Proteomics
www.proteomics-journal.com

The Role of Mass Spectrometry and Proteogenomics in the Advancement of HLA Epitope Prediction

*Amanda L. Creech, Ying S. Ting, Scott P. Goulding, John F.K. Sauld, Dominik Barthelme, Michael S. Rooney, Terri A. Addona, and Jennifer G. Abelin**

A challenge in developing personalized cancer immunotherapies is the prediction of putative cancer-specific antigens. Currently, predictive algorithms are used to infer binding of peptides to human leukocyte antigen (HLA) heterodimers to aid in the selection of putative epitope targets. One drawback of current epitope prediction algorithms is that they are trained on datasets containing biochemical HLA-peptide binding data that may not completely capture the rules associated with endogenous processing and presentation. The field of MS has made great improvements in instrumentation speed and sensitivity, chromatographic resolution, and proteogenomic database search strategies to facilitate the identification of HLA-ligands from a variety of cell types and tumor tissues. As such, these advances have enabled MS profiling of HLA-binding peptides to be a tractable, orthogonal approach to lower throughput biochemical assays for generating comprehensive datasets to train epitope prediction algorithms. In this review, we will highlight the progress made in the field of HLA-ligand profiling enabled by MS and its impact on current and future epitope prediction strategies.

1. Introduction

Cancer immunotherapy has become an attractive alternative to less specific treatment options such as chemotherapy and radiation.[1–7] Within this realm, there are several avenues that harness the capabilities of the immune system to both identify and eliminate cancer cells. The broadest treatments are antibodies that modulate the regulatory pathways of cytotoxic T lymphocytes (CTLs) by inhibiting immune checkpoint molecules such as programmed cell death protein 1, programmed cell death protein 1 ligand, and cytotoxic T-lymphocyte-associated protein 4.[4] Another class of immunotherapies are more personalized as they target endogenously processed and presented antigens such as tumor associated antigens (TAAs) that are often overexpressed self-antigens, and neoantigens, which are peptide antigens containing tumor-specific sequence variants (i.e., mutations).[1,2] Interestingly, improved outcomes among patients treated with checkpoint inhibitors have been linked to tumors with a high mutational load and increased numbers of predicted neoantigens.[8–12] Additionally, T cell responses against tumor neoantigens have been observed after immune checkpoint blockage with ipilimumab[13] and adoptive transfer T cell therapy[14] in the context of metastatic melanoma. Therefore, the accurate prediction of neoantigens has become a focus in the development of personalized cancer immunotherapeutics.

Personalized cancer vaccines are currently under development to further augment the adaptive immune response against tumor-specific neoantigens.[1,2] Historically, neoantigens were discovered by molecular cloning and laborious in vitro immune screening using cDNA libraries[15]. Currently, the combination of next generation sequencing and computational methods are used to identify putative neoantigens from a patient's tumor genome or transcriptome (**Figure 1**A).[1,2,16–21] These candidate neoantigens can then be manufactured into a vaccine and delivered to patients, aiming to induce anti-tumor responses (Figure 1B). More details regarding personalized neoantigen peptide-based vaccine modalities, production, and their use in cancer immunotherapy can be found in a review article recently published by Aldous et al.[22] In preclinical models, studies have shown that synthetic peptide vaccines against cancer neoantigens induce $CD8^+$ T cell responses and protect against tumor outgrowth.[6,17] Moreover, two recent clinical studies have demonstrated promising results using personalized neoantigen vaccines in small cohorts of melanoma patients.[1,2] After vaccination with neoantigens delivered by either a peptide-based[1] or RNA-based[2] vehicle, both $CD8^+$ and $CD4^+$ T cell responses were detected against vaccinated neoantigens, and protection against metastatic progression was observed. These successes further highlight the importance of developing personalized neoantigen vaccines for cancer immunotherapy.

The concept of producing personalized cancer vaccines appears straightforward, but the accurate prediction of tumor-specific peptide antigens for every patient remains a major obstacle. In humans, endogenous proteins are processed into

A. L. Creech, Dr. Y. S. Ting, Dr. S. P. Goulding, J. F. K. Sauld, Dr. D. Barthelme, Dr. M. S. Rooney, Dr. T. A. Addona, Dr. J. G. Abelin
Neon Therapeutics
Cambridge, MA, USA
E-mail: jabelin@neontherapeutics.com

The copyright line for this article was changed on 20 June 2018 after original online publication.

DOI: 10.1002/pmic.201700259

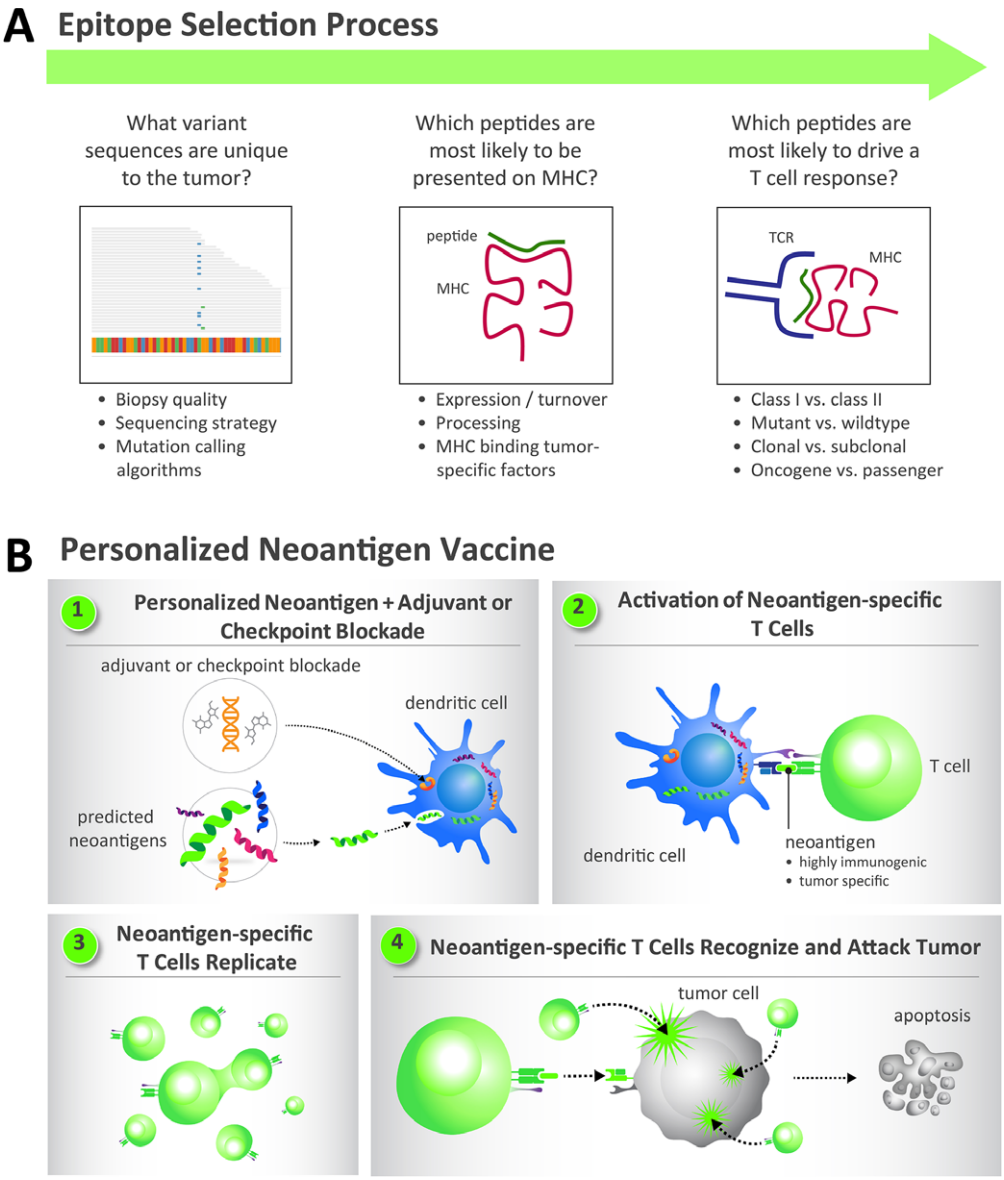

Figure 1. Development and function of personalized neoantigen vaccines. A) Overview of the epitope selection process in the development of personalized cancer vaccines. Tumor-specific variants are first identified using whole exome or transcriptome sequencing from the patient's tumor biopsy. The sequence variant containing peptides are ranked by an epitope selection pipeline. The resulting putative neoantigen epitopes could be further prioritized based on T cell response prediction or measurement. Predicted neoantigens containing high confident epitopes are selected for vaccine production. B) Schematic depicting the function of personalized neoantigen vaccines. Predicted neoantigens are administered in combination with adjuvants and/or checkpoint inhibitors to increase the patient's immune response. Upon vaccination, dendritic cells, and/or other antigen presenting cells uptake the neoantigens delivered by a vehicle, such as DNA, RNA, or long peptide form (box 1). After the dendritic cells or other antigen presenting cells process the neoantigens, they can present the resulting epitopes to naïve T cells, which are subsequently activated to become cytotoxic (box 2). These neoantigen-specific cytotoxic T cells replicate and circulate in the peripheral vascular system (box 3). When these T cells encounter tumor cells presenting the corresponding epitopes, they can identify and eliminate them by cytotoxicity and other immune functions (box 4).

