

The nature of self for T cells — a systems-level perspective

Diana P Granados^{1,2}, Céline M Laumont^{1,2}, Pierre Thibault^{1,3}
 and Claude Perreault^{1,2}



T-cell development and function are regulated by MHC-associated self peptides, collectively referred to as the immunopeptidome. Large-scale mass spectrometry studies have highlighted three key features of the immunopeptidome. First, it is not a mirror of the proteome or the transcriptome, and its content cannot be predicted with currently available bioinformatic tools. Second, the immunopeptidome is more plastic than previously anticipated, and is molded by several cell-intrinsic and cell-extrinsic factors. Finally, the complexity of the immunopeptidome goes beyond the 20-amino acids alphabet encoded in the germline, and is not restricted to canonical reading frames. The large amounts of 'dark matter' in the immunopeptidome, such as polymorphic, cryptic and mutant peptides, can now be explored using novel proteogenomic approaches that combine mass spectrometry and next-generation sequencing.

Addresses

¹ Institute for Research in Immunology and Cancer, Université de Montréal, P.O. Box 6128, Downtown Station, Montreal, QC, Canada H3C 3J7

² Department of Medicine, Université de Montréal, P.O. Box 6128, Downtown Station, Montreal, QC, Canada H3C 3J7

³ Department of Chemistry, Université de Montréal, P.O. Box 6128, Downtown Station, Montreal, QC, Canada H3C 3J7

Corresponding author: Perreault, Claude
 (claud.perreault@umontreal.ca, c.perreault@videotron.ca)

Current Opinion in Immunology 2015, 34:1–8

This review comes from a themed issue on **Antigen processing**

Edited by **Nilabh Shastri** and **Jonathan Yewdell**

<http://dx.doi.org/10.1016/j.coi.2014.10.012>

0952-7915/© 2014 Elsevier Ltd. All right reserved.

Introduction

Recognition of self has a pervasive influence on the development and functions of the immune system because the adaptive lymphocytes of jawed vertebrates are selected on self-molecules, sustained by self-molecules, and activated in the presence of self-molecules [1]. This raises the fundamental question: what is the molecular definition of self for the adaptive immune system? We will discuss how large-scale mass spectrometry (MS) studies have recently

allowed systems immunologists to unveil the molecular composition of the MHC-restricted immunopeptidome. Since most MS studies have been performed on peptides presented by MHC class I molecules (HLA class I in humans), they will represent the main focus of this article.

The origin and role of the immune self recognized by CD8T cells

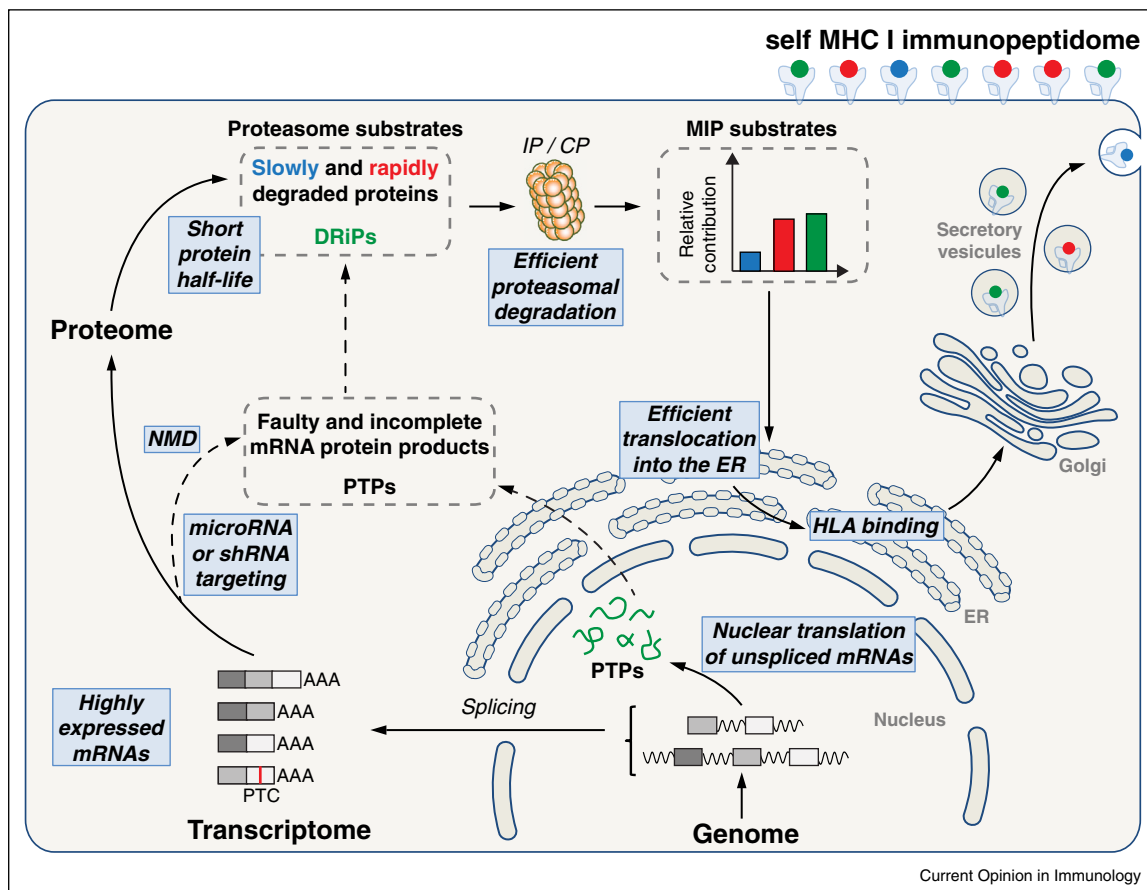
The T cell receptor (TCR) of classic adaptive CD8T cells recognizes MHC I-associated peptides (MIPs). In the absence of infection or phagocytosis, all MIPs derive from endogenous self proteins: these self MIPs are collectively referred to as the self MHC I immunopeptidome (SMII). The SMII regulates positive and negative selection in the thymus, as well as survival and reactivity of CD8T cells to non-self in the periphery [2]. The dominant paradigm holds that MIPs derive primarily from proteasomal digestion of rapidly degraded proteins (RDPs), which are degraded by the ubiquitin proteasome system during or within one hour after their synthesis [3^{••}]. In line with this, the majority of ubiquitination events in cells occurs on newly synthesized proteins [4]. A large proportion of RDPs are defective ribosomal products (DRiPs) (Figure 1) [5,6]. In theory, many processes can lead to DRiPs formation, but their relative importance remains elusive [3^{••},5]. Most, though not all authors [7], consider that the contribution of old proteins (retirees) to the SMII is relatively modest [5], except under two circumstances: (i) following irradiation, which leads to rapid degradation of old proteins [8], and (ii) in the case of cross-presented MIPs because they derive from endocytosis of stable proteins [9]. A schematic overview of the processes involved in the genesis of the SMII is shown in Figure 1.

