

Review

Uncovering the Tumor Antigen Landscape: What to Know about the Discovery Process

Sara Feola ^{1,2,3,4} , Jacopo Chiaro ^{1,2,3,4} , Beatriz Martins ^{1,2,3,4} and Vincenzo Cerullo ^{1,2,3,4,*}

¹ Drug Research Program (DRP) ImmunoViroTherapy Lab (IVT), Division of Pharmaceutical Biosciences, Faculty of Pharmacy, Viikinkaari 5E, University of Helsinki, 00790 Helsinki, Finland; sara.feola@helsinki.fi (S.F.); jacopo.chiaro@helsinki.fi (J.C.); beatriz.martins@helsinki.fi (B.M.)

² Helsinki Institute of Life Science (HiLIFE), Fabianinkatu 33, University of Helsinki, 00710 Helsinki, Finland

³ Translational Immunology Program (TRIMM), Faculty of Medicine, Haartmaninkatu 8, University of Helsinki, 00290 Helsinki, Finland

⁴ Digital Precision Cancer Medicine Flagship (iCAN), University of Helsinki, FI-00014 Helsinki, Finland

* Correspondence: vincenzo.cerullo@helsinki.fi; Tel.: +358-029-415-9328 or +358-050-318-5754

Received: 29 April 2020; Accepted: 20 June 2020; Published: 23 June 2020



Abstract: According to the latest available data, cancer is the second leading cause of death, highlighting the need for novel cancer therapeutic approaches. In this context, immunotherapy is emerging as a reliable first-line treatment for many cancers, particularly metastatic melanoma. Indeed, cancer immunotherapy has attracted great interest following the recent clinical approval of antibodies targeting immune checkpoint molecules, such as PD-1, PD-L1, and CTLA-4, that release the brakes of the immune system, thus reviving a field otherwise poorly explored. Cancer immunotherapy mainly relies on the generation and stimulation of cytotoxic CD8 T lymphocytes (CTLs) within the tumor microenvironment (TME), priming T cells and establishing efficient and durable anti-tumor immunity. Therefore, there is a clear need to define and identify immunogenic T cell epitopes to use in therapeutic cancer vaccines. Naturally presented antigens in the human leucocyte antigen-1 (HLA-I) complex on the tumor surface are the main protagonists in evocating a specific anti-tumor CD8+ T cell response. However, the methodologies for their identification have been a major bottleneck for their reliable characterization. Consequently, the field of antigen discovery has yet to improve. The current review is intended to define what are today known as tumor antigens, with a main focus on CTL antigenic peptides. We also review the techniques developed and employed to date for antigen discovery, exploring both the direct elution of HLA-I peptides and the *in silico* prediction of epitopes. Finally, the last part of the review analyses the future challenges and direction of the antigen discovery field.

Keywords: tumor antigens; immunopeptidome; cancer immunotherapy; epitope prediction

1. Introduction

The recent clinical success of antibodies targeting immune checkpoint molecules, such as programmed death receptor-1 (PD-1), its ligand PD-L1, and cytotoxic T cell-associated antigen 4 (CTLA-4), have led to a new and strong interest in the field of cancer immunotherapy [1,2]. Immune checkpoint inhibitors (ICIs) release the brakes of the immune system, reviving and boosting the effector function of specific anti-tumor T cells [3]. In 2018, James Patrick Allison and Tasuku Honjo received the Nobel prize for medicine “for their discovery of cancer therapy by inhibition of negative immune regulation” [4,5]. The overall response to ICIs is reportedly unsatisfactory for many types of cancer [6], highlighting the need to combine ICIs with cancer immunotherapeutic approaches, such as therapeutic cancer vaccines, with the artificial generation, stimulation, and tumor microenvironment

(TME) infiltration of cancer-specific CD8⁺ T cells called cytotoxic T lymphocytes (CTLs) [2,7–9]. CTLs patrol the whole organism, checking the major histocompatibility complex (MHC), called the human leucocytes antigen (HLA) in humans, that possesses an antigenic peptide that is typically 8–11 aminoacidic residues in length and expressed on the cellular surface. If an aberrant peptide is eventually spotted in the HLA-I complex, the CTLs kill the cell.

In this context, to create effective tumor protection and rejection strategies, the reliable identification of tumor antigens in the HLA-I complex plays a pivotal role. To date, the methodologies for the correct identification of antigens in the HLA-I complex rely on the direct isolation of peptides from the HLA-I complex and/or the *in silico* prediction of relevant antigens [10,11].

The direct identification of peptides from the HLA complex is accomplished using different techniques, of which immunoaffinity purification (HLA immunoprecipitation) and the extraction of HLA peptides are the most well-established [12]. Isolated peptides are identified by tandem mass-spectrometry (MS/MS). The entire process is highly laborious and time-consuming [13,14]. It has been reported that immunoaffinity purification produces a low yield (i.e., 1–3%) [15], even though additional experiments in different settings are needed to better determine how much of the sample is lost [16]. The peptide isolation procedure is universally recognized as the bottleneck of the entire procedure and therefore requires improvement [16].

In the antigen discovery process, *in silico* tools are used to predict immunologically relevant antigens. The identification of target candidates relies on bioinformatic approaches that allow a structural analysis of genomic and proteomic data. For instance, next generation sequencing (NGS) approaches can benefit from the use of epitope prediction tools to combine two kinds of information: the differential gene expression in cancer compared to matched healthy tissue and the probability of those candidates to be presented on the cell surface onto the HLA molecule [17,18]. To achieve this, the actual bioinformatic tools contain T cell epitope algorithms that are able to predict putative T cell epitopes based on the well-characterized rules to which HLA-I presented peptides adhere [19,20]. Indeed, several predictors of HLA binders have been developed [21]; these will be defined in Section 4. HLA binding is only one part of the story; today, the entire processing machinery can be taken into account (e.g., proteasomal cleavage, transporter-associated antigen processing (TAP) transport) by the bioinformatic algorithms to predict relevant T cell epitopes [21].

NGS and *in silico* tools are effective methods in antigen discovery; however, improvements are needed. These methods lack rich and experimentally validated datasets, thus decreasing the accuracy of the predictive algorithms. The bioinformatic predictions give hints, but the selection of relevant epitopes from among the candidates always requires experimental validation [18,21].

This paper presents the information necessary to understand the antigen discovery field. It describes the concept of antigens, focusing mainly on CTL antigenic peptides. It then provides a concise analysis of the methods endorsed for the antigen discovery process, including *in vitro* and *in silico* approaches. It also highlights the gaps and challenges in the field.

2. Antigens

The word “antigen” refers to the molecular structure seen by the antibodies or to any molecule or linear molecular fragment derived from the processing of the native antigen that can be recognized by T cell receptors (TCRs) [22]. This paper mainly examines CD8⁺ T cell-restricted antigens. The existence and role of tumor antigens have been discussed since 1940 and were reported on before the discovery of T cells, which occurred in the 1960s [23,24]. In 1943, Gross et al. performed a pivotal experiment that showed for the first time the role of the immune system in tumor rejection [25]. Briefly, mice were treated with methylcholanthrene and subsequently developed tumors. The tumors were then resected, and the tumor cells were implanted back into the mice. The mice then rejected the second tumor. This experiment demonstrated that acquired immunity was induced and directed against the tumor and that the rejection did not depend upon genetic differences between the inoculated mice and the mice that produced the tumor cells [25].

Later, Boon and colleagues achieved similar results using mutagen-treated murine cell lines that failed to form tumors in syngeneic mice; this research confirmed that tumors express antigens recognized by CTLs and confirmed the role of the immune system in rejecting malignant cells [26,27]. However, the molecular nature of the antigens expressed by tumors and recognized by CTLs was only discovered in 1989 by Lurquin et al. [28]. The researchers identified a single peptide recognized by CTLs that differed from the self-protein by single point mutation. This observation clearly showed that upon mutation, the tumors expressed altered proteins, thus labeling the cells for CTL recognition [28,29].

Today, based on the expression of the parental gene, tumor antigens can be classified into the general categories of tumor-associated antigens (TAAs) and tumor-specific antigens (TSAs) [30] (Figure 1).

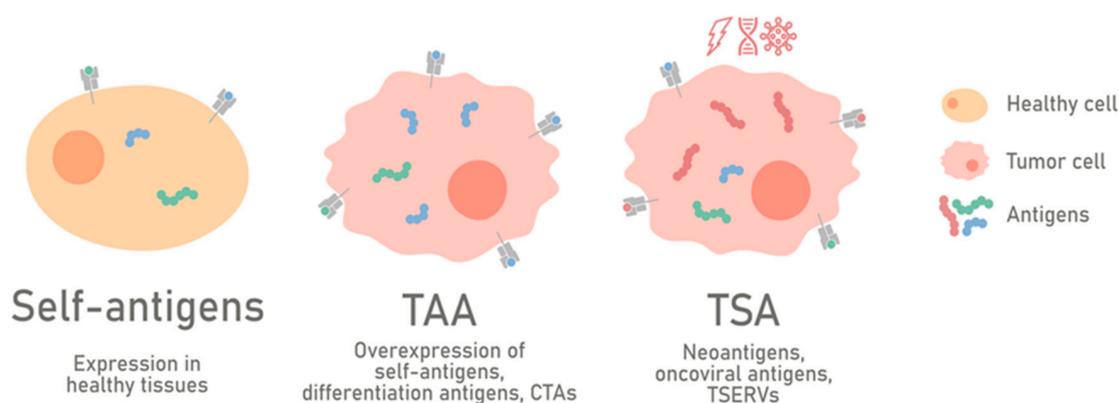


Figure 1. Classification of tumor antigens. Tumor antigens can generally be categorized into tumor-associated antigens (TAAs) and tumor-specific antigens (TSAs) based on the expression pattern of the parental gene. TAAs are self-proteins expressed in cancer cells; upon malignant transformation, the following consequences can be observed: the overexpression of normal proteins (gene overexpressed), the expression of proteins with tissue-specific gene patterns (differentiation antigens), or the expression of proteins derived from gene expression restricted to the testes (cancer germline/cancer testis antigens). TSAs are proteins expressed by tumor cells and can arise from mutations (neoantigens), from viruses involved in the oncogenic transformation (oncoviral antigens), or from the expression of tumor-specific endogenous retroviruses (*TSERVs*).

2.1. Tumor-Associated Antigens

TAAs are a large family of antigens that includes antigens derived from genes overexpressed in tumors, differentiation antigens, and cancer germline/cancer testis antigens [30,31]. Antigens derived from genes overexpressed in tumors comprise a class of normal self-proteins which are minimally expressed by healthy tissues but constitutively overexpressed in cancer cells as a result of their malignant profile. The proteins which are usually overexpressed (e.g., EGFR, hTERT, p53, carbonic anhydrase IX) are mainly involved in the survival of the cancer cells and therefore are not susceptible to downregulation mechanisms, making them an attractive target for cancer therapeutic approaches [31]. For instance, in 1996, Gaugler et al. discovered an overexpressed antigen in renal cell carcinoma (RCC) called renal antigen 1 (RAGE-1) [32]. RAGE-1 is the first example of an antigen recognized by autologous CTLs in RCC, and its expression is restricted to the retina and to different histological tumor types. As the retina does not express HLA-I [33], RAGE-1 becomes a possible candidate to target for cancer immunotherapy [34,35]. Another interesting example is HER2/NEU, which is overexpressed in epithelial tumors such as ovarian and breast tumors [36,37]. The use of trastuzumab, a monoclonal antibody targeting the extracellular domain of HER2, has revolutionized the treatment of breast cancer HER2+ [38]. In addition, HER2 is a suitable candidate for the peptide vaccine approach; indeed, the nonapeptide E75, derived from the HER2 sequence, has been described as being able to elicit four different ovarian CTL cell line responses [39]. E75 as a monotherapy (Nelipepimut-S) or in combination with granulocyte macrophage colony-stimulating factor (GM-CSF) (NeuVax) has been tested in clinical

trials [40–42]. The genes belonging to the apoptosis pathway are also upregulated in tumor cells compared to healthy tissues, representing a source of T cell epitopes. For instance, a peptide derived from survivin, an apoptosis inhibitor protein, was used to generate CTLs in vitro from healthy donors; matched cell lines and primary malignant cells from patients were then lysed by those T cells [43]. In addition, a p53 derived wild type peptide (L9V) was used to generate in vitro CTLs that were able to kill squamous carcinoma cell lines in an L9V-HLA-dependent fashion [44].

As TAAs have a higher expression level in tumors compared to normal tissue and are shared among several tumors, a safe use for them has been proposed for cancer therapeutic approaches. For instance, novel chimeric antigen receptor (CAR) T cells have largely been adopted to target this class of antigens [31]. CAR T cells consist of an extracellular domain made of the variable region of the antibody's heavy and light chains linked to the intracellular signaling domain (CD3-zeta, CD28, 41BB). This characteristic makes the CAR T cells able to kill the target in an HLA-independent manner [45]. The adoption of anti-CD19 CARs has showed good clinical outcomes in the treatment of B cell lymphomas and leukemias [46–48].

However, the main drawback in using TAAs in cancer immunotherapy is the feasibility of the expression analysis of these proteins in every single tissue under every physiological condition; this hinders a comprehensive safety profile of the TAAs [31]. Indeed, potential hazards (i.e., “on-target, off-tumor” toxicity, onset of autoimmune disease) associated with the clinical treatment based on these molecules have been reported. For instance, CAR T cells targeting carbonic anhydrase IX (CAIX) in RCC patients induced liver toxicity, requiring cessation of the treatment. Biopsies revealed CAIX expression in the bile duct epithelium with T cell infiltration, including CAR T cells; this is a typical example of “on-target, off-tumor” toxicity [49]. Moreover, the central and peripheral tolerance mechanisms eliminate the T and B cells' ability to recognize self-antigens. A TAA peptide-based vaccine must break the tolerance, thus stimulating the low-affinity and rare T cells still circulating. This could interfere with the development of a proper cancer therapeutic vaccine [30,50]. The identification and use of an effective vaccine adjuvant could overcome the problem, delivering benefits to cancer patients [51].

The differentiation antigens represent normal proteins which are expressed as a consequence of a specific function of the target tissue. These were first reported in melanoma in which proteins (e.g., tyrosinase, Melan-A/MART-1, gp100/Pmel17) involved in melanin production or melanosome generation were often observed to be targets for CTLs from melanoma patients and healthy donors [35,52,53]. In the 1993 work of Brichard et al., lymphocytes derived from two melanoma patients were stimulated with irradiated cells from the autologous melanoma, and CTLs that were able to lyse them were obtained. The authors then demonstrated that the lysis was antigen HLA-0201-specific, and upon using the cloning approach, they identified the gene encoding the antigen as tyrosinase [53]. Rosenberg and colleagues administered tumor-infiltrating lymphocytes (TILs) to treat metastatic melanoma patients and had good results. The authors used the TILs to clone and determine the antigens involved in the anti-tumor response. In 1994, Rosenberg and colleagues established a CTL cell line called TIL1200 from a metastatic melanoma patient and were able to lyse the autologous melanoma and other melanoma HLA-0201+ cell lines. The authors showed that TIL1200 recognized a self-antigen not mutated in the HLA-0201 context and that the antigen sequence belonged to a membrane glycoprotein known as gp100/Pmel17 [54]. In the same year, Rosenberg and colleagues identified a shared and commonly expressed HLA-0201-restricted melanoma antigen that was recognized by T cells 1 (MART-1), applying a similar approach to that used to isolate gp100/Pmel17 [55]. In the following years, other melanoma differentiation antigens were identified, such as TRP1/gp75 and TRP2 (tyrosinase-related proteins), in the context of the HLA-A31 molecule [56,57]. Later, the TRP2 case was reviewed to identify an HLA-0201-restricted peptide, as HLA-0201 occurs frequently in the population. Based on the peptide-binding motif for HLA-A0201 and experimental validation, TRP2 peptides 180–188 were reported, and the sequence was found to be identical to the one recognized by H2Kb-restricted B16 murine melanoma CTLs, paving the way for a murine tumor immunotherapy model [58]. Furthermore, differentiation antigens such as prostatic acid phosphatase (PAP) and prostate-specific antigen (PSA)

are used for prostate cancer immunotherapy. Naturally expressed HLA-02 restricted peptides from PAP were identified through sequence analysis and in vitro validation [59]. Moreover, two PSA HLA-02-restricted peptides which are capable of eliciting CTL responses have been reported by Correale et al. [60]. In addition, immunization experiments using HLA-0201 transgenic mice were employed to identify a peptide from carcinoembryonic antigens (CEA) [61]. This is a glycoprotein overexpressed in colon rectal cancer and other selected epithelial cancers, while in normal conditions, its expression is restricted to fetal development and in adults' tissue in the epithelial cells of the gastrointestinal tract [62,63].