peptides by the proteasome, cytosolic and endosomal/lysosomal proteases, and peptidases and presented by two classes of cell surface proteins encoded by the major histocompatibility complex (MHC). These cell surface proteins are referred to as human leukocyte antigens (HLA class I and class II) in humans and the group of peptides that bind them and elicit immune responses are termed HLA epitopes. HLA epitopes are one key component for the detection of danger signals by the immune system. Circulating $CD8^+$ T cells recognize class I MHC (HLA-A–C) epitopes derived from endogenous processing pathways and displayed on

almost all nucleated cells. $CD4^+$ T cells recognize class II MHC (HLA-DR, -DQ, and -DP) epitopes displayed on antigen presenting cells, such as dendritic cells and macrophages.

The genes coding for HLA proteins are highly polymorphic, with more than 12 000 class I and 4000 class II allele variants identified across the human population.[23] From maternal and paternal HLA haplotypes, an individual can inherit different alleles for each of the class I and class II HLA loci. Class I HLA molecules are heterodimers made up of an α-chain, encoded by class I HLA genes, and a β-2-microglobulin. Class II HLA molecules are α- and β-chain heterodimers, both encoded by the class II HLA genes. Because of the α- and β-chain pairing combinations, the population of HLA heterodimers is highly complex. In addition, each HLA heterodimer is estimated to bind thousands of peptides with allele-specific binding preferences.[24–28] Understanding the binding preferences of every HLA heterodimer is key to successfully predicting which neoantigens are likely to elicit tumor-specific T cell responses.

Biochemical methods such as stability and binding assays that leverage synthetic peptides bound to purified recombinant MHC molecules, are most commonly used to characterize HLA allele-specific peptide-binding motifs. However, these approaches are low throughput, biased by peptide (target) selection, and unable to capture the potential rules that govern endogenous HLA-ligand processing and presentation.[26–31] In contrast, LC coupled to MS/MS has been used to directly identify HLA epitopes and naturally processed ligands since the early 1990s.[25] LC–MS/MS approaches are high throughput, unbiased by target selection, and adaptive to experiments designed to characterize cellular HLA-ligand processing and presentation mechanisms.[24,31–33] Considering the high complexities of HLA polymorphisms and the repertoire of HLA-ligands, LC–MS/MS has been and will continue to be a major force to advance the development of personalized neoantigen vaccines. In this review, we will highlight (1) the current state of HLA-ligand profiling, (2) cancer TAA and neoantigen discovery enabled by LC–MS/MS, (3) the implementation of HLA-ligand profiling in epitope and neoantigen prediction strategies, and (4) current limitations and future applications of MS-based HLA epitope prediction methods in the development of personalized cancer immunotherapies.

1.1. Profiling of HLA-Ligands Using MS

Direct identification of endogenously processed and presented HLA-ligands by LC–MS/MS was pioneered by Hunt and colleagues in the early 1990s.[25,34] Concurrently, Rammensee and colleagues used HPLC and Edman degradation to sequence HLA ligands[35]. By using monoclonal antibodies specific to HLA heterodimers, HLA-peptide complexes can be immunoprecipitated from cellular lysate or culture media if HLA complexes are engineered to lack transmembrane regions (soluble HLA). Peptide ligands are typically dissociated from the HLA heterodimers by treatment with acid. The eluted peptide sample is further desalted or processed through a filter with a molecular weight cutoff so only HLA peptide ligands are introduced to the mass spectrometer. The exclusion of denatured HLA heterodimers can also be achieved by eluting only the HLA-bound peptides with a lower percentage of organic solvent from a RP material. HLA-peptide ligands can also be isolated from cell media via a mild-acid elution from the cell surface and desalted prior to LC–MS/MS analysis.

In early HLA-ligand profiling studies, only tens of endogenously processed and presented ligands could be identified by LC–MS/MS using immunoprecipitation techniques. Today, the numbers of HLA-binding peptides identified from a single experiment have increased by orders of magnitude. Studies now commonly report thousands of endogenously processed ligands identified directly from either cell lines or patient material.[24,32,33,36–39] The acceleration of HLA-ligand identification can be attributed to improvements in chromatography (e.g., high pressure LC systems with reproducible nanoliter flow rates and columns with higher resolution), increased sensitivity and speed of mass spectrometry instrumentation, as well as the increased confidence of peptide assignments resulting from high resolution MS/MS data. Other techniques, such as combining various fragmentation methods, have also been shown to increase the identification rate for specific populations of HLA-ligands, like those that have a basic amino acid residue in their anchor positions that interact with the peptide-binding groove.[37,40] HLA-ligand identification using MS will continue to improve as upstream sample processing techniques are enhanced, instrumentation scan rates and sensitivity advance, and data acquisition methods become more sophisticated.

The endogenous processing of HLA-ligands is a complex procedure and involves a variety of enzymes that are not all well characterized.[41] Thus, most studies identify HLA-ligands from LC–MS/MS data using database searching tools with no defined enzyme cleavage specificity (i.e., no-enzyme searches). Without an enzyme specificity, the database search space is at least an order of magnitude larger than that of a tryptic database search.[42,43] However, modern database searching algorithms can leverage high resolution data to improve their performance with no-enzyme searches.[42–44] While details of endogenous processing of ligands could be captured and potentially used to improve the confidence of peptide identifications,[33,41] such information is typically left out of the peptide identification process. Therefore, it is critical to use a stringent false discovery rate (FDR) estimation cutoff to control for false identifications and ensure high quality data is reported.[45] As such, most HLA-ligands profiling studies that utilize no-enzyme searches have reported identified ligands using a FDR threshold of 1% or less.[24,32,33,36,37,46] This FDR threshold does not prevent contaminant peptides (i.e., not HLA-bound) from being reported in HLA-ligandome datasets. One of the largest contributors to contaminant peptides in these datasets are tryptic peptides that are retained on LC columns after conditioning. However, this subset of contaminant peptides has been shown to be an average of 4% in a large mono-allelic dataset, and can be easily removed by using peptide motif clustering tools or by defining a tryptic peptide database as contaminants.[47] The implementation of de novo sequencing strategies[43,48] and HLA-allele specific database digestion parameters have also been used to identify spectra that were not identified using the no-enzyme search approach.[33] Moreover, proteogenomic and targeted database strategies have been applied and reported to increase rates of HLA-ligand identification from LC–MS/MS data.[36,39,44,49] However, caution should be used

with these peptide identification methods to ensure that FDR is accurately estimated and reported.[50] Future improvements to proteogenomic, no enzyme, and de novo peptide identification strategies will greatly boost confidence in HLA-ligands identified using MS.