The SMII is not a mirror of the proteome

Overview of the proteome and the SMII

Large-scale MS offers the sole direct approach to analyze the molecular composition of the proteome and the SMII [10–13]. The proteome of human cells contains about 10⁴ different proteins with a mean length of 449 amino acids [12]. The number of protein copies per cell spans a concentration range of seven orders of magnitude (up to 20 × 10⁶) [12]. In HeLa cells, the 40 most abundant proteins comprised 25% of the proteome and the most abundant 600 proteins constitute 75% of the proteome mass [11]. It has been estimated that a typical nucleated cell expresses at the cell surface about 1–5 × 10⁵ MHC class I molecules that present about 10⁴ different MIPs,

Figure 1



Overview of the major processes involved in the genesis of the SMII. Text in rectangular boxes indicates features whose contribution to MIP generation is demonstrated or suspected. The role of transporters, chaperones and non-proteasomal proteases is not depicted here. CP, constitutive proteasomes; ER, endoplasmic reticulum; IP, immunoproteasomes; NMD, nonsense-mediated decay; PTC, premature termination codons; PTP, pioneer translation products.

95% of which have a median length of 9 amino acids [10,14–16]. Hence, the SMII would represent at most 2% of the proteome: 10^4 MIPs @ 9 amino acids per MIP deriving from a proteome containing 10^4 proteins @ 449 amino acids/protein [10,12,14,15]. MIPs derive from a wide variety of proteins distributed in all cell compartments, encoded by genes located in all chromosomes [14,17,18^{••}]. Simply by virtue of their size, longer proteins generate more MIPs than shorter proteins [14,15,19]. For individual MIPs, the mean number of copies per cell is about 50 with extremes ranging from 1 to about 10,000 [14,20,21]. Estimates based on analyses of human EBV lymphoblastoid cell lines suggest that the five most abundant MIPs already ‘fill’ 6% of the HLA class I molecules, the top 50 fill 27%, the top 300 fill 59%, and the top 1000 fill 83% [14]. MHC allotypes display more or less stringent peptide binding motifs and on average, each MHC class I allotype presents more than 1000 different MIPs, although the size of the peptide repertoire is widely variable across allotypes [22]. Further

development of quantitative large-scale MS could allow the estimation of a diversity index for each particular SMII. In analogy to ecological sciences, the ‘SMII diversity index’ would represent a quantitative measure that reflects how many different MIPs compose the SMII of a given cell under specific conditions, and simultaneously take into account how evenly the number of copies per MIP are distributed.

Limited overlap between the proteome and the immunopeptidome

There is only a modest correlation between the amounts of MIPs and the relative abundance of their source protein: some MIPs derive from low abundance proteins whereas some highly abundant proteins do not generate MIPs [14,19,23–25]. Two related factors explain why the immunopeptidome is not a mirror of the proteome. First, the SMII preferentially derives from RDPs whereas the most abundant proteins in the proteome have a long half-life [26]. Although long-lived proteins contribute to the

SMII, their representation in the SMII is significantly lower than their relative proportions in the proteome. Indeed, each protein consists of two subpopulations: proteins that achieve a native conformation and are degraded at protein-specific rate, and DRiPs with a shorter half-life [27]. Irrespective of whether their native form has a long or short half-life, DRiPs are included in the SMII. Accordingly, by determining the relative rates of synthesis of MIPs and of their source proteins using large-scale stable isotope labeling by amino acids in cell culture, Admon's group concluded that a significant portion of the SMII is derived from DRiPs [13].

Second, different proteins have different DRiP rates. Though many processes may lead to DRiP formation, the relative importance of these processes as well as the physical nature of DRiPs remain ill-defined. In practice, various putative mechanisms of DRiP formation have been identified: nonsense-mediated decay of mRNAs carrying premature termination codons [28], and destabilization of mRNAs by miRNAs [29[•]] or shRNAs [30]. However, the relative impact of these processes on MIP generation could depend on the cell type or metabolic state, and have yet to be assessed in large-scale experiments.

Numerous factors enhance the complexity of the SMII

Cell lineage and metabolic state

Large-scale peptidomic studies have shown that the SMII is complex and plastic. Different tissues/cell types display different MIP repertoires [31,32]. Out of 614 MIPs detected on a mouse pancreatic β cell line, 314 (51%) were absent on thymocytes and splenocytes; treatment of β cells with interferon gamma (IFN γ) increased the number of MIPs to 883, of which 321 (36%) were absent on thymic and spleen cells [20]. Likewise, 40% of MIPs eluted from freshly harvested C57BL/6 thymocytes were absent on C57BL/6 dendritic cells (DCs) [31]. Adamopoulou *et al.* analyzed two populations of thymocyte-depleted thymic cells obtained from children undergoing corrective cardiac surgery: CD11c⁺ DCs and CD11c⁻ cells (a mixture of epithelial and lymphomyeloid cells) [32]. Strikingly, 83% of MIPs were unique to CD11c⁺ or CD11c⁻ cells. In B lymphocytes, the MIP-coding exome is enriched in genes implicated in immunoglobulin production [29[•]]; in LPS-treated myeloid DCs, it is enriched in genes involved in myeloid differentiation, antigen (Ag) processing and Toll-like receptor signaling [31]. Hence, the MIP repertoire conceals a cell-type-specific signature, which is closely linked to cell function.

Treatment with a selective mTORC1 inhibitor (rapamycin) led to significant changes in the abundance (2–15-fold) of 53% of MIPs present on EL4 cells [33]. Differentially expressed MIPs resulted from events at the transcriptional, translational and post-translational level, and derived from proteins tightly connected to the

mTORC1 signaling pathway [33]. Notably, six MIPs detected on rapamycin-treated cells were 'neo-MIPs', absent on untreated cells. Immunization against two of these neo-MIPs elicited cytotoxic T cell responses. This suggests that cells can communicate their metabolic status to the adaptive immune system and that CD8T cells are reactive to neo-MIPs expressed on metabolically stressed cells. Altogether, these studies demonstrate that the SMII projects at the cell surface a representation of biochemical networks and metabolic events regulated at multiple levels inside the cell [33].