The main disadvantage in exploiting differentiation antigens as vaccines for cancer therapy is the onset of autoimmune toxicity. In cases of targeting melanoma-melanocyte antigens, reactions such as severe skin rashes and vitiligo lesions have been reported [64,65]. CAR T cells targeting CEA that were used to treat three patients with metastatic colorectal cancer showed a promising outcome in at least one patient; however, severe transient colitis was induced, most probably due to the CEA presence in the colonic mucosa [62]. These observations highlight the importance of targeting tumor-specific antigens and/or antigens with limited expression in normal tissues (e.g., cancer germline/cancer testis antigens).

Cancer germline/cancer testis antigens (CTAs) are a large family (the CT database consists of 204 genes [66]) of tumor-associated antigens expressed in human tumors of different histological origins but not in normal tissue, except for testis and placenta tissue [67]. In 1991, Van der Bruggen and co-workers identified the first human gene that encoded for a cancer testis antigen [68]. The authors isolated the CTLs that were capable of lysing the autologous cell line derived from the melanoma patient MZ2. Using a clonal sub-line and a DNA cloning approach, they identified the gene which was able to sensitize the sub-line to the CTL lysis. This gene was named melanoma antigen family A, 1 (MAGEA1), and it is expressed in many human tumors of different histological types. No expression was detected in normal tissues, with the exception of testis and placenta tissue [68,69]. In the same work, another two MAGE members were reported, MAGE A2 and MAGE A3 [68]. Subsequently, the nonapeptides derived from HLA-A1-restricted MAGE A1 and MAGE A3 were described [70,71].

After this, many antigens were identified in diverse tumors with restricted expression in testis tissue; Old and Chen called these cancer testis (CT) antigens [72]. To date, the MAGE family is subdivided into type I and type II. Type I comprises three sub-families, MAGE-A, -B, and -C, and contains the relevant CTAs. Type II includes the MAGE-D, -E, -F, -G, -H, and -L sub-families and Necdin that are expressed in different adult tissues [73–75]. Several members have been identified in the MAGE group [76], and MAGE A3 is one of the most frequently expressed TAAs in many tumors, including melanoma, non-small cell lung carcinoma (NSCLC), and head and neck tumors [77]. In melanoma, the preferentially expressed antigen of melanoma (PRAME) is another example of a CTA; PRAME is defined as a testis-selective rather than a testis-restricted CTA as its expression has been observed in endometrial, ovarian, and adrenal gland tissues in addition to testis tissue [78]. Building on the work by Ikeda and co-workers [79], Kessler et al. subsequently identified four HLA-0201 peptides that were restricted in PRAME [80]. In 2011, Quintarelli and colleagues were able to generate PRAME-specific CTLs from both healthy donors and leukemia patients [81]. In 1997, the serological analysis of recombinant tumor cDNA expression libraries (SEREX) technique was used by Chen et al. to identify novel tumor antigens in esophageal squamous cell carcinoma patients; the screening revealed sequences belonging to eight genes, among them New York esophageal squamous cell carcinoma 1 (NY-ESO 1), that were expressed in normal tissue, such as testis and ovarian tissue, and in several tumors, such as melanoma, breast cancer, bladder cancer, prostate cancer, and hepatocellular carcinoma [82]. Since then, NY-ESO1 has been employed in several immunotherapy-based treatments [83]. For instance, recombinant NY-ESO1 protein was used in combination with the adjuvant ISCOMATRIX in 46 patients with resected NY-ESO1+ tumors. Overlapping peptides were then employed to verify the T cells' response to NY-ESO1 upon vaccination, showing CD4+ and CD8+ T cells' specific response to both known and uncharacterized peptides derived from NY-ESO1 [84]. In 1997, Türeci et al. used the

SEREX technique to investigate human melanoma and described a novel CTA gene called synovial sarcoma X chromosome breakpoint (SSX2) [85], to which a T cell response has been reported in patients with tumors of diverse histological origins [86]. An analysis of the sub-clones derived from the MZ2 melanoma patient allowed the identification of the B melanoma antigen (BAGE) gene and the HLA-Cw1601-restricted peptide [87]; BAGE expression was found only in testis and tumor (e.g., bladder cancer) tissues, and thus was found to be a member of the CTA family. Using a similar experimental procedure, other CTAs were identified—the gene G antigen 1 (GAGE) and the HLA-Cw0601-restricted peptide [88].

Being immunologically privileged sites, the testes and placenta do not express HLA molecules [89]; therefore, CTAs are promising targets for cancer immunotherapy. Indeed, MAGE A3 as a recombinant protein has been used in the largest ever phase III lung cancer clinical trial, the MAGE A3 as Adjuvant Non-Small Cell Lung Cancer Immunotherapy (MAGRIT) trial [90]. In phase II, the vaccination protocol was well tolerated and the results promising; however, MAGRIT failed to show any improvement in disease-free survival (DFS) when compared with a placebo [91]. In addition, TCR gene modified T cells to target the MAGEA3/A12 HLA-A0201 restricted peptide were used to treat nine patients with tumors expressing MAGE A3/A12 in a phase I/II clinical trial. The encouraging cancer regression that was observed was dampened by severe neurotoxicity that resulted in the deaths of two patients. Subsequent analysis revealed the expression of MAGE A12 in the brain, explaining the inflammation and neuronal degeneration that was observed [92]. The use of CTAs, such as MAGE family members, is promising; however, caution is required in further applications.

2.2. Tumor-Specific Antigens

TsAs are a category of antigens restricted to tumors and are not found in healthy cells; this is the result of malignant mutations or the expression of viral elements. Neoantigens, oncoviral antigens, and endogenous retroviral elements belong to this category [30,35,93].

Neoantigens are a subset of TsAs produced as a direct consequence of genetic alteration caused by tumor DNA mutations (e.g., non-synonymous single point mutations, frameshifts, insertions/deletions) and are patient-specific [94,95]. In 1995, Coulin et al. reported the first example of a neoantigen [96]. The authors identified the source gene of the antigen that was recognized by an autologous CTL and called it melanoma ubiquitous mutated (MUM-1). As the gene was expressed ubiquitously, the authors wondered whether a mutation had occurred in the gene to induce an anti-tumor T cell response. Actually, MUM-1 in the melanoma cells carried a single point mutation that resulted in an amino acid change in the nonapeptide HLA-B44 restricted, making it suitable for interaction with the T cells [96]. Several other neoantigens have since been identified. For instance, Rosenberg and colleagues discovered a single point mutation in the β catenin sequence in one melanoma patient that resulted in a change from a serine to a phenylalanine residue at position 37; the derived peptide (SYLDGSIHF) showed a high affinity binding for HLA-A24, the patient's HLA allele [97]. Moreover, a single point mutation in the gene encoding CDK4 resulted in the generation of an HLA-A0201-restricted peptide that was recognized by autologous CTLs in the melanoma patient, and the mutation altered the cell cycle regulation [98]. A mutation in the CASP-8 gene that reduced the function of the protein and generated an HLA-B3503-restricted peptide that was recognized by autologous CTLs was reported in a patient with squamous cell carcinoma of the oral cavity [99].

The neoantigens in cancer immunotherapy have the undiscussed advantage of being non self-antigens and hence the T cells are not affected by central tolerance, making them highly immunogenic antigens [94]. However, their main limitation is their great variability within and between tumors. Great variability within tumors can induce negative selection, allowing the survival of the cancer cells that no longer express the neoantigens. Great variability between tumors requires characterization and the development of a specific vaccine for each patient. In addition, the mutational burden plays a pivotal role in the generation of neoepitopes. Tumors with a high mutation frequency

are most likely to generate neoantigens. Tumors with a low mutation burden may be difficult to investigate in terms of identifying neoantigens and subsequent vaccine production [31,35].

Oncoviral antigens consist of proteins derived from the viruses driving the oncogenic transformation; these proteins are the source of peptides present on the cellular surface in the HLA context and recognized by T cells. As oncogenic viruses are shared by the same kinds of tumors, this class of antigen is not patient-specific [30,35]. Prophylactic vaccines have been produced and mainly rely on eliciting neutralizing antibodies to prevent the virus from entering the cells. The treatment of established tumors involves targeting T cell epitopes [30]. For instance, human papillomavirus (HPV) is associated with benign papilloma or warts and cancer of the cervix, anus, penis, and head and neck [100]. Ramos et al. isolated PBMCs from patients with HPV+ cancer and stimulated them in vitro with a mix of HLA-I-restricted and HLA-II restricted peptides, covering the E6/E7 proteins of HPV16; the authors were able to demonstrate that the patients had E6- and E7-specific T cells [101]. These have been called HPV-ST, and an on-going phase I clinical trial is evaluating the effect of HPV-ST cells in treating HPV+ tumors (NCT02379520). Another example is the development of therapeutic vaccines for cancers related to the Epstein–Barr virus (EBV) which is involved in the onset of several disorders, such as B-cell lymphoproliferative disorders and nasopharyngeal carcinoma [102,103]. In a phase I clinical trial, 16 patients with established HBV+ tumors were treated with modified vaccinia Ankara (MVA) encoding the full length of LMP2 and the C-terminal of EBNA1 proteins from EBV. A specific T cell response to LMP2 and/or EBNA1 was detected, showing the feasibility of boosting an EBV-specific immune response [104]. However, the extent of the clinical benefits is still being investigated in a phase II clinical trial (NCT01094405). Oncoviral antigens lack expression in healthy cells, making them highly tumor-specific. They are also common to patients. However, 15% of cancers have a viral etiopathology, limiting their clinical application [105,106].

Endogenous retroviral elements (ERVs) or human endogenous retroviral elements (HERVs) are fragments of genomic DNA derived from the integration of retrotranscribed retroviral RNAs that infected the germ line cells of humans' ancestors. Over time, ERVs have been vertically transmitted and, to date, they represent 8% of the human genome. ERVs have accumulated mutations over time, losing the capability of producing competent replicative viral particles [107]. Moreover, epigenetic mechanisms (e.g., methylation) suppress most of their expression in healthy cells. For instance, in the thymus, ERVs are partially epigenetically silenced, and ERVs which are reactive to T cells do not go through complete negative selection [108]. In cancer, ERVs are induced upon malignant transformation and/or epigenetic therapy, becoming targets for cancer therapeutic approaches [109]. In 2015, Rooney et al. investigated the cytolytic activity and expression pattern associated with 66 ERVs in tumors compared to healthy tissues. Surprisingly, they found three ERVs (ERVH5, ERVH48-1, and ERVE4) with minimal or undetectable expression in normal tissue and overexpression in tumors; these have been termed tumor-specific endogenous retroviruses (TSERVs) [93]. Interestingly, in regard to ERVE4, Rooney's data had already been experimentally validated. For instance, following hematopoietic stem cell transplantation (HSCT), RCC patients experienced disease regression. Child and colleagues found RCC-reactive CD8+T cells derived from the donor. Using a cDNA cloning approach, the authors identified the antigen in HERV-E gene products and described an HLA-A11-restricted 10-mer peptide (ATFLGSLTWK) as the target recognized by the tumor-reactive CTLs [110]. The same authors also reported the identification of three HLA-A0201-restricted peptides derived from HERV4_{env} that were able to elicit RCC-reactive CTL responses [111]. ERVH-5 has been reported in bladder, colorectal, head and neck, lung squamous, ovarian, stomach, and uterine cancers. ERVH48-1 is prominently expressed in bladder cancer and prostate cancer [109]. Schiavetti et al. described CTLs which were reactive against the peptide derived from HERV-K-MEL in two melanoma patients. The authors determined that the peptide sequence (MLAVISCAV) was HLA-A2-restricted, showing that the CTLs' reactivity against the peptide occurred only in the two patients and not in the healthy donors [112]. Based on the presence of HERV-K gag proteins in the cytoplasm of primary tumor cells and on the detection of antibodies to HERV-K gag in patients with seminoma, Rakoff–Nahoum investigated

the HERV-K-specific T cell-mediated immune response in the blood of those patients. The authors synthesized 15 HERV-K predicted peptides based on the HLA-I binding motif and proline-enriched region. Next, PBMCs from seminoma patients and healthy donors were screened with four pools of these peptides. The T cell reactivity was higher in at least three pools of peptides in the seminoma patients compared to the healthy donors [113].

Their high tumor specificity and expression [93] and incomplete T cell tolerance [108] make ERVs the ideal target for cancer immunotherapeutic approaches. In addition, autologous CTLs which are able to recognize HLA-restricted peptides have been reported [110–113], and ERVs are common to cancer patients. This allows for off-the-shelf therapy. However, the epitopes recognized by CTLs which are found in cancer patients are still few, and the expression of different HERV families in cancer is still limited. Therefore, future proteomic analysis, especially of the thymus, and an in-depth understanding of the mechanisms involved in HLA-I presentation will shed light on the use of ERVs in cancer immunotherapy [109].

3. In Vitro Methods for HLA Ligand Enrichment

3.1. Immunopeptidome

The existence and role of tumor antigens in eliciting a specific anti-cancer immune response, combined with the discovery of CD8+ T cell sub-populations which are able to recognize and kill tumor cells in an HLA I antigen-restricted manner, make the identification of epitopes recognized by CD8+ T cells a priority in the cancer therapeutic field. The peptides which are bound to the HLA complex and found on the cellular surface are referred as immunopeptidomes or ligandomes. The methods developed to study and analyze these are known as immunopeptidomics. The aim of immunopeptidomics is to reliably identify immunopeptidomes and thus guide the development of cancer therapeutic vaccines. The direct isolation of HLA peptides from the cell surfaces can be accomplished using different techniques. In this section, we describe the past and present approaches to the direct isolation of peptides in order to investigate the immunopeptidome landscape, highlighting the advantages and disadvantages of each (Figure 2).

3.1.1. Acid Stripping

In 1993, Storkus et al. published for the first time a method for the direct isolation of peptides from the HLA-I complex based on acid stripping. The method consisted of treating the cells with a citrate-phosphate buffer at pH 3.3 for a period as short as 15 s, thus allowing the cells to remain viable and become phenotypically class I-deficient [114]. After the treatment, a flow cytometry analysis of the acid-treated cells revealed the retention of class I heavy chains of the HLA complex in the cell membrane, while the class I light chain (β -2 microglobulin) was absent. Since β -2 microglobulin is essential to stabilize the binding of the peptide within the complex, its dissociation is directly associated with the release of HLA-bound peptides. Storkus et al. collected the supernatant from acid-treated, influenza A strain-infected cells and separated the fractions using reverse-phase high performance liquid chromatography (RP-HPLC). Next, the fractions were analyzed for their capacity to sensitize a B cell line to lysis which was mediated by a CTL line that was specific to the influenza A matrix peptide (Flu M1 57-68). Thus, the fractions that were able to induce the lysis were identified. The fractions contained a peptide with a sequence similar to the Flu M1 58-66 sequence [114].

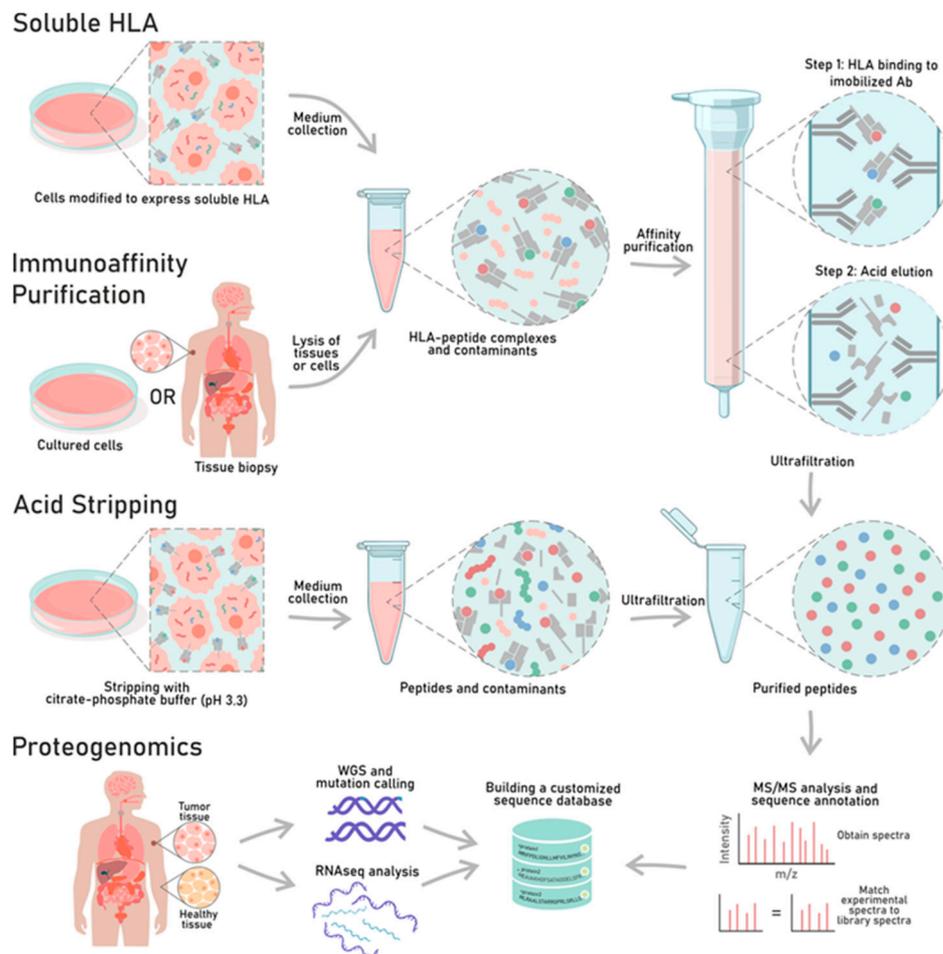


Figure 2. Workflow of the main approaches for the isolation of human leucocytes antigen (HLA)-I restricted peptides. In vitro methods for HLA enrichment are depicted. The direct isolation and identification of the peptides present in the HLA complex can be achieved through different approaches, such as the use of acid stripping, soluble HLA, immunoaffinity purification, and the proteogenomic. The purified peptides are then resolved by mass spectrometry analysis, and the spectra can be searched against a conventional database or a customized database (proteogenomic approach).