1.2. Discovery of Neoantigen and Post-translationally Modified TAA Enabled by LC-MS/MS

LC–MS/MS profiling of endogenously processed and presented HLA-ligands has successfully been applied to multiple cell lines and patient-derived material. An early application of LC–MS/MS to directly characterize immunogenic epitopes led to the discovery of a melanoma-restricted tyrosinase epitope that was shown to be recognized by CTLs from five melanoma patients.[34] Since then, several research groups have used MS to both profile class I and II HLA-ligands and even discover neoantigens, although at very low numbers, presented by various types of cancer cells.[34,38,49,51–56] For example, Yadav and colleagues applied whole exome sequencing to murine cell lines and used MS to profile endogenously processed and presented ligands and detected neoantigens.[19] In a murine sarcoma model, Gubin and colleagues investigated cancer neoantigens that are targets of T cells activated by immune checkpoint blockade, and confirmed neoantigen presentation by identifying peptides presented on tumor H-2K^b complexes.[6,17]

A few groups, including Bassani-Sternberg et al.[36] and Kalaora et al.[49], have demonstrated that neoantigens, along with a subset of well-known melanoma TAA's, can be directly identified from patient tumor material using MS. Bassani-Sternberg and colleagues profiled HLA-ligands from multiple melanoma patients to generate a dataset of >95 000 endogenously processed tumor ligands.[36] Using a proteogenomic strategy, the group was able to directly detect 11 neoantigens using LC–MS/MS from five of their patients where whole exome sequencing was available.[36] Kalaora and colleagues applied whole exome sequencing and HLA-peptide LC–MS/MS profiling to a primary cell line established from a patient's melanoma metastases. Using a similar proteogenomic strategy, the group identified two patient-specific neoantigens from the total population of 4956 HLA-peptides identified by LC–MS/MS. One of the LC–MS/MS identified mutated peptides but not its wild-type counterpart, was found to stimulate the patient's tumor infiltrating lymphocytes. Furthermore, Khodadoust and colleagues demonstrated the ability to directly interrogate the HLA-ligandome from a patient with B cell non-Hodgkin lymphoma and discovered neoantigens derived from immunoglobulin variable regions via LC–MS/MS [39]. It is important to note that in some of these examples multiple grams of material were required for LC–MS/MS peptide identification of neoantigens, which currently limits this approach to cancer types where tumor material is abundant. Another important consideration is that tumor tissue is heterogenous as it contains multiple cell types. In fact, melanoma tumors reported in The Cancer Genome Atlas(http://cancergenome.nih.gov/) show an average purity of 67%, and other tumor types likely have varying purities.[57] Therefore, HLA-ligands identified from tumor material represent a mixed population of HLA-ligandomes that may be difficult to assign as tumor specific.

In addition to recognizing neoantigens as immunotherapeutic targets, it has been postulated that aberrant post-translationally modified epitopes associated with cancer can be used to elicit tumor-specific T cell responses.[54,55,58,59] Using MS, identification of HLA ligands containing post-translational modifications such as phosphorylation, methylation, and glycosylation, have been reported.[54,55,58–62] Phosphorylation, which becomes dysregulated in cancer, has been reported in multiple HLA-ligand datasets and is the most well studied in this context.[36,54,55,60,63] For instance, Zarling and colleagues identified endogenously processed and presented phosphorylated HLA class I ligands by using immobilized metal affinity chromatography enrichment[64] and were able to elicit in vitro CTL responses against a subset of phosphorylated epitopes but not their unphosphorylated counterparts.[55,58] Hunt, Cobbold, and colleagues have also reported a subset of immunogenic class I-restricted phosphorylated HLA-epitopes from primary leukemia samples.[54] Additional clinical validation of HLA-ligand profiling by LC–MS/MS and in vivo immunogenicity investigations are likely required to fully understand the tumor specificity of post-translationally modified epitopes and their value as potential immunotherapeutic targets.

1.3. Implementation of HLA Ligand Profiling in Epitope and Neoantigen Prediction

To develop effective personalized immunotherapies, it is imperative that patient-specific neoantigens that can prime the adaptive immune response are accurately predicted and correctly prioritized. Discovery of putative neoantigens typically begins with whole exome sequencing of a matched set of patient tumor- and normal-cell DNA.[1,2] A comparison of these two samples can lead to the identification of tumor-specific somatic mutations. Further analysis of the transcriptome using RNA sequencing (RNA-Seq) can help to ascertain which DNA variants are expressed and confirm variants whose identification is challenging with DNA sequencing alone. There are several types of mutations taken into consideration when identifying putative neoantigens. Single nucleotide variants, which result in a single amino acid change in the subsequent protein, represent a majority of the mutations detected in tumor DNA that reportedly elicit T cell responses.[65] However, this type of mutation represents only one class for which mutation calling algorithms must account. Insertions/deletions (i.e., indels), fusion proteins, mRNA splice variants, and frame shift mutations can also produce unique, candidate neoantigens with predicted binding affinities for HLA heterodimers. Peptides that are erroneously translated from the 5′ and 3′ untranslated regions of mRNA predicted from RNA-Seq data have also been reported as sources of neoantigens.[66–68]

Somatic cancer variants can be detected using whole exome sequencing and translated into FASTA format to produce a patient-specific proteome to be used for spectral proteogenomic searches. During such workflows, tumor purity is an important concern. Although whole exome sequencing-based mutation calling can retain sensitivity at low purities (possibly as low as 5–10% at standard 150× sequencing depth),[69] the purity requirements for LC–MS/MS detection are not well defined and likely

much higher making instrument sensitivity a limiting factor for neoantigen detection. In addition to variants identified at the genetic and transcriptomic levels, recent studies have also shown that the proteasome can produce peptide splice variants during proteasome-mediated degradation of proteins that can be directly identified using LC–MS/MS.[61,70–72] MS and prediction methods for identifying these noncanonical HLA-peptides are emerging but require further validation before being implemented in clinical applications.[73]

For putative neoantigens to be prioritized, they must be able to bind to a patient's set of unique HLA heterodimers. Therefore, it is imperative that the HLA-ligand binding preferences of each expressed HLA allele are known. A vast amount of the HLA-epitope data generated to date resides in the Immune Epitope Database (IEDB), a public database that is funded by the National Institute of Allergy and Infectious Diseases.[28] IEDB has been an invaluable resource for the immunology community as it centralizes data characterizing antibody and T cell epitopes in humans, nonhuman primates, and other animal species that are involved in research of infectious disease, allergy, autoimmunity, and transplant rejection. Accordingly, many of the current neoantigen prediction pipelines either utilize IEDB data to train algorithms, such as machine learning-based approaches, that predict HLA-peptide binding, or use advanced epitope prediction tools that already incorporate IEDB data.[74–76]

NetMHC and NetMHCpan are two examples of state-of-the-art neural network-trained algorithms that perform HLA epitope prediction.[74–77] NetMHC covers HLA-A, -B, and -C alleles for which there are allele-specific data available for training, while NetMHCpan can predict HLA-peptide binding for HLA alleles that may not have well annotated peptide-binding motifs by using data from closely related alleles. IEDB contains >300 000 ligands with approximately half of these representing naturally processed ligands.[78] However, a large proportion of the reported HLA-ligands were also identified using biochemical binding assays that leverage synthetic peptide libraries tested against recombinant HLA heterodimers. These ligands may not be appropriate for neoantigen prediction because HLA-binding alone does not capture the complex processing and presentation rules that govern endogenously processed epitopes. Libraries of potential peptide ligands are also usually selected based on a hypothesis or prior knowledge, which can introduce bias in the collective knowledge deposited into IEDB. One example of peptide library bias is evident in the HLA-ligand lengths selected for biochemical HLA-binding screens. Historical data has shown that the most HLA heterodimers prefer peptide ligands that are nine amino acids in length. However, the generation of large ligand elution datasets via MS has shown that HLA class I heterodimers have varying preferences for ligand lengths.[33,77,79] Furthermore, this MS-based insight has been incorporated into current epitope prediction algorithms and increased their predictive value.[77] Thus, LC–MS/MS profiling of naturally processed HLA-ligands has received a great deal of attention.

Endogenously processed and presented HLA-ligands profiled from cell lines and patient-derived material are commonly multi-allelic, meaning that LC–MS/MS data generated from these samples contain a mixed population of ligands that can bind to one of the multiple simultaneously expressed HLA alleles, as shown in **Figure 2** (top). Multi-allelic datasets require deconvolution to ascertain which peptides bind to the different HLA heterodimers displayed by an individual. Thus, ligands from multi-allelic datasets have to be assigned to their corresponding HLA heterodimers using either (1) binding predictors trained with preexisting data[28] or (2) deconvolution algorithms that leverage overlap across HLA alleles represented in large ligand datasets.[74,80,81] It is important to note that only LC–MS/MS datasets with available HLA typing information can be confidently deconvoluted. In fact, nearly 40% of the naturally processed ligands bound to HLA class I complexes reported from multi-allelic studies in IEDB lack HLA allele-specific assignments either due to the lack of HLA typing information or inability to deconvolute, making it challenging to use this subset of data for allele-specific epitope prediction.[31] In addition, it is difficult to identify peptides bound to rare class I HLA heterodimers and many class II HLA heterodimers because there is not enough annotated data for deconvolution. The multi-allelic data generation approach also limits the discovery of novel binding motifs as deconvolution relies on preexisting knowledge. Though there are caveats to utilizing multi-allelic datasets for allele-specific epitope predictions, they are immensely valuable for determining patterns of ligand presentation that require co-expression of multiple alleles and for validating epitope prediction algorithms.

