A primer on recent developments in cancer immunotherapy, with a focus on neoantigen vaccines

Taku A. Tokuyasu¹, Jian-Dong Huang¹,²,³

¹Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, Guangdong, China.
²School of Biomedical Sciences, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong, China.
³Shenzhen Institute of Research and Innovation, University of Hong Kong, Hong Kong, China.

Correspondence to: Dr. Taku A. Tokuyasu, Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, Guangdong, China. E-mail: tokbio@fastmail.fm

How to cite this article: Tokuyasu TA, Huang JD. A primer on recent developments in cancer immunotherapy, with a focus on neoantigen vaccines. J Cancer Metastasis Treat 2018;4:2. http://dx.doi.org/10.20517/2394-4722.2017.52

Received: 8 Aug 2017  First Decision: 18 Oct 2017  Revised: 11 Dec 2017  Accepted: 11 Dec 2017  Published: 18 Jan 2018

Science Editor: Shuen-Kuei Liao  Copy Editor: Lu Liu  Production Editor: Huan-Liang Wu

Abstract
Cancer immunotherapy has now been conclusively shown to be capable of producing durable responses for a substantial number of patients. Adoptive cell transfer and checkpoint blockade therapies in particular both demonstrate that antigen-specific immune responses can be dramatically effective, even in previously refractory late stage disease. Such developments, together with advances in technology, have strongly encouraged revisiting the concept of neoantigen vaccines. Here we introduce basic ideas in the field to allow investigators from diverse backgrounds to understand these developments, grasp current issues, and contribute to further progress.

Keywords: Immunotherapy, cancer vaccine, immunoinformatics, precision medicine, combination therapy, theoretical models, systems biology

INTRODUCTION
In the late 1800s, Coley¹ pursued investigations of cancer regression in the context of bacterial disease. It has been clear since then that the immune system plays an important role in cancer. Over the ensuing century, strong arguments were put forward for both why cancer immunotherapy should work and why it should not, occasionally by the same investigator². The past decade has seen dramatic progress in cancer immunotherapies, such as checkpoint blockade, adoptive cell transfer, and vaccines. The success came on
two fronts: complete durable patient response was achieved in a substantial fraction of patients in the clinic, and the mechanism of action was T-cell antigen-specific. This spurred confidence that therapy approaching a “cure” was at hand, based on a rational extrapolation of current knowledge. The immune system is inextricably linked to both the phenomenon of cancer and its treatment. This represents a paradigm shift, where cancer is no longer seen as just a collection of aberrant cells, but rather a systemic disease.

While this new vista continues to capture the public imagination worldwide, we have learned enough over the years to understand that cancer immunotherapy, in its current form, is not a panacea. The central challenge facing cancer immunotherapies and neoantigen vaccines in particular is understanding resistance.

Integrating immunology and cancer research, already two of the most complex topics in biomedicine, is an interdisciplinary effort, drawing from fields such as biology, pharmacology, chemistry, physics, engineering, statistics, and mathematics. Our main aim in this primer is to lower the barrier to entry for readers who are not specialists in immunotherapy. We focus on neoantigen vaccines, which in some ways represent T-cell based cancer immunotherapy in its most elementary form. We also address general issues, enabling readers to quickly grasp other immunotherapies and future developments.

Background on the immune system and cancer
We embark first on a brief tour of immunology, with the caveat that the specifics and even the broad outlines may shift as the field advances. Many of the features described below have bearing on possible cancer resistance mechanisms.

In brief, the requirements for an effective immune system include mechanisms to recognize foreign invaders, the means to trigger and coordinate a potentially complex attack (“expansion”), then return to equilibrium (“contraction”), while not attacking normal tissue. This rests critically on the ability to distinguish self from non-self. In vertebrates, robust response also leads to the development of immune memory. Immunotherapy can be viewed as an attempt to shift the equilibrium point in a complex system that can actively amplify or suppress its effects.

Cancer cells can evade the immune system through a variety of routes, such as being viewed as self, hijacking suppressive mechanisms that prevent damage, attacking or subverting immune system agents, or simply growing at a rate beyond the capacity of an often aged and weakened immune system.

The vertebrate immune system is broadly divided into two arms. Innate immunity is encoded in the germline, while adaptive (“acquired”) immunity is mediated by B and T lymphocytes that undergo processes of diversification and selection. T cell selection relies on processes of central tolerance (at the thymus) and peripheral tolerance (on mature circulating T cells). The two arms interact, with some cell types having a role in both arms.

In the adaptive system, T cells play the key role in recognizing pathology via antigens. The core of this task involves three parts: a presenter (major histocompatibility complex molecule, MHC), an antigen fragment (peptide), and a recognizer (T cell receptor, TCR). Elaborate processes of MHC expression and maturation, antigen processing, peptide MHC loading, and generation of mature naive T cells through the thymus underlie their formation and interaction.

