The Human Immunopeptidome Project: A Roadmap to Predict and Treat Immune Diseases

Authors
Juan Antonio Vizcaíno, Peter Kubiniok, Kevin A. Kovalchik, Qing Ma, Jérôme D. Duquette, Ian Mongrain, Eric W. Deutsch, Bjoern Peters, Alessandro Sette, Isabelle Sirois, and Etienne Caron

Correspondence
etienne.caron@umontreal.ca

In Brief
Immunopeptidomics is a rapidly evolving field that is catalyzed by neoantigen discovery and cancer immunotherapy. Emulating the path of other omic disciplines, Vizcaíno et al. proposes that technological advances in the mass spectrometry field will lead to large immunopeptidomic cohort studies, and ultimately, the immunopeptidome-wide association study (IWAS) paradigm, which will provide profound insights into disease susceptibility and resistance.

Graphical Abstract

Highlights
• Immunopeptidomics bears the potential to link diseases to antigen representation.
• We suggest to achieve this by analyzing the immunopeptidomes of cohorts of patients.
• Current mass spectrometry-based techniques to analyze immunopeptidomes are described.
• We term the proposed approach “Immunopeptidome-wide association studies” (IWAS).
The science that investigates the ensembles of all peptides associated to human leukocyte antigen (HLA) molecules is termed “immunopeptidomics” and is typically driven by mass spectrometry (MS) technologies. Recent advances in MS technologies, neoantigen discovery and cancer immunotherapy have catalyzed the launch of the Human Immunopeptidome Project (HIPP) with the goal of providing a complete map of the human immunopeptidome and making the technology so robust that it will be available in every clinic. Here, we provide a long-term perspective of the field and we use this framework to explore how we think the completion of the HIPP will truly impact the society in the future. In this context, we introduce the concept of immunopeptidome-wide association studies (IWAS). We highlight the importance of large cohort studies for the future and how applying quantitative immunopeptidomics at population scale may provide a new look at individual predisposition to common immune diseases as well as responsiveness to vaccines and immunotherapies. Through this vision, we aim to provide a fresh view of the field to stimulate new discussions within the community, and present what we see as the key challenges for the future for unlocking the full potential of immunopeptidomics in this era of precision medicine. Molecular & Cellular Proteomics 19: 31–49, 2020. DOI: 10.1074/mcp.R119.001743.

“A RATIONALE FOR THE IWAS PARADIGM

The Human Genome Project was a major milestone in the life sciences (1, 2). First-completed twenty years ago, this mind-shifting project has changed and will continue to change the way we practice medicine. Historically, physicians have largely focused on treating disease already in progress, but modern medicine is now progressively shifting from disease treatment to disease prevention based on an individual’s risk (3). The past two decades have seen an enormous success of wide-scale studies in identifying genetic variants that predict an individual’s predisposition to common diseases (4). In fact, robust, rapid and inexpensive identification of functional genetic variants in individuals is now enabling predictive, preventive and personalized medicine approaches (5–7). Since the first report of single-nucleotide polymorphisms (SNPs)1 analyzed for association with myocardial infarction by genome-wide association studies (GWAS) in 2002 (8), the GWAS Catalogue resource has now grown to contain tens of thousands of SNPs associated with hundreds of common diseases (9). In this post-GWAS era, the HLA has been established as the region of the genome that is associated with the greatest number of human diseases (10). In fact, population studies of diverse ancestries have identified hundreds of susceptibility loci within the HLA region that predispose individuals to immune diseases (11–18). Most HLA disease associations that were reported over the last 50 years are related to the immune system, but the exact mechanisms driving the associations remain largely elusive.

The HLA is divided into two main subclasses: class I, which include the classical HLA-A, HLA-B and HLA-C molecules, as well as the nonclassical HLA-E, HLA-F and HLA-G molecules; and class II, which includes HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DQA2, HLA-DQB1, HLA-DQB2, HLA-DRB1, HLA-DRB2, HLA-DRB3, HLA-DRB4, and HLA-DRB5 molecules (Fig. 1A)(19). To date, more than 23,000 different classical HLA alleles have been identified (https://www.ebi.ac.uk/ipd/imgt/hla/stats.html) and this number may keep
climbing up to nearly 8–9 million HLA variants (Fig. 1B) (20). The role of HLA molecules is to present a repertoire of peptides at the cell surface for T-cell recognition. The non-random amino acid composition of those peptides is restricted by length, generally between 8 and 12 (can be up to 15) amino acids for class I and between 13 and 25 amino acids for class II, and by the presence of allele-specific binding motifs (21). Therefore, the nature of those peptides and the extreme diversity of HLA alleles at the population level greatly enhances the complexity of the peptide repertoires, which are collectively termed as the human immunopeptidome (22, 23). T cells scan the HLA immunopeptidome and seek for "abnormal" peptides that originate from metabolic perturbations, pathogenic sources or neoplastic transformations (24–28). Those abnormal cells are then eradicated once engaged by T cells. The extreme diversity of the human immunopeptidome maximizes the probability that at least some individuals within the world population can mount a T-cell attack against an emerging infection and survive (29). This is exemplified by the HIV epidemic and evidences that HLA-B57 individuals are more likely protected against HIV (30, 31). However, such evolutionary advantage comes with the price of disease susceptibility. In fact, individuals expressing specific HLA alleles are more susceptible to suffer from specific auto-inflammatory and autoimmune diseases. For instance, birdshot chorioretinopathy was shown to be associated with HLA-A29, ankylosing spondylitis with HLA-B27, Behçet’s disease with HLA-B51 and psoriasis with HLA-C06. Combinations of endoplasmic reticulum aminopeptidases (ERAP) 1 haplotypes, involved in trimming HLA-associated peptides, are also risk factors for these diseases in people that have specific HLAs (e.g. HLA-B27*05) (32–35). Moreover, an increasing number of GWAS results indicate that amino acid polymorphisms associated with immune diseases are likely affecting binding affinities of peptides within the groove of HLA proteins, and thus, affecting the repertoire of peptides presented to T cells that are capable of triggering or perpetuating human diseases (10, 12, 19, 36–45). Interestingly, those evidences obtained from GWAS, meta-analysis and HLA region fine-mapping studies point toward a fundamental role of the immunopeptidome in driving human diseases. We highlight below four cases of such studies: (1) in allergic rhinitis, the most common clinical presentation of allergy affecting 400 million people worldwide, a recent meta-analysis including 59,762 cases and 152,358 controls from European ancestry, confirmed in a replication phase of 60,720 cases and 618,527 controls, indi-
cated that the strongest associated amino acid variants in HLA-DQB1 and HLA-B were both located in the peptide-binding groove (36); (2) in lung cancer, analysis of HLA genetic variation among 26,044 lung cancer patients and 20,836 controls revealed amino acid variants in the HLA-B*08:01 peptide-binding groove (37); (3) in alopecia areata, one of the most prevalent autoimmune diseases, a meta-analysis including 3253 cases and 7543 controls highlighted HLA-DR as the primary risk factor of the disease with amino acid variants located in the peptide-binding groove (39); and (4) in common infections, a meta-analysis including over 200,000 individuals of European ancestry suggested important roles of specific amino acid polymorphisms in the peptide-binding groove, for eight common infectious diseases, in particular: i.e. chicken pox, shingles, cold sores, tuberculosis, scarlet fever, tinea, leprosy, pneumonia and plantar warts (38). GWAS studies have also shown the association of Parkinson’s disease with HLA-DRB5*01 and DRB1*15:01, which are present in ~15% of the general population (46). In this study, the authors showed that a defined set of peptides that are derived from alpha-synuclein, a protein aggregated in Parkinson’s disease, act as antigenic epitopes displayed by the alleles and drive helper and cytotoxic T-cell response in patients with Parkinson’s disease (47). Thus, if further tested and validated for other diseases, cohort studies may ultimately suggest a universal role of the immunopeptidome in disease susceptibility.

In addition to qualitative changes in the peptide repertoire, the level of expression and stability of HLA molecules at the cell surface have been found to be associated with diseases (11, 12, 19). For instance, a high level of expression of HLA-C increases the risk of Crohn’s disease but also promotes the control of HIV infection (13, 14); instability of HLA-DQ2 and HLA-DQ8 could increase the risk of type 1 diabetes whereas stability of HLA-DQ6 may confer protection (15–17). In cancer, a recent analysis of more than 1500 melanoma patients treated by immune checkpoint inhibitors (ICIs) showed that HLA-B44 alleles were associated with an improved survival rate whereas HLA-B62 alleles were associated with a decreased effectiveness of ICIs (48, 49). These data suggest that HLA-B44 immunopeptidomes may favor presentation of a broader, and perhaps more stable repertoire of tumor-specific antigens to trigger the function of cytotoxic T cells in response to ICIs. This is consistent with the recent observation that specific HLA genotypes are associated with the appearance of specific oncogenic mutations (50, 51).

GWAS can identify genetic variants at the gene level and can be complemented with genome-wide transcript quantitative trait loci (eQTL) and genome-wide protein quantitative trait loci (pQTL) to better understand genotype-phenotype associations (52–55). However, attempts to understand HLA disease associations by only measuring expression levels of intracellular genes and proteins is unlikely to work very accurately for immune diseases in which HLA-peptide—T-cell interactions play a critical role. In contrast, measuring HLA-bound peptides is highly relevant because those peptides are directly “seen” by T cells and are involved in disease progression (56). Therefore, technologies capable of robust and comprehensive measurements of immunopeptidomes at population-scale (from tens of thousands of samples) would be extremely powerful as they would provide direct physical evidence about the identity and quantity of the peptides that are directly “seen” and engaged by T cells. Such technologies could provide key information to better understand associations between HLA alleles and human diseases, or between HLA alleles and responsiveness to treatment. Thinking forward in this context, we anticipate that further development of “immunopeptidomics” (57) technologies will ultimately lead to the immunopeptidome-wide association study (IWAS) paradigm and will provide a new layer of information about disease susceptibility. Emulating the path of transcriptome-, proteome- and metabolome-wide association studies (TWAS, PWAS and MWAS, respectively) (55, 58–61), the road toward IWAS is technically very challenging, but conceptually straightforward: quantify the immunopeptidome of large human population cohorts and statistically link immunopeptidome variations to different clinical outcomes (Fig. 2A). IWAS coming from very large immunopeptidomic sample cohorts could be of great value to treat patients but may also provide highly precise information about the predisposition of an individual to get specific immune diseases, hence representing a new look at human disease risk factors.