An orthogonal approach to multi-allelic data generation and subsequent deconvolution is the creation of mono-allelic datasets from which peptide populations presented by a single HLA allele are identified (Figure 2, bottom). One method for generating mono-allelic data utilizes cell lines that are deficient in HLA expression. These cells can be transfected or transduced with single HLA alleles so that ligands can be profiled by LC–MS/MS to generate allele-specific ligand libraries.[32,33,82] For example, Wu and colleagues generated mono-allelic LC–MS/MS data by individually transducing 721.221 B cells with 16 HLA alleles of interest[33] while other research groups have independently utilized C1R cells to profile naturally processed ligands presented by individual HLA alleles including HLA-C and HLA-G.[82] Peptides bound to soluble HLA molecules can also be isolated from cell media and profiled by LC–MS/MS to produce mono-allelic data.[79,83,84] A major advantage of mono-allelic datasets is that they require no deconvolution and enable confident peptide-HLA allele assignments without preexisting data. Mono-allelic approaches also rapidly provide data for HLA alleles that have not been characterized previously—a task that multi-allelic data can do only if enough overlap is present amongst large datasets. Additionally, novel peptide-binding motifs can easily be discovered using mono-allelic systems as no previous knowledge is required for confident HLA-binding assignments. Mono-allelic data can even be leveraged to assign ligands from multi-allelic datasets when deconvolution methods fail to do so.

Although the mono-allelic approach is advantageous for rapidly learning HLA allele-specific peptide-binding motifs, both mono-allelic and multi-allelic LC–MS/MS datasets will likely be beneficial for learning the rules of HLA-ligand processing and presentation, as co-expression of different combinations of HLA alleles may impact the population of naturally processed and presented epitopes.[24,33,84] Rules relating to cleavage patterns, subcellular localization, and source protein attributes can also be better resolved and validated when both data types are combined and then embedded into multivariate

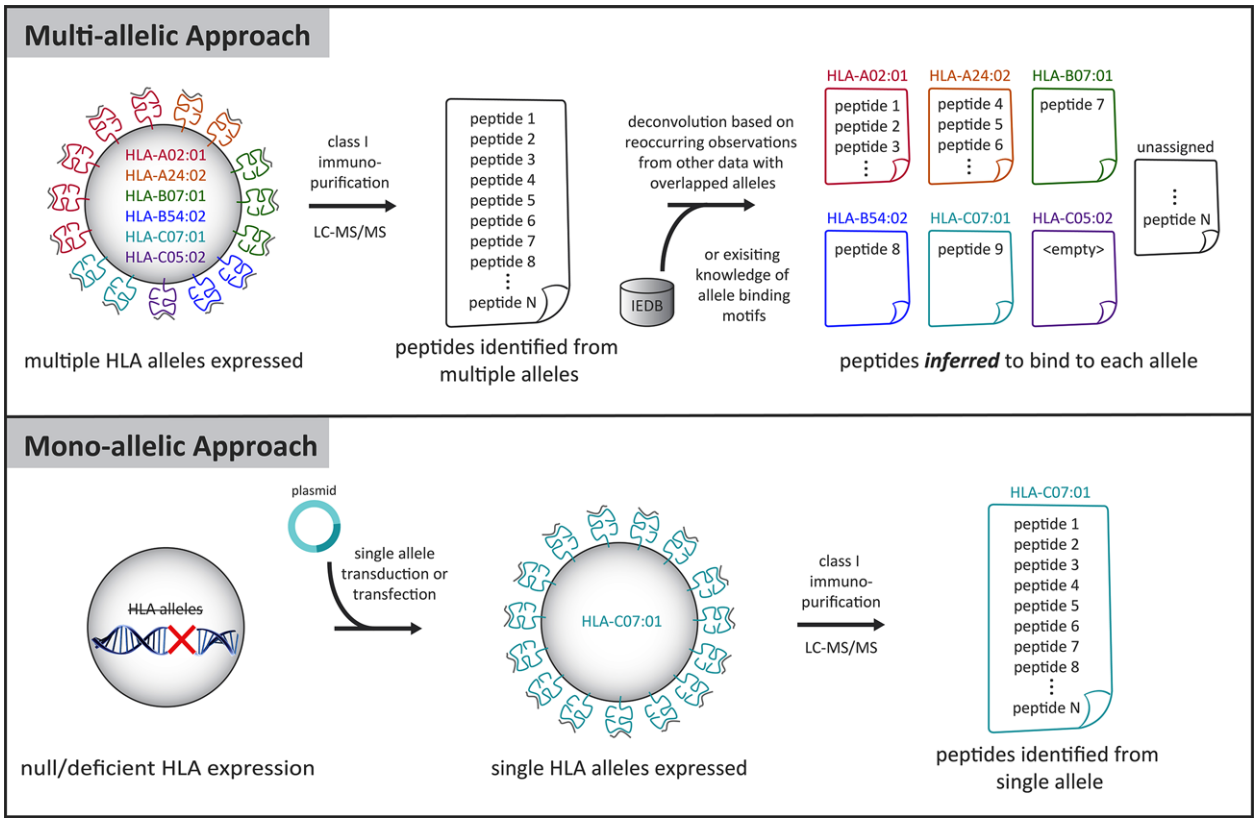

Figure 2. Multi-allelic and mono-allelic approaches in HLA ligand profiling. In a multi-allelic approach, the HLA ligands are co-immunoprecipitated with HLA heterodimers directly from patient material or cell lines (top). Because these cells naturally expressed multiple HLA alleles, peptides identified from such multi-allelic approaches must be deconvoluted to assign binding to a specific HLA heterodimer if the HLA types are known. In a mono-allelic approach, the HLA-ligands are co-immunoprecipitated with HLA heterodimers from cell lines genetically modified for expression of only a single HLA allele (bottom). Thus, peptides identified from mono-allelic approaches do not require deconvolution for HLA heterodimer binding assignments.

epitope prediction algorithms.[24,33,84] Overall, the combination of mono- and multi-allelic datasets will be highly valuable for advancing methods for the prediction of HLA epitopes and neoantigens.

1.4. Current Challenges and Future Perspectives: The Use of MS-Based Epitope Predictions in the Development of Personalized Immunotherapies

With the rise of personalized cancer immunotherapies, much work has been done to explore the endogenously processed ligands presented by diverse HLA heterodimers and learn HLA-specific binding preferences so that stable, and ideally immunogenic, epitopes can be predicted for individual patients. With the recent advances in instrumentation, data acquisition methods, and database searching algorithms, MS has already made enormous contributions towards this effort as evidenced by an increase in the amount of MS data deposited in IEDB.[28,78] Implementation of mono-allelic profiling workflows will likely be crucial for generating data for rare and low expressed HLA alleles, and to aid in deconvoluting multi-allelic datasets, whereas multi-allelic datasets will be valuable for supporting the development of neoantigen prediction algorithms if peptide HLA binding assignments can be made confidently. Because both mono-allelic and multi-allelic data are generated from endogenously processed ligands, future epitope prediction algorithms can only become more accurate as additional MS data are reported from diverse cell types representing less common HLA alleles.

In addition to data generation, MS data dissemination and standardization will be another important contribution to the field of HLA-ligand profiling. Data sharing amongst research groups generating endogenous HLA-ligandome and epitope immunogenicity datasets will play a critical role in improving neoantigen prediction pipelines. The Human Immuno-Peptidome Project (HIPP)[46,85] is already attempting to connect clinicians with research groups producing HLA-ligandome data, as sequencing HLA-ligands by LC–MS/MS is a complimentary and orthogonal approach to whole exome sequencing and current prediction methods. This effort coincides with the Tumor Neoantigen Selection Alliance (https://www.parkerici.org/research_project/tumor-neoantigen-selection-alliance/), which is evaluating the performance and accuracy of current predictive algorithms developed by multiple academic and industry laboratories by confirming the immunogenicity of predicted neoantigens with in vitro validation studies.

