). For example, PTM-derived TSA are not suitable for the development of DNA/mRNA-based vaccines. Similarly, private antigens would be less suitable for the development of a broad universal vaccine than shared non-mutated TSAs.

CONCLUSION

While the vast majority of the studies aiming to identify tumor antigens have concentrated their efforts in the detection of SNVs with limited therapeutic results, the possibility of enlarging the repertoire of targetable TSAs by looking at alternative classes of antigens opens new perspectives for the development of cancer immunotherapies. In line with recent improvements in both MAPs detection and prediction methods, our knowledge of these “alternative” sources of TSAs has remarkably increased over the past few years. From the maintenance of the genetic and epigenetic information at the genomic level to the ribosomal translation and PTM, every step of protein expression is susceptible to be dysregulated in cancer. While dysregulation may lead to the generation of specific types of TSA with their own features, they do not occur uniformly across malignancies. Therefore, more than “how”, the true question is now choosing “what” to identify. What class of neoantigen is the most likely to be predominant in this given type/subtype of tumor? What class of antigen is the most suitable for immunotherapy?

Given the fast-evolving nature of tumors and their genetic heterogeneity, it is very likely that future immunotherapies will need to target more than one TSA at once. While some neoantigens are derived from source proteins essential for tumor fitness, most TSAs that are now identified using conventional proteogenomic approaches are not necessarily required for tumor survival. To be efficient, immunotherapies must target multiple TSAs from different origins to cover the diversity of tumor subclones and prevent drug resistance.

AUTHOR CONTRIBUTIONS

All authors contributed to the design of the review. RM conducted literature search and wrote the first draft of the manuscript. All authors contributed to the article and approved the submitted version.

REFERENCES

- Couzin-Frankel J. Cancer Immunotherapy. *Science* (2013) 342:1432–3. doi: 10.1126/science.342.6165.1432
- Abbott M, Ustoyev Y. Cancer and the Immune System: The History and Background of Immunotherapy. *Semin Oncol Nurs* (2019) 35:150923. doi: 10.1016/j.soncn.2019.08.002
- Higano CS, Small EJ, Burch PA, Nemunaitis J, Yuh L, et al. Integrated data from 2 randomized, double-blind, placebo-controlled, phase 3 trials of active cellular immunotherapy with sipuleucel-T in advanced prostate cancer. *Cancer* (2009) 115:3670–9. doi: 10.1002/cncr.24429
- Grillo-López AJ, White CA, Varns C, Shen D, Wei A, McClure A, et al. Overview of the clinical development of rituximab: first monoclonal antibody approved for the treatment of lymphoma. *Semin Oncol* (1999) 26:66–73.
- Fukuhara H, Ino Y, Todo T. Oncolytic virus therapy: A new era of cancer treatment at dawn. *Cancer Sci* (2016) 107:1373–9. doi: 10.1111/cas.13027
- Sondak VK, Smalley KSM, Kudchadkar R, Gripton S, Kirkpatrick P. Ipilimumab. *Nat Rev Drug Discovery* (2011) 10:411–2. doi: 10.1038/nrd3463
- Wu Y, Chen W, Xu ZP, Gu W. PD-L1 Distribution and Perspective for Cancer Immunotherapy—Blockade, Knockdown, or Inhibition. *Front Immunol* (2019) 10:2022. doi: 10.3389/fimmu.2019.02022
- Smith CC, Selitsky SR, Chai S, Armistead PM, Vincent BG, Serody JS. Alternative tumour-specific antigens. *Nat Rev Cancer* (2019) 19:465–78. doi: 10.1038/s41568-019-0162-4
- Chan A, Anderson MS. Central tolerance to self revealed by the autoimmune regulator. *Ann N Y Acad Sci* (2015) 1356:80–9. doi: 10.1111/nyas.12960
- Hacohen N, Fritsch EF, Carter TA, Lander ES, Wu CJ. Getting personal with neoantigen-based therapeutic cancer vaccines. *Cancer Immunol Res* (2013) 1:11–5. doi: 10.1158/2326-6066.CIR-13-0022
- Laumont CM, Perreault C. Exploiting non-canonical translation to identify new targets for T cell-based cancer immunotherapy. *Cell Mol Life Sci* (2018) 75:607–21. doi: 10.1007/s00018-017-2628-4
- Zou H, Wu L-X, Tan L, Shang F-F, Zhou H-H. Significance of Single-Nucleotide Variants in Long Intergenic Non-protein Coding RNAs. *Front Cell Dev Biol* (2020) 8:347. doi: 10.3389/fcell.2020.00347
- Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SAJR, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. *Nature* (2013) 500:415–21. doi: 10.1038/nature12477
- Bassani-Sternberg M, Bräunlein E, Klar R, Engleitner T, Sinitcyn P, Audehm S, et al. Direct identification of clinically relevant neoepitopes presented on native human melanoma tissue by mass spectrometry. *Nat Commun* (2016) 7:1–16. doi: 10.1038/ncomms13404
- Haen SP, Löffler MW, Rammensee H-G, Brossart P. Towards new horizons: characterization, classification and implications of the tumour antigenic repertoire. *Nat Rev Clin Oncol* (2020) 17:595–610. doi: 10.1038/s41571-020-0387-x
- Alkallas R, et al. Multi-omic analysis reveals significantly mutated genes and DDX3X as a sex-specific tumor suppressor in cutaneous melanoma. *Nat Cancer* (2020) 1:635–52. doi: 10.1038/s43018-020-0077-8

FUNDING

This study was supported by grants from the Canadian Cancer Society (#705604) and the Leukemia and Lymphoma Society and Canada. RM is supported by a scholarship from MITACS. IRIC proteomics facility is a Genomics Technology platform funded in part by the Canadian Government through Genome Canada.

ACKNOWLEDGMENTS

We thank the Volff JN at from the École Normale Supérieure de Lyon for his critical reading of the manuscript and for valuable suggestions.