The road toward IWAS is probably very long and it is actually very likely that the feasibility of conducting IWAS from large human population cohorts will face a great deal of skepticism within the immunopeptidomics community as many technical issues can be considered at this point of time. In fact, the immunopeptidome is very different from one tissue to the other of every person and obvious questions will arise. For instance, (1) which tissues and cell types should be used?; (2) which disease should be prioritized while considering sample accessibility?; (3) what is the minimum number of samples (disease and control) that would be required?; and (4) which HLA alleles would be included and would they need to be shared across all samples to enable proper analysis of the IWAS data set? Addressing these questions would require the expertise of biostatisticians and are difficult to answer now. To our knowledge, the largest immunopeptidomic cohort study so far was conducted by the group of Arie Admon through the Glioma Actively Personalized Vaccine Consortium (GAPVAC). In this study, the authors have isolated HLA-associated peptides from 142 plasma samples collected from glioblastoma patients (62). Although impressive, 142 samples are not enough to conduct an IWAS. Akin to other association studies, it is likely that over 10,000 samples sharing specific HLA alleles (e.g. HLA-A02) would be required to enable IWAS. Accessibility to such cohorts of samples would probably need to be done through large multi-institutional networks and biobanking efforts, which may include both academia and
industrial partners. Such an endeavor may be viewed as over-optimistic by some experts in the field. Here, our intent is neither to answer all potential questions to make IWAS a reality, nor to provide a detailed roadmap to achieve this goal. Instead, we aim to emphasize the most pressing technical issues to move faster toward this long-term goal, and most importantly, stimulate a mind-shift within the community to move the field to a higher orbit. If the community acknowledges the impact of moving the field from small to large immunopeptidomic cohort studies, we predict that immunopeptidomics will emerge in the future as a central science of immunology to predict, diagnose, monitor and treat immune diseases, and will therefore profoundly impact medicine (Fig. 2).

The Human Immunopeptidome Project—Development of mass spectrometry (MS) technology platforms may pave the way toward the IWAS paradigm. In fact, MS technologies have recently attracted a high level of interest for providing direct measurements of immunopeptidomes given the ongoing revolution in immuno-oncology and the intense interest in identifying therapeutically relevant antigenic peptides that are actionable in the clinic for vaccine development (63–73). More specifically, advances in genomics and MS techniques have enabled the development of proteogenomic methods for...
the direct detection of immunogenic tumor-specific mutated and non-mutated peptides that originate from both coding and non-coding regions of the genome (74–77). Those studies have provided an important proof-of-concept that MS is an effective approach that can be deployed for direct identification of immunogenic tumor-specific antigens. If further developed and routinely applied in cohort studies, such analytical methods would be of great value for biomarker discovery in cancer immunotherapy to provide molecular signatures or epitope biomarkers of tumor rejection to discriminate patients who best respond to ICIs (Fig. 2B) (78). For the time being, imprecise predictions of HLA-associated peptides from genes and mRNA is still preferred over direct HLA-peptide measurements by MS (26, 79). This choice can be explained by the state and accessibility of the respective measurement techniques: whereas essentially complete genome and transcriptome analysis is readily available to immunologists through core facilities, MS-based analysis of immunopeptidomes is most effectively performed by expert labs and cannot easily reach the throughput, sensitivity and reproducibility of genome and transcriptome analysis. To solve this issue, the MS-based immunopeptidomics Human Immunopeptidome Project (HIPP) was launched under the umbrella of the Human Proteome Organization (HUPO) with the long-term goal of 1) mapping the entire composition of the human immunopeptidome and 2) making data and the robust experimental and computational techniques of immunopeptidomics accessible to any researcher and clinical investigators (22, 23). Ultimately, absolute quantification of immunopeptidomes should become as cheap, fast, sensitive and reproducible as it is nowadays for genome and transcriptome analysis. Only then, immunopeptidomics will become widely accessible to the broader immunology community to enable very large immunopeptidomic cohort studies, the development of epitope biomarker discovery platforms, and ultimately, IWAS.

**Historical Milestones in MS-based Immunopeptidomics**—Over the last five years, many excellent reviews have been published regarding the analysis of immunopeptidomes using MS technologies (71, 73, 80–93). To avoid redundancies with those reviews, we focus below on key historical milestones within three different timeframes, which we view as important toward the IWAS paradigm: (1) the groundbreaking work of Rammensee and Hunt/Engelhard, (2) the capability of MS technologies for large-scale sequencing of MHC-associated peptides, and (3) the renaissance of cancer immunotherapy (Fig. 3A).

In the early 1990s, the teams of Hans-Georg Rammensee, Victor Engelhard and Donald Hunt provided truly groundbreaking work (Fig. 3A). They isolated MHC/HLA class I-associated peptides by immunoaffinity purification and applied liquid chromatography (LC)-MS methods to provide the first physical evidence about the nature of MHC class I-bound peptides (94, 95). Specifically, in 1990, Rammensee’s group applied LC-MS to analyze naturally processed viral peptides recognized by cytotoxic T cells (95). In 1991, his group reported pooled sequencing of peptides eluted from several MHC class I molecules, with identification of allele-specific peptide binding motifs (21). The year after, the findings by Hunt et al. were consistent with the Rammensee motif (94). In fact, Hunt/Engelhard’s group quantified ∼200 peptide species by LC-MS and partially sequenced ∼19 of them, validating the motif concept and providing the first estimate that the total number of different peptides presented by HLA-A2 could easily exceed 1000, with most of the peptides present in 100 or fewer copies per cell, hence suggesting for the first time the existence of a large and complex immunopeptidome (94, 96–101). In 1992, Hunt/Engelhard provided physical evidences for non-canonical long peptides presented by HLA class I molecules (102), and in 1997, sequenced the very first mutated MHC class I peptides (Fig. 3A) (103), largely referred nowadays as tumor-specific neoantigens (26). Further, in 2006, Hunt/Engelhard’s group showed evidences for phosphopeptides as potential targets for cancer immunotherapy (104, 105). Since then, MS-based discovery of mutated (75, 76, 80, 106–108), non-canonical (86, 87, 109–111) and post-translationally modified HLA-associated peptides (phosphorylation (112–115), methylation (116), citrullination (117), deamidation (118–121), kynurenine (122), glycosylation (123–125), peptide splicing (126–130)) continues to be a topic of much interest in the field.

A key technical milestone that greatly impacted the immunopeptidomics field was the commercialization of the Orbitrap technology in 2005 (131). In fact, the launch of this instrument has enabled researchers to routinely perform large-scale sequencing of MHC-associated peptides, going from dozens to hundreds, to lately thousands of MHC-associated peptides identified in single experiments (87, 132). Computational algorithms were created and improved (91–93, 133–136) and different MS techniques developed by the proteomics community were used for the analysis of MHC peptides: data-dependent acquisition (DDA) for the discovery of new antigenic peptides, selected/multiple/parallel reaction monitoring (S/M/PRM) for targeted analysis of pre-defined sets of peptides with a high level of sensitivity, reproducibility and quantitative accuracy, and more recently, sequential window acquisition of all theoretical fragment ion spectra/data-independent acquisition (SWATH/DIA). SWATH/DIA makes use of high resolution qTOF instruments (SWATH) or Orbitrap instruments for high-throughput targeted analysis of large fractions of immunopeptidomes and is reviewed in (71). For three decades, a very small community of experts have deployed these MS techniques to generate MS data—generally shared nowadays through MS public data repositories such as PRIDE and MassIVE—in most cases through collaborations with immunologists and clinical investigators (Fig. 3B).

Until now, more than 500 original publications related to MS-based sequencing of MHC-associated peptides were reported (supplemental Table S1; see filtered PubMed search
**Next-generation Immunopeptidomics**

**A**

**#1 GROUND-BREAKING WORK**

First tumor-specific mutated MHC I neoantigens identified by MS

First MHC-associated peptides identified by MS

The term ‘immuno-peptidomics’ is used for the first time

**#2 LARGE-SCALE SEQUENCING OF MHC-ASSOCIATED PEPTIDES**

First large-scale sequencing of HLA class I peptides

A human immunopeptidome project is proposed

The magazine *Science* acknowledges cancer immunotherapy as Breakthrough of the Year

**#3 RENAISSANCE OF CANCER IMMUNOTHERAPY**

 (~1990-2013) The field is occupied by a handful of MS experts

First tumor-specific mutated MHC I neoantigens identified by combining MS and exome sequencing

First immunopeptidome-wide association study (IWAS)

1990-1992

2004

2011

2013

2014

2015-2019

(?)