Following this work, the same authors used the aforementioned method to extract peptides from human melanoma cell lines. The subsequent analysis revealed six peptides (P1–P6) that were recognized by HLA-A2-restricted TILs [115]. Subsequently, the method has been used to investigate HLA-bound peptides in melanoma [116] and to isolate the first leukemia-specific immunogenic peptide derived from the Bcr-Abl fusion protein [117] in order to develop peptide-based immunotherapy.

Acid stripping can be used to extract peptides from adherent and suspension cell lines. Moreover, the approach is quite simple and cost effective. However, acid-treated cells are often damaged in the process, releasing proteases that generate peptides from either abundant cell protein or from cytoplasmic protein. Consequently, the peptides are often contaminated with non-HLA-restricted peptides [11,118].

3.1.2. Soluble HLA Molecules

The use of soluble HLA molecules is based on cells engineered with the vector encoding the desired soluble HLA complex. This approach requires the generation and isolation of stable transfected cells to express the soluble HLA molecule. The latter lacks the transmembrane domain that is needed for it to be secreted. Hence, the soluble HLA is released in the cellular medium that is collected and immunopurified for the further characterization of peptides [11,119].

Different approaches for generating the soluble HLA-I complex have been used over the years. In 1986, McCluskey and co-workers developed a soluble HLA molecule by fusing the non-functional transmembrane carboxyl terminal C2 domain of Q10^b with the polymorphic amino-terminal N and C1 domains of H-2D^d, thus generating a chimeric H2D^D/Q10^b molecule [120]. In the following years, another group realized a soluble HLA as a fusion protein between the extracellular domain of H2Kb and the immunoglobulin heavy chain polypeptide [121]. A third approach to generating soluble HLA has been reported by Grumet et al. The authors removed the transmembrane and cytoplasmic regions from the B7 gene, creating soluble HLA-B07 (sHLA-B7) that in *in vivo* experiments suppressed humoral alloimmunization [122].

Soluble HLA molecules have been extensively used in investigating the immunopeptidome landscape in order to identify novel peptide candidates for therapeutic vaccines in the fields of cancer immunotherapy [123,124] and infectious diseases [125–128]. The use of soluble HLA for peptide isolation has been proposed as a valid alternative to immunoaffinity purification [119]. In some cases, the yield of retrieved peptides has been reported as improved compared to immunoaffinity purification [124]. Nevertheless, the approach is not feasible to study patient tumor samples or tissue because it requires the generation of stable transfected cells.

3.1.3. Immunoaffinity Purification

Immunoaffinity purification relies on the direct isolation of HLA-I complexes from solubilized samples that are subsequently applied to columns and previously coupled with monoclonal antibodies that are able to bind the desired HLA-complex. The peptides are isolated from the HLA complex through acid elution and purification. The sequence is resolved by tandem mass spectrometry (MS/MS) analysis and validated in *in vitro* and *in silico* assays (e.g., motif clustering analysis) [11].

Briefly, the protocol starts with selecting the material to process; this material can be cells, tissue, biopsies, or biological liquid (i.e., plasma) [129]. Cells are the simplest material to work with because it is possible to expand them in culture and preserve the pellets by freezing them at $-80\text{ }^{\circ}\text{C}$ for up to 6 months [11]. The amount is usually in the range of 5×10^7 and 1×10^9 cells [11]. Next, the material is lysed using a lysis buffer that is a combination of a 0.5–2% solubilization reagent (e.g., Igepal, sodium deoxycholate, CHAPS), salts, and protease inhibitors [11,14,118]. The cell lysate is then centrifugated and the supernatant applied to the columns coupled with a monoclonal antibody.

The choice of monoclonal antibodies varies depending on the experimental conditions according to the investigated material and to the desired HLA complex. After the HLA complex binding, the columns are washed extremely well to get rid of material that is not specifically bound, detergent that could interfere with the downstream mass spectrometry analysis, and salts that could form crystal precipitates. The final step is the elution of the peptides from the HLA complexes. This is achieved using an acid solution with a pH of 3.0 (e.g., 10% acetic acid). Purification of the peptides is carried out using C18 columns or low-protein binding with a molecular weight cut-off (MWCO) spin filter. The fraction containing the HLA peptides is then vacuum-dried to reduce the acetonitrile concentration and analyzed by mass spectrometry to resolve the amino acid sequences [11,14].

An early attempt to isolate HLA-I antigens using immunoaffinity purification was reported by Parham in the 1970s [130]. In this work, the author demonstrated the feasibility of isolating the HLA-A02 and HLA-B07 complexes' antigens, starting from the crude membrane of the JY cell line. The sample was applied to a series of columns, each one respectively coupled with the monoclonal antibodies anti-HLA-A02, anti-HLA-B07, and anti-pan HLA-I. This first attempt paved the way to studying the ligandome landscape using immunoaffinity purification. Immunoaffinity purification evolved profoundly over the subsequent years and has been used to systemically analyze the HLA-restricted peptides derived either from virus- or cancer-transformed cells, as shown in the work of Engelhard and Hunt [131–136], Natheson [137], and Rammensee [138]. For instance, immunoaffinity purification has been used to identify the peptides that are recognized by CTLs derived from patients, as shown by Cox et al. Cox et al. isolated peptides from melanoma cells via immunoaffinity purification;

nine peptides were found to be recognized by CTLs derived from different patients, showing the feasibility of identifying candidates for peptide-based vaccines [132]. Immunoaffinity purification has been used to investigate antigens in malignant and non-malignant samples in order to define a pool of tumor-restricted peptides. This approach led to the successful identification of the peptides which are exclusive to tumors that are able to activate pre-existing T cells in colorectal carcinoma patients [138]. Interestingly, immunoaffinity purification can be used to identify spontaneous T cell responses in malignancies with few described associated antigens; for example, in chronic lymphocytic leukemia [139].

Immunoaffinity purification results in highly specific HLA peptides; however, the number and quality of the retrieved sequences is strongly affected by the following circumstances. Firstly, the material to be processed has to be provided in large amounts (5×10^7 – 1×10^9 cells), preventing the analysis of small and often clinically relevant tissues (e.g., needle biopsies) [11,140]. Secondly, the expression level of the HLA varies according to the samples, thus impacting the final number of peptides. Moreover, specific anti-HLA antibodies are not always available or are not optimal for binding. Finally, the use of spin filters has been associated with polymer contamination and peptide loss [11].

3.2. Proteogenomics

Proteogenomics is a broad research area that combines knowledge from the proteomic and genomic fields. Indeed, it combines the rapidly developing MS/MS approaches in proteomics with the high-throughput sequencing methodologies of genomics. The aim of proteomics is to identify and characterize (e.g., cellular localization, signaling pathway) the proteome of a given species. From a technical point of view, this aim is achieved through the digestion of proteins to generate peptides. Regarding the analysis of the immunopeptidome landscape, peptides are isolated from HLA-I complexes according to one of the aforementioned methods (Section 3.1.1, Section 3.1.2, and Section 3.1.3). The peptides are then resolved by MS/MS analysis, and the generated spectra are searched against the theoretical spectra of all candidate peptides, as represented in conventional reference databases (e.g., EntrezProtein, UniProtKB). These databases contain all the protein-coding sequences in the genome. Mutations, products of novel coding proteins or of annotated non-coding regions, and frameshifts cannot be identified because the spectra do not match any references in the canonical database [141], hindering the identification of potential vaccine peptides such as neo-antigens.

Genomics examines the genome organization (e.g., gene–gene interactions) within a given organism and mainly relies on sequencing approaches (whole genome sequencing, RNA sequencing) and bioinformatic algorithms for the subsequent interpretation of data. The introduction of next-generation sequencing (NGS) and the advancement of bioinformatics have revolutionized our approach to DNA and RNA sequence analysis, paving the way for tailored therapeutic cancer treatment. With respect to HLA-restricted tumor antigens, sequencing methods direct the identification of patient-specific neoantigens in therapeutic settings, taking advantage of whole exome sequencing-based mutation calling [94,142]. In 2017, Sahin et al. reported the first example of neo-epitope prediction, which was applied to the development of a cancer vaccine for the treatment of 13 melanoma patients [143]. The authors identified nonsynonymous mutations by the comparative exome and RNA sequencing (RNA-seq) of tumor biopsies and healthy blood cells. They then prioritized the neo-epitopes with predicted high affinity binding to the autologous HLA and a high expression level of RNA. This led the authors to develop an RNA vaccine encoding poly-neoepitopes for personalized melanoma treatment [143]. Following a similar approach, Wu et al. created a vaccine to target personal neoantigens. The authors used a whole exome sequencing of a patient's tumor and normal cells to identify somatic mutations and RNA sequencing in order to confirm the mutation expression. Peptide selection was based on the predicted binding affinity to the autologous HLA, developing synthetic long peptides containing up to 20 neoantigens per patient [144]. Even though knowledge of the genome and/or the RNA expression level allows researchers to make inferences about the use of HLA-I peptides for

a therapeutic cancer vaccine, the genomic approach by its very nature does not take into account post-translational modifications such as methylation, phosphorylation, and glycosylation. These have reportedly been involved in eliciting T cell immune responses and can be identified solely using MS/MS approaches [145–147]. For instance, phosphorylated peptides, but not their unphosphorylated homologous peptides, were recognized by CD8+T cells [146]. In addition, sequencing analysis lacks authentic knowledge of the HLA presentation, as the complex pathway underlying the presentation machinery cannot be completely deduced from the sequencing methodology and downstream bioinformatic analysis.

In 2004, Church et al. for the first time combined the potential of proteomics with global genome annotation, developing a new method for mapping the peptides detected in *Mycoplasma pneumoniae*. The method was named “proteogenomic mapping” [148]. Thus, proteogenomics as a science was born in 2004 in order to integrate genome annotation and proteomic application [141].

Applied to HLA ligandome analysis, proteogenomics addresses two main issues—the actual presentation of HLA peptides in the genomic analysis and the need for patient- or sample-specific mutation databases in order to search for spectra data [149]. The main rationale is the use of exome sequencing mutation calling in order to identify mutations in tumor samples and RNA-seq to confirm their expression. The sequencing data are then used to assemble tailored databases to investigate the proteomic data in the form of MS/MS spectra which are derived from the immunopeptidome analysis of the same samples. This approach has been successfully applied to the identification of HLA-restricted tumor neoantigens [150–152].

Recently, Bassani–Sternberg and colleagues described a proteogenomic pipeline for the identification of non-canonical peptides (long non-coding genes, UTRs, transposable elements, pseudogenes) from the ligandome repertoire [153]. Using MS/MS, the authors compared the immunopeptidome landscape of seven patient-derived melanoma cell lines and two pairs of lung cancer samples with matched healthy tissues. Using the same sample, they performed exome sequencing mutation calling and RNA sequencing analysis. The MS data were then compared against a personalized database built from the translation of the transcripts acquired from the sequencing analysis. This approach identified hundreds of non-canonical HLA peptides [153]. The combination of exome sequencing and MS/MS was also used in Kalaora et al. to investigate immunogenic neoepitopes in human melanoma [154].

Proteogenomics is a promising method to shed light on the antigen landscape, especially to identify neoantigens which are otherwise not detectable with conventional approaches. However, the amount of samples required for proteomic analysis is still high, hampering the investigation of small amounts of material (e.g., needle biopsies) [155].

4. Prediction of T Cell Epitopes

As mentioned, CTLs control the health of an organism, recognizing linear peptides which are presented on the MHC or, in humans, the HLA molecule. This system allows T cells to scan intracellularly processed proteins and consequently have an indication of the metabolic state of the cells; for example, whether they are transformed (cancerous) or virally infected. Protein degradation is mainly performed by the proteasome; however, the contribution of other proteases at this point seems clear [156]. The cleaved peptides are then transported in the endoplasmic reticulum (ER) by an ATP-binding complex called transporter-associated antigen processing (TAP). Other mechanisms of peptide translocation into the ER are evident yet poorly understood [157]. In the ER, peptides bind to the MHC molecule and are transported to the cell’s surface, where the complex (pMHC) will be potentially recognized by the TCRs of CD8+ T cells (Figure 3).

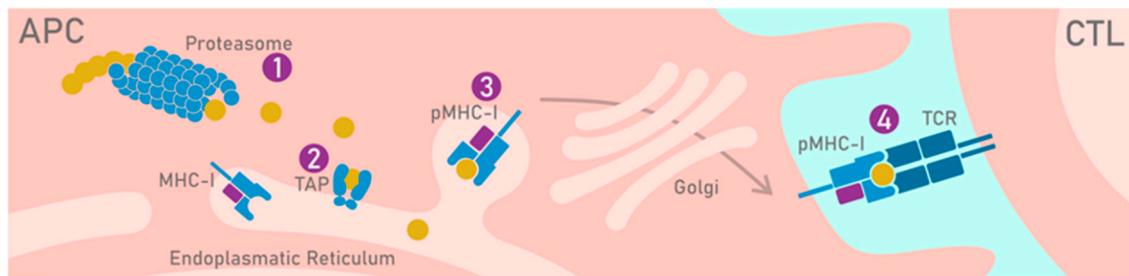


Figure 3. Processing and presentation pathway of HLA-I-restricted peptides. Intracellular proteins are processed by a proteasome (1). This cleavage determines the C-terminus of the peptides. The peptides bind transporter-associated antigen processing (TAP) that transports them into the endoplasmic reticulum (ER) (2). The peptides bind the HLA molecule that forms the pMHC complex (3) that is transported via the Golgi on the cell surface where CD8+ T cells can recognize them (4).

4.1. Prediction of Antigen Processing and Presentation

4.1.1. Proteasomal Cleavage Prediction

The proteasome is a multi-catalytic protease complex that is formed by different sub-units and represents the cellular main center of the degradation of misfolded or un-needed proteins. The core of the proteolytic machinery is the 20S complex that contains three catalytic sub-units ($\beta 1$, $\beta 2$, and $\beta 5$) that have specific catalytic activity: peptidyl-glutamyl-peptide-hydrolyzing activity (cleavage after acidic amino acids), trypsin-like activity (cleavage after basic amino acids), and chymotrypsin-like activity (cleavage after large, hydrophobic amino acids), respectively [158]. In the context of antigen presentation, proteasomes have been experimentally shown to be involved in the production of the C-termini of peptides [159,160].

Cells which are exposed to $\text{IFN}\gamma$ substitute the constitutive $\beta 1$, $\beta 2$, and $\beta 5$ with the homologs LMP2 ($\beta 1i$), MECL1 ($\beta 2i$), and LMP7 ($\beta 5i$) that are encoded in the same locus as the HLA molecules and for this reason are historically associated with “immune”-related activity. This newly formed protease complex is called immunoproteasome, and its catalytic activity is characterized by reduced cleavage after acidic amino acid residues and increased cleavage after hydrophobic and basic residues [158]. Proteasomal cleavage is the first step in the antigen processing machinery; therefore, significant effort has been made to develop models that are able to predict its activity.

Currently, the tool which is generally recognized as the best performing tool to predict both proteasomal and immuno-proteasomal cleavage is NetChop 3.1 [161]. This is an updated version of a previously developed artificial neural network (ANN) architecture [162]. Benchmarked with other proteasome cleavage prediction tools, NetChop has been shown to be better at predicting MHC I epitopes. It has been speculated that NetChop is better at predicting naturally processed epitopes because it is trained with HLA ligand data, unlike other tools that have been trained with in vitro proteasomal cleavage data only. This might have allowed the model to better generalize the different contributions coming from additional proteases [156]. Interestingly, proteasomes can also generate different kinds of epitopes by joining non-contiguous regions through a process called post-translational peptide splicing [163,164]. However, bioinformatic tools which are able to predict this phenomenon are currently lacking.