- Van den Eynden J, Jiménez-Sánchez A, Miller ML, Larsson E. Lack of detectable neoantigen depletion signals in the untreated cancer genome. *Nat Genet* (2019) 51:1741–8. doi: 10.1038/s41588-019-0532-6
- Hellmann MD, Nathanson T, Rizvi H, Creelan BC, Sanchez-Vega F, Ahuja A, et al. Genomic Features of Response to Combination Immunotherapy in Patients with Advanced Non-Small-Cell Lung Cancer. *Cancer Cell* (2018) 33:843–52.e4. doi: 10.1016/j.ccell.2018.03.018
- Tran E, Ahmadzadeh M, Lu Y-C, Gros A, Turcotte S, Robbins PF, et al. Immunogenicity of somatic mutations in human gastrointestinal cancers. *Science* (2015) 350:1387–90. doi: 10.1126/science.aad1253
- Gubin MM, Zhang X, Schuster H, Caron E, Ward JP, Nouguchi T, et al. Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens. *Nature* (2014) 515:577–81. doi: 10.1038/nature13988
- Tran E, Robbins PF, Rosenberg SA. ‘Final common pathway’ of human cancer immunotherapy: targeting random somatic mutations. *Nat Immunol* (2017) 18:255–62. doi: 10.1038/ni.3682
- Schumacher TN, Scheper W, Kvistborg P. Cancer Neoantigens. *Annu Rev Immunol* (2018) 37:173–200. doi: 10.1146/annurev-immunol-042617-053402
- Niavarani A, Shahrabi Farahani A, Sharafkhan M, Rassoulzadegan M. Pancancer analysis identifies prognostic high-APOBEC1 expression level implicated in cancer in-frame insertions and deletions. *Carcinogenesis* (2018) 39:327–35. doi: 10.1093/carcin/bgy005
- Turajlic S, Litchfield K, Xu H, Rosenthal R, McGranahan N, Reading JL, et al. Insertion-and-deletion-derived tumour-specific neoantigens and the immunogenic phenotype: a pan-cancer analysis. *Lancet Oncol* (2017) 18:1009–21. doi: 10.1016/S1470-2045(17)30516-8
- Linnebacher M, Gebert J, Rudy W, Woerner S, Yuan YP, Bork P, et al. Frameshift peptide-derived T-cell epitopes: a source of novel tumor-specific antigens. *Int J Cancer* (2001) 93:6–11. doi: 10.1002/ijc.1298
- Saeterdal I, Gjertsen MK, Straten P, Eriksen JA, Gaudernack G. A TGF betaRII frameshift-mutation-derived CTL epitope recognised by HLA-A2-restricted CD8+ T cells. *Cancer Immunol Immunother* (2001) 50:469–76. doi: 10.1007/s002620100222
- Tougeron D, Fauquembergue E, Rouquette A, Le Pessot F, Sesboué R, Laurent M, et al. Tumor-infiltrating lymphocytes in colorectal cancers with microsatellite instability are correlated with the number and spectrum of frameshift mutations. *Mod Pathol* (2009) 22:1186–95. doi: 10.1038/modpathol.2009.80
- Maby P, Tougeron D, Hamieh M, Mlecnik B, Kora H, Bindea G, et al. Correlation between Density of CD8+ T-cell Infiltrate in Microsatellite Unstable Colorectal Cancers and Frameshift Mutations: A Rationale for Personalized Immunotherapy. *Cancer Res* (2015) 75:3446–55. doi: 10.1158/0008-5472.CAN-14-3051
- Mandal R, Samstein RM, Lee K-W, Havel JJ, Wang H, Krishna C, et al. Genetic diversity of tumors with mismatch repair deficiency influences anti-PD-1 immunotherapy response. *Science* (2019) 364:485–91. doi: 10.1126/science.aau0447