Inflammation, infection and drugs

Inflammation and infection have cell-extrinsic and cell-intrinsic effects on the SMII. The inflammatory environment causes ER stress, largely via protein oxidation and deprivation of oxygen and nutrients. In turn, ER stress impinges on the SMII by decreasing protein synthesis and increasing protein degradation [33,34]. ER stress is linked to autophagy, and degradation of self proteins by autophagy generates MHC II-associated peptides [35]. Although autophagy has been reported to enhance the presentation of viral MIPs [36], it is unclear whether it can also contribute to the presentation of self MIPs. Cytokines in the inflammatory milieu can affect MIP generation. Thus, out of 883 MIPs detected in pancreatic β cells treated *in vitro* with IFN γ , 55% were absent on untreated cells [20]. The appearance of numerous MIPs after IFN γ treatment is probably due to the upregulation of molecules involved in MHC Ag processing, including immunoproteasomes. Indeed, immunoproteasomes significantly increase MIP abundance and diversity [31,37]. Other inflammatory cytokines most likely affect the SMII but their effect has yet to be assessed at the systems-level [38].

Viruses can affect the MHC Ag presentation pathway *per se* [39], and also have pleiotropic effects on synthesis and degradation of host proteins, two key processes that mold the SMII [33]. Hence, aside from expression of viral MIPs, infection leads to drastic changes in the profile of self MIPs expressed at the cell surface [40,41]. Wahl *et al.* observed that influenza-infected cells expressed 20 neo self MIPs (absent in uninfected cells) and over-expressed 347 other self MIPs [41]. Viral infections are frequently associated with autoimmune diseases, but the mechanistic underpinnings of this association remain controversial. An important unanswered question is whether virus-induced changes in the SMII are mechanistically linked to autoimmunity.

Finally, drugs can dramatically modify the SMII. Abacavir and carbamazepine can bind to the MIP-binding groove of specific HLA-B molecules (e.g., HLA-B*57:01 and HLA-B*15:02), thereby altering their MIP repertoire and triggering autoimmune-like reactions (reviewed in [42]). A totally different mechanism is involved in the case of drugs such as aminoglycosides.

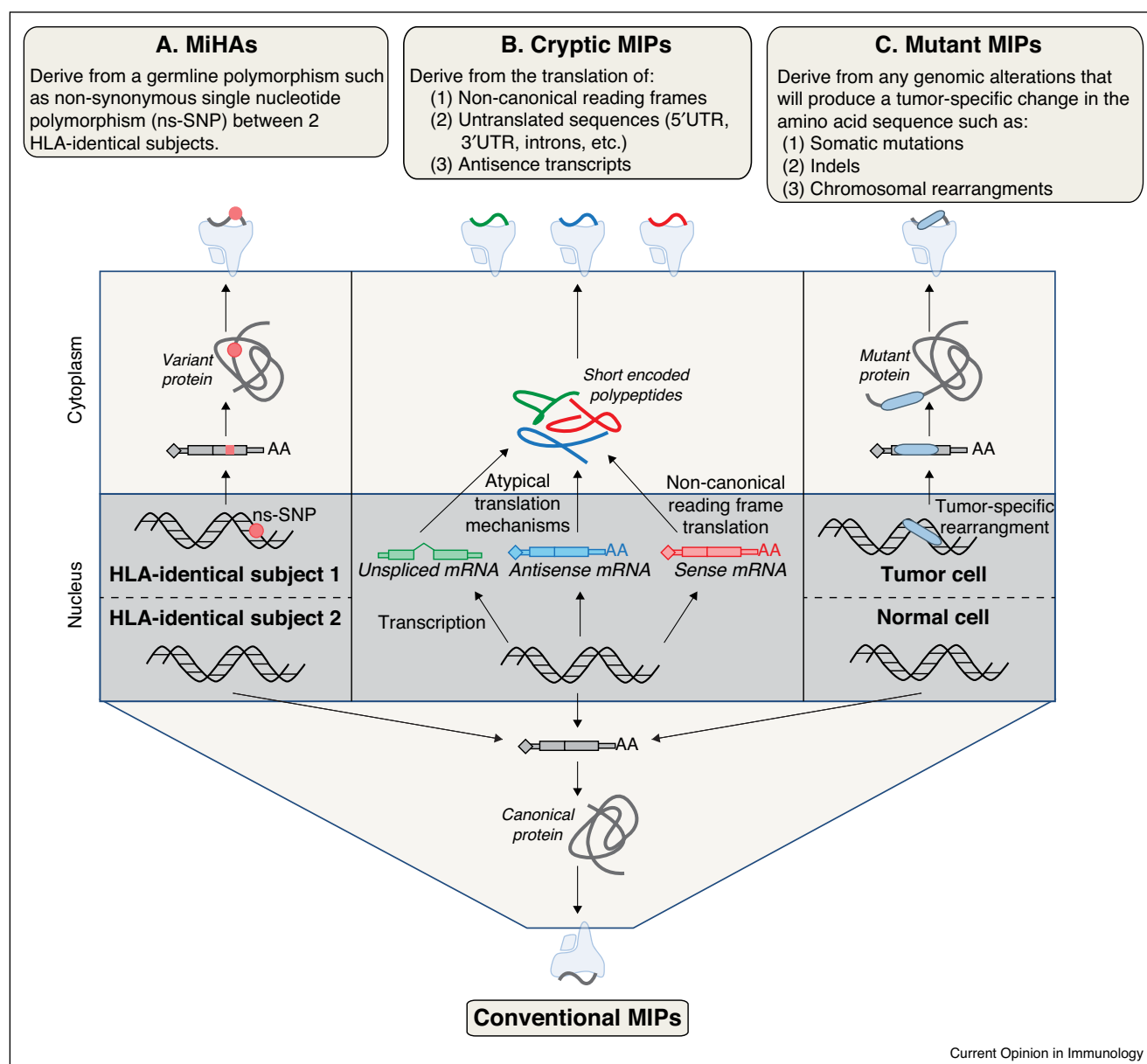
Indeed, aminoglycosides such as gentamicin induce stop codon read-through at a level sufficient to generate immunogenic MIPs derived from 3' untranslated regions [25].