Antigen recognition takes place when a receptor on a T cell encounters a cell presenting a cognate peptide-MHC (pMHC) complex on its surface. If a CD28 co-stimulatory receptor on the T cell simultaneously binds with CD80 or CD86 expressed on the presenting cell, an activation signal is propagated on the cytosolic side of the TCR, leading to cell proliferation, differentiation, and secretion of cytokines. A lack of a co-stimulatory signal leads to a hypo-responsive state known as T cell anergy. Inhibitory checkpoint molecules “put the
brakes” on adaptive immunity, where for example cytotoxic T-lymphocyte associated protein 4 (CTLA-4) competes with CD80/86 in binding CD28, thus suppressing activation.[11]

The detection limit for such T cell triggering is impressively low (four pMHC per TCR cluster)[12]. Note that the vast majority of the $10^4$ presented peptides in vivo are in fact normal “self” peptides, with only a few from foreign antigens, if any.[13]

MHC molecules and T cells come in two subtype pairs [Figure 1]. MHC class I (MHC-I) is normally expressed in all nucleated cells and presents intracellular (endogenous) antigen fragments. The pMHC-I complexes are recognized by CD8+ T cells, which are then activated and differentiate into cytotoxic T cells (CTLs) with direct cell killing capability. MHC class II (MHC-II) is expressed in “professional” antigen presenting cells (APCs), including dendritic cells, and presents exogenous antigens that have been engulfed by the APC. The resulting pMHC-II complexes are recognized by CD4+ T cells, which can differentiate e.g. into T helper cells whose primary role is to activate other immune system components.

The loaded peptides in the case of MHC-I are typically 8 to 12 residues in length and are loaded into a groove that is closed on both ends. The MHC-II-loaded peptides range in length from 12 to 25 residues and are loaded into a groove that is open on both ends [Figure 2].
Cells that do not express MHC-I on their surface are considered aberrant by the immune system. In normal environments, these are eliminated by natural killer cells, which are innate lymphoid cells with recognition receptors encoded in the germline [15].

Both the presentation (MHC) and recognition (TCR) components are highly diverse, although MHC diversity only appears at the population level. In humans, the MHC is known as the human leukocyte antigen (HLA) complex. Each individual inherits six MHC-I alleles from three loci, HLA-A, -B and -C (i.e. two parental alleles from each locus), and similarly, six MHC-II alleles from HLA-DP, -DQ, and -DR loci. Note that the HLA nomenclature was revised in 2010 [16]. As of Oct. 2015, there were 10,297 class I and 3543 class II known alleles [17,18]. Hence, the pMHC binding profile (“HLA peptidome”) varies broadly between individuals.

On the recognition side, there are in principle at least $10^{15}$ possible TCR variants. The number of T-cells in any individual is on the order of $10^{12}$, and the number of clonotypes possibly around $10^{7}$ [19,20]. The processes of receptor diversification and negative selection for immune self-tolerance is largely completed during youth. The TCR repertoire shows a linear loss of naive T cell diversity with age [21], although more subtle characterizations can be made [22]. Such age-related changes have been hypothesized to contribute to cancer susceptibility, although their impact is not yet clear [23]. The TCR repertoire continues to be a subject of intense research, which we touch on further below. The impact of age-related changes more generally is discussed in the context of checkpoint blockade therapy by Elias et al. [24].

Figure 2. Differences in recognition and downstream processes between CD8$^+$ and CD4$^+$ T cells. CD8$^+$ T cells recognize pMHC-I complexes, where the peptide is a fragment from an endogenous protein typically 8 to 11 AAs in length, which occupies a groove that is closed on both ends. CD4$^+$ T cells recognize pMHC-II complexes, where the peptide is derived from cells or antigens engulfed by the APC and is typically longer, 12 to 25 AAs in length. The MHC-II groove is open on both ends. After activation, CD8$^+$ T cells differentiate into cytotoxic T lymphocytes (CTLs), whereas CD4$^+$ T cells differentiate e.g. into T helper cells, depending on receipt of further cytokine signals. APC: antigen-presenting cell; MHC: major histocompatibility complex; TCR: T cell receptor
The idea that one clonal TCR recognizes one specific antigen has been supplanted by the notion that TCRs are cross-reactive. A discussion of how TCRs must be cross-reactive in principle is given by Sewell[25]. Indeed, TCR recognition that straddles the self/non-self boundary (e.g. between self and microbial peptides) underlies the theory of molecular mimicry, whereby bacterial antigens do not provoke attack or conversely may lead to autoimmune disease[26]. The specific mechanisms are now being worked out[27,28]. Similarly, mutant tumor proteins may avoid immunogenicity by being cross-reactive with self-proteins.

Each individual’s immune system will also have peptides that it cannot recognize, which can be characterized as “holes” or “blind spots”. These can arise both from gaps in presentation (lack of peptide-MHC binding) or recognition (absence from the TCR repertoire)[29]. A vaccine based solely on an antigen in such a hole will not work for that individual. Such phenomena are seen in the context of microbial immunity[30,31]. The concept of original antigenic sin[32] states that such a hole can paradoxically be created by initial exposure to an antigen, as the immune system does not mount a novel response when it encounters a slight variant.

Immunological research continues to reveal new features. Activated CD8+ T cells were found to require cross presentation, i.e. co-stimulation by dendritic cells that can present exogenous antigens on MHC-I, for full induction of cytotoxic response[33,34]. The CD4 lineage was resolved into four lines[35] and then a plastic set of more[36]. Some CD4+ T cells can acquire cytotoxic activity[37,38] (i.e. not only CD8+ T cells can be cytotoxic). More recently, a “second touch hypothesis”[39] suggests that the high-level picture for polarization of T cells may not yet be complete. New immune cell subtypes continue to be discovered[40]. As one consequence, mathematical modeling of the immune system is likely to remain a difficult endeavor for some time.