**B**

[Publishers of research (Left panel)]

[Publications from PRIDE and MassIVE (Right panel)]

**C**

[Technology (34)]

[Technology & Immunology (48)]

[Immunology (511)]

[Antigen processing & presentation]

[Autoimmunity & inflammation]

[Autoimmunity & inflammation]

[Autoimmunity & inflammation]

[Molecular & Cellular Proteomics 19.1]

**Fig. 3. Historical perspective from a systematic review of the literature.** A, Technical and cultural milestones in MS-based immunopeptidomics. The dotted arrow to the right indicates a hypothetical future point in time at which the first IWAS will have been successfully conducted. B, Global distribution of research groups that have shown productivity in the field of MS-based immunopeptidomics. (Left panel) A systematic PubMed search (2019/06/24) resulted in a total number of > 500 publications from 381 corresponding authors distributed across 180 cities and 36 countries, since the first peptide was sequenced in 1990. (Right panel) A systematic PRIDE and MassIVE (ProteomeXchange resources) search (2019/07/24) resulted in 94 publicly available MS immunopeptidomic datasets coming from 46 laboratory heads, distributed across 32 cities and 14 countries, since the first dataset was deposited in 2013. The circles indicate the location of the corresponding authors/laboratory heads. The size of the circles is proportional to the number of publications/datasets reported by the corresponding author/laboratory head. C, The pie charts show the proportion of publications from the systematic PubMed search. (Left) The pie chart shows...
The proportion of publications that have answered an immunology question and/or that have tackled a technological limitation. (Right) The pie chart shows the proportion of studies that have applied the available technology to answer an immunology question in the context of basic research (antigen processing and presentation) or applied research (cancer, infectious diseases and autoimmunity/inflammation). See supplemental Table S1 for details about the PubMed and PRIDE/MassIVE searches.
required to perform this procedure (i.e. ∼1 mg of antibody on average per sample). As of today, any laboratory aiming at conducting large-scale immunopeptidomic studies must generally produce and purify antibodies from hybridoma cell lines, which is relatively costly and time consuming, and therefore represents a considerable limitation for scaling up the procedure and for rapidly implementing immunopeptidomics workflows into new laboratories. Moreover, very little is known about the yield of the immunoaffinity purification procedure. Until now, the use of isotopically labeled peptide–HLA monomers represents the best strategy for the accurate quantification of stepwise yields of the procedure (154). In this regard, the group of Peter van Veelen has recently synthesized 24 different isotopically peptide–HLA monomers and showed that the yield was sequence-dependent. Specifically, the yield ranged between 2.1 and 15%, and was below 10% for 70% of the peptides tested (17 out of 24), indicating important losses and/or biases during sample preparation (155). Along these lines, other separation techniques were reported to create a bias in the detectable immunopeptidome, which can be explained by the enormous biochemical diversity of those peptides (156). For instance, highly hydrophilic peptides may not be captured by conventional reversed-phase C_{18} material whereas highly hydrophobic peptides may be preferentially lost during the procedure simply because they stick to plastic materials during the multiple handling steps.

To address the issues mentioned above, exploration of new approaches in digital microfluidics for immunoprecipitation may help increase the sample throughput (157); further development of acoustic technologies for contactless handling may accelerate unbiased analysis of immunopeptidomes (158, 159); and low-cost distribution of antibodies may accelerate the testing of new sample preparation protocols, which may ideally enable all sample processing steps to be carried out in a single tube, thereby enhancing sensitivity, throughput and scalability of immunopeptidomic analyses (160, 161). As a first step, HUPO-HIPP recently proposed to design a multi-laboratory study in which a large library of isotopically labeled peptide–HLA monomers would be distributed across different research groups for the accurate quantification of stepwise yields of the established immunoaffinity purification procedure (22). This type of work has been underappreciated over the years but is nevertheless invaluable to define the physicochemical properties of peptides that are preferentially detected or lost during the isolation procedure. Inspired by landmark studies from other -omic disciplines (162–167), conducting the proposed multi-laboratory study will clarify the uncertainty about the yield of the peptide isolation procedure, will determine the robustness of the methodology used by different groups and will possibly indicate the need to prioritize the development of innovative sample preparation protocols.

In summary, the field has greatly benefited from the traditional immunoaffinity purification method to make groundbreaking discoveries. The method will likely continue to be used for small-scale studies and relatively larger cohort studies in expert laboratories soon (62, 141). However, the development of new sample preparation protocols will become critical to scale up and bring the field to the next level. Isolation of MHC-associated peptides in a cheap, fast, unbiased, reproducible and high-throughput fashion has the potential to transform the field and increase the impact of immunopeptidomics in biomedical research through robust clinical applications, and ultimately, population-scale immunopeptidomic studies.

**Disruptive Technologies for Large-Scale Immunopeptidomics**—IWAS and population-scale studies will undoubtedly require immunopeptidomic technologies to be widely accessible and commercially well-supported. In genomics, high throughput generation of complete genomic maps became widely accessible only with the development and commercialization of technologies and methods that enabled the sequencing of millions of nucleic acid segments in parallel (168). Reaching this state in the field of immunopeptidomics for high throughput generation of complete digital immunopeptidomic maps is possible but embracing public-private partnerships will need to be prioritized in the coming years to accelerate commercialization of new immunopeptidomics technologies/methods and their wide-spread distribution. It is conceivable that MS technologies, such as SWATH/DIA-MS, will continue to improve toward this goal. However, it is also possible that disruptive new approaches will be developed for massively parallel peptide sequencing, making MS technologies potentially obsolete in the future. In this regard, Swaminathan et al. have recently introduced a groundbreaking peptide fluorescence sequencing technology which may accelerate the process toward the IWAS paradigm. Akin to massively parallel measurements of DNA using fluorescence as readout, they demonstrated that parallel fluorescence sequencing is also achievable for peptides (169). Their method is promising because it paves the way toward single-molecule peptide sequencing at very high throughput from minute amounts of samples (http://www.erisyon.com/) (170).

Their method builds on three well-characterized methods: Edman chemistry, massively parallel DNA sequencing and MS-based computational strategies for sequence database searching (171). Briefly, isolated peptides are first labeled with fluorophores for each amino acid residue; second, labeled peptides are immobilized on a glass surface and imaged by total internal reflection microscopy to monitor decreases in each molecule’s fluorescence after consecutive rounds of Edman degradation; third, the pattern of drops in fluorescence intensity is interpreted to provide a sequencing annotation for each peptide, which is matched and scored against a peptide sequence database to infer the most likely set of peptides present in the sample (169, 170, 172). In our opinion, this approach is very promising for proteins soon, but further development of the technology will be essential for sequenc-
ing HLA-associated peptides because it is currently impossible to label most of the amino acids with fluorophore. Although this proof of concept approach has limitations at this stage (e.g., dyes used, sample complexity, sequencing yield), it capitalizes on three established techniques that may accelerate maturation from proof of concept to routine application and wide adoption by the community.

If further developed, highly parallel single-molecule peptide sequencing may offer an improvement of more than one million-fold in sensitivity over conventional MS technologies and may allow for millions of distinct peptides to be sequenced in parallel, identified and digitally quantified (169, 171). Given the qualitative and quantitative complexity of the human immunopeptidome, this technology may radically transform the immunopeptidomics field. However, one has to keep in mind that the performance of this approach will always rely on the quality of upstream peptide isolation procedures. Hence, the shortcomings mentioned above related to the isolate of MHC-associated peptides by immunoaffinity purification will need to be overcome for a generally accessible, reliable and truly universal immunopeptidomic technology.

Data Analysis Challenges — The main approaches used currently for the analysis of immunopeptidomics experiments were developed originally for the analysis of MS proteomics datasets. Database searches are the main analysis method of choice for data-dependent acquisition workflows. Here, the acquired spectra are compared with generated (theoretical) ones coming from peptide sequences drawn from a given protein sequence database. Several limitations of this type of analyses can apply to immunopeptidomics. First of all, the size of the search database (and the related search space) can become extremely large. For instance, proteasome-generated spliced peptides have been detected by MS in recent studies (126, 129, 173, 174), so the inclusion of spliced variants in the search database could be required (129). However, this strategy has sparked intense discussions in the field over the last two years and remains actively debated (130, 175, 176). Additionally, in most proteomics experiments, trypsin is used as the digestion enzyme. Because “no enzyme specificity” must be used in immunopeptidome searches, the search space grows exponentially. As a result of these two related issues, the time required for performing the searches can increase enormously and even more, if variable post-translational modifications such as phosphorylation have to be considered in the search. More importantly, when making use of the decoy-target approach, it becomes very challenging for search engines to differentiate between true and false matches, considering the huge search space. Consequently, a compromise is often needed, and the accepted PSM false discovery rate can become 5% instead of the standard 1% used in standard MS proteomics experiments.

In parallel, the average percentage of identified spectra can be significantly lower in immunopeptidomics experiments, because of different factors (e.g., shortness of MHC class I peptides, lack of basic amino acid residues, low amount of samples, acquisition parameters, etc), as described previously (71). To address this issue, several approaches have been proposed, like the sequential use of different search engines or the use of Percolator. This tool performs machine learning on high-confidence matches to rescore database search results for lower-confidence peptides (177). Other approaches for performing rescoring have also been successfully used recently in the field (133).

Moreover, alternative analysis approaches to standard database searches can be also used, either alone or in combination. One example is the use of de novo search methods, where the spectra can be sequenced directly, without the need of a protein sequence database (178). The use of de novo algorithms has recently been recommended in immunopeptidomics protocols (72). Spectral searches approaches can also be employed (179). Here, the experimental spectra are compared with existing collections of spectra in the public domain called spectral libraries. However, this approach, although increasingly popular, has had a limited uptake so far. One of the reasons is that the generation of high-quality spectral libraries is not straightforward, and some researchers are reluctant to use publicly available spectral libraries generated by others.

Additionally, the increasing popularity of open modification searches makes them a promising approach. The new tools developed in the last few years (e.g., MSFragger, ANN-SoLo, TagGraph, among others (180–182) have solved the high computational cost of these methods. Additionally, in our opinion, a promising approach for targeting interesting spectra would be the use of clustering of MS/MS spectra (183–185). This approach can be used to select those spectra that remain unidentified and that are commonly found across MS runs in different samples. The hypothesis is that the corresponding peptides are potentially interesting and biologically relevant, because they are abundant. The resulting representatives of these clusters of unidentified spectra could then be subjected to those alternative analysis methods explained above. It is certainly a possibility that a significant fraction of unidentified spectra from such clusters includes contaminating peptides that show up in many samples. They could indeed be actual contaminants usually found in proteomics experiments (e.g., keratins), or not biologically relevant peptides as artifacts derived from the experimental protocol. If the latter case applies, information from such clusters could be used to feed a database of contaminating peptides that are constantly observed in immunopeptidomics. Such database does not exist yet but would be useful for the growing immunopeptidomics community and would be conceptually like the CRAPome, which serves as a repository of protein contaminants in protein-protein interaction studies (186).