30. Kim WK, Park M, Park M, Kim YJ, Shin N, Kim HK, et al. Identification and Selective Degradation of Neopeptide-Containing Truncated Mutant Proteins in the Tumors with High Microsatellite Instability. *Clin Cancer Res* (2013) 19:3369–82. doi: 10.1158/1078-0432.CCR-13-0684
31. Halvey PJ, Wang X, Wang J, Bhat AA, Dhawan P, Li M, et al. Proteogenomic analysis reveals unanticipated adaptations of colorectal tumor cells to deficiencies in DNA mismatch repair. *Cancer Res* (2014) 74:387–97. doi: 10.1158/0008-5472.CAN-13-2488
32. Motzer RJ, Escudier B, McDermott DF, George S, Hammers HJ, Srinivas S, et al. Nivolumab versus Everolimus in Advanced Renal-Cell Carcinoma. *N Engl J Med* (2015) 373:1803–13. doi: 10.1056/NEJMoa1510665
33. Hansen UK, Ramskov S, Bjerregaard A-M, Borch A, Andersen R, Draghi A, et al. Tumor-Infiltrating T Cells From Clear Cell Renal Cell Carcinoma Patients Recognize Neopeptides Derived From Point and Frameshift Mutations. *Front Immunol* (2020) 11:373. doi: 10.3389/fimmu.2020.00373
34. Chapman JR, Taylor MRG, Boulton SJ. Playing the End Game: DNA Double-Strand Break Repair Pathway Choice. *Mol Cell* (2012) 47:497–510. doi: 10.1016/j.molcel.2012.07.029
35. Pannunzio NR, Watanabe G, Lieber MR. Nonhomologous DNA end-joining for repair of DNA double-strand breaks. *J Biol Chem* (2018) 293:10512–23. doi: 10.1074/jbc.TM117.000374
36. Strickland KC, Howitt BE, Shukla SA, Rodig S, Ritterhouse LL, Liu JF, et al. Association and prognostic significance of BRCA1/2-mutation status with neoantigen load, number of tumor-infiltrating lymphocytes and expression of PD-1/PD-L1 in high grade serous ovarian cancer. *Oncotarget* (2016) 7:13587–98. doi: 10.18632/oncotarget.7277
37. Xu K, Yang S, Zhao Y. Prognostic significance of BRCA mutations in ovarian cancer: an updated systematic review with meta-analysis. *Oncotarget* (2016) 8:285–302. doi: 10.18632/oncotarget.12306
38. Nolan E, Savas P, Policheni AN, Darcy PK, Vaillant F, Mintoff CP, et al. Combined immune checkpoint blockade as a therapeutic strategy for BRCA1-mutated breast cancer. *Sci Transl Med* (2017) 9:eaal4922. doi: 10.1126/scitranslmed.aal4922
39. Boland CR, Koi M, Chang DK, Carethers JM. The biochemical basis of microsatellite instability and abnormal immunohistochemistry and clinical behavior in Lynch Syndrome: from bench to bedside. *Familial Cancer* (2008) 7:41–52. doi: 10.1007/s10689-007-9145-9
40. Ionov Y, Peinado MA, Malkhosyan S, Shibata D, Perucho M. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature* (1993) 363:558–61. doi: 10.1038/363558a0
41. Thibodeau SN, Bren G, Schaid D. Microsatellite instability in cancer of the proximal colon. *Science* (1993) 260:816–9. doi: 10.1126/science.8484122
42. Liu B, Nicolaides NC, Markowitz S, Willson JKV, Parsons RE, Jen J, et al. Mismatch repair gene defects in sporadic colorectal cancers with microsatellite instability. *Nat Genet* (1995) 9:48–55. doi: 10.1038/ng0195-48
43. Pal T, Permeth-Wey J, Kumar A, Sellers TA. Systematic Review and Meta-analysis of Ovarian Cancers: Estimation of Microsatellite-High Frequency and Characterization of Mismatch Repair Deficient Tumor Histology. *Clin Cancer Res* (2008) 14:6847–54. doi: 10.1158/1078-0432.CCR-08-1387
44. Boland CR, Thibodeau SN, Hamilton SR, Sidransky D, Eshleman JR, Burt RW, et al. A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res* (1998) 58:5248–57.
45. Sakurada K, Furukawa T, Kato Y, Kayama T, Huang S, Horii A. RIZ, the retinoblastoma protein interacting zinc finger gene, is mutated in genetically unstable cancers of the pancreas, stomach, and colorectum. *Genes Chromosomes Cancer* (2001) 30:207–11. doi: 10.1002/1098-2264(2000)9999:9999::AID-GCC1080>3.0.CO;2-V
46. Pritchard CC, Morrissey C, Kumar A, Zhang X, Smith C, Coleman I, et al. Complex MSH2 and MSH6 mutations in hypermutated microsatellite unstable advanced prostate cancer. *Nat Commun* (2014) 5:1–6. doi: 10.1038/ncomms5988
47. Hause RJ, Pritchard CC, Shendure J, Salipante SJ. Classification and characterization of microsatellite instability across 18 cancer types. *Nat Med* (2016) 22:1342–50. doi: 10.1038/nm.4191
48. Supek F, Lehner B. Differential DNA mismatch repair underlies mutation rate variation across the human genome. *Nature* (2015) 521:81–4. doi: 10.1038/nature14173
49. Le DT, Durham JN, Smith KN, Wang H, Bartlett BR, Aulakh LK, et al. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* (2017) 357:409–13. doi: 10.1126/science.aan6733
50. Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, et al. PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N Engl J Med* (2015) 372:2509–20. doi: 10.1056/NEJMoa1500596
51. Varley KE, Gertz J, Roberts BS, Davis NS, Bowling KM, Kirby MK, et al. Recurrent read-through fusion transcripts in breast cancer. *Breast Cancer Res Treat* (2014) 146:287–97. doi: 10.1007/s10549-014-3019-2
52. Nacu S, Yuan W, Kan Z, Bhatt D, Rivers CS, Stinson J, et al. Deep RNA sequencing analysis of readthrough gene fusions in human prostate adenocarcinoma and reference samples. *BMC Med Genomics* (2011) 4:11. doi: 10.1186/1755-8794-4-11
53. Jividen K, Li H. Chimeric RNAs generated by intergenic splicing in normal and cancer cells. *Genes Chromosomes Cancer* (2014) 53:963–71. doi: 10.1002/gcc.22207
54. Zhang Y, Gong M, Yuan H, Park HG, Frierson HF, Li H. Chimeric transcript generated by cis-splicing of adjacent genes regulates prostate cancer cell proliferation. *Cancer Discovery* (2012) 2:598–607. doi: 10.1158/2159-8290.CD-12-0042
55. Li H, Wang J, Ma X, Sklar J. Gene fusions and RNA trans-splicing in normal and neoplastic human cells. *Cell Cycle* (2009) 8:218–22. doi: 10.4161/cc.8.2.7358
56. Li X, Zhao L, Jiang H, Wang W. Short homologous sequences are strongly associated with the generation of chimeric RNAs in eukaryotes. *J Mol Evol* (2009) 68:56–65. doi: 10.1007/s00239-008-9187-0
57. Babiceanu M, Qin F, Xie Z, Jia Y, Lopez K, Janus N, et al. Recurrent chimeric fusion RNAs in non-cancer tissues and cells. *Nucleic Acids Res* (2016) 44:2859–72. doi: 10.1093/nar/gkw032
58. Rowley JD. A New Consistent Chromosomal Abnormality in Chronic Myelogenous Leukaemia identified by Quinacrine Fluorescence and Giemsa Staining. *Nature* (1973) 243:290–3. doi: 10.1038/243290a0
59. Alholle A, Karanian M, Brini AT, Morris MR, Kannappan V, Niada S, et al. Genetic analyses of undifferentiated small round cell sarcoma identifies a novel sarcoma subtype with a recurrent CRTC1-SS18 gene fusion. *J Pathol* (2018) 245:186–96. doi: 10.1002/path.5071
60. Kim RN, Moon H-G, Han W, Noh D-Y. Perspective Insight into Future Potential Fusion Gene Transcript Biomarker Candidates in Breast Cancer. *Int J Mol Sci* (2018) 19:502. doi: 10.3390/ijms19020502
61. Williams SV, Hurst CD, Knowles MA. Oncogenic FGFR3 gene fusions in bladder cancer. *Hum Mol Genet* (2013) 22:795–803. doi: 10.1093/hmg/dd5486
62. Seshagiri S, Stawiski EW, Durinck S, Modrusan Z, Storm EE, Conboy CB, et al. Recurrent R-spondin fusions in colon cancer. *Nature* (2012) 488:660–4. doi: 10.1038/nature11282
63. O'Regan L, Barone G, Adib R, Woo CG, Jeong HJ, Richardson EL, et al. EML4-ALK V3 oncogenic fusion proteins promote microtubule stabilization and accelerated migration through NEK9 and NEK7. *J Cell Sci* (2020). doi: 10.1242/jcs.241505
64. Xiao X, Garbutt CC, Hornicek F, Guo Z, Duan Z. Advances in chromosomal translocations and fusion genes in sarcomas and potential therapeutic applications. *Cancer Treat Rev* (2018) 63:61–70. doi: 10.1016/j.ctrv.2017.12.001
65. Abeshouse A, Adebamowo C, Adebamowo SN, Akbani R, Akeredolu T, Ally A, et al. Comprehensive and Integrated Genomic Characterization of Adult Soft Tissue Sarcomas. *Cell* (2017) 171:950–965.e28. doi: 10.1016/j.cell.2017.10.014
66. Worley BS, van den Broeke LT, Goletz TJ, Pendleton CD, Daschbach EM, Thomas EK, et al. Antigenicity of Fusion Proteins from Sarcoma-associated Chromosomal Translocations. *Cancer Res* (2001) 61:6868–75.
67. Vellichirammal NN, Albahrani A, Banwait JK, Mishra NK, Li Y, Roychoudhury S, et al. Pan-Cancer Analysis Reveals the Diverse Landscape of Novel Sense and Antisense Fusion Transcripts. *Mol Ther Nucleic Acids* (2020) 19:1379–98. doi: 10.1016/j.omtn.2020.01.023
68. Gao Q, Liang W-W, Foltz SM, Mutharasu G, Jayasinghe RG, Cao S, et al. Driver Fusions and Their Implications in the Development and Treatment of Human Cancers. *Cell Rep* (2018) 23:227–38.e3. doi: 10.1016/j.celrep.2018.03.050
69. Druker BJ, Talpaz M, Resta DJ, Peng B, Buchdunger E, Ford JM, et al. Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in