While the proteomics community has made great strides to advance our knowledge of HLA-ligand binding and prediction, several key areas of opportunity for improvement remain. Direct identification of neoantigens in tumor tissue, patient derived cell lines, and model systems by MS is the goal, but remains challenging due to observations that neoantigens represent a low percentage of the tumor's HLA ligandome.[6,19,36] Therefore, neoantigen discovery currently relies on whole exome sequencing and predominantly on the predictive power of algorithms that infer HLA-ligand binding. These studies do, however, highlight the clinical potential of MS-based sequencing of naturally processed HLA ligands, especially when combined with other "-omic" technologies such as next generation sequencing, and RNA-Seq. For example, multiple research groups have proposed general pipelines combining technologies to craft personalized immunotherapies across multiple cancer types.[86–90] Such proteogenomic methodologies incorporating MS sequencing to detect and validate immunotherapeutic epitopes have recently shown clinical promise.[91,92]

The large amount of patient material and the timeframe required for patient-specific multi-omic HLA-ligand and neoantigen profiling currently prevents its routine implementation in clinical settings. Tumor biopsy, HLA-typing, whole exome sequencing, and mutation calling can take 1–2 weeks. In parallel, HLA-ligand isolation, MS data acquisition, and data analysis can take anywhere from several days to weeks, while personalized vaccine design and manufacturing can take multiple weeks once neoantigen targets are identified. The combined timeframe for these processes is multiple months, which is currently too long for routine use in the clinic. However, if patients begin treatment with an off the shelf therapy at the time of biopsy, the use of personalized vaccines as a combination therapy becomes more tenable.[1,2] Additionally, as sequencing costs continue to decline, and various technologies employed in sample processing and data acquisition used for HLA-ligand profiling advance, patient-specific neoantigen identification using multi-omic approaches will become suitable for clinical timescales.

Regardless of the approach used to identify putative neoantigens, validation of epitope immunogenicity is crucial as neither direct epitope detection by LC–MS/MS nor putative neoantigen prediction pipelines can guarantee that a given epitope will elicit a T cell response. Even if the predicted epitope binds to its respective HLA heterodimer and is presented on the cell surface, many factors can prohibit a T cell from viewing the epitope as immunogenic. For example, the tumor microenvironment can inhibit the induction of T cell responses.[11,13,14] Strønen and colleagues have worked to investigate why autologous T cells recognize only a small subset of predicted neoantigens and observed that healthy donor T cells responded to a subset of neoantigens that did not elicit tumor infiltrating T cell responses. These data demonstrate that responses to personalized therapies that leverage accurately predicted neoantigens is also dependent on the status of each patient's immune systems.[93] Future studies investigating T cell responses to predicted neoantigens, or lack thereof, are needed to understand how a patient's adaptive immune response can be optimized and used to improve epitope prediction and prioritization.

For many predicted neoantigens applied to personalized immunotherapy, functional assays using a patient's autologous T cells are used to validate immunogenicity of both class I and class II epitopes. Predictions for candidate neoantigens in many recent studies[1,2,17,19,20,33,74,81] are predominantly made for class I HLA epitopes (given the availability of experimental data for class I prediction algorithms compared to class II), yet $CD4^+$ T cell responses are often observed in both preclinical and clinical personalized neoantigen vaccination studies.[1,2,20] These observations demonstrate that class II HLA epitope processing and presentation may also play a critical role in cancer treatment. Although HLA class II prediction algorithms exist, they are less accurate because the open-ended peptide-binding groove on class II HLA heterodimers allows for longer peptides (generally 15–25 amino acids) to bind, which increases the heterogeneity and complexity of epitope presentation.[39,40,63] Further work to better understand the characteristics of class II HLA peptide-binding cores and the cellular processes involved in class II epitope processing and presentation is therefore required.[94] The proteomics field is currently limited by the complexity of class II HLA heterodimer formation and the availability of immunoprecipitation grade antibodies for class II HLA-peptide complex isolation. Once these challenges are overcome, LC–MS/MS will likely play a large role in characterizing the class II HLA-ligandome, which will facilitate improvements to class II epitope prediction methods.

Direct identification of HLA-ligands by MS from mono-[33] and multi-allelic[32,36,39,81] systems has proven to be a faster and less biased technology than canonical methods used to study HLA-peptide binding, and have greatly contributed to our knowledge of the HLA-ligandome and epitope processing. As sample handling methods, instrumentation, data acquisition strategies, and data searching algorithms tailored to the HLA-ligandome improve, HLA-peptide sequencing will become more sensitive and comprehensive. With these improvements, detection of neoantigens directly from patient material and tumor-derived cell lines will become more feasible on clinical timescales. Until such advancements are made, integrating MS HLA-ligand profiling with other "-omic" technologies is critical to create pipelines for cancer vaccines targeting neoantigens, and enabling personalized immunotherapies with largescale clinical applications. Improved epitope prediction algorithms will continue to take advantage of increased coverage of ligands that bind rare HLA class I molecules and HLA class II molecules profiled by LC–MS/MS. Additionally, efforts to share and standardize LC–MS/MS data as outlined by HIPP and consortiums to benchmark epitope prediction algorithms like Tumor Neoantigen Selection Alliance will continue to push the field of personalized immunotherapy to improve the accuracy of neoantigen predictions. This will increase the number of candidate neoepitopes that are prioritized and, ideally, the magnitude of neoantigen specific antitumor responses in patients. With LC–MS/MS technology playing a central role in improving epitope prediction and our understanding of epitope processing and presentation, next generation personalized cancer immunotherapies have the potential to transform how cancer is treated in the clinic.

Abbreviations

CTLs, cytotoxic T lymphocytes; HIPP, the Human Immuno-Peptidome Project; HLA, human leukocyte antigen; IEDB, Immune Epitope

Database; MHC, major histocompatibility complex; TAAs, tumor associated antigens

Acknowledgements

M.S.R., T.A.A., and J.G.A. contributed equally to this work. The authors thank all the thought leaders for their work and contributions to the field of MHC epitope identification, specifically Donald F. Hunt and Jeffrey Shabanowitz. We thank the organizers of HIPP for bringing together research groups from around the globe to advance this field. The authors thank Ed Fritsch, Richard Gaynor, and Marit Van Buuren for useful discussion and critical feedback during preparation of the manuscript. The authors also thank Steven A. Carr and Karl Clauser for insightful conversations on the topic.

Conflict of Interest

A.L.C., Y.S.T., S.P.G., J.F.K.S., D.B., M.S.R., T.A.A., and J.G.A. are employees and shareholders of Neon Therapeutics, Inc.

Keywords

cancer, epitope prediction, immunoproteomics, neoantigens, peptidomics

Received: September 26, 2017
Revised: December 12, 2017

[1] P. A. Ott, Z. Hu, D. B. Keskin, S. A. Shukla, J. Sun, D. J. Bozym, W. Zhang, A. Luoma, A. Giobbie-Hurder, L. Peter, C. Chen, O. Olive, T. A. Carter, S. Li, D. J. Lieb, T. Eisenhaure, E. Gjini, J. Stevens, W. J. Lane, I. Javeri, K. Nellaiappan, A. M. Salazar, H. Daley, M. Seaman, E. I. Buchbinder, C. H. Yoon, M. Harden, N. Lennon, S. Gabriel, S. J. Rodig, D. H. Barouch, J. C. Aster, G. Getz, K. Wucherpfennig, D. Neuberg, J. Ritz, E. S. Lander, E. F. Fritsch, N. Hacohen, C. J. Wu, *Nature* **2017**, *547*, 217.

[2] U. Sahin, E. Derhovanessian, . Miller, B.-P. Kloke, P. Simon, M. Löwer, V. Bukur, A. D. Tadmor, U. Luxemburger, B. Schrörs, T. Omokoko, M. Vormehr, C. Albrecht, A. Paruzynski, A. N. Kuhn, J. Buck, S. Heesch, K. H. Schreeb, F. Müller, I. Ortseifer, I. Vogler, E. Godehardt, S. Attig, R. Rae, A. Breitkreuz, C. Tolliver, M. Suchan, G. Martic, A. Hohberger, P. Sorn, J. Diekmann, J. Ciesla, O. Waksmann, A.-K. Brück, M. Witt, M. Zillgen, A. Rothermel, B. Kasemann, D. Langer, S. Bolte, M. Diken, S. Kreiter, R. Nemecek, C. Gebhardt, S. Grabbe, C. Höller, J. Utikal, C. Huber, C. Loquai, Ö. Türeci, *Nature* **2017**, *547*, 222.