The dark matter in the immunopeptidome – the hidden side of self

The complexity of the SMII is spelled in words that go beyond a 20-amino acids alphabet encoded in the germline and restricted to canonical reading frames [43]. It is of utmost importance to understand that classically, peptide identification by MS is based on search algorithms that

compare observed peptide fragments with a reference database such as Uniprot [44]. Thus, MS softwares search for a match between sample peptides and the reference proteome (encoded by canonical reading frames). Hence, standard high-throughput MS is typically blind to a whole universe of peptides which are not included in protein databases (Figure 2): (i) MIPs encoded by regions containing genomic polymorphisms [minor histocompatibility Ags (MiHAs)], (ii) MIPs derived from unconventional translational mechanisms (cryptic MIPs), (iii) MIPs bearing somatic mutations (mutant MIPs). Therefore, MIPs of prime interest in transplantation and cancer immunology

Figure 2



The 'dark matter' of the SMII. Non-canonical MIPs that can be discovered using high-throughput proteogenomics include MiHAs, cryptic MIPs and mutant MIPs.

cannot be detected by standard MS approaches. This limitation can now be overcome with proteogenomic methods that combine transcriptomic data with MS [45,46].

MiHAs

For historical reasons, MHC-associated peptides coded by polymorphic genomic sequences are referred to as MiHAs [47]. MiHAs are essentially genetic polymorphisms viewed from a T-cell perspective. MiHA-coding alleles can be dominant or recessive at the peptide level. Thus, a non-synonymous single nucleotide polymorphism (ns-SNP) in a MIP-coding genomic sequence will either hinder MIP generation (recessive allele) or generate a variant MIP (dominant allele) [48]. MiHAs are medically relevant because they elicit three types of allo-immune response: graft rejection, graft-versus-host disease and the allogeneic graft-versus-leukemia effect, which currently represents the most widely effective strategy for cancer immunotherapy in humans [49]. Over the last three decades, 35 human MiHAs encoded by autosomes and presented by HLA class I molecules have been discovered using arduous reductionist approaches [47,48]. Yet, due to the lack of a suitable systems-level approach [50], the global impact of non-MHC genomic polymorphisms on the SMII (i.e., what proportion of MIPs are MiHAs) remained until recently a speculative question. We addressed this question using a combination of next-generation sequencing and high-throughput MS peptide identification [18^{••}]. Whole exome and transcriptome sequencing was performed on cells from two non-twin HLA-identical siblings. Next generation sequencing data were translated *in silico* to create personalized protein databases that were used for MS sequencing of MIPs. Out of 4468 MIPs sequenced by MS, 34 were MiHAs coded by dominant alleles: 24 of these alleles were shared by the two subjects and 10 were unique to one subject. The number of unshared MiHAs is expected to be 1.8-fold higher in unrelated (HLA-matched) subjects relative to siblings [47,48]. The dominant importance of the MHC genotype in molding the SMII is shown by the fact that, at the population level, 88% of the MIP-coding exome is invariant, and only 12% contains ns-SNPs [18^{••}]. Nonetheless, MiHAs are attractive targets for adoptive T-cell immunotherapy of cancer, particularly hematologic cancers. In mouse models injection of T cells targeted to a single MiHA can eradicate neoplastic cells without causing any graft-versus-host disease [21,51,52]. Because of the low number of molecularly defined human MiHAs, less than 30% of patients would be eligible for immunotherapy targeted to specific MiHAs [53]. Proteogenomic-based MiHA discovery should now rapidly overcome this caveat.

Cryptic MIPs

When studying the molecular bases of alloreactivity, Shastri *et al.* made a startling observation. Using a lacZ-inducible

alloreactive T cell clone as a probe, they screened a splenic cDNA library in transiently transfected Ag-presenting cells and isolated a cDNA clone that allowed expression of the MIP recognized by the T cell clone. The salient finding was that this peptide derived from a non-canonical reading frame initiated with a non-AUG start codon [54]. In a series of elegant studies, Shastri's group discovered that synthesis of this peptide was initiated with a CUG codon decoded as a leucine rather than a methionine [55^{••}]. More recently, studies using ribosome profiling have shown that the proteome is more complex than anticipated [56]. Indeed, ribosome profiling, an emerging technique that uses deep sequencing to monitor *in vivo* translation, has revealed that ribosomes occupy many regions of the transcriptome thought to be noncoding, including 5' UTRs and long noncoding RNAs [57]. These observations strongly suggest that translation is pervasive on cytosolic transcripts outside of conserved reading frames. Accordingly, recent studies suggest that MIPs can arise from untranslated regions (UTRs or introns) or from alternate translational reading frames [43,57,58^{••}]. In addition, seminal observations provide compelling evidence that MIPs arising from 'untranslated regions' may derive from immature mRNAs translated in the nucleus [6,58^{••},59^{••},60]. A few of these so-called 'cryptic MIPs' were validated by MS sequencing [25,61]. However, since most putative cryptic MIPs have not been validated by MS, some of them must be considered with skepticism because their identification relied primarily on methods fraught with high false discovery rates [62,63^{••},64]: bioinformatic predictions (that cannot predict key steps in MIP processing) and *in vitro* T-cell reactivity assays (which are tainted by T-cell cross-reactivity). Systems-level evaluation of the landscape of cryptic MIPs is possible but will absolutely require unbiased proteogenomic studies. Ideally, this type of study would entail transcriptome sequencing and ribosome profiling of relevant cells followed by MS analyses in which MIP spectra are assigned by searching the six-frame *in silico* translation of transcriptomic reads. This strategy has been successfully used for identification of proteins in non-model species as well as higher eukaryotes [46,65,66].