- chronic myeloid leukemia. *N Engl J Med* (2001) 344:1031–7. doi: 10.1056/NEJM200104053441401
70. Rossari F, Minutolo F, Orciuolo E. Past, present, and future of Bcr-Abl inhibitors: from chemical development to clinical efficacy. *J Hematol Oncol* (2018) 11:84. doi: 10.1186/s13045-018-0624-2
 71. Cocco E, Scaltriti M, Drilon A. NTRK fusion-positive cancers and TRK inhibitor therapy. *Nat Rev Clin Oncol* (2018) 15:731–47. doi: 10.1038/s41571-018-0113-0
 72. Cathcart K, Pinilla-Ibarz J, Korontsvit T, Schwartz J, Zakhaleva V, Papadopoulos EB, et al. A multivalent bcr-abl fusion peptide vaccination trial in patients with chronic myeloid leukemia. *Blood* (2004) 103:1037–42. doi: 10.1182/blood-2003-03-0954
 73. Mackall CL, Rhee EH, Read EJ, Khuu HM, Leitman SF, Bernstein D, et al. A pilot study of consolidative immunotherapy in patients with high-risk pediatric sarcomas. *Clin Cancer Res* (2008) 14:4850–8. doi: 10.1158/1078-0432.CCR-07-4065
 74. Bocchia M, Gentili S, Abruzzese E, Fanelli A, Iuliano F, Tabilio A, et al. Effect of a p210 multi-peptide vaccine associated with imatinib or interferon in patients with chronic myeloid leukaemia and persistent residual disease: a multicentre observational trial. *Lancet* (2005) 365:657–62. doi: 10.1016/S0140-6736(05)70931-4
 75. Rojas JM, Knight K, Wang L, Clark RE. Clinical evaluation of BCR-ABL peptide immunisation in chronic myeloid leukaemia: results of the EPIC study. *Leukemia* (2007) 21:2287–95. doi: 10.1038/sj.leu.2404858
 76. Jamal-Hanjani M, Wilson GA, McGranahan N, Birkbak NJ, Watkins TBK, Veeriah S, et al. Tracking the Evolution of Non-Small-Cell Lung Cancer. *New Engl J Med* (2017) 376:2109–21. doi: 10.1056/NEJMoa1616288
 77. Hicklin DJ, Wang Z, Arienti F, Rivoltini L, Parmiani G, Ferrone S. beta2-Microglobulin mutations, HLA class I antigen loss, and tumor progression in melanoma. *J Clin Invest* (1998) 101:2720–9. doi: 10.1172/JCI498
 78. McGranahan N, Rosenthal R, Hiley CT, Rowan AJ, Watkins TBK, Wilson GA, et al. Allele-Specific HLA Loss and Immune Escape in Lung Cancer Evolution. *Cell* (2017) 171:1259–71.e11. doi: 10.1016/j.cell.2017.10.001
 79. Rosenthal R, Cadieux EL, Salgado R, Bakir MA, Moore DA, Hiley CT, et al. Neoantigen-directed immune escape in lung cancer evolution. *Nature* (2019) 567:479–85. doi: 10.1038/s41586-019-1032-7
 80. Yang W, Lee K-W, Srivastava RM, Kuo F, Krishna C, Chowell D, et al. Immunogenic neoantigens derived from gene fusions stimulate T cell responses. *Nat Med* (2019) 25:767–75. doi: 10.3410/f.735590731.793566567
 81. Bannert N, Kurth R. Retroelements and the human genome: New perspectives on an old relation. *PNAS* (2004) 101:14572–9. doi: 10.1073/pnas.0404838101
 82. Attig J, Young GR, Hosie L, Perkins D, Encheva-Yokoya V, Stoye JP, et al. LTR retroelement expansion of the human cancer transcriptome and immunopeptidome revealed by de novo transcript assembly. *Genome Res* (2019) 29:1578–90. doi: 10.1101/gr.248922.119
 83. Burns KH, Boeke JD. Human Transposon Tectonics. *Cell* (2012) 149:740–52. doi: 10.1016/j.cell.2012.04.019
 84. Kassiotis G, Stoye JP. Immune responses to endogenous retroelements: taking the bad with the good. *Nat Rev Immunol* (2016) 16:207–19. doi: 10.1038/nri.2016.27
 85. Feschotte C, Gilbert C. Endogenous viruses: insights into viral evolution and impact on host biology. *Nat Rev Genet* (2012) 13:283–96. doi: 10.1038/nrg3199
 86. Mi S, Lee X, Li X, Veldman GM, Finnerty H, Racie L, et al. Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. *Nature* (2000) 403:785–9. doi: 10.1038/35001608
 87. Beyer U, Moll-Rocek J, Moll UM, Döbelstein M. Endogenous retrovirus drives hitherto unknown proapoptotic p63 isoforms in the male germ line of humans and great apes. *Proc Natl Acad Sci USA* (2011) 108:3624–9. doi: 10.1073/pnas.1016201108
 88. Chuong EB, Elde NC, Feschotte C. Regulatory evolution of innate immunity through co-option of endogenous retroviruses. *Science* (2016) 351:1083–7. doi: 10.1126/science.125497
 89. Larouche J-D, Trofimov A, Hesnard L, Ehx G, Zhao Q, Vincent K, et al. Widespread and tissue-specific expression of endogenous retroelements in human somatic tissues. *Genome Med* (2020) 12:1–16. doi: 10.1186/s13073-020-00740-7
 90. Florl AR, Löwer R, Schmitz-Dräger BJ, Schulz WA. DNA methylation and expression of LINE-1 and HERV-K provirus sequences in urothelial and renal cell carcinomas. *Br J Cancer* (1999) 80:1312–21. doi: 10.1038/sj.bjc.6690524
 91. Brocks D, et al. DNMT and HDAC inhibitors induce cryptic transcription start sites encoded in long terminal repeats. *Nat Genet* (2017) 49:1052–60. doi: 10.1038/ng.3889
 92. Chiappinelli KB, Strissel PL, Desrichard A, Li H, Henke C, Akman B, et al. Inhibiting DNA Methylation Causes an Interferon Response in Cancer via dsRNA Including Endogenous Retroviruses. *Cell* (2015) 162:974–86. doi: 10.1016/j.cell.2015.07.011
 93. Laumont CM, et al. Noncoding regions are the main source of targetable tumor-specific antigens. *Sci Transl Med* (2018) 10:eaa5516. doi: 10.1126/scitranslmed.aau5516
 94. Kong Y, Rose CM, Cass AA, Williams AG, Darwish M, Lianoglou S, et al. Transposable element expression in tumors is associated with immune infiltration and increased antigenicity. *Nat Commun* (2019) 10:5228. doi: 10.1038/s41467-019-13035-2
 95. Kassiotis G, Stoye JP. Making a virtue of necessity: the pleiotropic role of human endogenous retroviruses in cancer. *Philos Trans R Soc B: Biol Sci* (2017) 372:20160277. doi: 10.1098/rstb.2016.0277
 96. Cherkasova E, Scrivani C, Doh S, Weisman Q, Takahashi Y, Harashima N, et al. Detection of an Immunogenic HERV-E Envelope with Selective Expression in Clear Cell Kidney Cancer. *Cancer Res* (2016) 76:2177–85. doi: 10.1158/0008-5472.CAN-15-3139
 97. Stamm S, Ben-Ari S, Rafalska I, Tang Y, Zhang Z, Toiber D, et al. Function of alternative splicing. *Gene* (2005) 344:1–20. doi: 10.1016/j.gene.2004.10.022
 98. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* (2010) 28:511–5. doi: 10.1038/nbt.1621
 99. Wang B-D, Lee NH. Aberrant RNA Splicing in Cancer and Drug Resistance. *Cancers (Basel)* (2018) 10:458. doi: 10.3390/cancers10110458
 100. Omenn GS, Yocum AK, Menon R. Alternative splice variants, a new class of protein cancer biomarker candidates: findings in pancreatic cancer and breast cancer with systems biology implications. *Dis Markers* (2010) 28:241–51. doi: 10.1155/2010/705847
 101. Venables JP, Klinck R, Koh C, Gervais-Bird J, Bramard A, Inkel L, et al. Cancer-associated regulation of alternative splicing. *Nat Struct Mol Biol* (2009) 16:670–6. doi: 10.1038/nsmb.1608
 102. Venables JP. Unbalanced alternative splicing and its significance in cancer. *BioEssays* (2006) 28:378–86. doi: 10.1002/bies.20390
 103. Thorsen K, Sørensen KD, Brems-Eskildsen AS, Modin C, Gaustadnes M, Hein A-MK, et al. Alternative splicing in colon, bladder, and prostate cancer identified by exon array analysis. *Mol Cell Proteomics* (2008) 7:1214–24. doi: 10.1074/mcp.M700590-MCP200
 104. Rajan P, Elliott DJ, Robson CN, Leung HY. Alternative splicing and biological heterogeneity in prostate cancer. *Nat Rev Urol* (2009) 6:454–60. doi: 10.1038/nrurol.2009.125
 105. Misquitta-Ali CM, Cheng E, O'Hanlon D, Liu N, McGlade CJ, Tsao MS, et al. Global profiling and molecular characterization of alternative splicing events misregulated in lung cancer. *Mol Cell Biol* (2011) 31:138–50. doi: 10.1128/MCB.00709-10
 106. Jayasinghe RG, Cao S, Gao Q, Wendl MC, Vo NS, Reynolds SM, et al. Systematic Analysis of Splice-Site-Creating Mutations in Cancer. *Cell Rep* (2018) 23:270–281.e3. doi: 10.1016/j.celrep.2018.03.052
 107. Kahles A, Lehmann K-V, Toussaint NC, Hüser M, Stark SG, Sachsenberg T, et al. Comprehensive Analysis of Alternative Splicing Across Tumors from 8,705 Patients. *Cancer Cell* (2018) 34:211–224.e6. doi: 10.1016/j.ccell.2018.07.001
 108. Seiler M, Peng S, Agrawal AA, Palacino J, Teng T, Zhu P, et al. Somatic Mutational Landscape of Splicing Factor Genes and Their Functional Consequences across 33 Cancer Types. *Cell Rep* (2018) 23:282–296.e4. doi: 10.1016/j.celrep.2018.01.088
 109. Sotillo E, Barrett DM, Black KL, Bagashev A, Oldridge D, Wu G, et al. Convergence of Acquired Mutations and Alternative Splicing of CD19 Enables Resistance to CART-19 Immunotherapy. *Cancer Discovery* (2015) 5:1282–95. doi: 10.1158/2159-8290.CD-15-1020
 110. Berman E, Jhanwar S, Hedvat C, Arcila ME, Wahab OA-, Levine R, et al. Resistance to imatinib in patients with chronic myelogenous leukemia and