4.1.2. TAP Binding Prediction

The TAP complex is a heterodimer composed of two proteins, TAP1 and TAP2. It is an ATP-binding transporter delegated to the transport of peptides from the cytosol into the ER. Here, the peptides might be further processed by aminopeptidases or directly loaded onto the MHC and successively transported through the cell’s surface. It has been shown that the TAP transporter prioritizes peptides of certain lengths and carboxyl terminus residues commonly found to be HLA class I anchors, with minor contributions (positive and negative) from residues in other positions. It has therefore been

speculated that there is a coevolution of TAP with the MHC (despite its high polymorphism) [158]. Models to predict TAP's transport efficiency with peptides of arbitrary length have been produced over time [165,166], showing some improvement in epitope prediction efficiency. This approach has never been very successful alone; however, it has found a discrete use in combination with methods that will be discussed later.

Recently, it has been demonstrated that TAP-independent transport has a relevant role in CTL epitope production [157,167]. However, the possible mechanisms and the weight that it has globally on epitope presentation remain largely unknown [168].

4.1.3. Peptide–MHC Binding Prediction

HLA is considered the most polymorphic protein of the human genome (hyper-polymorphic protein). The class I region, located on chromosome 6, encodes for genes that form HLA class I molecules. Interestingly, just the HLA-A, -B, and -C gene polymorphisms account for the almost 17,000 different alleles which have been annotated so far [169]. Notably, despite the incredibly large number of alleles, the HLA class I molecules can be clustered into nine groups called “supertypes” that reflect close peptide-binding specificity [170]. However, each allele in a supertype has a unique and specific peptide-binding preference. Therefore, the fact that a peptide can bind to a given allele does not necessarily indicate that it can also bind to other alleles of the same supertype [171].

HLA Motif Deconvolution

Despite the constant increase in MS-eluted immune peptidome data [172–174], HLA peptide deconvolution—the process of associating each ligand to its presenting HLA molecule(s)—is still a critical task. The tool most commonly used to this end is GibbsCluster 2.0. It performs sequential unsupervised alignments and clustering tasks of peptide sequencing based on Gibbs sampling. The latest version can also handle variations in peptide length for the generation of a motif [175–177].

Interestingly, it has been shown that incorporating deconvoluted HLA peptidomics data can improve the accuracy of HLA-binding prediction tools for those HLA alleles with still few ligands in the existing databases [178]. This approach can be used to refine our understanding of peptide–HLA interaction when considering peptides which are longer than nine amino acids (generally the most common and studied length), as the motif can change slightly when longer ligands interact with HLAs [178]. This technique has been shown to help increase the yield of immunopeptidome runs by 20–30% [179].

MHC-Binding Affinity Prediction

MHC alleles may differ because of different amino acid substitutions, most of which are found within the binding site and are critical to determining peptide motif recognition [170]. One of the most crucial steps in the antigen presentation pathway is the ability of the peptide to form a complex with the HLA molecule, thus determining its probability of being presented. The ability to predict the peptide affinity for the HLA is generally recognized as a key factor in the selection of T cell epitopes. The first attempts to predict peptides' MHC-binding affinity relied on motif search techniques, but today, machine learning-based methods are preferred.

One of the most commonly used tools to predict MHC-binding affinity is NetMHC 4.0 [180], a feed-forward ANN ensemble with a single hidden layer which has been trained with a set of quantitative peptide–MHC class I affinity measurements from the Immune Epitope Database (IEDB) of peptides of different lengths. Another ANN-based tool is MHCflurry [181]. This is an open source package, and its architecture combines the use of locally connected and fully connected layers which have been trained with both experimental affinity measurements and MS-eluted peptides for allele-specific prediction.

Other methods are called pan-specific as they accept both peptide and HLA amino acid sequences and are able predict the affinity of any peptide to any HLA-I.

NetMHCpan is the most commonly used tool [182], and in its latest version (NetMHCpan 4.0) [183], an increased prediction accuracy has been reached by adding MS-eluted peptide data into the training set.

Interestingly, PSSMHCpan [184] is a pan-specific HLA-binding affinity method based on allele-specific position-specific scoring matrices (PSSM) which are generated by multiple sequence alignments of peptides binding to already characterized HLAs. When uncharacterized HLAs are encountered, the algorithm uses the similarity to the nearest HLA sequence as a weight that affects the binding score for the queried HLA.

In recent years, deep learning, a particular neural network architecture, has shown its power in classification and regression tasks and has attracted increasing attention. Zhao et al. [185] developed a convolutional neural network that uses different peptide properties, such as the order of the sequence, the hydropathy index, polarity, and the length of peptide needed to perform the prediction, as these are the key factors in determining the binding to the HLA molecule. Zeng and Gifford developed a deep learning-based method that consists of a binding affinity prediction module and a peptide embedding module [186]. The latter applies a deep language model to embed each peptide into a vector representation. The concatenated output of the two models represents the input for a fully connected neural network (NN). Zeng and Gifford also suggested a residual neural network (PUFFIN) that quantifies the uncertainty in peptide–HLA affinity prediction or, in other words, determines the confidence of the prediction [187]. Another recently developed pan-specific method is ACME [188], which consists of a deep convolutional neural network (CNN) implemented with an attention mechanism that returns interpretable patterns that facilitate the prediction. DeepHLApan is a recurrent neural network (RNN)-based method that takes neoepitopes as inputs and considers both the possibility of mutant peptide presentation and its potential immunogenicity [189].

Over time, sequence-based machine learning-driven methodologies have become increasingly popular because of their predictive power. However, they give poor structural insights into the dynamics of the interactions between the peptides and the HLA molecules. Modeling approaches are suitable tools to tackle this problem, but their use has been hindered for many years because of the high computational power required. In recent years, different tools have been developed for this purpose, or existing tools have been implemented to enhance their performance. One such tool is the web-server tool DockTope [190], currently hosted by IEDB (<http://tools.iedb.org/docktope/>). DockTope is able to iteratively model pMHC-I complexes until the best conformation is found, but given the small number of crystal structures of pMHC complexes that is available, the tool is only able to model the conformations of two human (HLA-A*02:01 and HLA-B*27:05) and two murine (H-2-Db and H-2-Kb) MHC-I allotypes. DINC [191] is another web-server tool based on the incremental meta-docking approach that has high accuracy in predicting the geometry of peptides in a complex with MHC molecules. GradDock [192] is a fast and accurate structure modeling algorithm for pMHC-I. The docking simulator was designed to be unbiased toward any MHC-I molecules by generating the initial unbound peptides in vacuo and successively inserting them into the desired MHC-I molecules. Finally, APE-Gen [193] is a very fast and accurate modeling method that has the potential to be useful because of its scalability (i.e., modeling thousands of pMHCs or non-canonical longer peptides) and flexibility.

4.1.4. Combination of Different Predicting Tools

In order to better mimic all the processing and presentation steps that a naturally presented epitope goes through, approaches that combine different single step predictors have been produced. Tools such as NetCTL and NetCTLpan have shown that the combination strategy is more effective than individual tools in predicting naturally presented epitopes [194,195]. NetMHCcons combines the predictive power of three state-of-the-art peptide–MHC binding predictors and has been shown to outperform individual tools [196]. The ultimate advantage of using a combination of the previously described methods is that the number of candidate peptides to be tested is significantly reduced

compared to predictions based on individual methods. Nevertheless, a comprehensive benchmarking of all these tools has never been performed.

4.2. Prediction of Epitope Immunogenicity

At this point, it is important to note that the ability of a peptide to bind the HLA molecule and be present on the cell's surface does not imply its ability to engage a T cell and promote its activation. There is a difference between naturally presented epitopes; some are able to elicit an immune response (and, hence, are immunogenic) and some are not.

4.2.1. Peptide Stability Prediction

A peptide's affinity to a specific HLA molecule is crucial for its presentation, yet it does not give any indication about its possible recognition by a T cell. A step forward in this sense has been taken by Jørgensen and Rasmussen, who respectively developed NetMHCstab and NetMHCstabpan, which are ANN-based methods for predicting the stability of the pHLA complex [197,198]. They have shown that the integration of peptide-HLA binding affinity predictors with pHLA stability prediction significantly improves the identification of CTL epitopes. Moreover, assuming that longer lasting epitope presentation increases the likelihood of T cell recognition, they show that pMHC stability better correlates with immunogenicity than HLA binding affinity.

4.2.2. Inherent Peptide Immunogenicity

Immunogenicity is defined as the ability of a given antigen—in this case, a pHLA complex—to elicit an immune response. There are different possible approaches to address this problem. One of these considers immunogenicity as an inherent property of the peptide. The first attempts to predict immunogenicity produced POPI (an SVM-based predictor) and POPSIK (an SVM-based method using the weighted degree string kernel). These models explored the physicochemical properties derived by the AA index [199] of the amino acid composing the peptides. Of the properties selected by the models, four were hydrophobicity-related and two were residue volume-related [200,201].

Calis et al. produced a scoring model derived by summing the log enrichment scores of amino acids found at non-anchor positions weighted by the importance of that position for immunogenic peptides compared to non-immunogenic peptides [202]. In the same year, Saethang et al. developed a method for predicting T cell reactivity based on peptide encoding using a combination of amino acid pairwise contact potentials (AAPPs) and quantum topological molecular similarity (QTMS) descriptors [203]. More recently, Zhang et al. attempted to predict immunogenic peptides based on their sequences. They analyzed the relationships between different features vis-à-vis immunogenicity and selected the optimal feature subset using a genetic algorithm [204].

Some interesting attempts to produce an integrated method are NepTepi [205], which comprises several of the HLA-binding affinity methods described above, the pMHC complex stability NetMHCStab, and the immunogenicity model produced by Calis et al.

Despite these efforts, however, immunogenicity remains a feature that is far more difficult to model and predict than mere MHC-binding affinity. On one hand, this might be a problem that is just due to the limited amount of data. On the other hand, it might reflect a more complex biological scenario; for example, genetic background (HLA haplotype), central tolerance, and TCR repertoire.

4.2.3. Interaction with T Cells

All existing approaches to predicting epitope presentation do not consider whether that epitope is going to be recognized by a CTL. Previously, it has been shown how immunogenicity can, to a certain extent, be predicted by analyzing the sequence-specific features of HLA-bound peptides. Nevertheless, the results have been poor, as only one side of the immunological synapsis has been considered. Recently, the more widely used multimers-based technology [206] and the development of single cell TCR sequencing have, to a certain extent, filled the gap between TCR sequences and their recognition

specificity, thus increasing the amount of experimental data available. This has given rise to different databases, such as McPAS-TCR [207], VDJdb [208,209], IEDB [210], STCRDab [211], and ATLAS [212].

An interesting work that brings peptide–TCR recognition to the same side is that of Gielis et al., who identified epitope-specific TCR sequences using a random forest algorithm [213]. However, it can only predict TCR sequences based on the peptides present in the training database.

Oghishi and Yotsuyanagi et al. abandoned the long-standing idea that the TCR repertoire is highly stochastic and individualized, showing that epitope recognition by T cells is to some extent predictable. The authors used a method to approximate the molecular scanning process of the presented peptides by the TCR repertoire, reducing the problem to an alignment-based approach [214].

Another valuable strategy for the future is to undertake structural modeling of the interaction between the pMHC and the TCR. With this concept in mind, Jensen et al. developed the TCRp-MHC model [215], a tool that accepts as input the amino acid sequence of both the peptide and the TCR of interest (alpha and beta chains). The tool automatically identifies the best structural templates and generates a structural model of the target using comparative modeling.

TCRp-MHC models the pMHC and the TCR separately using MODELLER and LYRA [216], respectively, after which they are assembled in an additional modeling step to form the full TCR-pMHC complex. Interestingly, the structural models generated by the tool have high quality and are generated within a computational time of only 2 min.

5. Conclusions

The advancements in antigen discovery have paved the way for a deeper understating of the complex interaction between peptides and the HLA-I complex on one side and between HLA-I-restricted peptides and CD8+ T cells on the other. This knowledge is the key to developing efficient strategies for cancer therapeutic vaccines. To reach a better outcome and to exploit the overall potential of the antigen discovery approaches, several challenges must be addressed. One is that the HLA *in vitro* methods suffer from a lack of lab-to-lab reproducibility, hampering the impartial comparison of the generated HLA-I ligand datasets. This issue demands the urgent standardization of the protocol used for the isolation of peptides. Moreover, the community would generally benefit from technical advancements that allow the analysis of small amounts of tumor material (e.g., needle biopsies). In addition, the availability of open access datasets encourages collaboration among the different labs involved in antigen discovery. This will ultimately facilitate reliable progress in the field of cancer immunotherapeutic treatment. Regarding the *in silico* approaches, the production of an increasing amount of experimental data surely facilitates the generation of models to predict given phenomena. However, the presence of sparse databases which sometimes contain redundant data makes the collection, processing, and analysis of data tedious and difficult. The continuous production of new prediction algorithms is a great achievement; however, benchmarking these tools is an increasingly difficult task because of the limited amount of data in the respective training sets. Moreover, the use of too many different tools and pipelines creates confusion about the best practices to follow.

We believe that the combination of *in vitro* methods and *in silico* approaches is the key to tackling the complex interaction between tumor antigens and the T cell immune response. This necessary interdisciplinary expertise would expand our knowledge and ultimately benefit cancer patients.

Author Contributions: S.F., J.C., B.M., and V.C. conceived the work, decided the topic, and drafted the outline. S.F., J.C., and B.M. wrote the bulk of the text, generated the figures, and found the references. V.C. supervised the work. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the European Research Council under the European Union's Horizon 2020 Framework program (H2020)/ERC-CoG-2015, Grant Agreement n. 681219; the Helsinki Institute of Life Science (HiLIFE); the Jane and Aatos Erkko Foundation (decision 19072019); and the Cancer Society of Finland (Syöpäjärjestöt).

Conflicts of Interest: Vincenzo Cerullo is a cofounder and shareholder at VALO Therapeutics. The other authors have no conflicts of interest.