Mutant MIPs

Several lines of evidence suggest that Ags expressed by neoplastic cells are able to elicit protective anti-tumor responses in humans and that T cells specific for these Ags are enriched in tumor-infiltrating lymphocytes (TILs) [67–69]. What is the nature of tumor Ags that can elicit protective anti-tumor responses? Tumor Ags belong to two main classes: (i) tumor-associated Ags (TAAs) which are qualitatively normal MIPs overexpressed on cancer cells and (ii) tumor-specific Ags (TSAs) which result from somatic gene mutations or translocations and are truly tumor-specific. Based on the assumption that T cells recognizing the most immunogenic tumor Ags are enriched in TILs, two groups assessed whether melanoma TILs recognized the 175 TAAs identified to

date on melanoma cells [70,71]. These comprehensive studies led to a striking conclusion: less than 1% of CD8 melanoma TILs recognized known TAAs. The inescapable corollary is that 99% of CD8 TILs are specific for unknown Ags. Since TILs recognize the most relevant (i.e., immunogenic) tumor Ags, it can be inferred that the most relevant tumor Ags have not been discovered yet and are most likely TSAs rather than TAAs [63^{••},72,73]. It is therefore sobering that the landscape of human TSAs remains practically unknown and that the contribution of MS to this field has been minimal [74,75]. A few TSAs have been discovered one at a time, usually by laborious screening of cDNA libraries with Ag-reactive T cells [74]. More recently, a total of ten human TSAs have been discovered by two groups using strategies based on whole-exome sequencing of melanoma tumor DNA from 4 subjects, MHC binding predictions, and *in vitro* T-cell assays [63^{••},76[•]]. Though these landmark studies represent a progress in the field of TSA discovery, they illustrate the limitations inherent to strategies that do not include high-throughput MS. First, the number of TSAs discovered (ten in four subjects) was low considering that melanoma is the type of cancer bearing the highest number of mutations [77]. Indeed, recent studies in patients with ovarian cancer or cholangiosarcoma suggest that the number of TSAs that can be discovered with this approach will be even lower in non-melanoma tumors [78[•],79[•]]. Second, more than 95% of TSA candidates identified with exome sequencing and MHC-binding predictions do not elicit T-cell responses [63^{••},76[•],78[•]], and are therefore probably false positives. Finally, identification of TSAs based on T-cell reactivity without validation by MS sequencing remains tentative [62].

Concluding remarks and perspectives

Recent progress in proteogenomics enables large-scale MS sequencing of all MIPs whose source transcript can be sequenced by RNA-seq (essentially all transcripts, including unprocessed mRNAs). In principle, the sole MIPs that can elude RNA-seq based proteogenomics are those derived from non-contiguous protein sequences via proteasome-mediated splicing. Indeed, after excision of intervening fragments, proteasomes can create peptide bonds between distant protein segments. Though proteasomal splicing is a low efficiency process, it can generate biologically relevant MIPs including MiHAs and TSAs [80]. Further studies are needed to determine whether refinements in *de novo* MS sequencing (i.e., without the assistance of database search) could allow high-throughput analysis of MIPs generated via proteasome splicing. Lastly, quantitative aspects of the SMII are still missing (number of copies for each individual MIPs) [10], and need to be studied in order to assess changes in the diversity index of the SMII under different conditions. With this in hand, and with the collaboration of chemists and bioinformaticians, systems immunologists can now explore a whole new universe: the dark matter in the immunopeptidome. It is particularly urgent

to qualitatively and quantitatively characterize the landscape of human TSAs by assessing the number of TSA copies per cancer cell and the extent of inter-tumor and intra-tumor heterogeneity in TSA expression.

Acknowledgements

Research performed in the authors' labs was supported by grants from the Canadian Institutes of Health Research (CIHR) (MOP 42384) and the Canadian Cancer Society (grant # 701564). C.P. and P.T. hold Canada Research Chairs in Immunobiology, and Proteomics and Bioanalytical Spectrometry, respectively. The Institute for Research in Immunology and Cancer is supported in part by the Canada Foundation for Innovation and the Fonds de Recherche Santé Québec. Special thanks to our colleague Sébastien Lemieux for making bioinformatics 'as simple as possible, but not simpler'.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Davis MM, Krogsgaard M, Huse M, Huppa J, Lillemeier BF, Li QJ: **T cells as a self-referential, sensory organ.** *Annu Rev Immunol* 2007, **25**:681-695.
2. Mandl JN, Monteiro JP, Vrsekooop N, Germain RN: **T cell-positive selection uses self-ligand binding strength to optimize repertoire recognition of foreign antigens.** *Immunity* 2013, **38**:263-274.
3. Anton LC, Yewdell JW: **Translating DRiPs: MHC class I immunosurveillance of pathogens and tumors.** *J Leukoc Biol* 2014, **95**:551-562.
- A timely review of the evolution of ideas on the biogenesis and importance of DRiPs.
4. Kim W, Bennett EJ, Huttlin EL, Guo A, Li J, Possemato A, Sowa ME, Rad R, Rush J, Comb MJ, Harper JW, Gygi SP: **Systematic and quantitative assessment of the ubiquitin-modified proteome.** *Mol Cell* 2011, **44**:325-340.
5. Yewdell JW: **DRiPs solidify: progress in understanding endogenous MHC class I antigen processing.** *Trends Immunol* 2011, **32**:548-558.
6. Baboo S, Bhushan B, Jiang H, Grovenor CR, Pierre P, Davis BG, Cook PR: **Most human proteins made in both nucleus and cytoplasm turn over within minutes.** *PLoS One* 2014, **9**:e99346.
7. Rock KL, Farfan-Arribas DJ, Colbert JD, Goldberg AL: **Re-examining class-I presentation and the DRiP hypothesis.** *Trends Immunol* 2014, **35**:144-152.
8. Reits EA, Hodge JW, Herberts CA, Groothuis TA, Chakraborty M, Wansley E, Camphausen K, Luiten RM, de Ru AH, Neijssen J *et al.*: **Radiation modulates the peptide repertoire, enhances MHC class I expression, and induces successful antitumor immunotherapy.** *J Exp Med* 2006, **203**:1259-1271.
9. Norbury CC, Basta S, Donohue KB, Tschärke DC, Princiotta MF, Berglund P, Gibbs J, Bennink JR, Yewdell JW: **CD8⁺ T cell cross-priming via transfer of proteasome substrates.** *Science* 2004, **304**:1318-1321.
10. Mester G, Hoffmann V, Stevanovic S: **Insights into MHC class I antigen processing gained from large-scale analysis of class I ligands.** *Cell Mol Life Sci* 2011, **68**:1521-1532.
11. Nagaraj N, Wisniewski JR, Geiger T, Cox J, Kircher M, Kelso J, Paabo S, Mann M: **Deep proteome and transcriptome mapping of a human cancer cell line.** *Mol Syst Biol* 2011, **7**:548.
12. Beck M, Schmidt A, Malmstroem J, Claassen M, Ori A, Szymborska A, Herzog F, Rinner O, Ellenberg J, Aebersold R: **The quantitative proteome of a human cell line.** *Mol Syst Biol* 2011, **7**:549.
13. Bourdetsky D, Schmelzer CE, Admon A: **The nature and extent of contributions by defective ribosome products to the HLA peptidome.** *Proc Natl Acad Sci U S A* 2014, **111**:E1591-E1599.