- the splice variant BCR-ABL1(35INS). *Leuk Res* (2016) 49:108–12. doi: 10.1016/j.leukres.2016.08.006
111. O'Hare T, Zabriskie MS, Eide CA, Agarwal A, Adrian LT, You H, et al. The BCR-ABL35INS insertion/truncation mutant is kinase-inactive and does not contribute to tyrosine kinase inhibitor resistance in chronic myeloid leukemia. *Blood* (2011) 118:5250–4. doi: 10.1182/blood-2011-05-349191
 112. Wang Y, Bernhardt AJ, Cruz C, Kraiss JJ, Nacson J, Nicolas E, et al. The BRCA1-Δ11q Alternative Splice Isoform Bypasses Germline Mutations and Promotes Therapeutic Resistance to PARP Inhibition and Cisplatin. *Cancer Res* (2016) 76:2778–90. doi: 10.1158/0008-5472.CAN-16-0186
 113. Avery-Kiejda KA, Zhang XD, Adams LJ, Scott RJ, Wojtesek B, Lane DP, et al. Small Molecular Weight Variants of p53 Are Expressed in Human Melanoma Cells and Are Induced by the DNA-Damaging Agent Cisplatin. *Clin Cancer Res* (2008) 14:1659–68. doi: 10.1158/1078-0432.CCR-07-1422
 114. Zhang X, Wang Z-Y. Estrogen Receptor-α Variant, ER-α36, is Involved in Tamoxifen Resistance and Estrogen Hypersensitivity. *Endocrinology* (2013) 154:1990–8. doi: 10.1210/en.2013-1116
 115. Nyström-Lahti M, Holmberg M, Fidalgo P, Salovaara R, de la Chapelle A, Jiricny J, et al. Missense and nonsense mutations in codon 659 of MLH1 cause aberrant splicing of messenger RNA in HNPCC kindreds. *Genes Chromosomes Cancer* (1999) 26:372–5. doi: 10.1002/(SICI)1098-2264(199912)26:4<372::AID-GCC12>3.0.CO;2-V
 116. Woolfe A, Mullikin JC, Elmtski L. Genomic features defining exonic variants that modulate splicing. *Genome Biol* (2010) 11:R20. doi: 10.1186/gb-2010-11-2-r20
 117. Zhang K, Nowak I, Rushlow D, Gallie BL, Lohmann DR. Patterns of missplicing caused by RB1 gene mutations in patients with retinoblastoma and association with phenotypic expression. *Hum Mutat* (2008) 29:475–84. doi: 10.1002/humu.20664
 118. Rice GI, Reijns MAM, Coffin SR, Forte GMA, Anderson BH, Szykiewicz M, et al. Synonymous mutations in RNASEH2A create cryptic splice sites impairing RNase H2 enzyme function in Aicardi-Goutières syndrome. *Hum Mutat* (2013) 34:1066–70. doi: 10.1002/humu.22336
 119. Pickrell JK, Pai AA, Gilad Y, Pritchard JK. Noisy Splicing Drives mRNA Isoform Diversity in Human Cells. *PLoS Genet* (2010) 6:e1001236. doi: 10.1371/journal.pgen.1001236
 120. Kracht MJL, van Lummel M, Nikolic T, Joosten AM, Laban S, van der Slik AR, et al. Autoimmunity against a defective ribosomal insulin gene product in type 1 diabetes. *Nat Med* (2017) 23:501–7. doi: 10.1038/nm.4289
 121. Liang H, Chen X, Yin Q, Ruan D, Zhao X, Zhang C, et al. PTENβ is an alternatively translated isoform of PTEN that regulates rDNA transcription. *Nat Commun* (2017) 8:14771. doi: 10.1038/ncomms14771
 122. Bullock TN, Eisenlohr LC. Ribosomal scanning past the primary initiation codon as a mechanism for expression of CTL epitopes encoded in alternative reading frames. *J Exp Med* (1996) 184:1319–29. doi: 10.1084/jem.184.4.1319
 123. Laumont CM, Daouda T, Laverdure J-P, Bonneil É, Caron-Lizotte O, Hardy M-P, et al. Global proteogenomic analysis of human MHC class I-associated peptides derived from non-canonical reading frames. *Nat Commun* (2016) 7:1–12. doi: 10.1038/ncomms10238
 124. Carbonnelle D, Vignard V, Schedic D, Moreau-Aubry A, Florenceau L, Charpentier M, et al. The Melanoma Antigens MELOE-1 and MELOE-2 Are Translated from a Bona Fide Polycistronic mRNA Containing Functional IRES Sequences. *PLoS One* (2013) 8:e75233. doi: 10.1371/journal.pone.0075233
 125. Namy O, Moran SJ, Stuart DI, Gilbert RJC, Brierley I. A mechanical explanation of RNA pseudoknot function in programmed ribosomal frameshifting. *Nature* (2006) 441:244–7. doi: 10.1038/nature04735
 126. Schueren F, Lingner T, George R, Hofhuis J, Dickel C, Gärtner J, et al. Peroxisomal lactate dehydrogenase is generated by translational readthrough in mammals. *Elife* (2014) 3:e03640. doi: 10.7554/eLife.03640
 127. Schwab SR, Li KC, Kang C, Shastri N. Constitutive Display of Cryptic Translation Products by MHC Class I Molecules. *Science* (2003) 301:1367–71. doi: 10.1126/science.1085650
 128. van Bergen CAM, van Luxemburg-Heijs SAP, de Wreede LC, Eefting M, von dem Borne PA, van Balen P, et al. Selective graft-versus-leukemia depends on magnitude and diversity of the alloreactive T cell response. *J Clin Invest* (2017) 127:517–29. doi: 10.1172/JCI86175
 129. Rosenberg SA, Tong-On P, Li Y, Riley JP, El-Gamil M, Parkhurst MR, et al. Identification of BING-4 Cancer Antigen Translated From an Alternative Open Reading Frame of a Gene in the Extended MHC Class II Region Using Lymphocytes From a Patient With a Durable Complete Regression Following Immunotherapy. *J Immunol* (2002) 168:2402–7. doi: 10.4049/jimmunol.168.5.2402
 130. Wang R-F, Johnston SL, Zeng G, Topalian SL, Schwartzentruber DJ, Rosenberg SA. A Breast and Melanoma-Shared Tumor Antigen: T Cell Responses to Antigenic Peptides Translated from Different Open Reading Frames. *J Immunol* (1998) 161:3596–606.
 131. Wang RF, Parkhurst MR, Kawakami Y, Robbins PF, Rosenberg SA. Utilization of an alternative open reading frame of a normal gene in generating a novel human cancer antigen. *J Exp Med* (1996) 183:1131–40. doi: 10.1084/jem.183.3.1131
 132. Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, et al. Landscape of transcription in human cells. *Nature* (2012) 489:101–8. doi: 10.1038/nature11233
 133. Gee MH, Han A, Lofgren SM, Beausang JF, Mendoza JL, Birnbaum ME, et al. Antigen Identification for Orphan T Cell Receptors Expressed on Tumor-Infiltrating Lymphocytes. *Cell* (2018) 172:549–563.e16. doi: 10.1016/j.cell.2017.11.043
 134. Probst P, Kopp J, Oxenius A, Colombo MP, Ritz D, Fugmann T, et al. Sarcoma Eradication by Doxorubicin and Targeted TNF Relies upon CD8+ T-cell Recognition of a Retroviral Antigen. *Cancer Res* (2017) 77:3644–54. doi: 10.1158/0008-5472.CAN-16-2946
 135. Prasad S, Starck SR, Shastri N. Presentation of Cryptic Peptides by MHC Class I Is Enhanced by Inflammatory Stimuli. *J Immunol* (2016) 197:2981–91. doi: 10.4049/jimmunol.1502045
 136. Goodenough E, Robinson TM, Zook MB, Flanigan KM, Atkins JF, Howard MT, et al. Cryptic MHC class I-binding peptides are revealed by aminoglycoside-induced stop codon read-through into the 3' UTR. *PNAS* (2014) 111:5670–5. doi: 10.1073/pnas.1402670111
 137. Ouspenskaia T, Law T, Clauser KR, Klaeger S, Sarkizova S, Aguet F, et al. Thousands of novel unannotated proteins expand the MHC I immunopeptidome in cancer. *bioRxiv* (2020) 2020.02.12.945840. doi: 10.1101/2020.02.12.945840
 138. Chong C, Müller M, Pak H, Harnett D, Huber F, Grun D, et al. Integrated proteogenomic deep sequencing and analytics accurately identify non-canonical peptides in tumor immunopeptidomes. *Nat Commun* (2020) 11:1293. doi: 10.1038/s41467-020-14968-9
 139. Witze ES, Old WM, Resing KA, Ahn NG. Mapping protein post-translational modifications with mass spectrometry. *Nat Methods* (2007) 4:798–806. doi: 10.1038/nmeth1100
 140. Archer TC, Ehrenberger T, Mundt F, Gold MP, Krug K, Mah CK, et al. Proteomics, Post-translational Modifications, and Integrative Analyses Reveal Molecular Heterogeneity within Medulloblastoma Subgroups. *Cancer Cell* (2018) 34:396–410.e8. doi: 10.1016/j.ccell.2018.08.004
 141. Jarrold J, Davies CC. PRMTs and Arginine Methylation: Cancer's Best-Kept Secret? *Trends Mol Med* (2019) 25:993–1009. doi: 10.1016/j.molmed.2019.05.007
 142. Krueger KE, Srivastava S. Posttranslational Protein Modifications: Current Implications for Cancer Detection, Prevention, and Therapeutics. *Mol Cell Proteomics* (2006) 5:1799–810. doi: 10.1074/mcp.R600009-MCP200
 143. Cho Y, Kang HG, Kim S-J, Lee S, Jee S, Ahn SG, et al. Post-translational modification of OCT4 in breast cancer tumorigenesis. *Cell Death Diff* (2018) 25:1781–95. doi: 10.1038/s41418-018-0079-6
 144. Xu H, Wang Y, Lin S, Deng W, Peng D, Cui Q, et al. PTMD: A Database of Human Disease-associated Post-translational Modifications. *Genom Proteomics Bioinf* (2018) 16:244–51. doi: 10.1016/j.gpb.2018.06.004
 145. Reimand J, Wagih O, Bader GD. The mutational landscape of phosphorylation signaling in cancer. *Sci Rep* (2013) 3:2651. doi: 10.1038/srep02651
 146. Andersen MH, Bonfill JE, Neisig A, Arsequell G, Sondergaard I, Valencia G, et al. Phosphorylated peptides can be transported by TAP molecules, presented by class I MHC molecules, and recognized by phosphopeptide-specific CTL. *J Immunol* (1999) 163:3812–8.
 147. Zarling AL, Ficarro SB, White FM, Shabanowitz J, Hunt DF, Engelhard VH, et al. Phosphorylated peptides are naturally processed and presented by major histocompatibility complex class I molecules in vivo. *J Exp Med* (2000) 192:1755–62. doi: 10.1084/jem.192.12.1755