References

1. Sharma, P.; Wagner, K.; Wolchok, J.D.; Allison, J.P. Novel cancer immunotherapy agents with survival benefit: Recent successes and next steps. *Nat. Rev. Cancer* **2011**, *11*, 805–812. [[CrossRef](#)] [[PubMed](#)]
2. Sharma, P.; Allison, J.P. The future of immune checkpoint therapy. *Science* **2015**, *348*, 56–61. [[CrossRef](#)]
3. Pardoll, D.M. The blockade of immune checkpoints in cancer immunotherapy. *Nat. Rev. Cancer* **2012**, *12*, 252–264. [[CrossRef](#)]
4. Krummel, M.F.; Allison, J.P. CD28 and CTLA-4 have opposing effects on the response of T cells to stimulation. *J. Exp. Med.* **1995**, *182*, 459–465. [[CrossRef](#)]
5. Ishida, Y.; Agata, Y.; Shibahara, K.; Honjo, T. Induced expression of PD-1, a novel member of the immunoglobulin gene superfamily, upon programmed cell death. *EMBO J.* **1992**, *11*, 3887–3895. [[CrossRef](#)] [[PubMed](#)]
6. Rotte, A. Combination of CTLA-4 and PD-1 blockers for treatment of cancer. *J. Exp. Clin. Cancer Res.* **2019**, *38*, 255. [[CrossRef](#)]
7. Bonaventura, P.; Shekarian, T.; Alcazer, V.; Valladeau-Guilemond, J.; Valsesia-Wittmann, S.; Amigorena, S.; Caux, C.; Depil, S. Cold Tumors: A Therapeutic Challenge for Immunotherapy. *Front. Immunol.* **2019**, *10*, 168. [[CrossRef](#)] [[PubMed](#)]
8. Ribas, A.; Dummer, R.; Puzanov, I.; VanderWalde, A.; Andtbacka, R.H.I.; Michielin, O.; Olszanski, A.J.; Malvehy, J.; Cebon, J.; Fernandez, E.; et al. Oncolytic Virotherapy Promotes Intratumoral T Cell Infiltration and Improves Anti-PD-1 Immunotherapy. *Cell* **2017**, *170*, 1109–1119.e10. [[CrossRef](#)] [[PubMed](#)]
9. Sharma, P.; Allison, J.P. Immune checkpoint targeting in cancer therapy: Toward combination strategies with curative potential. *Cell* **2015**, *161*, 205–214. [[CrossRef](#)] [[PubMed](#)]
10. Trolle, T.; McMurtrey, C.P.; Sidney, J.; Bardet, W.; Osborn, S.C.; Kaefer, T.; Sette, A.; Hildebrand, W.H.; Nielsen, M.; Peters, B. The Length Distribution of Class I-Restricted T Cell Epitopes Is Determined by Both Peptide Supply and MHC Allele-Specific Binding Preference. *J. Immunol.* **2016**, *196*, 1480–1487. [[CrossRef](#)] [[PubMed](#)]
11. Purcell, A.W.; Ramarathinam, S.H.; Ternette, N. Mass spectrometry-based identification of MHC-bound peptides for immunopeptidomics. *Nat. Protoc.* **2019**, *14*, 1687–1707. [[CrossRef](#)] [[PubMed](#)]
12. Bassani-Sternberg, M.; Coukos, G. Mass spectrometry-based antigen discovery for cancer immunotherapy. *Curr. Opin. Immunol.* **2016**, *41*, 9–17. [[CrossRef](#)] [[PubMed](#)]
13. Kowalewski, D.J.; Stevanovic, S. Biochemical large-scale identification of MHC class I ligands. *Methods Mol. Biol.* **2013**, *960*, 145–157. [[CrossRef](#)] [[PubMed](#)]
14. Bassani-Sternberg, M. Mass Spectrometry Based Immunopeptidomics for the Discovery of Cancer Neoantigens. *Methods Mol. Biol.* **2018**, *1719*, 209–221. [[CrossRef](#)]
15. Hassan, C.; Kester, M.G.; Oudgenoeg, G.; de Ru, A.H.; Janssen, G.M.; Drijfhout, J.W.; Spaapen, R.M.; Jimenez, C.R.; Heemskerk, M.H.; Falkenburg, J.H.; et al. Accurate quantitation of MHC-bound peptides by application of isotopically labeled peptide MHC complexes. *J. Proteom.* **2014**, *109*, 240–244. [[CrossRef](#)]
16. Caron, E.; Aebbersold, R.; Banaei-Esfahani, A.; Chong, C.; Bassani-Sternberg, M. A Case for a Human Immuno-Peptidome Project Consortium. *Immunity* **2017**, *47*, 203–208. [[CrossRef](#)] [[PubMed](#)]
17. Backert, L.; Kohlbacher, O. Immunoinformatics and epitope prediction in the age of genomic medicine. *Genome Med.* **2015**, *7*, 119. [[CrossRef](#)]
18. Gubin, M.M.; Artyomov, M.N.; Mardis, E.R.; Schreiber, R.D. Tumor neoantigens: Building a framework for personalized cancer immunotherapy. *J. Clin. Investig.* **2015**, *125*, 3413–3421. [[CrossRef](#)]
19. Rotzschke, O.; Falk, K.; Stevanovic, S.; Jung, G.; Walden, P.; Rammensee, H.G. Exact prediction of a natural T cell epitope. *Eur. J. Immunol.* **1991**, *21*, 2891–2894. [[CrossRef](#)]
20. Falk, K.; Rotzschke, O.; Stevanovic, S.; Jung, G.; Rammensee, H.G. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature* **1991**, *351*, 290–296. [[CrossRef](#)]
21. Soria-Guerra, R.E.; Nieto-Gomez, R.; Govea-Alonso, D.O.; Rosales-Mendoza, S. An overview of bioinformatics tools for epitope prediction: Implications on vaccine development. *J. Biomed. Inf.* **2015**, *53*, 405–414. [[CrossRef](#)] [[PubMed](#)]
22. Murphy, K.; Weaver, C. *Janeway's Immunobiology*, 9th ed.; Garland Science/Taylor & Francis Group, LLC: New York, NY, USA, 2016; p. 14, 904 pages.

23. Miller, J.F.; Mitchell, G.F. The thymus and the precursors of antigen reactive cells. *Nature* **1967**, *216*, 659–663. [[CrossRef](#)] [[PubMed](#)]
24. Miller, J.F. Events that led to the discovery of T-cell development and function—A personal recollection. *Tissue Antigens* **2004**, *63*, 509–517. [[CrossRef](#)]
25. Gross, L. Intradermal Immunization of C3H Mice against a Sarcoma That Originated in an Animal of the Same Line. *Cancer Res.* **1943**, *3*, 326–333.
26. Boon, T.; Kellermann, O. Rejection by syngeneic mice of cell variants obtained by mutagenesis of a malignant teratocarcinoma cell line. *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 272–275. [[CrossRef](#)]
27. Van Pel, A.; Boon, T. Protection against a nonimmunogenic mouse leukemia by an immunogenic variant obtained by mutagenesis. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 4718–4722. [[CrossRef](#)]
28. Lurquin, C.; Van Pel, A.; Mariame, B.; De Plaen, E.; Szikora, J.P.; Janssens, C.; Reddehase, M.J.; Lejeune, J.; Boon, T. Structure of the gene of tum- transplantation antigen P91A: The mutated exon encodes a peptide recognized with Ld by cytolytic T cells. *Cell* **1989**, *58*, 293–303. [[CrossRef](#)]
29. Vigneron, N.; Van den Eynde, B.J. Insights into the processing of MHC class I ligands gained from the study of human tumor epitopes. *Cell. Mol. Life Sci.* **2011**, *68*, 1503–1520. [[CrossRef](#)] [[PubMed](#)]
30. Hollingsworth, R.E.; Jansen, K. Turning the corner on therapeutic cancer vaccines. *NPJ Vaccines* **2019**, *4*, 7. [[CrossRef](#)]
31. Ilyas, S.; Yang, J.C. Landscape of Tumor Antigens in T Cell Immunotherapy. *J. Immunol.* **2015**, *195*, 5117–5122. [[CrossRef](#)]
32. Gaugler, B.; Brouwenstijn, N.; Vantomme, V.; Szikora, J.P.; Van der Spek, C.W.; Patard, J.J.; Boon, T.; Schrier, P.; Van den Eynde, B.J. A new gene coding for an antigen recognized by autologous cytolytic T lymphocytes on a human renal carcinoma. *Immunogenetics* **1996**, *44*, 323–330. [[CrossRef](#)] [[PubMed](#)]
33. Abi-Hanna, D.; Wakefield, D.; Watkins, S. HLA antigens in ocular tissues. I. In vivo expression in human eyes. *Transplantation* **1988**, *45*, 610–613. [[CrossRef](#)] [[PubMed](#)]
34. Oehlrich, N.; Devitt, G.; Linnebacher, M.; Schwitalle, Y.; Grosskinski, S.; Stevanovic, S.; Zoller, M. Generation of RAGE-1 and MAGE-9 peptide-specific cytotoxic T-lymphocyte lines for transfer in patients with renal cell carcinoma. *Int. J. Cancer* **2005**, *117*, 256–264. [[CrossRef](#)] [[PubMed](#)]
35. Vigneron, N. Human Tumor Antigens and Cancer Immunotherapy. *Biomed. Res. Int.* **2015**, *2015*, 948501. [[CrossRef](#)]
36. Kraus, M.H.; Popescu, N.C.; Amsbaugh, S.C.; King, C.R. Overexpression of the EGF receptor-related proto-oncogene erbB-2 in human mammary tumor cell lines by different molecular mechanisms. *EMBO J.* **1987**, *6*, 605–610. [[CrossRef](#)]
37. Pils, D.; Pinter, A.; Reibenwein, J.; Alfan, A.; Horak, P.; Schmid, B.C.; Hefler, L.; Horvat, R.; Reinthaller, A.; Zeillinger, R.; et al. In ovarian cancer the prognostic influence of HER2/neu is not dependent on the CXCR4/SDF-1 signalling pathway. *Br. J. Cancer* **2007**, *96*, 485–491. [[CrossRef](#)]
38. Hudis, C.A. Trastuzumab—mechanism of action and use in clinical practice. *N. Engl. J. Med.* **2007**, *357*, 39–51. [[CrossRef](#)]
39. Fisk, B.; Blevins, T.L.; Wharton, J.T.; Ioannides, C.G. Identification of an immunodominant peptide of HER-2/neu protooncogene recognized by ovarian tumor-specific cytotoxic T lymphocyte lines. *J. Exp. Med.* **1995**, *181*, 2109–2117. [[CrossRef](#)]
40. Mittendorf, E.A.; Lu, B.; Melisko, M.; Price Hiller, J.; Bondarenko, I.; Brunt, A.M.; Sergii, G.; Petrakova, K.; Peoples, G.E. Efficacy and Safety Analysis of Nelipecimut-S Vaccine to Prevent Breast Cancer Recurrence: A Randomized, Multicenter, Phase III Clinical Trial. *Clin. Cancer Res.* **2019**, *25*, 4248–4254. [[CrossRef](#)]
41. Benavides, L.C.; Gates, J.D.; Carmichael, M.G.; Patil, R.; Holmes, J.P.; Hueman, M.T.; Mittendorf, E.A.; Craig, D.; Stojadinovic, A.; Ponniah, S.; et al. The impact of HER2/neu expression level on response to the E75 vaccine: From U.S. Military Cancer Institute Clinical Trials Group Study I-01 and I-02. *Clin. Cancer Res.* **2009**, *15*, 2895–2904. [[CrossRef](#)]
42. Clifton, G.T.; Peoples, G.E.; Mittendorf, E.A. The development and use of the E75 (HER2 369-377) peptide vaccine. *Future Oncol.* **2016**, *12*, 1321–1329. [[CrossRef](#)] [[PubMed](#)]
43. Schmidt, S.M.; Schag, K.; Muller, M.R.; Weck, M.M.; Appel, S.; Kanz, L.; Grunebach, F.; Brossart, P. Survivin is a shared tumor-associated antigen expressed in a broad variety of malignancies and recognized by specific cytotoxic T cells. *Blood* **2003**, *102*, 571–576. [[CrossRef](#)] [[PubMed](#)]

44. Ropke, M.; Hald, J.; Guldberg, P.; Zeuthen, J.; Norgaard, L.; Fugger, L.; Svejgaard, A.; Van der Burg, S.; Nijman, H.W.; Melief, C.J.; et al. Spontaneous human squamous cell carcinomas are killed by a human cytotoxic T lymphocyte clone recognizing a wild-type p53-derived peptide. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 14704–14707. [[CrossRef](#)] [[PubMed](#)]
45. Rosenberg, S.A.; Restifo, N.P. Adoptive cell transfer as personalized immunotherapy for human cancer. *Science* **2015**, *348*, 62–68. [[CrossRef](#)]
46. Kochenderfer, J.N.; Dudley, M.E.; Feldman, S.A.; Wilson, W.H.; Spaner, D.E.; Maric, I.; Stetler-Stevenson, M.; Phan, G.Q.; Hughes, M.S.; Sherry, R.M.; et al. B-cell depletion and remissions of malignancy along with cytokine-associated toxicity in a clinical trial of anti-CD19 chimeric-antigen-receptor-transduced T cells. *Blood* **2012**, *119*, 2709–2720. [[CrossRef](#)]
47. Maude, S.L.; Teachey, D.T.; Porter, D.L.; Grupp, S.A. CD19-targeted chimeric antigen receptor T-cell therapy for acute lymphoblastic leukemia. *Blood* **2015**, *125*, 4017–4023. [[CrossRef](#)]
48. Kochenderfer, J.N.; Wilson, W.H.; Janik, J.E.; Dudley, M.E.; Stetler-Stevenson, M.; Feldman, S.A.; Maric, I.; Raffeld, M.; Nathan, D.A.; Lanier, B.J.; et al. Eradication of B-lineage cells and regression of lymphoma in a patient treated with autologous T cells genetically engineered to recognize CD19. *Blood* **2010**, *116*, 4099–4102. [[CrossRef](#)]
49. Lamers, C.H.; Sleijfer, S.; van Steenbergen, S.; van Elzakker, P.; van Krimpen, B.; Groot, C.; Vulto, A.; den Bakker, M.; Oosterwijk, E.; Debets, R.; et al. Treatment of metastatic renal cell carcinoma with CAIX CAR-engineered T cells: Clinical evaluation and management of on-target toxicity. *Mol. Ther.* **2013**, *21*, 904–912. [[CrossRef](#)]
50. Pedersen, S.R.; Sorensen, M.R.; Buus, S.; Christensen, J.P.; Thomsen, A.R. Comparison of vaccine-induced effector CD8 T cell responses directed against self- and non-self-tumor antigens: Implications for cancer immunotherapy. *J. Immunol.* **2013**, *191*, 3955–3967. [[CrossRef](#)]
51. Overwijk, W.W. Cancer vaccines in the era of checkpoint blockade: The magic is in the adjuvant. *Curr. Opin. Immunol.* **2017**, *47*, 103–109. [[CrossRef](#)]
52. Pittet, M.J.; Valmori, D.; Dunbar, P.R.; Speiser, D.E.; Lienard, D.; Lejeune, F.; Fleischauer, K.; Cerundolo, V.; Cerottini, J.C.; Romero, P. High frequencies of naive Melan-A/MART-1-specific CD8(+) T cells in a large proportion of human histocompatibility leukocyte antigen (HLA)-A2 individuals. *J. Exp. Med.* **1999**, *190*, 705–715. [[CrossRef](#)] [[PubMed](#)]
53. Brichard, V.; Van Pel, A.; Wolfel, T.; Wolfel, C.; De Plaen, E.; Lethe, B.; Coulie, P.; Boon, T. The tyrosinase gene codes for an antigen recognized by autologous cytolytic T lymphocytes on HLA-A2 melanomas. *J. Exp. Med.* **1993**, *178*, 489–495. [[CrossRef](#)] [[PubMed](#)]
54. Kawakami, Y.; Eliyahu, S.; Delgado, C.H.; Robbins, P.F.; Sakaguchi, K.; Appella, E.; Yannelli, J.R.; Adema, G.J.; Miki, T.; Rosenberg, S.A. Identification of a human melanoma antigen recognized by tumor-infiltrating lymphocytes associated with in vivo tumor rejection. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 6458–6462. [[CrossRef](#)]
55. Kawakami, Y.; Eliyahu, S.; Delgado, C.H.; Robbins, P.F.; Rivoltini, L.; Topalian, S.L.; Miki, T.; Rosenberg, S.A. Cloning of the gene coding for a shared human melanoma antigen recognized by autologous T cells infiltrating into tumor. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 3515–3519. [[CrossRef](#)] [[PubMed](#)]
56. Wang, R.F.; Appella, E.; Kawakami, Y.; Kang, X.; Rosenberg, S.A. Identification of TRP-2 as a human tumor antigen recognized by cytotoxic T lymphocytes. *J. Exp. Med.* **1996**, *184*, 2207–2216. [[CrossRef](#)] [[PubMed](#)]
57. Wang, R.F.; Robbins, P.F.; Kawakami, Y.; Kang, X.Q.; Rosenberg, S.A. Identification of a gene encoding a melanoma tumor antigen recognized by HLA-A31-restricted tumor-infiltrating lymphocytes. *J. Exp. Med.* **1995**, *181*, 799–804. [[CrossRef](#)]
58. Parkhurst, M.R.; Fitzgerald, E.B.; Southwood, S.; Sette, A.; Rosenberg, S.A.; Kawakami, Y. Identification of a shared HLA-A*0201-restricted T-cell epitope from the melanoma antigen tyrosinase-related protein 2 (TRP2). *Cancer Res.* **1998**, *58*, 4895–4901.
59. Olson, B.M.; Frye, T.P.; Johnson, L.E.; Fong, L.; Knutson, K.L.; Disis, M.L.; McNeel, D.G. HLA-A2-restricted T-cell epitopes specific for prostatic acid phosphatase. *Cancer Immunol. Immunother.* **2010**, *59*, 943–953. [[CrossRef](#)] [[PubMed](#)]
60. Correale, P.; Walmsley, K.; Nieroda, C.; Zaremba, S.; Zhu, M.; Schlom, J.; Tsang, K.Y. In vitro generation of human cytotoxic T lymphocytes specific for peptides derived from prostate-specific antigen. *J. Natl. Cancer Inst.* **1997**, *89*, 293–300. [[CrossRef](#)]