[3] N. Hacohen, E. F. Fritsch, T. A. Carter, E. S. Lander, C. J. Wu, *Cancer Immunol. Res.* **2013**, *1*, 11.

[4] P. Sharma, J. P. Allison, *Science* **2015**, *348*, 56.

[5] D. M. Pardoll, *Nat. Rev. Cancer* **2012**, *12*, 252.

[6] M. M. Gubin, X. Zhang, H. Schuster, E. Caron, J. P. Ward, T. Noguchi, Y. Ivanova, J. Hundal, C. D. Arthur, W.-J. Krebber, G. E. Mulder, M. Toebes, M. D. Vesely, S. S. Lam, A. J. Korman, J. P. Allison, G. J. Freeman, A. H. Sharpe, E. L. Pearce, T. N. Schumacher, R. Aebersold, H.-G. Rammensee, C. J. M. Melief, E. R. Mardis, W. E. Gillanders, M. N. Artyomov, R. D. Schreiber, *Nature* **2014**, *515*, 577.

[7] P. G. Coulie, B. J. Van den Eynde, P. van der Bruggen, T. Boon, *Nat. Rev. Cancer* **2014**, *14*, 135.

[8] A. Miller, L. Cattaneo, Y. W. Asmann, E. Braggio, J. J. Keats, D. Auclair, S. Lonial, T. MMRF CoMMpass Network, S. J. Russell, A. K. Stewart, *Blood* **2016**, *128*, 193.

[9] M. S. Rooney, S. A. Shukla, C. J. Wu, G. Getz, N. Hacohen, *Cell* **2015**, *160*, 48.

[10] E. M. Van Allen, D. Miao, B. Schilling, S. A. Shukla, C. Blank, L. Zimmer, A. Sucker, U. Hillen, M. H. Geukes Foppen, S. M. Goldinger, J. Utikal, J. C. Hassel, B. Weide, K. C. Kaehler, C. Loquai, P. Mohr, R. Gutzmer, R. Dummer, S. Gabriel, C. J. Wu, D. Schadendorf, L. A. Garraway, *Science* **2015**, *350*, 207.

[11] E. M. E. Verdegaal, N. F. C. C. de Miranda, M. Visser, T. Harryvan, M. M. van Buuren, R. S. Andersen, S. R. Hadrup, C. E. van der Minne, R. Schotte, H. Spits, J. B. A. G. Haanen, E. H. W. Kapiteijn, T. N. Schumacher, S. H. van der Burg, *Nature* **2016**, *536*, 91.

[12] J. E. Rosenberg, J. Hoffman-Censits, T. Powles, M. S. van der Heijden, A. V. Balar, A. Necchi, N. Dawson, P. H. O'Donnell, A. Balmanoukian, Y. Loriot, S. Srinivas, M. M. Retz, P. Grivas, R. W. Joseph, M. D. Galsky, M. T. Fleming, D. P. Petrylak, J. L. Perez-Gracia, H. A. Burris, D. Castellano, C. Canil, J. Bellmunt, D. Bajorin, D. Nickles, R. Bourgon, G. M. Frampton, N. Cui, S. Mariathasan, O. Abidoye, G. D. Fine, R. Dreicer, *Lancet* **2016**, *387*, 1909.

[13] N. van Rooij, M. M. van Buuren, D. Philips, A. Velds, M. Toebes, B. Heemskerk, L. J. van Dijk, S. Behjati, H. Hilkmann, D. el Atmioui, M. Nieuwland, M. R. Stratton, R. M. Kerkhoven, C. Kesmir, J. B. Haanen, P. Kvistborg, T. N. Schumacher, *J. Clin. Oncol.* **2013**, *31*, e439.

[14] P. F. Robbins, Y.-C. Lu, M. El-Gamil, Y. F. Li, C. Gross, J. Gartner, J. C. Lin, J. K. Teer, P. Cliften, E. Tycksen, Y. Samuels, S. A. Rosenberg, *Nat. Med.* **2013**, *19*, 747.

[15] *J. Exp. Med.* **1983**, *157*, 1040.

[16] T. N. Schumacher, R. D. Schreiber, *Science* **2015**, *348*, 69.

[17] M. M. Gubin, M. N. Artyomov, E. R. Mardis, R. D. Schreiber, *J. Clin. Invest.* **2015**, *125*, 3413.

[18] J. Scholtalbers, S. Boegel, T. Bukur, M. Byl, S. Goerges, P. Sorn, M. Loewer, U. Sahin, J. C. Castle, *Genome Med.* **2015**, *7*, 118.

[19] M. Yadav, S. Jhunjhunwala, Q. T. Phung, P. Lupardus, J. Tanguay, S. Bumbaca, C. Franci, T. K. Cheung, J. Fritsche, T. Weinschenk, Z. Modrusan, I. Mellman, J. R. Lill, L. Delamarre, *Nature* **2014**, *515*, 572.

[20] S. Kreiter, M. Vormehr, N. van de Roemer, M. Diken, M. Löwer, J. Diekmann, S. Boegel, B. Schrörs, F. Vascotto, J. C. Castle, A. D. Tadmor, S. P. Schoenberger, C. Huber, Ö. Türeci, U. Sahin, *Nature* **2015**, *520*, 692.

[21] S. Bobisse, P. G. Foukas, G. Coukos, A. Harari, *Ann. Transl. Med.* **2016**, *4*, 262.

[22] A. R. Aldous, J. Z. Dong, *Bioorg. Med. Chem.* **2017**, https://doi.org/10.1016/j.bmc.2017.10.021.

[23] J. Robinson, J. A. Halliwell, J. D. Hayhurst, P. Flicek, P. Parham, S. G. E. Marsh, *Nucleic Acids Res.* **2015**, *43*, D423.

[24] M. Bassani-Sternberg, S. Pletscher-Frankild, L. J. Jensen, M. Mann, *Mol. Cell. Proteomics* **2015**, *14*, 658.

[25] D. Hunt, R. Henderson, J. Shabanowitz, K. Sakaguchi, H. Michel, N. Sevilir, A. Cox, E. Appella, V. Engelhard, *Science* **1992**, *255*, 1261.

[26] H.-G. Rammensee, T. Friede, S. Stevanović, *Immunogenetics* **1995**, *41*, 178.

[27] H.-G. Rammensee, J. Bachmann, N. P. N. Emmerich, O. A. Bachor, S. Stevanović, *Immunogenetics* **1999**, *50*, 213.

[28] R. Vita, J. A. Overton, J. A. Greenbaum, J. Ponomarenko, J. D. Clark, J. R. Cantrell, D. K. Wheeler, J. L. Gabbard, D. Hix, A. Sette, B. Peters, *Nucleic Acids Res.* **2015**, *43*, D405.

[29] E. Boen, A. R. Crownover, M. McIlhaney, A. J. Korman, J. Bill, *J. Immunol.* **2000**, *165*, 2040.

[30] M. V. Larsen, C. Lundegaard, K. Lamberth, S. Buus, O. Lund, M. Nielsen, *BMC Bioinformatics* **2007**, *8*, 424.

[31] K. Vaughan, X. Xu, E. Caron, B. Peters, A. Sette, *Expert Rev. Proteomics* **2017**, *14*, 729.