14. Hassan C, Kester MG, de Ru AH, Hombrink P, Drijfhout JW, Nijveen H, Leunissen JA, Heemskerk MH, Falkenburg JH, van Veelen PA: **The human leukocyte antigen-presented ligandome of B lymphocytes.** *Mol Cell Proteomics* 2013, **12**:1829-1843.
 15. Mommen GP, Frese CK, Meiring HD, van Gaans-van den Brink J, de Jong AP, van Els CA, Heck AJ: **Expanding the detectable HLA peptide repertoire using electron-transfer/higher-energy collision dissociation (ET^hCD).** *Proc Natl Acad Sci U S A* 2014, **111**:4507-4512.
 16. Berlin C, Kowalewski DJ, Schuster H, Mirza N, Walz S, Handel M, Schmid-Horch B, Salih HR, Kanz L, Rammensee HG, Stevanovic S, Stickel JS: **Mapping the HLA ligandome landscape of acute myeloid leukemia: a targeted approach towards peptide based immunotherapy.** *Leukemia* 2014 <http://dx.doi.org/10.1038/leu.2014.233>. [Epub ahead of print].
 17. Fortier MH, Caron E, Hardy MP, Voisin G, Lemieux S, Perreault C, Thibault P: **The MHC class I peptide repertoire is molded by the transcriptome.** *J Exp Med* 2008, **205**:595-610.
 18. Granados DP, Sriranganadane D, Daouda T, Zieger A, Laumont CM, Caron-Lizotte O, Boucher G, Hardy MP, Gendron P, Côté C et al.: **Impact of genomic polymorphism on the repertoire of human MHC class I-associated peptides.** *Nat Commun* 2014, **5**:3600.
- This work presents a novel proteogenomic approach for high-throughput discovery of polymorphic MIPs and assesses for the first time the impact of non-HLA genomic polymorphisms on the SMII.
19. Hoof I, van Baarle D, Hildebrand WH, Kesmir C: **Proteome sampling by the HLA class I antigen processing pathway.** *PLoS Comput Biol* 2012, **8**:e1002517.
 20. Dudek NL, Tan CT, Gorasia DG, Croft NP, Illing PT, Purcell AW: **Constitutive and inflammatory immunopeptidome of pancreatic beta-cells.** *Diabetes* 2012, **61**:3018-3025.
 21. Vincent K, Hardy MP, Trofimov A, Laumont CM, Sriranganadane D, Hadj-Mimoun S, Fourati IS, Soudeyns H, Thibault P, Perreault C: **Rejection of leukemic cells requires antigen-specific T cells with high functional avidity.** *Biol Blood Marrow Transplant* 2014, **20**:37-45.
 22. Paul S, Weiskopf D, Angelo MA, Sidney J, Peters B, Sette A: **HLA class I alleles are associated with peptide-binding repertoires of different size, affinity and immunogenicity.** *J Immunol* 2013, **191**:5831-5839.
 23. Milner E, Barnea E, Beer I, Admon A: **The turnover kinetics of MHC peptides of human cancer cells.** *Mol Cell Proteomics* 2006, **5**:357-365.
 24. Croft NP, Smith SA, Wong YC, Tan CT, Dudek NL, Flesch IEA, Tschärke DC, Purcell AW: **Kinetics of antigen expression and epitope presentation during virus infection.** *PLoS Pathog* 2013, **9**:e1003129.
 25. Goodenough E, Robinson TM, Zook MB, Flanagan KM, Atkins JF, Howard MT, Eisenlohr LC: **Cryptic MHC class I-binding peptides are revealed by aminoglycoside-induced stop codon read-through into the 3'UTR.** *Proc Natl Acad Sci U S A* 2014, **111**:5670-5675.
 26. Boisvert FM, Ahmad Y, Gierlinski M, Charrière F, Lamont D, Scott M, Barton G, Lamond AI: **A quantitative spatial proteomics analysis of proteome turnover in human cells.** *Mol Cell Proteomics* 2012, **11**:M111.
 27. Qian S-B, Reits E, Neefjes J, Deslich JM, Bennink JR, Yewdell JW: **Tight linkage between translation and MHC class I peptide ligand generation implies specialized antigen processing for defective ribosomal products.** *J Immunol* 2006, **177**:227-233.
 28. Apcher S, Daskalogianni C, Lejeune F, Manoury B, Imhoos G, Heslop L: **Major source of antigenic peptides for the MHC class I pathway is produced during the pioneer round of mRNA translation.** *Proc Natl Acad Sci U S A* 2011, **108**:11572-11577.
 29. Granados DP, Yahyaoui W, Laumont CM, Daouda T, Muratore-Schroeder TL, Côté C, Laverdure JP, Lemieux S, Thibault P, Perreault C: **MHC I-associated peptides preferentially derive from transcripts bearing miRNA response elements.** *Blood* 2012, **119**:e181-e191.