We mention here briefly the once-dominant view of cancer as an autonomous genetic disease, as captured by the original “hallmarks of cancer”[41]. The cancer phenotype arises as a result of selection pressure on genome mutations, leading to acquisition of limitless growth and survival potential, with genome instability as an “enabling characteristic”. These mechanisms also underlie cancer’s uncanny ability to acquire additional phenotypes such as eliciting immune tolerance and angiogenesis. A recent proposal that epigenetics alone may be sufficient to generate the hallmarks of cancer[42] may, amongst other things, alter our understanding of the time scales involved in tumor response[43].

The careful examination of tumor cell evolution and its therapeutic implications are in its beginning stages[44-47]. Principles such as antagonistic pleiotropy[48], where reproductive fitness in youth is played off against fitness in old age, are also sometimes raised as setting fundamental biological limits.

For further background on cancer and immunology, the reader can consult reference books[49,50], a three-volume series[51], and a broad history from a contrarian perspective[52].

**RECENT DEVELOPMENTS IN CANCER IMMUNOTHERAPY**

Modalities of T-cell based immunotherapy

The design of currently popular T-cell based immunotherapies can be described as follows:

- **Release the brakes**: checkpoint blockade[53];
- **Boost instruction, via antigens**: cancer vaccines;
- **Boost instruction, via cell transfer, bypassing presentation**: adoptive dendritic cell therapy[54,55];
- **Boost recognition, via cell transfer, bypassing instruction**: adoptive T cell therapy[56];
- **Boost recognition, via cell transfer, bypassing instruction and MHC restriction**: adoptive chimeric antigen receptor T-cell (CAR-T) therapy[57,58].

All of these therapies are based on T cells. Checkpoint blockade therapy is distinguished by not targeting cancer, relying instead on the host immune system training (or having already trained) itself to target tumors.
At the other end of the spectrum, CAR-T therapy does not rely on the host immune system for tumor killing. These span so-called active to passive therapies. Passive therapies do not necessarily induce immune memory, although T cell proliferation may allow extended response. The various immunotherapies can be visualized in an informative hierarchy\textsuperscript{[59]}. The 2014 Society for Immunotherapy of Cancer (SITC) primer provides an unhurried perspective on many of these developments\textsuperscript{[60]}.

While the current wave of immunotherapies was heralded by dendritic cell therapy (sipuleucel-T)\textsuperscript{[55]}, the most notable breakthrough was probably the development of anti-CTLA4 checkpoint blockade, which utilizes antibodies to block receptors that inhibit T cell activation. This treatment allowed some of the first demonstrations in humans of the therapeutic efficacy of neoantigen-specific T cells\textsuperscript{[61]}. Checkpoint therapies based on programmed cell death protein 1 (PD-1) and programmed death-ligand 1 (PD-L1) blockade have further demonstrated improved efficacy with reduced toxicity.

Impacts of immunotherapy on standard practice

The mainstream acceptance of cancer immunotherapy has stimulated efforts to modify clinical trial reporting\textsuperscript{[62]}, with the introduction of “immune-related” adverse events (irAE) and response criteria (irRC). Progression criteria must now allow for pseudo-progression, i.e. the appearance of growing or new lesions that indicate T cell infiltration. A call for “assay harmonization” seeks to reduce variability in cellular immune response reporting. Survival criteria must account for time-dependent hazard ratios, with agent-specific delays in Kaplan-Meier survival curve separation ranging from four to eight months.

Clinical trial design itself is evolving, a process that began in response to targeted therapies (precision oncology) and is now accelerating\textsuperscript{[63]}. This has seen the advent of expansion cohorts, and platform, bucket, adaptive\textsuperscript{[64]}, and seamless trials. It will be increasingly important to understand the cohort and trial design to interpret results.

We note in passing the recent reports of hyperprogression\textsuperscript{[65]}. Tumor size has been observed to dramatically increase with anti-PD-1/PD-L1 treatment, although whether this is more than a statistical fluctuation has been questioned\textsuperscript{[66]}. It is nevertheless safe to say immunotherapies behave differently than previous standard therapies.

The effort to go beyond tumor cell-based staging has begun with the proposal of an Immunoscore \textsuperscript{[67]}, which quantifies the density of CD3\textsuperscript{+} and CD8\textsuperscript{+} T cells in solid tumors. Due to its prognostic value, it has been proposed to augment traditional tumor size/nodal status/distant metastasis (TNM) staging\textsuperscript{[68]}.

Recent advances have triggered a reconsideration of the effect of conventional therapies (surgery, chemotherapy, radiation) and of molecularly targeted therapies\textsuperscript{[69,70]}. Oncogenes such as Myc have been found to also regulate immune response. When such oncogenes are inactivated, immune response is restored and plays a role in the subsequent “oncogene withdrawal”\textsuperscript{[71]}. Chemotherapy perhaps surprisingly also appears to rely in part on the immune system for cytotoxic effect\textsuperscript{[72]}

Cancer immunotherapies can in principle have much milder side effects compared to radiotherapy and chemotherapy. In practice, they are associated with their own spectrum of adverse events\textsuperscript{[73,74]}. In particular, cytokine release syndrome (“cytokine storm”) can lead to organ failure and death. Both treatment efficacy and adverse events are associated with proliferative and persistent cellular responses, which can vary significantly between individuals, thus requiring careful monitoring\textsuperscript{[75]}. Adverse events associated with neoantigen vaccines appear to be relatively mild, compared to adoptive cell transfer, checkpoint blockade, and tumor-associated antigen (TAA) vaccine therapies\textsuperscript{[76]}.
NEOANTIGEN VACCINES

Introduction

We now turn to neoantigen-based cancer vaccines. The objective of a vaccine is to introduce a small amount of material to instruct T and B cells to eliminate invaders that present the cognate antigen. Vaccines in general can be prophylactic (preventative) or therapeutic (cure or control of observable disease). Current neoantigen vaccines are therapeutic, with the goal of restoring immune surveillance of a tumor that has likely already evolved to evade the immune system (e.g., through immunoediting; see below).