Further, artificial intelligence approaches have the potential to improve the analysis dramatically. In MS-based proteomics, deep-learning techniques have been recently deployed.
using millions of high quality MS2 spectra generated from hundreds of thousands of synthetic tryptic peptides, to build new algorithms to predict fragmentation patterns of peptides, thereby enabling successful generation of in-silico peptide spectral libraries for high throughput targeted analysis of DIA data (187–189). These developments should also greatly benefit the further improvement of de novo search algorithms. Deep-learning was also recently applied for de novo identification of HLA-associated peptides using DIA-MS (190, 191).

Finally, it is worth highlighting that a significant number of challenges would need to be considered in addition with regards to the quantification of immunopeptidomes. This quantitative aspect is much less mature at present. Analogous challenges in quantitative MS proteomics approaches have been described before (192).

The SysteMHC Atlas Project—Open and comprehensive reference atlas in life sciences are increasingly beneficial for the scientific community (193–198). Similarly, the creation of a comprehensive atlas of the immunopeptidome in human, mouse, and other species would be of great value for both understanding health and diagnosing, monitoring and treating immune diseases (22, 23, 69). In this context, participants of the HIPP initiative recognized the need for an open immunopeptidomics atlas/repository in which output files of mass spectrometric measurements of immunopeptidome samples would be annotated, stored and shared without restriction. In response to this need, the SysteMHC Atlas project was created—a project fully dedicated to the public dissemination and analysis of immunopeptidomic data generated by MS (https://systemhcatlas.org) (Fig. 4) (199). The SysteMHC Atlas uploads raw immunopeptidomics MS data originally deposited into public proteomics databases (mainly the PRIDE database (200), which is the leading ProteomeXchange repository (201, 202)) along with the metadata associated with the data.

**Fig. 4.** The vision of the SysteMHC Atlas project. The SysteMHC Atlas project was launched in 2018 with the main goal of collecting, harmonizing and sharing immunopeptidomic data generated by the community. Over the next years, we anticipate that the growing MS community working on various immunopeptidomic-related research topics (e.g. cancer, autoimmunity, infectious diseases) will contribute to the expansion of the SysteMHC Atlas by uploading data into existing repositories of the ProteomeXchange consortium (mainly PRIDE and MassIVE; the others shown are PeptideAtlas, Panorama, iProX and jPOST). Those data will be reprocessed and harmonized using a uniform and open source computational pipeline developed under the SysteMHC Atlas project. We envision that ProteomeXchange resources will be integrated better with the SysteMHC Atlas in the near future, thereby facilitating the application of deep-learning approaches from quality-controlled immunopeptidomic “Big Data” to improve software tools for MHC peptide identification and quantification by DDA- and DIA-MS techniques.
Inspired by the vision and achievements of pioneers in genomics (213–216) and proteomics (217, 218), we envision that the ongoing enthusiasm in MS profiling of the human immunopeptidome will lead to the development of the IWAS paradigm, in which absolute quantification of immunopeptidomes at population-scale will become a reality. If tested and validated, we envision that IWAS will radically change the way we think about immunopeptidomics and will represent a new milestone in the history of the field. Achieving this goal will require tremendous technological development and community efforts; from highly standardized techniques for high throughput isolation of peptides to advanced knowledge about the baseline and dynamics of the human immunopeptidome, as well as appropriate data and computational resources to handle the analysis of very large immunopeptidomic datasets. With the advent of numerous new “biobank” type of studies, great opportunities will exist to collect samples for immunopeptidomic analyses—together with genomic, transcriptomic, proteomic and metabolomic analyses—that will impact personalised health care management and public health care policy in the future.

Acknowledgments — We thank the members of the Human Immunopeptidome Project for insightful discussions.

This work was supported by funding from the Fonds de recherche du Québec – Santé (FRQS), the Coe Foundation and the CHU Sainte-Justine and the Charles-Bruneau Foundations. JAV wants to acknowledge funding from Welcome Trust [grant number 208391/2/177/Z]), BBSRC [grant number BB/P024599/1] and the EPIC-XS project [grant number 823839], funded by the Horizon 2020 programme of the European Union. The authors declare that they have no conflicts of interest with the contents of this article.

Author contributions: J.A.V. and E.C. wrote the paper; J.A.V., P.K., K.A.K., J.D.D., I.M., B.P., E.W.D., and A.S. commented the manuscript; P.K. created the Graphical Abstract; Q.M. created the Figure 2B; I.S. and Q.M. commented the manuscript. Created the Figure 2B and supplemental Table S1.

REFERENCES

1. Venter, C. J., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., Gocayne, J. D., Amanatides, P., Ballew, R. M., Huson, D. H., Wortman, J., Zhang, Q., Kodira, C. D., Zheng, X. H., Chen, L., Skupsky, M., Subramanian, G., Thomas, P. D., Zhang, J., Miklos, G. L., Nelson, C., Broder, S., Clark, A. G., Nadeau, J., McKusick, V. A., Zinder, N., Levine, A. J., Roberts, R. J., Simon, M., Sleiman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., dhannenhall, Kravitz, S., Levy, S., Maberry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Francesco, V., Dunn, P., Elbeek, K., Evangelista, C., Gabrieli, A., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T. J., Higgins, M. E., Ji, R.-R., Ke, Z., Ketchum, K. A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G. V., Milshina, N., Moore, H. M., Naik, A. K., Narayan, V. A., Neelam, B., Nusskern, D., Rusch, D. B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M.-H., Wides, R., Xiao, C., Yan, C., Yao, A., Ye, J., Zhan, K.A.K. created the Figure 2B; I.S. and Q.M. commented the manuscript. Created the Figure 2B and supplemental Table S1.

CONCLUSION

Cohort studies of the post-GWAS era are increasingly impactful and an important driving force for revolutionizing healthcare (4). Recent GWAS-based evidences indicate that generation of digital immunopeptidome maps from thousands of individuals may unveil a vast array of immunopeptidomic signatures (or epitope biomarkers) that could eventually be used in the clinic to predict the susceptibility or resistance of an individual to immune diseases as well as an individual’s predisposition to respond to vaccines and immunotherapies.
Next-generation Immunopeptidomics

Suh, E., Thomas, R., Tint, N., Tse, S., Vech, C., Wang, G., Wetter, J., Moy, M., Moy, L., Murphy, B., Nelson, K., Pfannkoch, C., Pratts, E., Puri, Love, A., Mann, F., May, D., McCawley, S., McIntosh, T., McMullen, I., Cheng, M., Curry, L., Danaher, S., Davenport, L., Desilets, R., Dietz, H., Barnstead, M., Barrow, I., Beeson, K., Busam, D., Carver, A., Center, Zhu, S. C., Zhao, S., Gilbert, D., Baumhueter, S., Spier, G., Carter, C., Zhang, W., Zhang, H., Zhao, Q., Zheng, L., Zhong, F., Zhong, W., Zandieh, A., and Zhu, X. (2001) The sequence of the human genome. Science 291, 133-138.

2. Lander, E., Linton, L., Birren, B., Nusbaum, C., Zody, M., Baldwin, J., Devon, K., Dewar, K., et al. (2001) Initial sequencing and analysis of the human genome. Nature 409, 860-921.

3. Gambhir, S., Ge, J. T., Vemuganti, R., and Spitzer, R. (2018) Toward achieving precision health. Sci. Transl. Med. 10, eaao3612.

4. Wijmenga, C., and Zhernakova, A. (2018) The importance of cohort studies in the post-GWAS era. Nat. Genet. 50, 322-328.

5. Tardif, J.-C., Rheumàe, E., Perreault, L.-P., Grégoire, J. C., Zada, Y., Asselin, G., Provost, S., Barhdadi, A., Rhains, D., L’Allier, P. L., Ibrahim, R., Ulimanyi, R., Nesor, E. J., Bengeholi, R., Suchankova, G., Laghrissi-Thode, F., Guertin, M.-C., Olson, A. G., Mongrain, I., Kutikova, K. G., Courchesne, R., and Depouilly, M.-P. (2018) Pharmacogenomic determinants of the cardiovascular effects of Dalcetrapib. Circulation Cardiovasc. Genet. 11, 372-382.

6. Tardif, J.-C., Rhains, A., Brodeur, M., Zada, Y., Fouedjio, R., Provost, S., Boulé, M., Alem, S., Grégoire, J. C., L’Allier, P. L., Ibrahim, R., Guertin, M.-C., Mongrain, I., Olson, A. G., Schwartz, G. W., Rhéaume, E., and Dubé, M.-P. (2018) Genotype-dependent effects of dalcetrapib on cholesterol efflux and inflammation. Circulation Cardiovasc. Genet. 11, 340-348.

7. Hayden, E. (2017) The rise and fall and rise again of 23andMe. Nat. News 550, 174.

8. Ozaki, K., Ohnishi, Y., Iida, A., Sekine, A., Yamada, R., Tsunoda, T., Sat0, H., Sat0, H., Horii, M., Nakamura, Y., and Tanaka, T. (2002) Functional SNPs in the lymphotoxin-α gene that are associated with susceptibility to myocardial infarction and myocardial expression level on HIV control. Science 340, 87-91.

9. Weller, D., MacArthur, J., Morales, J., Burdett, T., Hall, J., Binkins, P., Klemm, I., Flicek, P., Manoilo, T., Hindorff, L., and Parkinson, H. (2014) The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res. 42, D1001–D1006.

10. Trowsdale, J., and Knight, J. C. (2012) Major histocompatibility complex genomics and human disease. Annu. Rev. Genom. Hum. Genet. 14, 301–323.