61. Parkhurst, M.R.; Joo, J.; Riley, J.P.; Yu, Z.; Li, Y.; Robbins, P.F.; Rosenberg, S.A. Characterization of genetically modified T-cell receptors that recognize the CEA:691-699 peptide in the context of HLA-A2.1 on human colorectal cancer cells. *Clin. Cancer Res.* **2009**, *15*, 169–180. [[CrossRef](#)]
62. Parkhurst, M.R.; Yang, J.C.; Langan, R.C.; Dudley, M.E.; Nathan, D.A.; Feldman, S.A.; Davis, J.L.; Morgan, R.A.; Merino, M.J.; Sherry, R.M.; et al. T cells targeting carcinoembryonic antigen can mediate regression of metastatic colorectal cancer but induce severe transient colitis. *Mol. Ther.* **2011**, *19*, 620–626. [[CrossRef](#)] [[PubMed](#)]
63. Hammarstrom, S. The carcinoembryonic antigen (CEA) family: Structures, suggested functions and expression in normal and malignant tissues. *Semin. Cancer Biol.* **1999**, *9*, 67–81. [[CrossRef](#)] [[PubMed](#)]
64. Yee, C.; Thompson, J.A.; Roche, P.; Byrd, D.R.; Lee, P.P.; Piepkorn, M.; Kenyon, K.; Davis, M.M.; Riddell, S.R.; Greenberg, P.D. Melanocyte destruction after antigen-specific immunotherapy of melanoma: Direct evidence of t cell-mediated vitiligo. *J. Exp. Med.* **2000**, *192*, 1637–1644. [[CrossRef](#)] [[PubMed](#)]
65. Jager, E.; Maeurer, M.; Hohn, H.; Karbach, J.; Jager, D.; Zidianakis, Z.; Bakhshandeh-Bath, A.; Orth, J.; Neukirch, C.; Necker, A.; et al. Clonal expansion of Melan A-specific cytotoxic T lymphocytes in a melanoma patient responding to continued immunization with melanoma-associated peptides. *Int. J. Cancer* **2000**, *86*, 538–547. [[CrossRef](#)]
66. Almeida, L.G.; Sakabe, N.J.; deOliveira, A.R.; Silva, M.C.; Mundstein, A.S.; Cohen, T.; Chen, Y.T.; Chua, R.; Gurung, S.; Gnjatic, S.; et al. CTdatabase: A knowledge-base of high-throughput and curated data on cancer-testis antigens. *Nucleic Acids Res.* **2009**, *37*, D816–D819. [[CrossRef](#)] [[PubMed](#)]
67. Fratta, E.; Coral, S.; Covre, A.; Parisi, G.; Colizzi, F.; Danielli, R.; Nicolay, H.J.; Sigalotti, L.; Maio, M. The biology of cancer testis antigens: Putative function, regulation and therapeutic potential. *Mol. Oncol.* **2011**, *5*, 164–182. [[CrossRef](#)]
68. van der Bruggen, P.; Traversari, C.; Chomez, P.; Lurquin, C.; De Plaen, E.; Van den Eynde, B.; Knuth, A.; Boon, T. A gene encoding an antigen recognized by cytolytic T lymphocytes on a human melanoma. *Science* **1991**, *254*, 1643–1647. [[CrossRef](#)]
69. De Plaen, E.; Arden, K.; Traversari, C.; Gaforio, J.J.; Szikora, J.P.; De Smet, C.; Brasseur, F.; van der Bruggen, P.; Lethe, B.; Lurquin, C.; et al. Structure, chromosomal localization, and expression of 12 genes of the MAGE family. *Immunogenetics* **1994**, *40*, 360–369. [[CrossRef](#)]
70. Gaugler, B.; Van den Eynde, B.; van der Bruggen, P.; Romero, P.; Gaforio, J.J.; De Plaen, E.; Lethe, B.; Brasseur, F.; Boon, T. Human gene MAGE-3 codes for an antigen recognized on a melanoma by autologous cytolytic T lymphocytes. *J. Exp. Med.* **1994**, *179*, 921–930. [[CrossRef](#)]
71. Traversari, C.; van der Bruggen, P.; Luescher, I.F.; Lurquin, C.; Chomez, P.; Van Pel, A.; De Plaen, E.; Amar-Costesec, A.; Boon, T. A nonapeptide encoded by human gene MAGE-1 is recognized on HLA-A1 by cytolytic T lymphocytes directed against tumor antigen MZ2-E. *J. Exp. Med.* **1992**, *176*, 1453–1457. [[CrossRef](#)]
72. Old, L.J.; Chen, Y.T. New paths in human cancer serology. *J. Exp. Med.* **1998**, *187*, 1163–1167. [[CrossRef](#)]
73. Chomez, P.; De Backer, O.; Bertrand, M.; De Plaen, E.; Boon, T.; Lucas, S. An overview of the MAGE gene family with the identification of all human members of the family. *Cancer Res.* **2001**, *61*, 5544–5551. [[PubMed](#)]
74. Simpson, A.J.; Caballero, O.L.; Jungbluth, A.; Chen, Y.T.; Old, L.J. Cancer/testis antigens, gametogenesis and cancer. *Nat. Rev. Cancer* **2005**, *5*, 615–625. [[CrossRef](#)] [[PubMed](#)]
75. Lian, Y.; Meng, L.; Ding, P.; Sang, M. Epigenetic regulation of MAGE family in human cancer progression-DNA methylation, histone modification, and non-coding RNAs. *Clin. Epigenetics* **2018**, *10*, 115. [[CrossRef](#)] [[PubMed](#)]
76. Weon, J.L.; Potts, P.R. The MAGE protein family and cancer. *Curr. Opin. Cell. Biol.* **2015**, *37*, 1–8. [[CrossRef](#)]
77. Van den Eynde, B.J.; van der Bruggen, P. T cell defined tumor antigens. *Curr. Opin. Immunol.* **1997**, *9*, 684–693. [[CrossRef](#)]
78. Al-Khadairi, G.; Decock, J. Cancer Testis Antigens and Immunotherapy: Where Do We Stand in the Targeting of PRAME? *Cancers* **2019**, *11*, 984. [[CrossRef](#)]
79. Ikeda, H.; Lethe, B.; Lehmann, F.; van Baren, N.; Baurain, J.F.; de Smet, C.; Chambost, H.; Vitale, M.; Moretta, A.; Boon, T.; et al. Characterization of an antigen that is recognized on a melanoma showing partial HLA loss by CTL expressing an NK inhibitory receptor. *Immunity* **1997**, *6*, 199–208. [[CrossRef](#)]

80. Kessler, J.H.; Beekman, N.J.; Bres-Vloemans, S.A.; Verdijk, P.; van Veelen, P.A.; Kloosterman-Joosten, A.M.; Vissers, D.C.; ten Bosch, G.J.; Kester, M.G.; Sijts, A.; et al. Efficient identification of novel HLA-A(*)0201-presented cytotoxic T lymphocyte epitopes in the widely expressed tumor antigen PRAME by proteasome-mediated digestion analysis. *J. Exp. Med.* **2001**, *193*, 73–88. [[CrossRef](#)]
81. Quintarelli, C.; Dotti, G.; Hasan, S.T.; De Angelis, B.; Hoyos, V.; Errichiello, S.; Mims, M.; Luciano, L.; Shafer, J.; Leen, A.M.; et al. High-avidity cytotoxic T lymphocytes specific for a new PRAME-derived peptide can target leukemic and leukemic-precursor cells. *Blood* **2011**, *117*, 3353–3362. [[CrossRef](#)]
82. Chen, Y.T.; Scanlan, M.J.; Sahin, U.; Tureci, O.; Gure, A.O.; Tsang, S.; Williamson, B.; Stockert, E.; Pfreundschuh, M.; Old, L.J. A testicular antigen aberrantly expressed in human cancers detected by autologous antibody screening. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 1914–1918. [[CrossRef](#)] [[PubMed](#)]
83. Thomas, R.; Al-Khadairi, G.; Roelands, J.; Hendrickx, W.; Dermime, S.; Bedognetti, D.; Decock, J. NY-ESO-1 Based Immunotherapy of Cancer: Current Perspectives. *Front. Immunol.* **2018**, *9*, 947. [[CrossRef](#)] [[PubMed](#)]
84. Davis, I.D.; Chen, W.; Jackson, H.; Parente, P.; Shackleton, M.; Hopkins, W.; Chen, Q.; Dimopoulos, N.; Luke, T.; Murphy, R.; et al. Recombinant NY-ESO-1 protein with ISCOMATRIX adjuvant induces broad integrated antibody and CD4(+) and CD8(+) T cell responses in humans. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 10697–10702. [[CrossRef](#)]
85. Tureci, O.; Sahin, U.; Schobert, I.; Koslowski, M.; Scmitt, H.; Schild, H.J.; Stenner, F.; Seitz, G.; Rammensee, H.G.; Pfreundschuh, M. The SSX-2 gene, which is involved in the t(X;18) translocation of synovial sarcomas, codes for the human tumor antigen HOM-MEL-40. *Cancer Res.* **1996**, *56*, 4766–4772. [[PubMed](#)]
86. Smith, H.A.; McNeel, D.G. The SSX family of cancer-testis antigens as target proteins for tumor therapy. *Clin. Dev. Immunol.* **2010**, *2010*, 150591. [[CrossRef](#)] [[PubMed](#)]
87. Boel, P.; Wildmann, C.; Sensi, M.L.; Bresseur, R.; Renauld, J.C.; Coulie, P.; Boon, T.; van der Bruggen, P. BAGE: A new gene encoding an antigen recognized on human melanomas by cytolytic T lymphocytes. *Immunity* **1995**, *2*, 167–175. [[CrossRef](#)]
88. Van den Eynde, B.; Peeters, O.; De Backer, O.; Gaugler, B.; Lucas, S.; Boon, T. A new family of genes coding for an antigen recognized by autologous cytolytic T lymphocytes on a human melanoma. *J. Exp. Med.* **1995**, *182*, 689–698. [[CrossRef](#)]
89. Fiszer, D.; Kurpisz, M. Major histocompatibility complex expression on human, male germ cells: A review. *Am. J. Reprod. Immunol.* **1998**, *40*, 172–176. [[CrossRef](#)]
90. Tyagi, P.; Mirakhur, B. MAGRIT: The largest-ever phase III lung cancer trial aims to establish a novel tumor-specific approach to therapy. *Clin. Lung Cancer* **2009**, *10*, 371–374. [[CrossRef](#)]
91. Ruiz, R.; Hunis, B.; Raez, L.E. Immunotherapeutic agents in non-small-cell lung cancer finally coming to the front lines. *Curr. Oncol. Rep.* **2014**, *16*, 400. [[CrossRef](#)]
92. Morgan, R.A.; Chinnsamy, N.; Abate-Daga, D.; Gros, A.; Robbins, P.F.; Zheng, Z.; Dudley, M.E.; Feldman, S.A.; Yang, J.C.; Sherry, R.M.; et al. Cancer regression and neurological toxicity following anti-MAGE-A3 TCR gene therapy. *J. Immunother.* **2013**, *36*, 133–151. [[CrossRef](#)] [[PubMed](#)]
93. Rooney, M.S.; Shukla, S.A.; Wu, C.J.; Getz, G.; Hachohen, N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* **2015**, *160*, 48–61. [[CrossRef](#)] [[PubMed](#)]
94. Schumacher, T.N.; Schreiber, R.D. Neoantigens in cancer immunotherapy. *Science* **2015**, *348*, 69–74. [[CrossRef](#)]
95. Schumacher, T.N.; Schepers, W.; Kvistborg, P. Cancer Neoantigens. *Annu. Rev. Immunol.* **2019**, *37*, 173–200. [[CrossRef](#)] [[PubMed](#)]
96. Coulie, P.G.; Lehmann, F.; Lethe, B.; Herman, J.; Lurquin, C.; Andrawiss, M.; Boon, T. A mutated intron sequence codes for an antigenic peptide recognized by cytolytic T lymphocytes on a human melanoma. *Proc Natl. Acad. Sci. USA* **1995**, *92*, 7976–7980. [[CrossRef](#)]
97. Robbins, P.F.; El-Gamil, M.; Li, Y.F.; Kawakami, Y.; Loftus, D.; Appella, E.; Rosenberg, S.A. A mutated beta-catenin gene encodes a melanoma-specific antigen recognized by tumor infiltrating lymphocytes. *J. Exp. Med.* **1996**, *183*, 1185–1192. [[CrossRef](#)]
98. Wolfel, T.; Hauer, M.; Schneider, J.; Serrano, M.; Wolfel, C.; Klehmann-Hieb, E.; De Plaen, E.; Hankeln, T.; Meyer zum Buschenfelde, K.H.; Beach, D. A p16INK4a-insensitive CDK4 mutant targeted by cytolytic T lymphocytes in a human melanoma. *Science* **1995**, *269*, 1281–1284. [[CrossRef](#)]
99. Mandruzzato, S.; Bresseur, F.; Andry, G.; Boon, T.; van der Bruggen, P. A CASP-8 mutation recognized by cytolytic T lymphocytes on a human head and neck carcinoma. *J. Exp. Med.* **1997**, *186*, 785–793. [[CrossRef](#)]
100. Tashiro, H.; Brenner, M.K. Immunotherapy against cancer-related viruses. *Cell Res.* **2017**, *27*, 59–73. [[CrossRef](#)]

101. Ramos, C.A.; Narala, N.; Vyas, G.M.; Leen, A.M.; Gerdemann, U.; Sturgis, E.M.; Anderson, M.L.; Savoldo, B.; Heslop, H.E.; Brenner, M.K.; et al. Human papillomavirus type 16 E6/E7-specific cytotoxic T lymphocytes for adoptive immunotherapy of HPV-associated malignancies. *J. Immunother.* **2013**, *36*, 66–76. [[CrossRef](#)]
102. Carbone, A.; Gloghini, A.; Dotti, G. EBV-associated lymphoproliferative disorders: Classification and treatment. *Oncologist* **2008**, *13*, 577–585. [[CrossRef](#)] [[PubMed](#)]
103. Javier, R.T.; Butel, J.S. The history of tumor virology. *Cancer Res.* **2008**, *68*, 7693–7706. [[CrossRef](#)] [[PubMed](#)]
104. Taylor, G.S.; Jia, H.; Harrington, K.; Lee, L.W.; Turner, J.; Ladell, K.; Price, D.A.; Tanday, M.; Matthews, J.; Roberts, C.; et al. A recombinant modified vaccinia ankara vaccine encoding Epstein-Barr Virus (EBV) target antigens: A phase I trial in UK patients with EBV-positive cancer. *Clin. Cancer Res.* **2014**, *20*, 5009–5022. [[CrossRef](#)]
105. Plummer, M.; de Martel, C.; Vignat, J.; Ferlay, J.; Bray, F.; Franceschi, S. Global burden of cancers attributable to infections in 2012: A synthetic analysis. *Lancet Glob. Health* **2016**, *4*, e609–e616. [[CrossRef](#)]
106. He, Q.; Liu, Z.; Liu, Z.; Lai, Y.; Zhou, X.; Weng, J. TCR-like antibodies in cancer immunotherapy. *J. Hematol. Oncol.* **2019**, *12*, 99. [[CrossRef](#)] [[PubMed](#)]
107. Bannert, N.; Hofmann, H.; Block, A.; Hohn, O. HERVs New Role in Cancer: From Accused Perpetrators to Cheerful Protectors. *Front. Microbiol.* **2018**, *9*, 178. [[CrossRef](#)]
108. Kassiotis, G.; Stoye, J.P. Immune responses to endogenous retroelements: Taking the bad with the good. *Nat. Rev. Immunol.* **2016**, *16*, 207–219. [[CrossRef](#)]
109. Attermann, A.S.; Bjerregaard, A.M.; Saini, S.K.; Gronbaek, K.; Hadrup, S.R. Human endogenous retroviruses and their implication for immunotherapeutics of cancer. *Ann. Oncol.* **2018**, *29*, 2183–2191. [[CrossRef](#)]
110. Takahashi, Y.; Harashima, N.; Kajigaya, S.; Yokoyama, H.; Cherkasova, E.; McCoy, J.P.; Hanada, K.; Mena, O.; Kurlander, R.; Tawab, A.; et al. Regression of human kidney cancer following allogeneic stem cell transplantation is associated with recognition of an HERV-E antigen by T cells. *J. Clin. Investig.* **2008**, *118*, 1099–1109. [[CrossRef](#)]
111. Cherkasova, E.; Scrivani, C.; Doh, S.; Weisman, Q.; Takahashi, Y.; Harashima, N.; Yokoyama, H.; Srinivasan, R.; Linehan, W.M.; Lerman, M.I.; et al. Detection of an Immunogenic HERV-E Envelope with Selective Expression in Clear Cell Kidney Cancer. *Cancer Res.* **2016**, *76*, 2177–2185. [[CrossRef](#)] [[PubMed](#)]
112. Schiavetti, F.; Thonnard, J.; Colau, D.; Boon, T.; Coulie, P.G. A human endogenous retroviral sequence encoding an antigen recognized on melanoma by cytolytic T lymphocytes. *Cancer Res.* **2002**, *62*, 5510–5516.
113. Rakoff-Nahoum, S.; Kuebler, P.J.; Heymann, J.J.; Sheehy, M.E.; Ortiz, G.M.; Ogg, G.S.; Barbour, J.D.; Lenz, J.; Steinfeld, A.D.; Nixon, D.F. Detection of T lymphocytes specific for human endogenous retrovirus K (HERV-K) in patients with seminoma. *Aids Res. Hum. Retrovir.* **2006**, *22*, 52–56. [[CrossRef](#)] [[PubMed](#)]
114. Storkus, W.J.; Zeh, H.J., 3rd; Salter, R.D.; Lotze, M.T. Identification of T-cell epitopes: Rapid isolation of class I-presented peptides from viable cells by mild acid elution. *J. Immunother. Emphas. Tumor Immunol.* **1993**, *14*, 94–103. [[CrossRef](#)]
115. Storkus, W.J.; Zeh, H.J., 3rd; Maeurer, M.J.; Salter, R.D.; Lotze, M.T. Identification of human melanoma peptides recognized by class I restricted tumor infiltrating T lymphocytes. *J. Immunol.* **1993**, *151*, 3719–3727. [[PubMed](#)]
116. Castelli, C.; Storkus, W.J.; Maeurer, M.J.; Martin, D.M.; Huang, E.C.; Pramanik, B.N.; Nagabhusan, T.L.; Parmiani, G.; Lotze, M.T. Mass spectrometric identification of a naturally processed melanoma peptide recognized by CD8+ cytotoxic T lymphocytes. *J. Exp. Med.* **1995**, *181*, 363–368. [[CrossRef](#)]
117. Clark, R.E.; Dodi, I.A.; Hill, S.C.; Lill, J.R.; Aubert, G.; Macintyre, A.R.; Rojas, J.; Bourdon, A.; Bonner, P.L.; Wang, L.; et al. Direct evidence that leukemic cells present HLA-associated immunogenic peptides derived from the BCR-ABL b3a2 fusion protein. *Blood* **2001**, *98*, 2887–2893. [[CrossRef](#)]
118. Disis, M.L. Immunotherapy of cancer. In *Cancer Drug Discovery and Development*; Humana Press: Totowa, NJ, USA, 2006.
119. Scull, K.E.; Dudek, N.L.; Corbett, A.J.; Ramarathinam, S.H.; Gorasia, D.G.; Williamson, N.A.; Purcell, A.W. Secreted HLA recapitulates the immunopeptidome and allows in-depth coverage of HLA A*02:01 ligands. *Mol. Immunol.* **2012**, *51*, 136–142. [[CrossRef](#)]
120. Margulies, D.H.; Ramsey, A.L.; Boyd, L.F.; McCluskey, J. Genetic engineering of an H-2Dd/Q10b chimeric histocompatibility antigen: Purification of soluble protein from transformant cell supernatants. *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 5252–5256. [[CrossRef](#)]