[32] E. Caron, L. Espona, D. J. Kowalewski, H. Schuster, N. Ternette, A. Alpízar, R. B. Schittenhelm, S. H. Ramarathinam, C. S. Lindestam Arlehamn, C. Chiek Koh, L. C. Gillet, A. Rabsteyn, P. Navarro, S. Kim, H. Lam, T. Sturm, M. Marcilla, A. Sette, D. S. Campbell, E. W. Deutsch, R. L. Moritz, A. W. Purcell, H.-G. Rammensee, S. Stevanovic, R. Aebersold, *Elife* **2015**, *4*, e07661.
[33] J. G. Abelin, D. B. Keskin, S. Sarkizova, C. R. Hartigan, W. Zhang, J. Sidney, J. Stevens, W. Lane, G. L. Zhang, T. M. Eisenhaure, K. R. Clauser, N. Hacohen, M. S. Rooney, S. A. Carr, C. J. Wu, *Immunity* **2017**, *46*, 315.
[34] A. Cox, J. Skipper, Y. Chen, R. Henderson, T. Darrow, J. Shabanowitz, V. Engelhard, D. Hunt, C. Slingluff, *Science* **1994**, *264*, 716.
[35] K. Falk, O. Rotzschke, S. Stevanovie, G. Jung, H.-G. Rammensee, *Nature* **1991**, *351*, 290.
[36] M. Bassani-Sternberg, E. Bräunlein, R. Klar, T. Engleitner, P. Sinitcyn, S. Audehm, M. Straub, J. Weber, J. Slotta-Huspenina, K. Specht, M. E. Martignoni, A. Werner, R. Hein, D. H. Busch, C. Peschel, R. Rad, J. Cox, M. Mann, A. M. Krackhardt, *Nat. Commun.* **2016**, *7*, 13404.
[37] G. P. M. Mommen, C. K. Frese, H. D. Meiring, J. van Gaans-van den Brink, A. P. J. M. de Jong, C. A. C. M. van Els, A. J. R. Heck, *Proc. Natl. Acad. Sci.* **2014**, *111*, 4507.
[38] F.-R. Schumacher, L. Delamarre, S. Jhunjhunwala, Z. Modrusan, Q. T. Phung, J. E. Elias, J. R. Lill, *Proteomics* **2017**, *17*, 1600061.
[39] M. S. Khodadoust, N. Olsson, L. E. Wagar, O. A. W. Haabeth, B. Chen, K. Swaminathan, K. Rawson, C. L. Liu, D. Steiner, P. Lund, S. Rao, L. Zhang, C. Marceau, H. Stehr, A. M. Newman, D. K. Czerwinski, V. E. H. Carlton, M. Moorhead, M. Faham, H. E. Kohrt, J. Carette, M. R. Green, M. M. Davis, R. Levy, J. E. Elias, A. A. Alizadeh, *Nature* **2017**, *543*, 723.
[40] G. P. M. Mommen, F. Marino, H. D. Meiring, M. C. M. Poelen, J. A. M. van G. den Brink, S. Mohammed, A. J. R. Heck, C. A. C. M. van Els, *Mol. Cell. Proteomics* **2016**, *15*, 1412.
[41] M. Basler, C. J. Kirk, M. Groettrup, *Innate Immun. Antigen Process.* **2013**, *25*, 74.
[42] B. Ma, *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 1885.
[43] J. Zhang, L. Xin, B. Shan, W. Chen, M. Xie, D. Yuen, W. Zhang, Z. Zhang, G. A. Lajoie, B. Ma, *Mol. Cell. Proteomics* **2012**, *11*, https://doi.org/10.1074/mcp.M111.010587.
[44] J. P. Murphy, P. Konda, D. J. Kowalewski, H. Schuster, D. Clements, Y. Kim, A. M. Cohen, T. Sharif, M. Nielsen, S. Stevanovic, P. W. Lee, S. Gujar, *J. Proteome Res.* **2017**, *16*, 1806.
[45] A. I. Nesvizhskii, *J. Proteomics* **2010**, *73*, 2092.
[46] E. Caron, R. Aebersold, A. Banaei-Esfahani, C. Chong, M. Bassani-Sternberg, *Immunity* **2017**, *47*, 203.
[47] B. Alvarez, C. Barra, M. Nielsen, M. Andreatta, *bioRxivorg* **2017**, https://doi.org/10.1101/210336.
[48] T. Muth, B. Y. Renard, *Brief. Bioinform.* **2017**, bbx033.
[49] S. Kalaora, E. Barnea, E. Merhavi-Shoham, N. Qutob, J. K. Teer, N. Shimony, J. Schachter, S. A. Rosenberg, M. J. Besser, A. Admon, Y. Samuels, *Oncotarget* **2016**, *7*, 5110.
[50] A. I. Nesvizhskii, *Nat. Methods* **2014**, *11*, 1114.
[51] C. N. Boß, F. Grünebach, K. Brauer, M. Häntschel, V. Mirakaj, T. Weinschenk, S. Stevanovic, H.-G. Rammensee, P. Brossart, *Clin. Cancer Res.* **2007**, *13*, 3347.
[52] O. Chornoguz, A. Gapeev, M. C. O'Neill, S. Ostrand-Rosenberg, *Mol. Cell. Proteomics* **2012**, *11*, 1457.
[53] S. Jarmalavicius, Y. Welte, P. Walden, *J. Biol. Chem.* **2012**, *287*, 33401.
[54] M. Cobbold, H. De La Peña, A. Norris, J. M. Polefrone, J. Qian, A. M. English, K. L. Cummings, S. Penny, J. E. Turner, J. Cottine, J. G. Abelin, S. A. Malaker, A. L. Zarling, H.-W. Huang, O. Goodyear, S. D. Freeman, J. Shabanowitz, G. Pratt, C. Craddock, M. E. Williams, D. F. Hunt, V. H. Engelhard, *Sci. Transl. Med.* **2013**, *5*, 203ra125.
[55] A. L. Zarling, J. M. Polefrone, A. M. Evans, L. M. Mikesh, J. Shabanowitz, S. T. Lewis, V. H. Engelhard, D. F. Hunt, *Proc. Natl. Acad. Sci. USA.* **2006**, *103*, 14889.
[56] M. Bassani-Sternberg, E. Barnea, I. Beer, I. Avivi, T. Katz, A. Admon, *Proc. Natl. Acad. Sci. USA.* **2010**, *107*, 18769.
[57] K. Yoshihara, M. Shahmoradgoli, E. Martínez, R. Vegesna, H. Kim, W. Torres-Garcia, V. Treviño, H. Shen, P. W. Laird, D. A. Levine, S. L. Carter, G. Getz, K. Stemke-Hale, G. B. Mills, R. G. W. Verhaak, *Nat. Commun.* **2013**, *4*, 2612.
[58] A. L. Zarling, S. B. Ficarro, F. M. White, J. Shabanowitz, D. F. Hunt, V. H. Engelhard, *J. Exp. Med.* **2000**, *192*, 1755.
[59] S. A. Malaker, S. A. Penny, L. G. Steadman, P. T. Myers, J. C. Loke, M. Raghavan, D. L. Bai, J. Shabanowitz, D. F. Hunt, M. Cobbold, *Cancer Immunol. Res.* **2017**, *5*, 376.
[60] F. Mohammed, M. Cobbold, A. L. Zarling, M. Salim, G. A. Barrett-Wilt, J. Shabanowitz, D. F. Hunt, V. H. Engelhard, B. E. Willcox, *Nat. Immunol.* **2008**, *9*, 1236.
[61] F. Marino, M. Bern, G. P. M. Mommen, A. C. Leney, J. A. M. van Gaans-van den Brink, A. M. J. J. Bonvin, C. Becker, C. A. C. M. van Els, A. J. R. Heck, *J. Am. Chem. Soc.* **2015**, *137*, 10922.
[62] M. Marcilla, A. Alpízar, M. Lombardía, A. Ramos-Fernandez, M. Ramos, J. P. Albar, *Mol. Cell. Proteomics* **2014**, *13*, 462.
[63] F. R. Depontieu, J. Qian, A. L. Zarling, T. L. McMiller, T. M. Salay, A. Norris, A. M. English, J. Shabanowitz, V. H. Engelhard, D. F. Hunt, S. L. Topalian, *Proc. Natl. Acad. Sci. USA.* **2009**, *106*, 12073.
[64] J. G. Abelin, P. D. Trantham, S. A. Penny, A. M. Patterson, S. T. Ward, W. H. Hildebrand, M. Cobbold, D. L. Bai, J. Shabanowitz, D. F. Hunt, *Nat. Protoc.* **2015**, *10*, 1308.
[65] L. B. Alexandrov, S. Nik-Zainal, D. C. Wedge, S. A. J. R. Aparicio, S. Behjati, A. V. Biankin, G. R. Bignell, N. Bolli, A. Borg, A.-L. Borresen-Dale, S. Boyault, B. Burkhardt, A. P. Butler, C. Caldas, H. R. Davies, C. Desmedt, R. Eils, J. E. Eyfjord, J. A. Foekens, M. Greaves, F. Hosoda, B. Hutter, T. Ilicic, S. Imbeaud, M. Imielinsk, N. Jager, D. T. W. Jones, D. Jones, S. Knappskog, M. Kool, S. R. Lakhani, C. Lopez-Otin, S. Martin, N. C. Munshi, H. Nakamura, P. A. Northcott, M. Pajic, E. Papaemmanuil, A. Paradiso, J. V. Pearson, X. S. Puente, K. Raine, M. Ramakrishna, A. L. Richardson, J. Richter, P. Rosenstiel, M. Schlesner, T. N. Schumacher, P. N. Span, J. W. Teague, Y. Totoki, A. N. J. Tutt, R. Valdes-Mas, M. M. van Buuren, L. van /'t Veer, A. Vincent-Salomon, N. Waddell, L. R. Yates, Australian Pancreatic Cancer Genome Initiative, ICGC Breast Cancer Consortium, ICGC MMML-Seq Consortium, ICGC PedBrain, J. Zucman-Rossi, P. Andrew Futreal, U. McDermott, P. Lichter, M. Meyerson, S. M. Grimmond, R. Siebert, E. Campo, T. Shibata, S. M. Pfister, P. J. Campbell, M. R. Stratton, *Nature* **2013**, *500*, 415.
[66] A. Bansal, T. Mann, S. Sterrett, B. J. Peng, A. Bet, J. M. Carlson, P. A. Goepfert, *J. Acquir. Immune Defic. Syndr.* **1999**, *70*, 1.
[67] E. Goodenough, T. M. Robinson, M. B. Zook, K. M. Flanigan, J. F. Atkins, M. T. Howard, L. C. Eisenlohr, *Proc. Natl. Acad. Sci. USA.* **2014**, *111*, 5670.
[68] C. M. Laumont, T. Daouda, J.-P. Laverdure, É. Bonneil, O. Caron-Lizotte, M.-P. Hardy, D. P. Granados, C. Durette, S. Lemieux, P. Thibault, C. Perreault, *Nat. Commun.* **2016**, *7*, 10238.
[69] K. Cibulskis, M. S. Lawrence, S. L. Carter, A. Sivachenko, D. Jaffe, C. Sougnez, S. Gabriel, M. Meyerson, E. S. Lander, G. Getz, *Nat. Biotechnol.* **2013**, *31*, 213.
[70] J. Liepe, F. Marino, J. Sidney, A. Jeko, D. E. Bunting, A. Sette, P. M. Kloetzel, M. P. H. Stumpf, A. J. R. Heck, M. Mishto, *Science* **2016**, *354*, 354.
[71] C. R. Berkers, A. de Jong, K. G. Schuurman, C. Linnemann, J. A. J. Geenevasen, T. N. M. Schumacher, B. Rodenko, H. Ovaa, *J. Immunol.* **2015**, *195*, 4075.