148. Mohammed F, Stones DH, Zarling AL, Willcox CR, Shabanowitz J, Cummings KL, et al. The antigenic identity of human class I MHC phosphopeptides is critically dependent upon phosphorylation status. *Oncotarget* (2017) 8:54160–72. doi: 10.18632/oncotarget.16952
149. Olsson N, Schultz LM, Zhang L, Khodadoust MS, Narayan R, Czerwinski DK, et al. T-Cell Immunopeptidomes Reveal Cell Subtype Surface Markers Derived From Intracellular Proteins. *Proteomics* (2018) 18:e1700410. doi: 10.1002/pmic.201700410
150. Zarling AL, Polefrone JM, Evans AM, Mikesh LM, Shabanowitz J, Lewis ST, et al. Identification of class I MHC-associated phosphopeptides as targets for cancer immunotherapy. *Proc Natl Acad Sci* (2006) 103:14889–94. doi: 10.1073/pnas.0604045103
151. Meyer VS, Drews O, Günder M, Hennenlotter J, Rammensee H-G, Stevanovic S, et al. Identification of natural MHC class II presented phosphopeptides and tumor-derived MHC class I phospholigands. *J Proteome Res* (2009) 8:3666–74. doi: 10.1021/pr800937k
152. Petersen J, Wurzbacher SJ, Williamson NA, Ramarathinam SH, Reid HH, Nair AKN, et al. Phosphorylated self-peptides alter human leukocyte antigen class I-restricted antigen presentation and generate tumor-specific epitopes. *Proc Natl Acad Sci USA* (2009) 106:2776–81. doi: 10.1073/pnas.0812901106
153. Lin M-H, Shen K-Y, Liu B-S, Chen I-H, Sher Y-P, Tseng G-C, et al. Immunological evaluation of a novel HLA-A2 restricted phosphopeptide of tumor associated Antigen, TRAP1, on cancer therapy. *Vaccine: X* (2019) 1:100017. doi: 10.1016/j.jvax.2019.100017
154. Solleder M, Guillaume P, Racle J, Michaux J, Pak H-S, Müller M, et al. Mass Spectrometry Based Immunopeptidomics Leads to Robust Predictions of Phosphorylated HLA Class I Ligands. *Mol Cell Proteomics* (2020) 19:390–404. doi: 10.1074/mcp.TIR119.001641
155. Buitinga M, Callebaut A, Marques Câmara Sodré F, Crèvecoeur I, Blahnik-Fagan G, Yang M-L, et al. Inflammation-Induced Citrullinated Glucose-Regulated Protein 78 Elicits Immune Responses in Human Type 1 Diabetes. *Diabetes* (2018) 67:2337–48. doi: 10.2337/db18-0295
156. McGinty JW, Chow I-T, Greenbaum C, Odegard J, Kwok WW, James EA, et al. Recognition of posttranslationally modified GAD65 epitopes in subjects with type 1 diabetes. *Diabetes* (2014) 63:3033–40. doi: 10.2337/db13-1952
157. Yang L, Tan D, Piao H. Myelin Basic Protein Citrullination in Multiple Sclerosis: A Potential Therapeutic Target for the Pathology. *Neurochem Res* (2016) 41:1845–56. doi: 10.1007/s11064-016-1920-2
158. James EA, Rieck M, Pieper J, Gebe JA, Yue BB, Tatum M, et al. Citrulline-specific Th1 cells are increased in rheumatoid arthritis and their frequency is influenced by disease duration and therapy. *Arthritis Rheumatol (Hoboken NJ)* (2014) 66:1712–22. doi: 10.1002/art.38637
159. Feitsma AL, van der Voort EIH, Franken KLMC, el Bannoudi H, Elferink BG, Drijfhout JW, et al. Identification of citrullinated vimentin peptides as T cell epitopes in HLA-DR4-positive patients with rheumatoid arthritis. *Arthritis Rheumatol* (2010) 62:117–25. doi: 10.1002/art.25059
160. Jiang Z, Cui Y, Wang L, Zhao Y, Yan S, Chang X, et al. Investigating citrullinated proteins in tumour cell lines. *World J Surg Oncol* (2013) 11:260. doi: 10.1186/1477-7819-11-260
161. Brentville VA, Vankemmelbeke M, Metheringham RL, Durrant LG. Post-translational modifications such as citrullination are excellent targets for cancer therapy. *Semin Immunol* (2020) 47:101393. doi: 10.1016/j.smim.2020.101393
162. Chang X, Han J, Pang L, Zhao Y, Yang Y, Shen Z, et al. Increased PADI4 expression in blood and tissues of patients with malignant tumors. *BMC Cancer* (2009) 9:40. doi: 10.1186/1471-2407-9-40
163. Wang L, Song G, Zhang X, Feng T, Pan J, Chen W, et al. PADI2-Mediated Citrullination Promotes Prostate Cancer Progression. *Cancer Res* (2017) 77:5755–68. doi: 10.1158/0008-5472.CAN-17-0150
164. Scally SW, Petersen J, Law SC, Dudek NL, Nel HJ, Loh KL, et al. A molecular basis for the association of the HLA-DRB1 locus, citrullination, and rheumatoid arthritis. *J Exp Med* (2013) 210:2569–82. doi: 10.1084/jem.20131241
165. Hill JA, Southwood S, Sette A, Jevnikar AM, Bell DA, Cairns E, et al. Cutting edge: the conversion of arginine to citrulline allows for a high-affinity peptide interaction with the rheumatoid arthritis-associated HLA-DRB1*0401 MHC class II molecule. *J Immunol* (2003) 171:538–41. doi: 10.4049/jimmunol.171.2.538