121. Dal Porto, J.; Johansen, T.E.; Catipovic, B.; Parfiit, D.J.; Tuveson, D.; Gether, U.; Kozlowski, S.; Fearon, D.T.; Schneek, J.P. A soluble divalent class I major histocompatibility complex molecule inhibits alloreactive T cells at nanomolar concentrations. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 6671–6675. [[CrossRef](#)]
122. Grumet, F.C.; Krishnaswamy, S.; See-Tho, K.; Filvaroff, E.; Hiraki, D.D. Soluble form of an HLA-B7 class I antigen specifically suppresses humoral alloimmunization. *Hum. Immunol.* **1994**, *40*, 228–234. [[CrossRef](#)]
123. Yoshida, M.; Ishioka, Y.; Ozawa, T.; Okuyama, H.; Iguchi, M.; Ota, T.; Ito, T.; Nagira, M.; Morita, A.; Tanaka, H.; et al. Soluble HLA-associated peptide from PSF1 has a cancer vaccine potency. *Sci. Rep.* **2017**, *7*, 11137. [[CrossRef](#)] [[PubMed](#)]
124. Barnea, E.; Beer, I.; Patoka, R.; Ziv, T.; Kessler, O.; Tzeheval, E.; Eisenbach, L.; Zavazava, N.; Admon, A. Analysis of endogenous peptides bound by soluble MHC class I molecules: A novel approach for identifying tumor-specific antigens. *Eur. J. Immunol.* **2002**, *32*, 213–222. [[CrossRef](#)]
125. Gilchuk, P.; Spencer, C.T.; Conant, S.B.; Hill, T.; Gray, J.J.; Niu, X.; Zheng, M.; Erickson, J.J.; Boyd, K.L.; McAfee, K.J.; et al. Discovering naturally processed antigenic determinants that confer protective T cell immunity. *J. Clin. Investig.* **2013**, *123*, 1976–1987. [[CrossRef](#)]
126. Wahl, A.; Schafer, F.; Bardet, W.; Buchli, R.; Air, G.M.; Hildebrand, W.H. HLA class I molecules consistently present internal influenza epitopes. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 540–545. [[CrossRef](#)] [[PubMed](#)]
127. Yaciuk, J.C.; Skaley, M.; Bardet, W.; Schafer, F.; Mojsilovic, D.; Cate, S.; Stewart, C.J.; McMurtrey, C.; Jackson, K.W.; Buchli, R.; et al. Direct interrogation of viral peptides presented by the class I HLA of HIV-infected T cells. *J. Virol.* **2014**, *88*, 12992–13004. [[CrossRef](#)] [[PubMed](#)]
128. Norcross, M.A.; Luo, S.; Lu, L.; Boyne, M.T.; Gomarteli, M.; Rennels, A.D.; Woodcock, J.; Margulies, D.H.; McMurtrey, C.; Vernon, S.; et al. Abacavir induces loading of novel self-peptides into HLA-B*57: 01: An autoimmune model for HLA-associated drug hypersensitivity. *AIDS* **2012**, *26*, F21–F29. [[CrossRef](#)]
129. Bassani-Sternberg, M.; Barnea, E.; Beer, I.; Avivi, I.; Katz, T.; Admon, A. Soluble plasma HLA peptidome as a potential source for cancer biomarkers. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 18769–18776. [[CrossRef](#)]
130. Parham, P. Purification of immunologically active HLA-A and -B antigens by a series of monoclonal antibody columns. *J. Biol. Chem.* **1979**, *254*, 8709–8712.
131. Hunt, D.F.; Henderson, R.A.; Shabanowitz, J.; Sakaguchi, K.; Michel, H.; Sevilir, N.; Cox, A.L.; Appella, E.; Engelhard, V.H. Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. *Science* **1992**, *255*, 1261–1263. [[CrossRef](#)]
132. Cox, A.L.; Skipper, J.; Chen, Y.; Henderson, R.A.; Darrow, T.L.; Shabanowitz, J.; Engelhard, V.H.; Hunt, D.F.; Slingluff, C.L., Jr. Identification of a peptide recognized by five melanoma-specific human cytotoxic T cell lines. *Science* **1994**, *264*, 716–719. [[CrossRef](#)]
133. Skipper, J.C.; Hendrickson, R.C.; Gulden, P.H.; Brichard, V.; Van Pel, A.; Chen, Y.; Shabanowitz, J.; Wolfel, T.; Slingluff, C.L., Jr.; Boon, T.; et al. An HLA-A2-restricted tyrosinase antigen on melanoma cells results from posttranslational modification and suggests a novel pathway for processing of membrane proteins. *J. Exp. Med.* **1996**, *183*, 527–534. [[CrossRef](#)] [[PubMed](#)]
134. Meadows, L.; Wang, W.; den Haan, J.M.; Blokland, E.; Reinhardus, C.; Drijfhout, J.W.; Shabanowitz, J.; Pierce, R.; Agulnik, A.I.; Bishop, C.E.; et al. The HLA-A*0201-restricted H-Y antigen contains a posttranslationally modified cysteine that significantly affects T cell recognition. *Immunity* **1997**, *6*, 273–281. [[CrossRef](#)]
135. Crotzer, V.L.; Christian, R.E.; Brooks, J.M.; Shabanowitz, J.; Settlege, R.E.; Marto, J.A.; White, F.M.; Rickinson, A.B.; Hunt, D.F.; Engelhard, V.H. Immunodominance among EBV-derived epitopes restricted by HLA-B27 does not correlate with epitope abundance in EBV-transformed B-lymphoblastoid cell lines. *J. Immunol.* **2000**, *164*, 6120–6129. [[CrossRef](#)] [[PubMed](#)]
136. Zarling, A.L.; Ficarro, S.B.; White, F.M.; Shabanowitz, J.; Hunt, D.F.; Engelhard, V.H. Phosphorylated peptides are naturally processed and presented by major histocompatibility complex class I molecules in vivo. *J. Exp. Med.* **2000**, *192*, 1755–1762. [[CrossRef](#)] [[PubMed](#)]
137. Van Bleek, G.M.; Nathenson, S.G. Isolation of an endogenously processed immunodominant viral peptide from the class I H-2Kb molecule. *Nature* **1990**, *348*, 213–216. [[CrossRef](#)]
138. Loffler, M.W.; Kowalewski, D.J.; Backert, L.; Bernhardt, J.; Adam, P.; Schuster, H.; Dengler, F.; Backes, D.; Kopp, H.G.; Beckert, S.; et al. Mapping the HLA Ligandome of Colorectal Cancer Reveals an Imprint of Malignant Cell Transformation. *Cancer Res.* **2018**, *78*, 4627–4641. [[CrossRef](#)]

139. Kowalewski, D.J.; Schuster, H.; Backert, L.; Berlin, C.; Kahn, S.; Kanz, L.; Salih, H.R.; Rammensee, H.G.; Stevanovic, S.; Stickle, J.S. HLA ligandome analysis identifies the underlying specificities of spontaneous antileukemia immune responses in chronic lymphocytic leukemia (CLL). *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E166–E175. [[CrossRef](#)] [[PubMed](#)]
140. Chong, C.; Marino, F.; Pak, H.; Racle, J.; Daniel, R.T.; Muller, M.; Gfeller, D.; Coukos, G.; Bassani-Sternberg, M. High-throughput and Sensitive Immunopeptidomics Platform Reveals Profound Interferon-gamma-Mediated Remodeling of the Human Leukocyte Antigen (HLA) Ligandome. *Mol. Cell Proteom.* **2018**, *17*, 533–548. [[CrossRef](#)]
141. Nesvizhskii, A.I. Proteogenomics: Concepts, applications and computational strategies. *Nat. Methods* **2014**, *11*, 1114–1125. [[CrossRef](#)]
142. Segal, N.H.; Parsons, D.W.; Peggs, K.S.; Velculescu, V.; Kinzler, K.W.; Vogelstein, B.; Allison, J.P. Epitope landscape in breast and colorectal cancer. *Cancer Res.* **2008**, *68*, 889–892. [[CrossRef](#)]
143. Sahin, U.; Derhovanessian, E.; Miller, M.; Kloke, B.P.; Simon, P.; Lower, M.; Bukur, V.; Tadmor, A.D.; Luxemburger, U.; Schrors, B.; et al. Personalized RNA mutanome vaccines mobilize poly-specific therapeutic immunity against cancer. *Nature* **2017**, *547*, 222–226. [[CrossRef](#)]
144. Ott, P.A.; Hu, Z.; Keskin, D.B.; Shukla, S.A.; Sun, J.; Bozym, D.J.; Zhang, W.; Luoma, A.; Giobbie-Hurder, A.; Peter, L.; et al. An immunogenic personal neoantigen vaccine for patients with melanoma. *Nature* **2017**, *547*, 217–221. [[CrossRef](#)] [[PubMed](#)]
145. Malaker, S.A.; Penny, S.A.; Steadman, L.G.; Myers, P.T.; Loke, J.C.; Raghavan, M.; Bai, D.L.; Shabanowitz, J.; Hunt, D.F.; Cobbold, M. Identification of Glycopeptides as Posttranslationally Modified Neoantigens in Leukemia. *Cancer Immunol. Res.* **2017**, *5*, 376–384. [[CrossRef](#)]
146. Zarling, A.L.; Polefrone, J.M.; Evans, A.M.; Mikesch, L.M.; Shabanowitz, J.; Lewis, S.T.; Engelhard, V.H.; Hunt, D.F. Identification of class I MHC-associated phosphopeptides as targets for cancer immunotherapy. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 14889–14894. [[CrossRef](#)]
147. Malaker, S.A.; Ferracane, M.J. Mass Spectrometric Identification and Molecular Modeling of Glycopeptides Presented by MHC Class I and II Processing Pathways. *Methods Mol. Biol.* **2019**, *2024*, 269–285. [[CrossRef](#)]
148. Jaffe, J.D.; Berg, H.C.; Church, G.M. Proteogenomic mapping as a complementary method to perform genome annotation. *Proteomics* **2004**, *4*, 59–77. [[CrossRef](#)]
149. Kanaseki, T.; Tokita, S.; Torigoe, T. Proteogenomic discovery of cancer antigens: Neoantigens and beyond. *Pathol. Int.* **2019**, *69*, 511–518. [[CrossRef](#)]
150. Bassani-Sternberg, M.; Braunlein, E.; Klar, R.; Engleitner, T.; Sinitcyn, P.; Audehm, S.; Straub, M.; Weber, J.; Slotta-Huspenina, J.; Specht, K.; et al. Direct identification of clinically relevant neoepitopes presented on native human melanoma tissue by mass spectrometry. *Nat. Commun.* **2016**, *7*, 13404. [[CrossRef](#)]
151. Khodadoust, M.S.; Olsson, N.; Wagar, L.E.; Haabeth, O.A.; Chen, B.; Swaminathan, K.; Rawson, K.; Liu, C.L.; Steiner, D.; Lund, P.; et al. Antigen presentation profiling reveals recognition of lymphoma immunoglobulin neoantigens. *Nature* **2017**, *543*, 723–727. [[CrossRef](#)]
152. Creech, A.L.; Ting, Y.S.; Goulding, S.P.; Sauld, J.F.K.; Barthelme, D.; Rooney, M.S.; Addona, T.A.; Abelin, J.G. The Role of Mass Spectrometry and Proteogenomics in the Advancement of HLA Epitope Prediction. *Proteomics* **2018**, *18*, e1700259. [[CrossRef](#)]
153. Chong, C.; Muller, M.; Pak, H.; Harnett, D.; Huber, F.; Grun, D.; Leleu, M.; Auger, A.; Arnaud, M.; Stevenson, B.J.; et al. Integrated proteogenomic deep sequencing and analytics accurately identify non-canonical peptides in tumor immunopeptidomes. *Nat. Commun.* **2020**, *11*, 1293. [[CrossRef](#)] [[PubMed](#)]
154. Kalaora, S.; Barnea, E.; Merhavi-Shoham, E.; Qutob, N.; Teer, J.K.; Shimony, N.; Schachter, J.; Rosenberg, S.A.; Besser, M.J.; Admon, A.; et al. Use of HLA peptidomics and whole exome sequencing to identify human immunogenic neo-antigens. *Oncotarget* **2016**, *7*, 5110–5117. [[CrossRef](#)] [[PubMed](#)]
155. Kanaseki, T.; Torigoe, T. Proteogenomics: Advances in cancer antigen research. *Immunol. Med.* **2019**, *42*, 65–70. [[CrossRef](#)] [[PubMed](#)]
156. Calis, J.J.A.; Reinink, P.; Keller, C.; Kloetzel, P.M.; Keşmir, C. Role of peptide processing predictions in T cell epitope identification: Contribution of different prediction programs. *Immunogenetics* **2015**, *67*, 85–93. [[CrossRef](#)] [[PubMed](#)]
157. Vigneron, N.; Ferrari, V.; Van den Eynde, B.J.; Cresswell, P.; Leonhardt, R.M. Cytosolic Processing Governs TAP-Independent Presentation of a Critical Melanoma Antigen. *J. Immunol.* **2018**, *201*, 1875–1888. [[CrossRef](#)] [[PubMed](#)]