Cancer cells are genomically unstable, which leads to the expression of novel proteins due to non-synonymous mutations. Many of these are likely to be immunogenic and are termed neoantigens. Vaccines that precisely target such neoantigens (also known as tumor-specific antigens, or TSAs) would prime an immune response that rejects tumors while sparing normal tissues, leading to optimal therapies with mild if any toxicity. A timeline that traces the foundations of this idea back to 1943 is provided by Coulie et al.

Types of antigen-based cancer vaccines

Prior to the advent of next-generation sequencing, cancer vaccines were developed based on TAAs or cancer germline antigens. These self-antigens are overexpressed in tumors, or normally expressed only during development but re-expressed in tumors. Vaccines targeting these can be produced in advance at lower cost and applied across a range of tumors that share expression of the target. As the targets are self-antigens, such vaccines are possibly limited by self-tolerance and adverse events. Tumor resistance mechanisms, many of which are shared with neoantigen vaccines, are also a prominent concern. Another class of targets are shared tumor neoantigens, which are commonly found across a subtype of tumor. As in TAAs, the vaccine can be produced beforehand, and treatment progress can be easily followed, as the neoantigen epitopes (i.e., recognized peptide fragments) are typically well known. Such epitopes however may not be the most effective for any given tumor.

With massively parallel sequencing and MHC binding and functional prediction software tools, the key hurdle to developing personalized neoantigen vaccines can now be overcome. Vaccines custom designed for each patient represents a paradigm shift in cancer treatment.

Some of the strengths and weaknesses of the neoantigen vaccine approach are summarized in Table 1 and are discussed further below.

Vaccine formulation and administration

In addition to the selection of epitopes, a number of other considerations can strongly influence the success or failure of neoantigen vaccines. Cancer vaccines can be formulated as whole cells, peptides/proteins, RNA, DNA, and glycolipids. Vaccines are typically formulated as peptides, due to ease of construction and low cost, although these are often observed to be weakly immunogenic. They can be modified to enhance delivery to immune cells and improve pMHC binding stability. Synthetic long peptides require dendritic cell processing, argued as essential for durable response. Protein vaccines are more immunogenic but have a higher risk of anaphylaxis. The robust discussion about designing and assessing peptide vaccines has been reviewed by Kumai et al. DNA vaccines introduce DNA coding for antigenic fragments into host cells, where they are expressed and lead to the presentation of epitopes via the MHC-I pathway. They are generally safe, stable, and easy to produce at low cost, although currently weakly immunogenic. The vaccine or delivery vehicle itself can be attacked by the host immune system. RNA vaccines can encode several epitopes on a single molecule, can trigger the innate immune system, and are not at risk of integrating into the genome. Whole cell vaccines that employ weakened or killed tumor cells can trigger immune response with the entire complement of tumor antigens, without specific instruction of the immune system, reducing time and expense. They may however induce immune response to self-proteins.
Table 1. Strengths and weakness of the neoantigen vaccine approach

<table>
<thead>
<tr>
<th>Strengths</th>
<th>Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precise targeting</td>
<td>Need for tumor biopsy (in general)</td>
</tr>
<tr>
<td>Mild adverse events</td>
<td>Need to overcome tumor defenses</td>
</tr>
<tr>
<td>Few constraints on dosage</td>
<td>Slow induction of immune response</td>
</tr>
<tr>
<td>Better profile than TAA vaccines</td>
<td>May not be applicable to tumors with few mutations</td>
</tr>
<tr>
<td>No need for T cell extraction and ex vivo growth</td>
<td>Unreliable epitope binding prediction algorithms</td>
</tr>
<tr>
<td>Many opportunities to optimize/combine formulations</td>
<td>Time lag from biopsy to vaccine</td>
</tr>
<tr>
<td>Multi-epitope designs can compensate for inaccurate binding predictions, tumor heterogeneity and evolution</td>
<td>Cost</td>
</tr>
</tbody>
</table>

Other considerations are choice of carrier, delivery vehicle[91] (including bacteria[92] or viral vectors[93]), and administration route (intravenous, intratumoral, subcutaneous, intra-lymph node, nasal, ingested). Further afield, cancer vaccine engineering has emerged to offer benefits such as lymph node targeting, reduced systemic toxicity, elimination of ex vivo expansion requirements, and controlled release of immunomodulators while ignoring suppressive signals[94-96].

Neo-epitope binding prediction

Neoantigen vaccines are produced by first inspecting the patient’s tumor for immunogenic peptides, specifically epitopes[97]. TCRs recognize linear epitopes, i.e. a continuous fragment of an antigen. Note that pMHC binding is a necessary but not sufficient condition for immunogenicity.