11. Kaur, G., Gras, S., Mobbs, J. I., Vivian, J. P., Cortes, A., Barber, T., Beksac, M., Gragert, L., Fingerson, S., Maiers, M., Zhang, M.-J., Albrecht, M., Tardif, J.-C., Rhainds, D., Provost, S., Barhdadi, A., Busam, D., Carver, A., Center, Zhu, S. C., Zhao, S., Gilbert, D., Baumhueter, S., Spier, G., Carter, C., Zhang, W., Zhang, H., Zhao, Q., Zheng, L., Zhong, F., Zhong, W., Zandieh, A., and Zhu, X. (2001) The sequence of the human genome. Science 291, 133-138.

12. Aponte, Q. Y., Carlsso, J. M., Chen, H., Gao, X., Thomas, R., Yuki, Y., Prete, G. O., Goulard, P., Brumme, Z. L., Brumme, C. J., John, M., Mallal, S., Nelson, G., Bosch, R., Heckerman, D., Stein, J. L., Soderberg, K. A., Moody, A. M., Denny, N. T., Zeng, X., Fang, J., Moffett, A., Lifson, J. D., Goedert, J. J., Buchbinder, S., Kirk, G. D., Fellay, J., McLaren, P., Deeks, S. G., Pereyra, F., Walker, B., Michael, N. L., Weinrott, A., Wolinsky, S., Liao, W., and Carrington, M. (2013) Influence of HLA-C expression level on HIV control. Science 340, 87-91.

13. Jung, E., Cheon, J., Lee, J., Park, S., Jang, H., Chung, S., Park, M., Kim, T.-G., Oh, H.-B., Bang, S.-K., Park, S., Han, J., Hong, S., Kim, T., Kim, W., and Lee, M. (2016) HLA-C*01 is a risk factor for Crohn’s disease. Inflamm. Bowel Dis. 22, 796–806.

14. Miyadera, H., Ohashi, Y., Lernmark, Å., Kitamura, T., and Tokunaga, K. (2015) Additive and interaction effects at three amino acid positions in HLA-DQ and HLA-DR molecules drive type 1 diabetes susceptibility. Nat. Genet. 47, 322–328.

15. Jung, E., Cheon, J., Lee, J., Park, S., Jang, H., Chung, S., Park, M., Kim, T.-G., Oh, H.-B., Bang, S.-K., Park, S., Han, J., Hong, S., Kim, T., Kim, W., and Lee, M. (2016) HLA-C*01 is a risk factor for Crohn’s disease. Inflamm. Bowel Dis. 22, 796–806.

16. Miyadera, H., Ohashi, Y., Lernmark, Å., Kitamura, T., and Tokunaga, K. (2015) Additive and interaction effects at three amino acid positions in HLA-DQ and HLA-DR molecules drive type 1 diabetes susceptibility. Nat. Genet. 47, 322–328.

17. Zhou, Z., and Jensen, P. E. (2013) Structural characteristics of HLA-DQ that may impact DM editing and susceptibility to type-1 diabetes. Front. Immunol. 4, 262.

18. Hu, X., Deutsch, A. J., Lenz, T. L., Onengut-Gumucio, S., Han, B., Chen, W.-M., Howson, J. M., Todd, J. A., de Bakker, P. I., Rich, S. S., and Raychaudhuri, S. (2015) Additive and interaction effects at three amino acid positions in HLA-DQ and HLA-DR molecules drive type 1 diabetes risk. Nat. Genet. 47, ng3353.

19. Bekes, M., Gragert, L., Fingerson, S., Maiers, M., Zhang, M.-J., Albrecht, M., Zhong, X., Cozen, W., Dispensieri, A., Lonial, S., and Hari, P. (2016) HLA polymorphism and risk of multiple myeloma. Leukemia 30, 2260.

20. Dendrou, C. A., Petersen, J., Rossjohn, J., and Fugger, L. (2018) HLA variation and disease. Nat. Rev. Immunol. 8, 193.
Next-generation Immunopeptidomics

Jöckel, K.-H., von Strandmann, E., Lightfoot, T., Kane, E., Roman, E., Lake, A., Montgomery, D., Jarrett, R. F., Sverdlov, A. J., Engert, A., Orr, N., Hemminki, K., and Houlston, R. S. (2017) Genome-wide association study of classical Hodgkin lymphoma identifies key regulators of disease susceptibility. Nat. Commun. 8, 1892

45. Okada, Y., Momozawa, Y., Sakaue, S., Kanai, M., Ishigaki, K., Akiyama, M., Kishikawa, T., Arai, Y., Sasaki, T., Kosaki, K., Suematsu, M., Matsumoto, K., Yamamoto, K., Kubo, M., Hirose, N., and Kamatani, Y. (2018) Deep-genome sequence mapping reveals recent selection signatures linked to evolution and disease risk of Japanese. Nat. Commun. 9, 1631

46. Wissmann, W. T., Hill-Burns, E. M., Zabetian, C. P., Factor, S. A., Kvistborg, P., and Yewdell, J. W. (2018) Enhancing responses to cancer therapy. Nat. Rev. Immunol. 18, 705–716

47. Chick, J. M., Munger, S. C., Simecek, P., Huttlin, E. L., Choi, K., Gatti, R., Thompson, W., Lem, R., Font-Burgada, J., Zanetti, M., and Williams, E. G., Wu, Y., Jha, P., Dubuis, S., Blattmann, P., Argmann, C. A., Wainberg, M., Sinnott-Armstrong, N., Mancuso, N., Barbeira, A. N., Nicholson, J. K., Holmes, E., and Elliott, P. (2008) The metabolome-wide association study of classical Hodgkin lymphoma identifies key regulators of disease susceptibility. Nat. Commun. 11, 592–599

48. Schenkelberg, T., Kieny, M.-P., Bianco, A., and Koff, W. C. (2015) Building the Human Vaccines Project: A roadmap for cancer vaccine development. Sci. Transl. Med. 7, 303ps7–303ps7

49. Koff, W. C., Burton, D. R., Johnson, P. R., Walker, B. D., King, C. R., Nabel, G. J., Ahmed, R., Bhan, M. K., and Plotkin, S. A. (2013) Accelerating next-generation vaccine development for global disease prevention. Expert Rev. Vaccines 14, 3105–3117

50. Caron, E., Kowalewski, D. J., Koh, C., Sturm, T., Schuster, H., and Aebersold, R. (2013) Analysis of Major Histocompatibility Complex (MHC) Immunopeptidomes using mass spectrometry. Mol. Cell. Proteomics 14, 629–636

51. Arlehamn, C. S., and Sette, A. (2017) T cells from patients with Parkinson disease are restricted by MHC-I molecules linked to evolution and disease risk of Japanese. Proc. Natl. Acad. Sci. U.S.A. 114, 656–661

52. Romero, P., Bachmayer, J., Bhadrirajna, M. P., Cockett, M., Disis, M. L., Dranoff, G., Gilboa, E., Hammond, S. A., Hershberg, R., Korman, A. J., Kvistborg, P., Miele, C., Mellman, I., Palucka, K. A., Reddenkopf, I., Schenkelberg, T., Schoenberg, S., Sosman, J., Türeci, Ö., den Eynede, B., Koff, W., and Coukos, G. (2016) The Human Vaccines Project: A roadmap for cancer vaccine development. Nat. Rev. Genet. 17, 582–587

53. Caron, E., Kowalewski, D. J., Koh, C., Sturm, T., Schuster, H., and Aebersold, R. (2013) Analysis of Major Histocompatibility Complex (MHC) Immunopeptidomes using mass spectrometry. Mol. Cell. Proteomics 14, 3105–3117

54. Purcell, A. W., McCluskey, J., and Rossjohn, J. (2007) More than one reason to rethink the use of peptides in vaccine design. Nat. Rev. Drug Discov. 6, nd2224

55. Hu, Z., Ott, P. A., and Wu, C. J. (2018) Towards personalized, tumour-specific, therapeutic vaccines for cancer. Nat. Rev. Immunol. 18, 168–182

56. Koff, W. C., Gust, I. D., and Plotkin, S. A. (2014) Toward a human vaccines project. Nat. Immunol. 15, 589–592

57. Caron, E., Kowalewski, D. J., Koh, C., Sturm, T., Schuster, H., and Aebersold, R. (2013) Accelerating next-generation vaccine development for global disease prevention. Clin. Infect. Dis. 54, 1329–132910

58. Sreenivasa, P., Haddad, E., Lemieux, S., Thibault, P., and Perreault, C. (2018) Noncodelimited antigen discovery for cancer immunotherapy. Curr. Opin. Immunol. 41, 1–17

59. Laumont, C. M., Vincent, K., Hesnard, L., Audemard, E., Bonnell, É., Lavender, J.-P., Gendron, P., Courcelles, M., Hardy, M.-P., Côté, C., Durette, C., St-Pierre, C., Benhammadi, M., Lanoix, J., Vobecky, S., Haddad, E., Lemieux, S., Thibault, P., and Perreault, C. (2018) Noncoding regions are the main source of targetable tumor-specific antigens. Nat. Genet. 50, 592–599

60. Tschakar, D. C., Croft, N. P., Doherty, P. C., and Gruta, N. L. (2015) Sizing up the key determinants of the CD8+ T cell response. Nat. Rev. Immunol. 15, 705–716

61. Marshall, S. S., Hallez, J.-L., Van Damme, J., Vandekerckhove, P., Van Damme, J., and Sire, S. (2014) Comparative immunoepitodomal and their pathogens. Proc. Natl. Acad. Sci. U.S.A. 111, 13268–13272

62. Nicholson, J. K., Holmes, E., and Elliott, P. (2008) The metabolome-wide association study of classical Hodgkin lymphoma identifies key regulators of disease susceptibility. J. Proteome Res. 7, 3637–3638

63. Holmes, E., Loo, R., Stamler, J., Bictash, M., Yap, I. K., Chan, O., Ebels, T., Iorio, M., Brown, J. I., Velkov, K. A., Daviglus, M. L., Kesteloot, H., Ueshima, H., Zhao, L., Nicholson, J. K., and Elliott, P. (2008) Human metabolic phenotype diversity and its association with diet and blood pressure. Nature 453, 396