- This study establishes a link between the SMII and microRNAs and suggests that interactions between mRNAs and miRNAs may generate DRiPs.
30. Gu W, Cochrane M, Leggatt GR, Payne E, Choyce A, Zhou F, Tindle R, McMillan NAJ: **Both treated and untreated tumors are eliminated by short hairpin RNA-based induction of target-specific immune responses.** *Proc Natl Acad Sci U S A* 2009, **106**:8314-8319.
 31. de Verteuil D, Muratore-Schroeder TL, Granados DP, Fortier MH, Hardy MP, Bramoullé A, Caron E, Vincent K, Mader S, Lemieux S, Thibault P, Perreault C: **Deletion of immunoproteasome subunits imprints on the transcriptome and has a broad impact on peptides presented by major histocompatibility complex I molecules.** *Mol Cell Proteomics* 2010, **9**:2034-2047.
 32. Adamopoulou E, Tenzer S, Hillen N, Klug P, Rota IA, Tietz S, Gebhardt M, Stevanovic S, Schild H, Tolosa E, Melms A, Stoeckle C: **Exploring the MHC-peptide matrix of central tolerance in the human thymus.** *Nat Commun* 2013, **4**:2039.
 33. Caron E, Vincent K, Fortier MH, Laverdure JP, Bramoullé A, Hardy MP, Voisin G, Roux P, Lemieux S, Thibault P, Perreault C: **The MHC I immunopeptidome conveys to the cell surface an integrative view of cellular regulation.** *Mol Syst Biol* 2011, **7**:533.
 34. Granados DP, Tanguay PL, Hardy MP, Caron E, De Verteuil D, Meloche S, Perreault C: **ER stress affects processing of MHC class I-associated peptides.** *BMC Immunol* 2009, **10**:10.
 35. Neefjes J, Jongsma MLM, Paul P, Bakke O: **Towards a systems understanding of MHC class I and MHC class II antigen presentation.** *Nat Rev Immunol* 2011, **11**:823-836.
 36. English L, Chemali M, Duron J, Rondeau C, Laplante A, Gingras D, Alexander D, Leib D, Norbury C, Lippé R, Desjardins M: **Autophagy enhances the presentation of endogenous viral antigens on MHC class I molecules during HSV-1 infection.** *Nat Immunol* 2009, **10**:480-487.
 37. Kincaid EZ, Che JW, York I, Escobar H, Reyes-Vargas E, Delgado JC, Welsh RM, Karow ML, Murphy AJ, Valenzuela DM, Yancopoulos GD, Rock KL: **Mice completely lacking immunoproteasomes show major changes in antigen presentation.** *Nat Immunol* 2012, **13**:129-135.
 38. Morandi F, Di CE, Ferrone S, Petretto A, Pistoia V, Airolidi I: **IL-27 in human secondary lymphoid organs attracts myeloid dendritic cells and impairs HLA class I-restricted antigen presentation.** *J Immunol* 2014, **192**:2634-2642.
 39. Hansen TH, Bouvier M: **MHC class I antigen presentation: learning from viral evasion strategies.** *Nat Rev Immunol* 2009, **9**:503-513.
 40. Hickman HD, Luis AD, Bardet W, Buchli R, Battson CL, Shearer MH, Jackson KW, Kennedy RC, Hildebrand WH: **Cutting edge: class I presentation of short peptides following HIV infection.** *J Immunol* 2003, **171**:22-26.
 41. Wahl A, Schafer F, Bardet W, Hildebrand WH: **HLA class I molecules reflect an altered host proteome after influenza virus infection.** *Hum Immunol* 2010, **71**:14-22.
 42. Illing PT, Vivian JP, Purcell AW, Rossjohn J, McCluskey J: **Human leukocyte antigen-associated drug hypersensitivity.** *Curr Opin Immunol* 2013, **25**:81-89.
 43. Starck SR, Shastri N: **Non-conventional sources of peptides presented by MHC class I.** *Cell Mol Life Sci* 2011, **68**:1471-1479.
 44. Steen H, Mann M: **The ABC's (and XYZ's) of peptide sequencing.** *Nat Rev Mol Cell Biol* 2004, **5**:699-711.
 45. Woo S, Cha SW, Merrihew G, He Y, Castellana N, Guest C, MacCoss M, Bafna V: **Proteogenomic database construction driven from large scale RNA-seq data.** *J Proteome Res* 2014, **13**:21-28.
 46. Branca RM, Orre LM, Johansson HJ, Granholm V, Huss M, Perez-Bercoff A, Forshed J, Kall L, Lehtio J: **HiRIEF LC-MS enables deep proteome coverage and unbiased proteogenomics.** *Nat Methods* 2014, **11**:59-62.
 47. Warren EH, Zhang XC, Li S, Fan W, Storer BE, Chien JW, Boeckh MJ, Zhao LP, Martin PJ, Hansen JA: **Effect of MHC and**