166. Nguyen H, James EA. Immune recognition of citrullinated epitopes. *Immunology* (2016) 149:131–8. doi: 10.1111/imm.12640
167. Snir O, Rieck M, Gebe JA, Yue BB, Rawlings CA, Nepom G, et al. Identification and functional characterization of T cells reactive to citrullinated-vimentin in HLA-DRB1*0401 humanized mice and RA patients. *Arthritis Rheum* (2011) 63:2873–83. doi: 10.1002/art.30445
168. Brentville VA, Metheringham RL, Gunn B, Symonds P, Daniels I, Gijon M, et al. Citrullinated Vimentin Presented on MHC-II in Tumor Cells Is a Target for CD4+ T-Cell-Mediated Antitumor Immunity. *Cancer Res* (2016) 76:548–60. doi: 10.1158/0008-5472.CAN-15-1085
169. Golks A, Guerini D. The O-linked N-acetylglucosamine modification in cellular signalling and the immune system. *EMBO Rep* (2008) 9:748–53. doi: 10.1038/embor.2008.129
170. Hart GW, Housley MP, Slawson C. Cycling of O-linked β -N-acetylglucosamine on nucleocytoplasmic proteins. *Nature* (2007) 446:1017–22. doi: 10.1038/nature05815
171. de Queiroz RM, Carvalho E, Dias WB. O-GlcNAcylation: The Sweet Side of the Cancer. *Front Oncol* (2014) 4:132. doi: 10.3389/fonc.2014.00132
172. Haurum JS, Arsequell G, Lellouch AC, Wong SY, Dwek RA, McMichael AJ, et al. Recognition of carbohydrate by major histocompatibility complex class I-restricted, glycopeptide-specific cytotoxic T lymphocytes. *J Exp Med* (1994) 180:739–44. doi: 10.1084/jem.180.2.739
173. Haurum JS, Tan L, Arsequell G, Frodsham P, Lellouch AC, Moss PAH, et al. Peptide anchor residue glycosylation: effect on class I major histocompatibility complex binding and cytotoxic T lymphocyte recognition. *Eur J Immunol* (1995) 25:3270–6. doi: 10.1002/eji.1830251211
174. Glithero A, Tormo J, Haurum JS, Arsequell G, Valencia G, Edwards J, et al. Crystal Structures of Two H-2Db/Glycopeptide Complexes Suggest a Molecular Basis for CTL Cross-Reactivity. *Immunity* (1999) 10:63–74. doi: 10.1016/S1074-7613(00)80007-2
175. Malaker SA, Penny SA, Steadman LG, Myers PT, Loke JC, Raghavan M, et al. Identification of Glycopeptides as Posttranslationally Modified Neoantigens in Leukemia. *Cancer Immunol Res* (2017) 5:376–84. doi: 10.1158/2326-6066.CIR-16-0280
176. Swatek KN, Komander D. Ubiquitin modifications. *Cell Res* (2016) 26:399–422. doi: 10.1038/cr.2016.39
177. Milhollen M, Sappal D, Duffy J, Hoar K, Huck J, Sha P, et al. 577 Characterization of the cellular mechanism of action of the first in class investigational inhibitor of the Ubiquitin Activating Enzyme, MLN7243. *Eur J Cancer* (2014) 50:186. doi: 10.1016/S0959-8049(14)70703-8
178. Hundal J, Carreno BM, Petti AA, Linette GP, Griffith OL, Mardis ER, et al. pVAC-Seq: A genome-guided in silico approach to identifying tumor neoantigens. *Genome Med* (2016) 8:1–11. doi: 10.1186/s13073-016-0264-5
179. Kim S, Kim HS, Kim E, Lee MG, Shin E-C, Paik S, et al. Neopepsee: accurate genome-level prediction of neoantigens by harnessing sequence and amino acid immunogenicity information. *Ann Oncol: Off J Eur Soc Med Oncol* (2018) 29:1030–6. doi: 10.1093/annonc/mdy022
180. Bjerregaard A-M, Nielsen M, Hadrup SR, Szallasi Z, Eklund AC. MuPeXI: prediction of neo-epitopes from tumor sequencing data. *Cancer Immunol Immunother: CII* (2017) 66:1123–30. doi: 10.1007/s00262-017-2001-3
181. Rubinsteyn A, Kodysh J, Hodes I, Mondet S, Aksoy BA, Finnigan JP, et al. Computational Pipeline for the PGV-001 Neoantigen Vaccine Trial. *Front Immunol* (2018) 8:1807. doi: 10.3389/fimmu.2017.01807
182. Rech AJ, Balli D, Mantero A, Ishwaran H, Nathanson KL, Stanger BZ, et al. Tumor Immunity and Survival as a Function of Alternative Neopeptides in Human Cancer. *Cancer Immunol Res* (2018) 6:276–87. doi: 10.1158/2326-6066.CIR-17-0559
183. Zhang J, Mardis ER, Maher CA. INTEGRATE-neo: a pipeline for personalized gene fusion neoantigen discovery. *Bioinf (Oxford England)* (2017) 33:555–7. doi: 10.1093/bioinformatics/btw674
184. Smith CC, Beckermann KE, Bortone DS, De Cubas AA, Bixby LM, Lee SJ, et al. Endogenous retroviral signatures predict immunotherapy response in clear cell renal cell carcinoma. *J Clin Invest* (2018) 128:4804–20. doi: 10.1172/JCI121476
185. Smit A, Hubley R, Green P. *RepeatMasker Open-4.0*. (2013–2015). Available at: <http://www.repeatmasker.org>.
186. Reynisson B, Alvarez B, Paul S, Peters B, Nielsen M. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res* (2020) 48:W449–54. doi: 10.1093/nar/gkaa379