64. Janssen, J., van der Burg, S. H., Straten, P., Poulsen, H., Ponsati, O., Okada, H., Ramussen, H.-G., Veerle, M., and van der Burg, S. H. (2016) Proteome-wide association studies identify biochemical modules associated with a wing-size phenotype in Drosophila melanogaster. Nat. Commun. 7, 12649

65. Butter, F., Davison, L., Viturawong, T., Scheibe, M., Vermeulen, M., Todd, J. A., and Mann, M. (2012) Proteome-wide analysis of disease-associated SNPs that show allele-specific transcription factor binding. Plos Genet. 8, e1002982

66. Shraibman, B., Bamea, E., Kadoch, D., Haimovich, Y., Slobodin, G., Rosner, I., López-Larrea, C., Hilt, N., Kuttruff, S., Song, C., Britten, C., Castle, J., Kreiter, S., Frenzel, K., Tatagiba, M., Tabatabai, G., Dietrich, P.-Y., Dutoit, V., Wick, W., Platten, M., Winkler, F., and Deimling, A., Krop, J., Sahuquillo, J., Martinez-Ricarte, F., Rodon, J., Lassen, U., Ottenmeier, C., van der Burg, S. H., Stratton, P., Poulshen, L., Ponsati, O., Okada, H., Ramussen, H.-G., Sahin, U., Singh, H., and Admon, A. (2019) Identification of tumor antigens among the HLA peptides of cancerous melanomas and plasma. Mol. Cell. Proteomics 18, 1265–1268
and van Hall, T. (2018) Identification of non-mutated neoantigens presented by TAP-deficient tumors. J. Exp. Med. 215, 2325–2337

78. Havel, J. J., Chowell, D., and Chan, T. A. (2019) The evolving landscape of biomarkers for checkpoint inhibitor immunotherapy. Nat. Rev. Cancer 19, 133–150

79. Etx, G., and Perreault, C. (2019) Discovery and characterization of actionable tumor antigens. Genome Med. 11, 29

80. Wilson, E. A., and Anderson, K. S. (2018) Lost in the crowd: identifying targetable MHC class I neoepitopes for cancer immunotherapy. Expert Rev. Proteomic 15, 1065–1077

81. Ramarathnam, H., Croft, N. P., Illing, P. T., Faridi, P., and Purcell, A. W. (2018) Employing proteomics in the study of antigen presentation: an update. Expert Rev. Proteomic 124, 637–645

82. Bräunlein, E., and Krackhardt, A. M. (2017) Tools to define the melanoma-associated phosphopeptide identification by mass spectrometry. Immunology 154, 331–345

83. Schmidt, M., and Lill, J. R. (2018) MHC class I presented antigens from malignancies: A perspective on analytical characterization & immunogenicity. J. Proteomics 191, 48–57

84. Grancos, D. P., Hermont, C. M., Thibaut, P., and Perreault, C. (2015) The nature of self for T cells—a systems-level perspective. Curr. Opin. Immunol. 34, 1–8

85. Laumont, C. M., and Perreault, C. (2018) Exploiting non-canonical translation to identify new targets for T cell-based cancer immunotherapy. Cell. Mol. Life Sci. 75, 607–621

86. Gfeller, D., and Bassani-Sternberg, M. (2018) Predicting antigen presentation by the MHC II tetramer: how to learn from a million peptides? Front. Immunol. 9, 1716

87. Purcell, A. W., Croft, N. P., and Tscharke, D. C. (2016) Immunology by numbers: quantitation of antigen presentation completes the quantitative milieu of systems immunology. Curr. Opin. Immunol. 40, 88–95

88. Croft, N. P., Purcell, A. W., and Tscharke, D. C. (2015) Quantifying epitope presentation using mass spectrometry. Mol. Immunol. 68, 1–4

89. Stern, L. J., and Santambrogio, L. (2016) The melting pot of the MHC II tetramer. Curr. Opin. Immunol. 40, 70–77

90. Faridi, P., Purcell, A. W., and Croft, N. P. (2018) In immunoproteomics we need a sniper instead of a shotgun. Proteomics 18, 1700464

91. Alvarez, B., Barra, C., Nielsen, M., and Andreatta, M. (2018) Computational tools for the identification and interpretation of sequence motifs in immunoproteomics. Proteomics 18, e1700252

92. Creech, A. L., Ting, Y. S., Goulding, S. P., Sauld, J. F., Barthelme, D., Roetzschke, O., Falk, K., Deres, K., Schild, H., Norda, M., Metzger, J., Jung, D., and van Hall, T. (2018) Identification of non-mutated neoantigens presentable tumor antigens by the presentation of peptides phosphorylated at their main anchor. J. Immunol. 189, 133–150

93. Marcilla, M., Alpízar, A., Lombardía, M., and Gfeller, D. (2018) The C-terminal extension landscape of naturally presented HLA-I ligands. Nat. Communications 9, 348–357

94. Abramovich, S., Cummings, K. L., Penny, S., Turner, J. E., Cottine, J., Abelin, J. G., Cummings, K. L., Penny, S., Turner, J. E., Cottine, J., Abelin, J. G., Parmentier, M., and Albar, J. (2014) Increased diversity of the HLA-B40 ligandome facilitated by the presentation of phosphorylated peptides by HLA-B40. J. Immunol. 192, 348–357

95. Gubin, M. M., Zhang, X., Schuster, H., Caron, E., Ward, J. P., Noguchi, T., Ivanova, Y., Hundal, J., Arthur, C. D., Krebber, W.-J., Mulder, G. E., Toebes, M., Vesely, M. D., Lam, S. S., Korman, A. J., Allison, J. P., Freeman, G. J., Sharpe, A. H., Pearce, E. L., Schumacher, T. N., Aeberold, S., Rammensee, H.-G., Meleif, C. J., Mardis, E. R., Gillanders, W. E., Artyomov, M. N., and Schreiber, R. D. (2014) Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens. Nature 515, 577

96. Adeva, M., Juñhunwala, S., Phung, O. T., Lupardus, P., Tanguay, J., Bumbaca, S., Franci, C., Cheung, T. K., Fritsche, J., Weinschenk, T., Modrusan, Z., Meillan, I., Lill, J. R., and Delamarre, L. (2015) Predicting immunogenic tumor mutations by combining mass spectrometry and exome sequencing. Nature 515, 572–576

97. Carreno, B. M., Magrini, V., Becker-Hapak, M., Kasibinejadian, S., Hundal, J., Pett, A. A., Ly, A., Lie, W.-R., Hildebrand, W. H., Mardis, E. R., and Schreiber, R. D. (2015) A dendritic cell vaccine increases the breadth and diversity of melanoma neoantigen-specific T cells. Science 348, 803–808

98. Pynn, P., Illing, P. T., Ramarathnam, H., O’Connor, G. M., Hughes, V. A., Hitchen, C., Price, D. A., Ho, B. K., McVicar, D. W., Brooks, A. G., Purcell, A. W., Rossjohn, J., and Vivian, J. P. (2017) MHC-I peptides get out of the groove and enable a novel mechanism of HIV-1 escape. Nat. Struct. Mol. Biol. 24, 849–856

99. McMurtrey, C., Trolle, T., Sansom, T., Remesh, S. G., Kaeve, T., Bardet, W., Jackson, K., McLeod, R., Sette, A., Nielsen, M., Zajonc, D. M., Blader, I. J., Peters, B., and Hildebrand, W. (2016) Toxoplasma gondii peptide ligands open the gate of the HLA class I binding groove. Elife 5, e12556

100. Guillaume, P., Picaud, S., Baumgaertner, P., Montandont, N., Schmidt, J., Speiser, D. E., Coukos, G., Rosenberg-Stemberg, M., Filipakopoulos, P., and Gfeller, D. (2018) The C-terminal extension landscape of naturally presented HLA-I ligands. Proc. Natl. Acad. Sci. USA 115, 5083–5088

101. Alpizar, A., Marino, F., Ramos-Fernandez, A., Lombardia, M., Jeko, A., Pazos, F., Padarla, A., Santiago, C., Heck, A. J., and Marcilla, M. (2017) A molecular basis for the presentation of phosphorylated peptides by HLA-B antigens. Mol. Cell. Proteomics 16, 181–193

102. Marcilla, M., Alpizar, A., Lombardia, M., Ramos-Fernandez, A., and Albar, J. (2014) Increased diversity of the HLA-B40 ligandome by the presentation of peptides phosphorylated at their main anchor residue. Mol. Cell. Proteomics 13, 462–474

103. Abelin, J. G., Trantham, P. D., Penny, S. A., Patterson, A. M., Ward, S. T., Hildebrand, W. H., Cobbold, M., Bai, D. L., Shabanowitz, J., and Hunt, D. F. (2015) Complementary IMAC enrichment methods for HLA-associated phosphopeptide identification by mass spectrometry. Nat. Protoc. 10, 1308–1318

104. Cobbold, M., Peña, H., Norris, A., Polefrone, J. M., Gian, J., English, A., Cummings, K. L., Penny, S., Turner, J. E., Cottine, J., Abelin, J. G., Malaker, S. A., Zarling, A. L., Huang, H.-W., Goodyear, O., Freeman, S. D., Shabanowitz, J., Pratt, G., Craddock, C., Williams, M. E., Hunt, D. F., and Engelhard, V. H. (2013) MHC class I-associated phospho-
Next-generation Immunopeptidomics

peptides are the targets of memory-like immunity in leukemia. Sci. Transl. Med. 5, 203ra125

116. Marino, F., Mommense, G. P., Jako, E., Meiring, H. D., van den Brink, J. A., Schelter, R. A., van Els, C. A., and Heck, A. J. (2016) Arginine (d)metylated human leukocyte antigen class I peptides are favorably presented by HLA-B^*07. J. Proteome Res. 16, 34–44