- non-MHC donor/recipient genetic disparity on the outcome of allogeneic HCT.** *Blood* 2012, **120**:2796-2806.
48. Spierings E, Hendriks M, Absi L, Canossi A, Chhaya S, Crowley J, Dolstra H, Eliaou JF, Ellis T, Enczmann J *et al.*: **Phenotype frequencies of autosomal minor histocompatibility antigens display significant differences among populations.** *PLoS Genet* 2007, **3**:e103.
 49. Vincent K, Roy DC, Perreault C: **Next-generation leukemia immunotherapy.** *Blood* 2011, **118**:2951-2959.
 50. Hombrink P, Hassan C, Kester MG, de Ru AH, van Bergen CA, Nijveen H, Drijfhout JW, Falkenburg JH, Heemskerk MH, van Veelen PA: **Discovery of T cell epitopes implementing HLA-peptidomics into a reverse immunology approach.** *J Immunol* 2013, **190**:3869-3877.
 51. Fontaine P, Roy-Proulx G, Knafo L, Baron C, Roy DC, Perreault C: **Adoptive transfer of T lymphocytes targeted to a single immunodominant minor histocompatibility antigen eradicates leukemia cells without causing graft-versus-host disease.** *Nat Med* 2001, **7**:789-794.
 52. Meunier MC, Delisle JS, Bergeron J, Rineau V, Baron C, Perreault C: **T cells targeted against a single minor histocompatibility antigen can cure solid tumors.** *Nat Med* 2005, **11**:1222-1229.
 53. Bleakley M, Otterud BE, Richardt JL, Mollerup AD, Hudecek M, Nishida T, Chaney CN, Warren EH, Leppert MF, Riddell SR: **Leukemia-associated minor histocompatibility antigen discovery using T-cell clones isolated by in vitro stimulation of naive CD8⁺ T cells.** *Blood* 2010, **115**:4923-4933.
 54. Malarkannan S, Afkarian M, Shastri N: **A rare cryptic translation product is presented by Kb major histocompatibility complex class I molecule to alloreactive T cells.** *J Exp Med* 1995, **182**:1739-1750.
 55. Starck SR, Jiang V, Pavon-Eternod M, Prasad S, McCarthy B, Pan T, Shastri N: **Leucine-tRNA initiates at CUG start codons for protein synthesis and presentation by MHC class I.** *Science* 2012, **336**:1719-1723.
- This landmark work revealed that in addition to the canonical AUG-methionine codon, CUG-leucine codons can be used to initiate protein synthesis and supply cryptic MIPs.
56. Ingolia NT: **Ribosome profiling: new views of translation, from single codons to genome scale.** *Nat Rev Genet* 2014, **15**: 205-213.
 57. Ingolia NT, Brar GA, Stern-Ginossar N, Harris MS, Talhouarne GJ, Jackson SE, Wills MR, Weissman JS: **Ribosome profiling reveals pervasive translation outside of annotated protein-coding genes.** *Cell Rep* 2014, **8**:1365-1379.
 58. Apcher S, Millot G, Daskalogianni C, Scherl A, Manoury B, Fahraeus R: **Translation of pre-spliced RNAs in the nuclear compartment generates peptides for the MHC class I pathway.** *Proc Natl Acad Sci U S A* 2013, **110**:17951-17956.
 59. David A, Dolan BP, Hickman HD, Knowlton JJ, Clavarino G, Pierre P, Binnink JR, Yewdell JW: **Nuclear translation visualized by ribosome-bound nascent chain puromycylation.** *J Cell Biol* 2012, **197**:45-57.
- Refs. [58**,59**] provide compelling evidence that some MIPs are generated within the nucleus during the early steps of mRNA maturation (and before RNA splicing) via a noncanonical translation mechanism.
60. Yewdell JW, David A: **Nuclear translation for immunosurveillance.** *Proc Natl Acad Sci U S A* 2013, **110**: 17612-17613.
 61. Weinzierl AO, Maurer D, Altenberend F, Schneiderhan-Marra N, Klingel K, Schoor O, Wernet D, Joos T, Rammensee HG, Stevanovic S: **A cryptic vascular endothelial growth factor T-cell epitope: identification and characterization by mass spectrometry and T-cell assays.** *Cancer Res* 2008, **68**: 2447-2454.
 62. Popovic J, Li LP, Kloetzel PM, Leisegang M, Uckert W, Blankenstein T: **The only proposed T-cell epitope derived from the TEL-AML1 translocation is not naturally processed.** *Blood* 2011, **118**:946-954.
 63. Robbins PF, Lu YC, El-Gamil M, Li YF, Gross C, Gartner J, Lin JC, Teer JK, Clifton P, Tycksen E, Samuels Y, Rosenberg SA: **Mining exomic sequencing data to identify mutated antigens recognized by adoptively transferred tumor-reactive T cells.** *Nat Med* 2013, **19**:747-752.
- This work demonstrates the value of exome sequencing for the identification of TSAs in melanoma.
64. Sewell AK: **Why must T cells be cross-reactive?** *Nat Rev Immunol* 2012, **12**:669-677.
 65. Castellana N, Bafna V: **Proteogenomics to discover the full coding content of genomes: a computational perspective.** *J Proteomics* 2010, **73**:2124-2135.
 66. Evans VC, Barker G, Heesom KJ, Fan J, Bessant C, Matthews DA: **De novo derivation of proteomes from transcriptomes for transcript and protein identification.** *Nat Methods* 2012, **9**:1207-1211.
 67. Restifo NP, Dudley ME, Rosenberg SA: **Adoptive immunotherapy for cancer: harnessing the T cell response.** *Nat Rev Immunol* 2012, **12**:269-281.
 68. Galon J, Angell HK, Bedognetti D, Marincola FM: **The continuum of cancer immunosurveillance: prognostic, predictive, and mechanistic signatures.** *Immunity* 2013, **39**:11-26.
 69. Page DB, Postow MA, Callahan MK, Allison JP, Wolchok JD: **Immune modulation in cancer with antibodies.** *Annu Rev Med* 2014, **65**:185-202.
 70. Andersen RS, Thure CA, Junker N, Lyngaa R, Donia M, Ellebaek E, Svane IM, Schumacher TN, Thor Straten P, Hadrup SR: **Dissection of T-cell antigen specificity in human melanoma.** *Cancer Res* 2012, **72**:1642-1650.
 71. Kvistborg P, Shu CJ, Heemskerk B, Fankhauser M, Thure CA, Toebes M, van Rooij N, Linnemann C, van Buuren MM, Urbanus JH *et al.*: **TIL therapy broadens the tumor-reactive CD8(+) T cell compartment in melanoma patients.** *Oncoimmunology* 2012, **1**:409-418.
 72. Heemskerk B, Kvistborg P, Schumacher TN: **The cancer antigenome.** *EMBO J* 2013, **32**:194-203.
 73. Hinrichs CS, Restifo NP: **Reassessing target antigens for adoptive T-cell therapy.** *Nat Biotechnol* 2013, **31**:999-1008.
 74. Coulie PG, Van den Eynde BJ, van der Bruggen P, Boon T: **Tumor antigens recognized by T lymphocytes: at the core of cancer immunotherapy.** *Nat Rev Cancer* 2014, **14**:135-146.
 75. Rosenberg SA: **Finding suitable targets is the major obstacle to cancer gene therapy.** *Cancer Gene Ther* 2014, **21**:45-47.
 76. van Rooij N, van Buuren MM, Philips D, Velds A, Toebes M, Heemskerk B, van Dijk LJ, Behjati S, Hilkmann H, El AD *et al.*: **Tumor exome analysis reveals neoantigen-specific T-cell reactivity in an ipilimumab-responsive melanoma.** *J Clin Oncol* 2013, **31**:e439-e442.
- This study supports and extends data in Ref. [63**] showing that exome data can guide discovery of melanoma TSAs.
77. Lawrence MS, Stojanov P, Mermel CH, Robinson JT, Garraway LA, Golub TR, Meyerson M, Gabriel SB, Lander ES, Getz G: **Discovery and saturation analysis of cancer genes across 21 tumour types.** *Nature* 2014, **505**:495-501.
 78. Wick DA, Webb JR, Nielsen JS, Martin SD, Kroeger DR, Milne K, Castellarin M, Twumasi-Boateng K, Watson PH, Holt RA, Nelson BH: **Surveillance of the tumor mutanome by T cells during progression from primary to recurrent ovarian cancer.** *Clin Cancer Res* 2014, **20**:1125-1134.
- This work validates the utility of exomic data for TSAs discovery and suggests that ovarian cancers may display less TSAs than melanomas.
79. Tran E, Turcotte S, Gros Aa, Robbins PF, Lu YC, Dudley ME, Wunderlich JR, Somerville RP, Hogan K, Hinrichs CS *et al.*: **Cancer immunotherapy based on mutation-specific CD4⁺ T cells in a patient with epithelial cancer.** *Science* 2014, **344**:641-645.
- Although the present review focuses on MHC class I peptides, the work of Tan *et al.* shows that TSAs can also be presented by MHC class II molecules.
80. Vigneron N, Van den Eynde BJ: **Proteasome subtypes and the processing of tumor antigens: increasing antigenic diversity.** *Curr Opin Immunol* 2012, **24**:84-91.