[72] C. R. Berkers, A. de Jong, K. G. Schuurman, C. Linnemann, H. D. Meiring, L. Janssen, J. J. Neefjes, T. N. M. Schumacher, B. Rodenko, H. Ovaa, *J. Immunol.* **2015**, *195*, 4085.
[73] A. C. M. Platteel, J. Liepe, K. Textoris-Taube, C. Keller, P. Henklein, H. H. Schalkwijk, R. Cardoso, P. M. Kloetzel, M. Mishto, A. J. A. M. Sijts, *Cell Rep.* **2017**, *20*, 1242.
[74] E. Karosiene, M. Rasmussen, T. Blicher, O. Lund, S. Buus, M. Nielsen, *Immunogenetics* **2013**, *65*, 711.
[75] C. Lundegaard, K. Lamberth, M. Harndahl, S. Buus, O. Lund, M. Nielsen, *Nucleic Acids Res.* **2008**, *36*, W509.
[76] I. Hoof, B. Peters, J. Sidney, L. E. Pedersen, A. Sette, O. Lund, S. Buus, M. Nielsen, *Immunogenetics* **2009**, *61*, 1.
[77] V. Jurtz, S. Paul, M. Andreatta, P. Marcatili, B. Peters, M. Nielsen, *J. Immunol.* **2017**, *199*, 3360.
[78] W. Fleri, S. Paul, S. K. Dhanda, S. Mahajan, X. Xu, B. Peters, A. Sette, *Front. Immunol.* **2017**, *8*, 278.
[79] T. Trolle, C. P. McMurtrey, J. Sidney, W. Bardet, S. C. Osborn, T. Kaever, A. Sette, W. H. Hildebrand, M. Nielsen, B. Peters, *J. Immunol.* **2016**, *196*, 1480.
[80] M. Bassani-Sternberg, D. Gfeller, *J. Immunol.* **2016**, *197*, 2492.
[81] M. Bassani-Sternberg, C. Chong, P. Guillaume, M. Solleder, H. Pak, P. O. Gannon, L. E. Kandalaft, G. Coukos, D. Gfeller, *PLoS Comput. Biol.* **2017**, *13*, e1005725.
[82] M. Di Marco, H. Schuster, L. Backert, M. Ghosh, H.-G. Rammensee, S. Stevanović, *J. Immunol.* **2017**, *199*, 2369.
[83] O. E. Hawkins, R. S. VanGundy, A. M. Eckerd, W. Bardet, R. Buchli, J. A. Weidanz, W. H. Hildebrand, *J. Proteome Res.* **2008**, *7*, 1445.
[84] R. Marty, S. Kaabinejadian, D. Rossell, M. J. Slifker, J. van de Haar, H. B. Engin, N. de Prisco, T. Ideker, W. H. Hildebrand, J. Font-Burgada, H. Carter, *Cell* **2017**, *171*, 1272.
[85] A. Admon, M. Bassani-Sternberg, *Mol. Cell. Proteomics* **2011**, *10*, https://doi.org/10.1074/mcp.O111.011833.
[86] M. Bassani-Sternberg, G. Coukos, *Curr. Opin. Immunol.* **2016**, *41*, 9.
[87] H.-G. Rammensee, H. Singh-Jasuja, *Expert Rev. Vaccines* **2013**, *12*, 1211.
[88] H. Singh-Jasuja, N. P. N. Emmerich, H.-G. Rammensee, *Cancer Immunol. Immunother.* **2004**, *53*, 187.
[89] W. Zhao, M. Fitzgibbon, L. Bergan, N. Clegg, D. Crispin, G. B. Mills, M. McIntosh, *Cancer J.* **2017**, *23*.
[90] T.-C. Chang, R. A. Carter, Y. Li, Y. Li, H. Wang, M. N. Edmonson, X. Chen, P. Arnold, T. L. Geiger, G. Wu, J. Peng, M. Dyer, J. R. Downing, D. R. Green, P. G. Thomas, J. Zhang, *Genome Med.* **2017**, *9*, 78.
[91] M. W. Löffler, P. A. Chandran, K. Laske, C. Schroeder, I. Bonzheim, M. Walzer, F. J. Hilke, N. Trautwein, D. J. Kowalewski, H. Schuster, M. Günder, V. A. Carcamo Yañez, C. Mohr, M. Sturm, H.-P. Nguyen, O. Riess, P. Bauer, S. Nahnsen, S. Nadalin, D. Zieker, J. Glatzle, K. Thiel, N. Schneiderhan-Marra, S. Clasen, H. Bösmüller, F. Fend, O. Kohlbacher, C. Gouttefangeas, S. Stevanović, A. Königsrainer, H.-G. Rammensee, *J. Hepatol.* **2016**, *65*, 849.
[92] V. Dutoit, C. Herold-Mende, N. Hilf, O. Schoor, P. Beckhove, J. Bucher, K. Dorsch, S. Flohr, J. Fritsche, P. Lewandrowski, J. Lohr, H.-G. Rammensee, S. Stevanovic, C. Trautwein, V. Vass, S. Walter, P. R. Walker, T. Weinschenk, H. Singh-Jasuja, P.-Y. Dietrich, *Brain* **2012**, *135*, 1042.
[93] E. Strønen, M. Toebes, S. Kelderman, M. M. van Buuren, W. Yang, N. van Rooij, M. Donia, M.-L. Böschen, F. Lund-Johansen, J. Olweus, T. N. Schumacher, *Science* **2016**, *352*, 1337.
[94] A. Kim, T. N. Boronina, R. N. Cole, E. Darrah, S. Sadegh-Nasseri, *Sci. Rep.* **2017**, *7*, 46418.