158. Uebel, S.; Tampé, R. Specificity of the proteasome and the TAP transporter. *Curr. Opin. Immunol.* **1999**, *11*, 203–208. [[CrossRef](#)]
159. Mo, X.Y.; Cascio, P.; Lemerise, K.; Goldberg, A.L.; Rock, K. Distinct proteolytic processes generate the C and N termini of MHC class I-binding peptides. *J. Immunol.* **1999**, *163*, 5851–5859. [[PubMed](#)]
160. Craiu, A.; Akopian, T.; Goldberg, A.; Rock, K.L. Two distinct proteolytic processes in the generation of a major histocompatibility complex class I-presented peptide. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 10850–10855. [[CrossRef](#)]
161. Nielsen, M.; Lundegaard, C.; Lund, O.; Keşmir, C. The role of the proteasome in generating cytotoxic T-cell epitopes: Insights obtained from improved predictions of proteasomal cleavage. *Immunogenetics* **2005**, *57*, 33–41. [[CrossRef](#)]
162. Keşmir, C.; Nussbaum, A.K.; Schild, H.; Detours, V.; Brunak, S. Prediction of proteasome cleavage motifs by neural networks. *Protein Eng.* **2002**, *15*, 287–296. [[CrossRef](#)]
163. Mishto, M.; Liepe, J. *Post-Translational Peptide Splicing and T Cell Responses*; Elsevier Ltd: Amsterdam, The Netherlands, 2017; Volume 38, pp. 904–915.
164. Mylonas, R.; Beer, I.; Iseli, C.; Chong, C.; Pak, H.S.; Gfeller, D.; Coukos, G.; Xenarios, I.; Müller, M.; Bassani-Sternberg, M. Estimating the contribution of proteasomal spliced peptides to the HLA-I ligandome. *Mol. Cell. Proteom.* **2018**, *17*, 2347–2357. [[CrossRef](#)] [[PubMed](#)]
165. Brusic, V.; van Endert, P.; Zeleznikow, J.; Daniel, S.; Hammer, J.; Petrovsky, N. A neural network model approach to the study of human TAP transporter. *Silico Biol.* **1999**, *1*, 109–121.
166. Peters, B.; Bulik, S.; Tampe, R.; van Endert, P.M.; Holzhütter, H.-G. Identifying MHC Class I Epitopes by Predicting the TAP Transport Efficiency of Epitope Precursors. *J. Immunol.* **2003**, *171*, 1741–1749. [[CrossRef](#)] [[PubMed](#)]
167. Weinzierl, A.O.; Rudolf, D.; Hillen, N.; Tenzer, S.; van Endert, P.; Schild, H.; Rammensee, H.G.; Stevanovic, S. Features of TAP-independent MHC class I ligands revealed by quantitative mass spectrometry. *Eur. J. Immunol.* **2008**, *38*, 1503–1510. [[CrossRef](#)]
168. Del Val, M.; Anton, L.C.; Ramos, M.; Munoz-Abad, V.; Campos-Sanchez, E. Endogenous TAP-independent MHC-I antigen presentation: Not just the ER lumen. *Curr. Opin. Immunol.* **2020**, *64*, 9–14. [[CrossRef](#)]
169. Robinson, J.; Barker, D.J.; Georgiou, X.; Cooper, M.A.; Fliceck, P.; Marsh, S.G.E. IPD-IMGT/HLA Database. *Nucleic Acids Res.* **2020**, *48*, D948–D955. [[CrossRef](#)]
170. Sidney, J.; Peters, B.; Frahm, N.; Brander, C.; Sette, A. HLA class I supertypes: A revised and updated classification. *BMC Immunol.* **2008**, *9*, 1. [[CrossRef](#)]
171. Sidney, J.; Southwood, S.; Mann, D.L.; Fernandez-Vina, M.A.; Newman, M.J.; Sette, A. Majority of peptides binding hla-a*0201 with high affinity crossreact with other a2-supertype molecules. *Hum. Immunol.* **2001**, *62*, 1200–1216. [[CrossRef](#)]
172. Gfeller, D.; Bassani-Sternberg, M. Predicting Antigen Presentation-What Could We Learn From a Million Peptides? *Front. Immunol.* **2018**, *9*, 1716. [[CrossRef](#)]
173. Shao, W.; Caron, E.; Pedrioli, P.; Aebersold, R. The SystemMHC Atlas: A Computational Pipeline, a Website, and a Data Repository for Immuno-peptidomic Analyses. *Methods Mol. Biol.* **2020**, *2120*, 173–181. [[CrossRef](#)]
174. Vizcaino, J.A.; Kubiniok, P.; Kovalchik, K.A.; Ma, Q.; Duquette, J.D.; Mongrain, I.; Deutsch, E.W.; Peters, B.; Sette, A.; Sirois, I.; et al. The Human Immuno-peptidome Project: A Roadmap to Predict and Treat Immune Diseases. *Mol. Cell. Proteom.* **2020**, *19*, 31–49. [[CrossRef](#)] [[PubMed](#)]
175. Andreatta, M.; Lund, O.; Nielsen, M. Simultaneous alignment and clustering of peptide data using a Gibbs sampling approach. *Bioinformatics* **2013**, *29*, 8–14. [[CrossRef](#)] [[PubMed](#)]
176. Andreatta, M.; Alvarez, B.; Nielsen, M. GibbsCluster: Unsupervised clustering and alignment of peptide sequences. *Nucleic Acids Res.* **2017**, *45*, W458–W463. [[CrossRef](#)] [[PubMed](#)]
177. Alvarez, B.; Reynisson, B.; Barra, C.; Buus, S.; Ternette, N.; Connelley, T.; Andreatta, M.; Nielsen, M. NNAlign-MA; MHC peptidome deconvolution for accurate MHC binding motif characterization and improved t-cell epitope predictions. *Mol. Cell. Proteom.* **2019**, *18*, 2459–2477. [[CrossRef](#)]
178. Bassani-Sternberg, M.; Gfeller, D. Unsupervised HLA Peptidome Deconvolution Improves Ligand Prediction Accuracy and Predicts Cooperative Effects in Peptide–HLA Interactions. *J. Immunol.* **2016**, *197*, 2492–2499. [[CrossRef](#)]

179. Andreatta, M.; Nicastrì, A.; Peng, X.; Hancock, G.; Dorrell, L.; Ternette, N.; Nielsen, M. MS-Rescue: A Computational Pipeline to Increase the Quality and Yield of Immunopeptidomics Experiments. *Proteomics* **2019**, *19*, e1800357. [[CrossRef](#)]
180. Andreatta, M.; Nielsen, M. Gapped sequence alignment using artificial neural networks: Application to the MHC class I system. *Bioinformatics* **2016**, *32*, 511–517. [[CrossRef](#)]
181. O'Donnell, T.J.; Rubinsteyn, A.; Bonsack, M.; Riemer, A.B.; Laserson, U.; Hammerbacher, J. MHCflurry: Open-Source Class I MHC Binding Affinity Prediction. *Cell Syst.* **2018**, *7*, 129–132.e4. [[CrossRef](#)]
182. Nielsen, M.; Lundegaard, C.; Blicher, T.; Lamberth, K.; Harndahl, M.; Justesen, S.; Røder, G.; Peters, B.; Sette, A.; Lund, O.; et al. NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and -B locus protein of known sequence. *PLoS ONE* **2007**. [[CrossRef](#)]
183. Jurtz, V.; Paul, S.; Andreatta, M.; Marcatili, P.; Peters, B.; Nielsen, M. NetMHCpan-4.0: Improved Peptide–MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. *J. Immunol.* **2017**, *199*, 3360–3368. [[CrossRef](#)]
184. Liu, G.; Li, D.; Li, Z.; Qiu, S.; Li, W.; Chao, C.C.; Yang, N.; Li, H.; Cheng, Z.; Song, X.; et al. PSSMHCPan: A novel PSSM-based software for predicting class I peptide–HLA binding affinity. *GigaScience* **2017**. [[CrossRef](#)]
185. Zhao, T.; Cheng, L.; Zang, T.; Hu, Y. Peptide–Major Histocompatibility Complex Class I Binding Prediction Based on Deep Learning with Novel Feature. *Front. Genet.* **2019**, *10*, 1191. [[CrossRef](#)] [[PubMed](#)]
186. Zeng, H.; Gifford, D.K. DeepLigand: Accurate prediction of MHC class I ligands using peptide embedding. *Bioinformatics* **2019**, *35*, i278–i283. [[CrossRef](#)] [[PubMed](#)]
187. Zeng, H.; Gifford, D.K. Quantification of Uncertainty in Peptide–MHC Binding Prediction Improves High-Affinity Peptide Selection for Therapeutic Design. *Cell Syst.* **2019**, *9*, 159–166.e3. [[CrossRef](#)]
188. Hu, Y.; Wang, Z.; Hu, H.; Wan, F.; Chen, L.; Xiong, Y.; Wang, X.; Zhao, D.; Huang, W.; Zeng, J. ACME: Pan-specific peptide–MHC class I binding prediction through attention-based deep neural networks. *Bioinformatics* **2019**, *35*, 4946–4954. [[CrossRef](#)] [[PubMed](#)]
189. Wu, J.; Wang, W.; Zhang, J.; Zhou, B.; Zhao, W.; Su, Z.; Gu, X.; Wu, J.; Zhou, Z.; Chen, S. DeepHLApan: A Deep Learning Approach for Neoantigen Prediction Considering Both HLA–Peptide Binding and Immunogenicity. *Front. Immunol.* **2019**, *10*. [[CrossRef](#)] [[PubMed](#)]
190. Rigo, M.M.; Antunes, D.A.; Vaz de Freitas, M.; Fabiano de Almeida Mendes, M.; Meira, L.; Sinigaglia, M.; Vieira, G.F. DockTope: A Web-based tool for automated pMHC-I modelling. *Sci. Rep.* **2015**, *5*, 18413. [[CrossRef](#)]
191. Antunes, D.A.; Devaurs, D.; Moll, M.; Lizee, G.; Kavraki, L.E. General Prediction of Peptide–MHC Binding Modes Using Incremental Docking: A Proof of Concept. *Sci. Rep.* **2018**, *8*, 4327. [[CrossRef](#)]
192. Kyeong, H.H.; Choi, Y.; Kim, H.S. GradDock: Rapid simulation and tailored ranking functions for peptide–MHC Class I docking. *Bioinformatics* **2018**, *34*, 469–476. [[CrossRef](#)]
193. Abella, J.R.; Antunes, D.A.; Clementi, C.; Kavraki, L.E. APE-Gen: A Fast Method for Generating Ensembles of Bound Peptide–MHC Conformations. *Molecules* **2019**, *24*, 881. [[CrossRef](#)]
194. Larsen, M.V.; Lundegaard, C.; Lamberth, K.; Buus, S.; Brunak, S.; Lund, O.; Nielsen, M. An integrative approach to CTL epitope prediction: A combined algorithm integrating MHC class I binding, TAP transport efficiency, and proteasomal cleavage predictions. *Eur. J. Immunol.* **2005**, *35*, 2295–2303. [[CrossRef](#)] [[PubMed](#)]
195. Stranzl, T.; Larsen, M.V.; Lundegaard, C.; Nielsen, M. NetCTLpan: Pan-specific MHC class I pathway epitope predictions. *Immunogenetics* **2010**, *62*, 357–368. [[CrossRef](#)] [[PubMed](#)]
196. Karosiene, E.; Lundegaard, C.; Lund, O.; Nielsen, M. NetMHCcons: A consensus method for the major histocompatibility complex class I predictions. *Immunogenetics* **2012**, *64*, 177–186. [[CrossRef](#)] [[PubMed](#)]
197. Jørgensen, K.W.; Rasmussen, M.; Buus, S.; Nielsen, M. NetMHCstab—Predicting stability of peptide–MHC-I complexes; impacts for cytotoxic T lymphocyte epitope discovery. *Immunology* **2014**, *141*, 18–26. [[CrossRef](#)]
198. Rasmussen, M.; Fenoy, E.; Harndahl, M.; Kristensen, A.B.; Nielsen, I.K.; Nielsen, M.; Buus, S. Pan-Specific Prediction of Peptide–MHC Class I Complex Stability, a Correlate of T Cell Immunogenicity. *J. Immunol.* **2016**, *197*, 1517–1524. [[CrossRef](#)]
199. Kawashima, S.; Pokarowski, P.; Pokarowska, M.; Kolinski, A.; Katayama, T.; Kanehisa, M. AAindex: Amino acid index database, progress report 2008. *Nucleic Acids Res.* **2008**, *36*, D202–D205. [[CrossRef](#)] [[PubMed](#)]
200. Tung, C.W.; Ho, S.Y. POPI: Predicting immunogenicity of MHC class I binding peptides by mining informative physicochemical properties. *Bioinformatics* **2007**, *23*, 942–949. [[CrossRef](#)]

201. Tung, C.W.; Ziehm, M.; Kämper, A.; Kohlbacher, O.; Ho, S.Y. POPISK: T-cell reactivity prediction using support vector machines and string kernels. *BMC Bioinform.* **2011**, *12*, 446. [[CrossRef](#)] [[PubMed](#)]
202. Calis, J.J.A.; Maybeno, M.; Greenbaum, J.A.; Weiskopf, D.; De Silva, A.D.; Sette, A.; Keşmir, C.; Peters, B. Properties of MHC Class I Presented Peptides That Enhance Immunogenicity. *PLoS Comput. Biol.* **2013**, *9*, e1003266. [[CrossRef](#)] [[PubMed](#)]
203. Saethang, T.; Hirose, O.; Kimkong, I.; Tran, V.A.; Dang, X.T.; Nguyen, L.A.T.; Le, T.K.T.; Kubo, M.; Yamada, Y.; Satou, K. PAAQD: Predicting immunogenicity of MHC class I binding peptides using amino acid pairwise contact potentials and quantum topological molecular similarity descriptors. *J. Immunol. Methods* **2013**, *387*, 293–302. [[CrossRef](#)]
204. Zhang, W.; Niu, Y.; Zou, H.; Luo, L.; Liu, Q.; Wu, W. Accurate prediction of immunogenic T-cell epitopes from epitope sequences using the genetic algorithm-based ensemble learning. *PLoS ONE* **2015**, *10*, e0128194. [[CrossRef](#)]
205. Trolle, T.; Nielsen, M. NetTepi: An integrated method for the prediction of T cell epitopes. *Immunogenetics* **2014**, *66*, 449–456. [[CrossRef](#)]
206. Bentzen, A.K.; Hadrup, S.R. Evolution of MHC-based technologies used for detection of antigen-responsive T cells. *Cancer Immunol. Immunother.* **2017**, *66*, 657–666. [[CrossRef](#)]
207. Tickotsky, N.; Sagiv, T.; Prilusky, J.; Shifrut, E.; Friedman, N. McPAS-TCR: A manually curated catalogue of pathology-associated T cell receptor sequences. *Bioinformatics* **2017**, *33*, 2924–2929. [[CrossRef](#)]
208. Shugay, M.; Bagaev, D.V.; Zvyagin, I.V.; Vroomans, R.M.; Crawford, J.C.; Dolton, G.; Komech, E.A.; Sycheva, A.L.; Koneva, A.E.; Egorov, E.S.; et al. VDJdb: A curated database of T-cell receptor sequences with known antigen specificity. *Nucleic Acids Res.* **2018**, *46*, D419–D427. [[CrossRef](#)]
209. Bagaev, D.V.; Vroomans, R.M.A.; Samir, J.; Stervbo, U.; Rius, C.; Dolton, G.; Greenshields-Watson, A.; Attaf, M.; Egorov, E.S.; Zvyagin, I.V.; et al. VDJdb in 2019: Database extension, new analysis infrastructure and a T-cell receptor motif compendium. *Nucleic Acids Res.* **2020**, *48*, D1057–D1062. [[CrossRef](#)] [[PubMed](#)]
210. Vita, R.; Mahajan, S.; Overton, J.A.; Dhanda, S.K.; Martini, S.; Cantrell, J.R.; Wheeler, D.K.; Sette, A.; Peters, B. The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Res.* **2019**, *47*, D339–D343. [[CrossRef](#)] [[PubMed](#)]
211. Leem, J.; De Oliveira, S.H.P.; Krawczyk, K.; Deane, C.M. STCRDab: The structural T-cell receptor database. *Nucleic Acids Res.* **2018**, *46*, D406–D412. [[CrossRef](#)] [[PubMed](#)]
212. Borrmann, T.; Ciments, J.; Cosiano, M.; Purcaro, M.; Pierce, B.G.; Baker, B.M.; Weng, Z. ATLAS: A database linking binding affinities with structures for wild-type and mutant TCR-pMHC complexes. *Proteins: Struct. Funct. Bioinform.* **2017**, *85*, 908–916. [[CrossRef](#)] [[PubMed](#)]
213. Gielis, S.; Moris, P.; Bittremieux, W.; De Neuter, N.; Ogunjimi, B.; Laukens, K.; Meysman, P. Detection of Enriched T Cell Epitope Specificity in Full T Cell Receptor Sequence Repertoires. *Front. Immunol.* **2019**, *10*, 2820. [[CrossRef](#)] [[PubMed](#)]
214. Ogishi, M.; Yotsuyanagi, H. Quantitative prediction of the landscape of T cell epitope immunogenicity in sequence space. *Front. Immunol.* **2019**, *10*, 827. [[CrossRef](#)] [[PubMed](#)]
215. Jensen, K.K.; Rantos, V.; Jappe, E.C.; Olsen, T.H.; Jespersen, M.C.; Jurtz, V.; Jessen, L.E.; Lanzarotti, E.; Mahajan, S.; Peters, B.; et al. TCRpMHCmodels: Structural modelling of TCR-pMHC class I complexes. *Sci. Rep.* **2019**, *9*, 1–12. [[CrossRef](#)] [[PubMed](#)]
216. Klausen, M.S.; Anderson, M.V.; Jespersen, M.C.; Nielsen, M.; Marcatili, P. LYRA, a webserver for lymphocyte receptor structural modeling. *Nucleic Acids Res.* **2015**, *43*, W349–W355. [[CrossRef](#)] [[PubMed](#)]