187. Nijveen H, Kester MGD, Hassan C, Viars A, de Ru AH, de Jager M, et al. HSPVdb—the Human Short Peptide Variation Database for improved mass spectrometry-based detection of polymorphic HLA-ligands. *Immunogenetics* (2011) 63:143–53. doi: 10.1007/s00251-010-0497-1
188. Jawa V, Cousens LP, Awwad M, Wakshull E, Kropshofer H, De Groot AS, et al. T-cell dependent immunogenicity of protein therapeutics: Preclinical assessment and mitigation. *Clin Immunol* (2013) 149:534–55. doi: 10.1016/j.clim.2013.09.006
189. Smith CC, Chai S, Washington AR, Lee SJ, Landoni E, Field K, et al. Machine-Learning Prediction of Tumor Antigen Immunogenicity in the Selection of Therapeutic Epitopes. *Cancer Immunol Res* (2019) 7:1591–604. doi: 10.1158/2326-6066.CIR-19-0155
190. Zhao Q, Laverdure J-P, Lanoix J, Durette C, Côté C, Bonneil E, et al. Proteogenomics Uncovers a Vast Repertoire of Shared Tumor-Specific Antigens in Ovarian Cancer. *Cancer Immunol Res* (2020) 8:544–55. doi: 10.1158/2326-6066.CIR-19-0541
191. Andreatta M, Nielsen M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinformatics* (2016) 32:511–7. doi: 10.1093/bioinformatics/btv639
192. Andersen MH, Søndergaard I, Zeuthen J, Elliott T, Haurum JS. An assay for peptide binding to HLA-Cw*0102. *Tissue Antigens* (1999) 54:185–90. doi: 10.1034/j.1399-0039.1999.540210.x
193. Ross P, Holmes JC, Gojanovich GS, Hess PR. A cell-based MHC stabilization assay for the detection of peptide binding to the canine classical class I molecule, DLA-88. *Vet Immunol Immunopathol* (2012) 150:206–12. doi: 10.1016/j.vetimm.2012.08.012
194. Kranz LM, Diken M, Haas H, Kreiter S, Loquai C, Reuter KC, et al. Systemic RNA delivery to dendritic cells exploits antiviral defence for cancer immunotherapy. *Nature* (2016) 534:396–401. doi: 10.1038/nature18300
195. Ott PA, Hu Z, Keskin DB, Shukla SA, Sun J, Bozym DJ, et al. An immunogenic personal neoantigen vaccine for patients with melanoma. *Nature* (2017) 547:217–21. doi: 10.1038/nature22991
196. Cintolo JA, Datta J, Mathew SJ, Czerniecki BJ. Dendritic cell-based vaccines: barriers and opportunities. *Future Oncol* (2012) 8:1273–99. doi: 10.2217/fon.12.125
197. Koski GK, Koldovsky U, Xu S, Mick R, Sharma A, Fitzpatrick E. A novel dendritic cell-based immunization approach for the induction of durable Th1-polarized anti-HER-2/neu responses in women with early breast cancer. *J Immunother* (2012) 35:54–65. doi: 10.1097/CJI.0b013e318235f512

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer AA declared a past co-authorship with two of the authors CP and PT to the handling editor.

Copyright © 2020 Minati, Perreault and Thibault. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.