117. Sidney, J., Becart, S., Zhou, M., Duffy, K., Lindvall, M., Moore, E. C., Moore, E. L., Rao, T., Rao, N., Nielsen, M., Peters, B., and Sette, A. (2017) Cytodestruction only infrequently impacts peptide binding to HLA class II MHC. Plos One 12, e0177140

118. Molberg Ø., Mcadam, S. N., Körner, R., Quarsten, H., Kristiansen, C., Sidney, J., Becart, S., Zhou, M., Duffy, K., Lindvall, M., Moore, E. C., Moore, E. L., Rao, T., Rao, N., Nielsen, M., Peters, B., and Sette, A. (2017) Cytodestruction only infrequently impacts peptide binding to HLA class II MHC. Plos One 12, e0177140

119. Bodd, M., Kim, C., Lundin, K., and Sollid, L. M. (2012) T-cell response to glut in patients with HLA-DQ2.2 reveals requirement of peptide-MHC stability in celiac disease. Gastroenterology 142, 552–561

120. Fallall, L.-E., Bergseng, E., Hotta, K., Berg-Larsen, A., Kim, C.-Y., and Ild, L. (2009) Differences in the risk of celiac disease associated with HLA-DQ2.3 or HLA-DQ2.2 are related to sustained gluten antigen presentation. Nat. Immunol. 10, 1096–1101

121. Hovhannisyan, Z., Weiss, A., Martin, A., Wiesner, M., Tollefsen, S., Young, A., Derksen, M., Koning, F., Teyton, L., and Jabri, B. (2008) The role of HLA-DQ8 peptide-MHC in celiac disease. Nature 456, 534–538

122. Ramarathinam, H., Gras, S., Alcantara, S., Yeung, A., Mifsud, N. A., Mishto, M., and Liepe, J. (2017) Post-translational peptide splicing and T cell peptides are the targets of memory-like immunity in leukemia. J. Immunol. 198, 10922–10925

123. Andreatta, M., Nicasira, A., Peng, X., Hancock, G., Dorrell, L., Ternette, N., and Nielsen, M. (2019) MS-rescue: a computational pipeline to increase the quality and yield of immunopeptidomics experiments. Proteomics 19, 1800357

124. Nielsen, M., Lundegaard, C., Blicher, T., Lambeth, K., Hamdal, M., Justesen, S., Røder, G., Peters, B., Seette, A., Lund, O., and Buus, S. (2007) NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and -B locus protein of known sequence. Plos One 2, e796

125. Dhand, S., Mahajan, S., Paul, S., Yan, Z., Kim, H., Jespersen, M., Jurtz, V., Andreatta, M., Greenbaum, J. A., Marcutili, P., Sette, A., Nielsen, M., and Peters, B. (2019) IEDB-AR: immune epitope database—analysis resource in 2019. Nucleic Acids Res. 47, W502–W506

126. O’Donnell, T. J., Rubinstein, A., Bonsack, M. A., Laserson, U., and Hammerbacher, J. (2018) MHCFlurry: open-source class I MHC interaction affinity predictor. Biol. Direct. 13, 23

127. Couzin-Frankel, J. (2013) Breakthrough of the year 2013. Cancer immunother. Science 342, 1432–1433

128. Ribas, A., and Wolchok, J. D. (2018) Cancer immunotherapy using checkpoint blockade. Science 359, 1350–1355

129. Schuster, H., Peiper, J. K., Bösömer, H.-C., Röhlé, K., Backert, L., Billich, T., Ney, B., Löfler, M. W., Kowalewski, D. J., Trautwein, N., Rabenstein, N., Mifsud, L., Rau, T., Braun, S., Haen, S. P., Waltz, J. S., Schmid-Horch, B., Brucker, S. Y., Walliënner, D., Kohlbacher, O., Fend, F., Rammensee, H.-G., Stevanovic, S., Staeblé, A., and Wagner, P. (2017) The immunopeptidomics landscape of ovarian carcinomas. Proc. Natl. Acad. Sci. U.S.A. 114, E9942–E9951

130. Hilt, N., Kuttruff-Coqui, S., Srenzen, K., Bukur, V., x001007, S., coute Gouttefanges, C., Platten, M., Tabatahab, G., Dutot, V., Burg, S. H., thon, H., et al. (2015) Identification of native and posttranslationally modified HLA-B^*07:01-restricted HIV envelope derived epitopes using immunoproteomics. Proteomics 18, 1700253

131. Malaker, S. A., Ferracane, M. J., Depontieu, F. R., Zarling, A. L., Shabanowitz, J., Bai, D. L., Topalian, S. L., Engelhard, V. H., and Hunt, D. F. (2016) Identification and characterization of complex glycosylated peptides presented by the MHC class II processing pathway in melanoma. J. Proteome Res. 16, 228–237

132. Dengel, J., Ramnæs, H., and Stevanovic, S. (2005) Glycin side chains on naturally presented MHC class II ligands. J. Mass Spectrom. 40, 100–104

133. Marino, F., Bemm, M., Mommense, G. P., Leney, A. C., van den Brink, J. A., Bonvin, A. M., Becker, C., van Els, C. A., and Heck, A. J. (2015) Extended O-GlcNAc on HLA class-I-bound peptides. J. Am. Chem. Soc. 137, 10922–10925

134. Faridi, P., Liu, C., Ramarathinam, H., Vivian, J. P., Illing, T., Mifsud, N. A., Ayala, R., Song, J., Gearing, L. J., Hertzog, P. J., Ternette, N., Thomas, S. R., Kent, S. J., Ternette, N., Purcell, D. F., Rossjohn, J., and Purcell, A. W. (2018) Identification of native and posttranslationally modified HLA-B^*07:01-restricted HIV envelope derived epitopes using immunoproteomics. Proteomics 18, 1700253

135. Le, D. T., Durham, J. N., Smith, K. N., Wang, H., Bartlett, B. R., Aulakh, L. K., Lu, S., Kembering, H., Witt, C., Luber, B. S., Wong, F., Azad, N. S., Rucki, A. A., Laheru, D., Donehower, R., Zaheer, A., Fisher, G. A., Crocenzi, T. S., Lee, J. J., Greten, T. F., Duffy, A. G., Combror, K. K., Eyring, A. D., Lam, B. H., Joe, A., Kang, P. S., Holdhoff, M., Danilova, L., Cope, L., Meyer, C., Zhou, S., Goldberg, R. M., Armstrong, K. D., Bever, K. M., Fader, A. N., Taube, J., Houssseau, F., Spetzler, D., Xiao, N., Pardoll, D. M., Papadopoulos, N., Kinzler, K. W., Eshleman, J. R., Vogelstein, B., Anders, R. A., and Diaz, L. A. (2017) Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. Science 357, 409–413

136. Le, D. T., Uram, J. N., Wang, H., Bartlett, B. R., Kemberling, H., Eyring, A. D., Skora, A. D., Luber, B. S., Azad, N. S., Laheru, D., Biedrzycki, B., Donehower, R. C., Zaheer, A., Fisher, G. A., Crocenzi, T. S., Lee, J. J., Greten, T. F., Duffy, A. G., Combror, K. K., Eyring, A. D., Lam, B. H., Joe, A., Kang, P. S., Holdhoff, M., Danilova, L., Cope, L., Meyer, C., Zhou, S., Goldberg, R. M., Armstrong, K. D., Bever, K. M., Fader, A. N., Taube, J., Houssseau, F., Spetzler, D., Xiao, N., Pardoll, D. M., Papadopoulos, N., Kinzler, K. W., Eshleman, J. R., Vogelstein, B., Anders, R. A., and Diaz, L. A. (2017) Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. Science 357, 409–413
Rosenberg, J., Iyer, G., Bochner, B. H., Bajorin, D. F., Al-Ahmade, H. A.,
Chaff, J. E., Rudin, C. M., Riely, J. G., Baxi, S., Ho, A. L., Wong, R. J.,
Pfister, D. G., Wolcho, J. D., Barker, C. A., Gutin, P. H., Brennan, C. W.,
Tabar, V., Mellinghoff, I. K., DeAngelis, L. M., Arjan, C. E., Lee, N., Tap,
W. D., Gounder, M. M., D’Angelo, S. P., Saltz, L., Stadler, Z. K., Scher,
H. I., Baserga, J., Razavi, P., Klebanoff, C. A., Yager, R., Segal, N. H.,
Ko, G. Y., DeMatteo, R. P., Ladanyi, M., Riciz, N. A., Berger, M. F., Riaz,
Solit, D. B., Chan, T. A., and Morris, L. G. (2019) Tumor mutation load
predicts survival after immunotherapy across multiple cancer types. Nat.
Genet. 51, 202–206

145. June, C. H., O’Connor, R. S., Kawalekar, O. U., Ghassemi, S., and Milone,
M. C. (2018) CAR T cell immunotherapy for human cancer. Science 359,
1361–1365

146. Kelly, P. N. (2018) The cancer immunotherapy revolution. Science 359,
Proteome Res. 18, 1634–1643

147. Bouchie, A., and DeFrancesco, L. (2016) Nature Biotechnology’s aca-
tademic spinouts of 2015. Nat. Biotechnol. 34, 484

148. Ternette, N., and Purcell, A. W. (2018) Immunopeptidomics special issue.

149. Marino, F., Chong, C., Michaux, J., and Bassani-Sternberg, M. (2019)
Science 365, 1634–1643

150. Bijen, H. M., Hassan, C., Kester, M. G., Janssen, G. M., Hombrink, P., Ru,
Drijfhout, J. W., Spaapen, R. M., Jimenez, C. R., Heemskerk, M., Falk-
ner, C., and van Veeelen, P. A. (2014) Accurate quantification of
MHC-bound peptides by application of isotopically labeled peptide
MHCPlexj. J. Proteomics 109, 240–244