The neo-epitope selection problem can thus be reduced to finding mutant peptides that bind well to the patient’s MHC alleles. This is amenable to computational treatment and is one of the most prominent applications of machine learning to immunology[98]. The realization that such bioinformatic approaches can reveal a “gold mine” of targets and that neoantigen vaccines were feasible can be traced back to a 2008 paper[99].

A simplified neo-epitope selection pipeline can be described as follows:

- Perform exome sequencing of tumor and normal tissue to identify non-synonymous single nucleotide variants and generate an initial list of candidate genes;
- Perform RNA-Seq to confirm expression;
- Use informatics tools to predict neoantigen-derived peptides that bind to the patient's set of HLA alleles;
- Filter candidates based on survival or growth function (“driver genes”);
- Choose the top 10 or 20 epitopes.

Proteasomal cleavage predictions[100] can also be incorporated into the workflow, although the predictive value is rather low, due to the lack of sufficient training data[98].

Numerous excellent reviews of the available tools are available[82,98,101]. The Immune Epitope Database[102] is probably the most prominent epitope database and analysis resource, freely available on the Web. TANTIGEN[103] is a database of tumor-tissue derived antigens with experimentally validated HLA binding. Step-by-step instructions on the use of a prominent suite of tools is available[104]. Mutant Peptide Extractor and Informer[105] is a web-based tool that attempts to integrate best practices and simplify neo-epitope analysis and selection for non-bioinformaticians (limited to MHC-I epitopes). ImmunoNodes[106] is a software framework for building complex immunoinformatics workflows, such as those for neo-epitope selection.

Amongst other challenges, prediction of MHC-II peptide binding lags behind MHC-I prediction, partly due to the greater length of loaded peptides that interact in flanking regions with highly polymorphic alleles.
Also, binding data does not exist for many less common MHC alleles, which has given rise to “pan MHC” algorithms with somewhat reduced performance. Therapeutic strategies based on so-called promiscuous epitopes that bind to several MHC alleles may place less stringent requirements on the accuracy of binding predictions\(^{107}\).

The extent to which epitope binding scores are a good surrogate for immunogenicity remains unclear. Peptide binding stability rather than affinity has been proposed as a better predictor of immunogenicity\(^{29}\). Numerous factors can affect antigen presentation and recognition processes, such as pH, inflammation, and peptide post-translational modifications\(^{108}\).

Many of the structural aspects of peptide-MHC binding and TCR recognition are reviewed by Hudrisier and Gairin\(^{109}\). Important aspects of the problem formulation can be found e.g. in the references cited by Meydan \textit{et al}..\(^{14}\) A recent examination of empirical TCR-pMHC kinetic constants measured in three-dimensional assays suggests these may not accurately reflect dynamics in a two-dimensional context, such as T cell scanning of the APC surface\(^{116}\). This could suggest that some of the data underlying current epitope binding prediction algorithms needs to be re-measured.

In general, while immunogenic antigens tend to have high binding scores, the converse does not hold\(^{111}\). In addition, indels and gene fusions are typically not chosen, due to the difficulty of predicting binding. Snyder and Chan\(^{110}\) caution that current prediction tools on their own are not ready for routine clinical use.

**Choice of epitope candidates**

In tumors with a large number of mutations, the candidate filtering step is essential to avoid being overwhelmed by false positives\(^{112}\). Mass spectrometry has been effectively used for this task by identifying MHC-bound peptides\(^{113}\). Indeed, it can be used to generate candidates on its own\(^{114,115}\). There remain possible issues with sensitivity and translation into a clinical setting\(^{112}\). Combining functional analysis and T cell detection via multimers can help in the search for tumor rejection epitopes\(^{116}\). Proximity ligation assays can assess whether an antigen is presented \textit{in situ}, although this requires a mutant-specific antibody\(^{117}\). Another approach tests epitopes experimentally in MHC-transgenic mice\(^{118,119}\). Further work is necessary to validate the efficacy of such workflows\(^{120}\). An interesting suggestion is that PD-1\(^+\) peripheral blood cells are enriched in tumor neoantigens, from which candidate epitopes can be derived\(^{121}\).

There is a general exhortation to prioritize genes that target essential tumor “driver” functions such as growth and survival. This however may not be too helpful, as only a small percentage of neoantigens are of this type in e.g. melanoma\(^{112}\), the vast majority being “passenger” mutants not associated with cell transformation. Efforts to expand and/or refine the list of functional cancer genes may help in this regard\(^{112,123}\).

Current vaccine strategy employs several epitopes to address tumor heterogeneity and reduce acquisition of resistance, while also compensating for the imperfect predictive value of pMHC binding tools. The phenomena of immunodominance\(^{124-127}\) and T cell cross-reactivity\(^{128}\) suggests that simply increasing the number of epitopes in a vaccine may not be advisable, as a suboptimal epitope may interfere with the others in a dominance hierarchy, and auto-immunity remains an issue. Indeed, pioneering efforts in cancer epitope selection\(^{111,129}\) found possible instances of immunodominance. Initial experience with long peptides on the other hand suggested this may not be an issue\(^{130}\). Further work is required to understand how to choose the number of epitopes to include in a vaccine, which could be e.g. cancer type-specific. The thinking behind many current vaccine approaches is examined by Kumai \textit{et al}..\(^{133}\) who also describe four steps to developing cancer vaccines and five ways of monitoring the response.

Initial effort e.g. in adoptive cell transfer was focused on MHC-I restricted epitopes to elicit direct tumor cell killing. Attention has now shifted to MHC-II restricted epitopes, in part due to the fuller realization that
CTLs require CD4+ help\textsuperscript{[132-135]}. Indeed, adoptive cell transfer of CD4+ T cells was enough to induce tumor regression in a mouse model of melanoma\textsuperscript{[136]} and in a human patient\textsuperscript{[137]}.

**Initial human trials**

The personalized neoantigen vaccine strategy has begun to reach the clinic with the recent reports from two Phase I trials\textsuperscript{[76,138]}. These trials, on patients with advanced melanoma, demonstrate that such therapies are safe and induce a targeted neoantigen-specific response as designed. Ott et al.\textsuperscript{[76]} enrolled 6 patients who had no evidence of disease after surgery, with 4 remaining tumor-free after 25 months. Sahin et al.\textsuperscript{[138]} enrolled 13 patients, and 8 patients remained tumor free after 23 months. The time to develop personalized vaccines (weeks to months) remains a key obstacle, especially for patients with advanced disease.

In an apparent pattern, both trials utilized MHC-I binding scores to select neo-epitopes (Sahin et al.\textsuperscript{[138]} combined these predictions with MHC-II binding scores). The vaccines were then seen to activate CD4+ T cells, possibly because MHC-II binding is less restrictive\textsuperscript{[139]}. In more detail, Ott et al.\textsuperscript{[76]} selected neo-epitopes using predicted binding to HLA-A and HLA-B, and employed long peptides (15-30 amino acids) in several pools targeting up to 20 neoantigens for five priming and two boosting vaccinations injected subcutaneously. They observed CD4+ response almost immediately and a peak response generally at 16 weeks, and found two to four immunogenic peptides per patient. Sahin et al.\textsuperscript{[138]} ranked mutant epitopes on a combination of predicted MHC-I and MHC-II binding, plus high expression and variant allele frequency, and chose 10 epitopes per patient. Two synthetic RNAs were used to encode five 27mer peptides with the single nucleotide variant (SNV) at position 14. Patients were treated with at least eight and up to twenty neo-epitope vaccine doses injected into inguinal lymph nodes, and T cells were developed against at least three mutations per patient, with the majority being exclusively CD4+ responses. Neo-epitope specific CD8+ T cells expanded within two to four weeks.

**WHO BENEFITS?**

We now return to a more general discussion of immunotherapies. The seminal studies in checkpoint blockade therapy have primarily been done in melanoma and lung cancer. A survey of solid tumor types that have been studied with immunotherapy, with an emphasis on understudied cancers, has been made by Young\textsuperscript{[140]}.

In an emerging consensus, the cancers that are best indicated for immunotherapy are slow growing, exhibit high mutational load and low burden at the start of therapy, and are inflamed\textsuperscript{[141]}. This suggests immunotherapies may be more effective in the early stage disease setting. A correlation has been observed in checkpoint blockade between somatic mutations per megabase and objective response rates\textsuperscript{[111]}. Mutagen induced cancers such as melanoma and lung cancer subtypes with high mutational loads were some of the first to show durable complete response. Estimating mutational load using custom reduced gene panels\textsuperscript{[142]} or pre-existing ones\textsuperscript{[143]} may aid quick assessments within the clinic. Cancer types are characterized by a wide range of mutational loads\textsuperscript{[144]}.

Recent work has sharpened focus on mismatch-repair deficiency as a biomarker to identify patients who can benefit from PD-1 blockade, independent of tissue type\textsuperscript{[145]}. The immunological phenotype of microsatellite instability-high (MSI-H) colorectal tumors in particular may be unique\textsuperscript{[146]}. In a literature review of anti-PD-1 clinical trials, atypical responses appeared in all cancer types except tumors with mismatch-repair deficiency and head and neck squamous cell carcinomas\textsuperscript{[147]}.

While checkpoint inhibitors are often presumed to exacerbate the symptoms of patients with inflammatory/autoimmune diseases, anecdotal reports suggest this may not be the case\textsuperscript{[74]}.
RESISTANCE AND ESCAPE

The fact that immunotherapies fail to produce durable responses in a majority of patients has spurred intensive investigations of resistance. Both primary resistance, where no beneficial response is observed, and secondary resistance, where initial benefit is followed by relapse, are observed.

Before proceeding, we first ask why neoantigen vaccines work at all, as they would appear catastrophically prone to failure due to antigen loss. Such loss has been seen in checkpoint blockade therapy, where both chromosomal deletion of clonal neoantigens and negative selection of tumor subclones were observed.

Accumulating evidence suggests part of the answer lies in the phenomenon of antigen spreading, aka cascade. T cells are initially “instructed” by the vaccine epitopes, but effector activity need not be limited to these. This hypothesis suggests that the role of the vaccine is to nucleate immunity to an iteratively growing cascade of antigens, the epitopes of which are then committed to T cell memory. This could be key to a durable response that can also target new tumor mutations. The time required to generate such a cascade could also explain the lag often seen between vaccine administration and objective response. A related idea suggests that the initial vaccine-induced attack reverses immuno-suppressive mechanisms, allowing preexisting CTLs that already recognize other tumor neoantigens to be unleashed in a cascade.