151. Bertos, T., Bergeron, B. A., Chappell, T., Meredith, G., Sheffield, P.,
Gray, P., Hajimoto, M., Pope, M., Predki, P., Kollori, M., Hincapie, M.,
Hancock, W. S., Jia, W., Song, L., Li, L., Wei, J., Yang, B.,
Wang, J., Ying, W., Zhang, Y., Cai, Y., Qian, X., He, F., Meyer, H. E.,
Stephan, C., Eisenacher, M., Marcus, K., Langenfeld, E., May, C.,
Carr, S. A., Ahmad, R., Zhu, W., Smith, J. W., Hanash, S. M., Struthers, J. J.,
Wang, Z., Zhang, Q., An, Y., Goldman, R., Carls, E. R., van der Post,
S., Hung, K. E., Sarracino, A. D., Parker, K., Krastiras, B., Kucherlapati,
P., Brasara, S., Poirier, E. S., Kapp, E., Rassenti, L., Moritz, R.,
Simpson, R., Houle, B., LaBoissiere, S., Metalinov, P., Nguyen,
Pawson, T., Wong, C. C., Cociorva, D., III, J. R., Ellison, M. J., Lopez-
Campistol, A., Semchuk, P., Wang, Y., Ping, P., Elia, G., Dunn, M. J.,
Wynne, K., Walker, A. K., Strahler, J. R., Andrews, P. C., Hood, B. L.,
Bigbee, W. L., Conrads, T. P., Smith, D., Borchers, C. H., Lajoie, G. A.,
Bennall, S. C., Speicher, K. D., Speicher, D. W., Fujimoto, M., Naka-
nur, K., Pak, Y. K., Krmpotic, A., Vázquez, S. K., Zäng, A., Miller, C. A.,
Chung, A., Miller, C. A., Grimm, R., Williams, K., Dorschel, C., Falkner,
J. A. M., Martens, L., and Vizcaíno, J. (2009) A HUPO test sample study
reveals common problems in mass spectrometry–based proteomics.
Nat. Methods 6, 423–430

152. Ziegenhain, C., Vieth, B., Parekh, S., Reinuis, B., Guillamet-Adkins, A.,
Smet, M., Leonhardt, H., Heyn, H. H., Hellmann, I., and Enard, W. (2017)
Comparative analysis of single-cell RNA sequencing methods. Mol. Cell
63, 631–643.e4

153. Svensson, V., Natarajan, K., Ly, L.-H., Miragaia, R. J., Labalette, C.,
Macaulay, I. C., Cvejic, A., and Teichmann, S. A. (2017) Power analysis of
single-cell RNA-sequencing experiments. Nat. Methods 14, 381–387

154. Collins, B. C., Hunter, C. G., Liu, Y., Schilling, B., Rosenberger, G., Bader,
S. L., Chan, D. W., Gibson, B. W., Gingeras, A.-C., Held, J. M., Hirayama-
Kurogi, M., Hou, G., Kris, C., Larsen, B., Lin, L., Liu, S., Molloy, M. P.,
Moritz, R. L., Ohtsuji, S., Schlapbach, R., Selevsek, N., Thomas, S. N.,
Tseng, S.-C., Zhang, H., and Aebersold, R. (2017) Multi-laboratory assessment of reproducibility, qualitative and quantitative performance of
SWATH-mass spectrometry. Nat. Commun. 8, 1–11

155. Goodwin, S., McPherson, J. D., and McCombie, R. W. (2016) Coming of
age: ten years of next-generation sequencing technologies. Nat. Rev.
Genet. 17, 333–351

156. Swaminathan, J., Bouglakov, A. A., Hernandez, E. T., Bardo, A. M.,
Bachman, J. L., Marotta, J., Johnson, A. M., Anslyn, E. V., and Marcotte,
E. M. (2018) Highly parallel single-molecule identification of proteins in
tozeome-scale mixtures. Nat. Biotechnol. 36, 1–16

157. Collins, B. C., and Aebersold, R. (2018) Proteomics goes parallel. Nat.
Biotechnol. 36, 1051

Molecular & Cellular Proteomics 19.1
Purcell, A. W., Marcilla, M., Paradela, A., Wang, Q., Costello, C. E., Ternette, N., van Veelen, P. A., van Els, C. A. C. M., Heck, A. J. R., de Souza, G. A., Soloid, L. M., Admon, A., Stevanovic, S., Rammensee, H. G., Thibault, P., Perreault, C., Bassani-Sternberg, M., Aebersold, R., and Caron, E. (2018) The SysteMHC Atlas project. *Nucleic Acids Res.* **46**, D1237–D1247

200. Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D. J., Inuganti, A., Griss, J., Mayer, G., Eisenacker, M., Pérez, E., Uszkoreit, J., Pfeuffer, J., Sachsenberg, T., Yilmaz, S., Tiwary, S., Cox, J., Audain, E., Walzer, M., Jamuczak, A. F., Ternent, T., Brazma, A., and Vizcaíno, J. (2018) The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* **47**, D442–D450

201. Vizcaíno, J. A., Deutsch, E. W., Wang, R., Csordas, A., Reisinger, F., Rios, D., Dianes, J. A., Sun, Z., Farrah, T., Bandeira, N., Binz, P.-A., Xenarios, I., Eisenacker, M., Mayer, G., Gatto, L., Campos, A., Chakley, R. J., Kraus, H.-J., Albar, J., Martinez-Bartolomé, S., Apweiler, R., Omenn, G. S., Martens, L., Jones, A. R., and Hermjakob, H. (2014) ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat. Biotechnol.* **32**, 223–226

202. Deutsch, E., Csordas, A., and Sun, Z. (2016) The ProteomeXchange consortium in 2017: supporting the cultural change in proteomics public data deposition. *Nucleic Acids Res.* **45**, D1100–D1106

203. Lill, J. R., van Veelen, P. A., Tenzer, S., Admon, A., Caron, E., Elias, J. E., Heck, A. J., Marcilla, M., Marino, F., Müller, M., Peters, B., Purcell, A., Sette, A., Sturm, T., Ternette, N., Vizcaíno, J., and Bassani-Sternberg, M. (2018) Minimal information about an immuno-peptidomics experiment (MIAIPE). *Proteomics* **18**, e1800110

204. Karosiene, E., Lundegaard, C., Lund, O., and Nielsen, M. (2012) NetMHCons: a consensus method for the major histocompatibility complex class I predictions. *Immunogenetics* **64**, 177–186

205. Stteynberg, D., Nesvizhskii, A., Moritz, R., and Deutsch, E. (2013) Combining Results of Multiple Search Engines in Proteomics. *Mol. Cell. Proteomics* **12**, 2383–2393

206. Stteynberg, D., Deutsch, E., Lam, H., Eng, J., Sun, Z., Tasman, N., Mendoza, L., Moritz, R., Aebersold, R., and Nesvizhskii, A. (2011) iProphet: multi-level integrative analysis of shotgun proteomic data improves peptide and protein identification rates and error estimates. *Mol. Cell. Proteomics* **10**, M111.007690

207. Lam, H., Deutsch, E. W., Eddes, J. S., Eng, J. K., Stein, S. E., and Aebersold, R. (2008) Building consensus spectral libraries for peptide identification in proteomics. *Nat. Methods* **5**, 873–875

208. Asquith, M., Sternes, P. R., Costello, M., Karstens, L., Diamond, S., Martin, T. M., Li, Z., Marshall, M. S., Spector, T. D., le Cao, K., Rosenbaum, J. T., and Brown, M. A. (2019) HLA alleles associated with risk of ankylosing spondylitis and rheumatoid arthritis influence the gut microbiome. *Arthritis Rheumatol.* **71**, 1642–1650

209. Caron, E., Espona, L., Kowalewski, D. J., Schuster, H., Ternette, N., Alpizar, A., Schittenhelm, R. B., Ramaratham, H., Arienhamn, C. S., Koh, C., Gillet, L. C., Rabsteyn, A., Navarro, P., Kim, S., Lam, H., Sturm, T., Marcilla, M., Sette, A., Campbell, D. S., Deutsch, E. W., Moritz, R. L., Purcell, A. W., Rammensee, H.-G., Stevanovic, S., and Aebersold, R. (2015) An open-source computational and data resource to analyze digital maps of immunopeptidomes. *Elife* **4**

210. Ritz, D., Kirzi, J., Neri, D., and Fugmann, T. (2017) Data-independent acquisition of HLA class I peptide analyses using data-dependent and independent acquisition. *Proteomics* **18**, e1700246

211. Ritz, D., Sani, E., Debiec, H., Ronco, P., Neri, D., and Fugmann, T. (2018) Membranal and blood-soluble HLA class II peptide analyses using data-dependent and independent acquisition. *Proteomics* **18**, e1700246

212. Schuster, H., Shao, W., Weiss, T., Pedrioli, P. G., Roth, P., Weller, M., Campbell, D. S., Deutsch, E. W., Moritz, R. L., Planz, O., Rammensee, H.-G., Aebersold, R., and Caron, E. (2018) A tissue-based draft map of the murine MHC class I immunopeptidome. *Sci. Data* **5**, 180157

213. Hood, L., and Price, N. D. (2014) Demystifying disease, democratizing health care. *Sci. Transl. Med.* **6**, 225ed5

214. Hood, L., and Friend, S. H. (2011) Predictive, personalized, preventive, participatory (P4) cancer medicine. *Nat. Rev. Clin. Oncol.* **8**, 184–187

215. Auffray, C., Charron, D., and Hood, L. (2010) Predictive, preventive, personalized and participatory medicine: back to the future. *Genome Med.* **2**, 57

216. Green, E. D., Watson, J. D., and Collins, F. S. (2015) Human Genome Project: twenty-five years of big biology. *Nat. News* **526**, 29–31

217. Aebersold, R., and Mann, M. (2016) Mass-spectrometric exploration of proteome structure and function. *Nature* **537**, 347–355

218. Ahrens, C. H., Brunner, E., Qeli, E., Basler, K., and Aebersold, R. (2010) Generating and navigating proteome maps using mass spectrometry. *Nat. Rev. Mol. Cell. Biol.* **11**, 789–801

---

**Next-generation Immunopeptidomics**