Cancers can resist therapy by circumvention of each of the immune system roles mentioned previously:

- Disrupt presentation:
  - Inhibit MHC-I expression;
  - Disrupt dendritic cell trafficking to lymph nodes (i.e. T cell priming);
  - Disrupt peptide processing;
  - Disrupt recognition (prevent T-cell trafficking from lymph nodes back to tumor, exploit holes in TCR repertoire);
- Exploit immune suppressive mechanisms.

In addition, tumor cells employ explicit defense mechanisms, e.g. downregulation of pro-apoptotic pathways, and counterattack by secreting FasL death ligands, resulting in CTL death.

The tumor microenvironment (TME), i.e. the nearby cellular, vascular, and extracellular matrix environment remodeled by the tumor, is the focus of much research into resistance. Here tumors are seen to induce host self-protective mechanisms, through recruitment of suppressive cells such as MDSCs, regulatory T cells, and tumor associated macrophages. The tumor creates an immune privileged site, akin to the eye and brain, that excludes T cells. Recent work provides a detailed picture of effector T cell exclusion based on a β-catenin signaling mechanism.

The TME is a metabolically demanding place, with competition for oxygen and nutrients. Tumor cells can outlast T cells through the induction of T cell anergy or exhaustion, part of a class of phenomena termed T cell dysfunction. There also remains the possibility that the tumor simply grows faster than an often aged and weakened immune system can eliminate it. A careful 2011 discussion of the TME in which CTLA4 operates is given by Quezada et al.

As antigen-specific vaccines seek to activate the adaptive system, harnessing the innate immune system, and in particular natural killer (NK) cells, would appear to be an attractive complementary approach. Unlike naive CD8+ T cells that require time to acquire cytotoxic activity, NK cells are “ready to kill” NK and dendritic cells engage in mutual activation, and the former can “edit” the latter population. Inhibitory receptors that bind MHC-I allow NK cells to recognize and eliminate cells that do not present MHC-I, thus closing one avenue of tumor cell escape. The activating receptor NKG2D has attracted particular interest, as
its ligands (NKG2DL) are commonly expressed by tumors. Tumor cells however can also express NKG2D and hijack NKG2DL signaling to drive stem-cell like behavior\textsuperscript{170}. NK cells participate in tumor-induced polarization, acquiring a pro-tumorigenic and pro-angiogenic phenotype\textsuperscript{171}.

Dammeijer \textit{et al}.\textsuperscript{172} provide a thorough review of primary and secondary resistance mechanisms and treatment options for re-sensitizing tumors. Guo \textit{et al}.\textsuperscript{80} provide a compact review of the wide variety of immunosuppressive mechanisms employed by tumors. Chen and Mellman\textsuperscript{173} describe these mechanisms in the context of the cancer immunity cycle. A concise table of many elements that underlie tumor escape is given by Accolla and Tosi\textsuperscript{174}. Seliger\textsuperscript{150} and Seliger \textit{et al}.\textsuperscript{175} review MHC-I and MHC-II-based evasion mechanisms, respectively. A report of HLA allele-specific risk of metastasis in papillary thyroid cancer\textsuperscript{176} provides evidence that MHC allele status impacts cancer progression.

\textbf{Frameworks for understanding tumor-immune system interactions}

Reducing therapeutic resistance is closely tied to our understanding of how cancer arises in the context of the immune system, which we briefly discuss here. The primary framework for understanding the interplay between cancer and the immune system is known as immunoediting\textsuperscript{177}. This posits that selection pressure from the immune system “edits” the tumor, forcing it to find a custom response to the local and systemic state of the immune system in order to escape immune pressure after many years of genetic changes.

Therapeutic success is often defined by reduction in the incidence and impact of metastatic cancer. The origin and nature of metastases is a dynamic research area, with much remaining to be discovered. Do metastases represent dissemination of cancer cells from a primary tumor in late stage disease (Halsted-Meyer theory)\textsuperscript{178}, or do they reflect the outgrowth of pre-existing cells that disseminated early on? TNM staging\textsuperscript{180} encourages the former perspective. Weichselbaum and Hellman\textsuperscript{181} posit the existence of cancers with intermediate metastatic potential. The hypothesis of cancer dormancy posits that tumor cells may disseminate early and are forced into dormancy in order to survive immune surveillance\textsuperscript{182,183}. From this perspective, one goal of immunotherapy is to keep such cells dormant, as opposed to attempting to eliminate them all\textsuperscript{184,185}. Such topics have been covered in a chapter-length review, including the different kinds of dormancy, the role of circulating tumor cells and of innate and adaptive immune cells, and ideas for keeping dormant tumor cell indolent\textsuperscript{186}. It is evident that this research area is both difficult and in its early stages.

\textbf{NEXT STEPS}

A number of authors have sought to identify the most urgent and interesting near-term trends. Whiteside \textit{et al}.\textsuperscript{187} and Hoos\textsuperscript{188} foresee a focus on, amongst other topics, understanding PD-1 nonresponders, targeting the tumor microenvironment, improving therapy of tumors with few mutations or low tumor infiltrating lymphocyte count, better tumor and patient assessments, combination therapies, and biomarkers of response. This includes the proposed acknowledgment that stable chronic disease (“functional cure”) is a worthwhile endpoint.

Combination therapy is hailed as possibly the best way to increase the percentage of responders. Current examples of combination therapy tend to have a reactive character, applying an additional therapy in response to failure of an initial one. A strong call was issued in 2015 for increased funding of trials to study how to combine molecularly targeted and immuno-therapies\textsuperscript{189}.

In general, it is hoped that progress can be made through examination of rational combinations\textsuperscript{172}. The meaning of “rational synergy” has been dissected in the context of cytotoxic drug combinations\textsuperscript{190}. The diversity of currently available modalities may allow combinations where treatments are carefully scheduled to act as e.g. “mutual adjuvants”. Adjuvants\textsuperscript{191-193} continue to be topics of active research, with the line
between adjuvant and therapeutic agent blurring. New adjuvants and combinations thereof show promise and are proposed as the focus of clinical trials\textsuperscript{194}.

Improved outcomes are being reported for combined checkpoint therapies, though at the cost of more adverse events\textsuperscript{195,196}. Even those withdrawing from combination treatment due to severe adverse events may still be receiving benefit\textsuperscript{197}. Emens \textit{et al.}\textsuperscript{198} present a list of clinical development priorities to push forward the state of the cancer immunotherapy art.

The literature on combination therapies is expanding rapidly. Dunn and Rao\textsuperscript{199} have reviewed the combination of epigenetics and immunotherapies. Increasing attention is being paid to traditional targets of the innate immune system. Expression of endogenous viral peptides\textsuperscript{200} and dsRNA have been shown to lead to innate and adaptive responses. The role of microbes, both as commensal microbiota that are modulatory targets\textsuperscript{201} and as therapeutic agents\textsuperscript{202}, are subjects of active research. The universe of immunomodulator targets is rapidly growing\textsuperscript{203}, expanding the scope of possible combinations.

We note that therapy modalities in combination are not necessarily additive, e.g. the combination of chemotherapy (tyrosine kinase inhibitor) and a TAA vaccine required careful scheduling to avoid failure in a mouse model\textsuperscript{204}. A literature review of combination (mostly targeted) therapies in metastatic renal cancer expresses both the promise and challenges of extracting benefit from such studies\textsuperscript{205}. Careful reporting as captured e.g. in the Consolidated Standards of Reporting Trials (CONSORT) guidelines\textsuperscript{206} will be essential for trial data to have maximum value.

**MONITORING AND MODELING**

Immunological research is increasingly driven by the ability to gather data, often in a high throughput manner. This opens new vistas that will allow therapy to be properly targeted and monitored, and may eventually alter the character of therapy itself.

The ease with which data can now be generated highlights the responsibility to employ best practices in experimental and trial design, data acquisition (including patient metadata), and downstream analysis. The
accumulation of high quality datasets should ideally go hand in hand with the ability to model the data, with the ultimate goal of defining optimal interventions to reach a desired outcome (e.g. disease stabilization or cure) \[77\].

An initial goal is the discovery of prognostic and predictive biomarkers. These can be used for treatment selection \[207\], e.g. high PD-L1 expression level for pembrolizumab treatment \[208\]. At a more rigorous level, biomarkers indicate system state of the immune system, cancer, or both. The importance of prospective studies for data collection and analysis has long been emphasized. REMARK guidelines (REporting recommendations for tumour MARKer prognostic studies) attempt to capture the minimal information needed to objectively assess the import of a given biomarker study \[209\]. This baseline however is still commonly not met \[210\].

The field of immune system-related prognostic and predictive biomarkers is complex and rapidly advancing. Urgent efforts are now being made to translate current knowledge and capabilities to the understanding of baseline immunity and response monitoring, and thence the choice of predictive biomarkers \[68,211\]. Magnetic resonance imaging (MRI) based biomarkers of response to immunotherapy have been recently proposed \[212\]. Systemic immune response coordinated across tissues has been observed to be essential to tumor rejection \[213\]. The fraction of tumor-infiltrating partially exhausted cytotoxic T lymphocytes (peCTLs) correlates with response to anti-PD-1 monotherapy, with a low fraction indicating the use of combination checkpoint blockade therapy \[214\]. A possible implication is that checkpoint blockade therapy is most effective when the immune system has already mounted a tumor-specific if suppressed response.

TCR repertoire profiling shows promise in immune monitoring and perhaps response prediction \[215,216\]. Checkpoint blockade is seen to induce diversification of T cell receptor repertoire \[217\], which has been suggested as a biomarker for PD-1 inhibitor disease stabilization \[218\]. The assessment of TCR repertoire diversity is becoming increasingly accessible \[219\]. Important choices such as library preparation method, in-house versus service provider, output data type (raw and/or analyzed), and the use or not of unique molecular identifiers must first be matched to project goals \[220\]. Basic features of the T cell receptor repertoire are still being revealed, e.g. unexpectedly biased distributions of TCR receptors (CDR3 sequence similarity networks) that change in stereotyped ways with aging, immunization, and antigen selection \[221\]. Progress has been reported in developing statistical means of “reading” T cell memory, as relates e.g. to cytomegalovirus status and HLA typing \[222\].

As we dissect components and interactions in more detail, the research enterprise can begin to embrace variation to learn better from animal models \[213,214\] and humans \[215,216\], including with respect to age \[223\]. Data sharing can help ensure technical advances are employed towards broad evidence-based progress \[224\]. In this regard, standards for reporting neo-antigens, HLA alleles, and TCR repertoires may need to be developed.

Adoption of high throughput technologies such as massively parallel sequencing, immunosequencing, microarrays, mass cytometry, and DNA-barcoded pMHC multimers has led to the advent of systems immunology \[20\]. Rather than dissect mechanistic relationships between a few actors, systems methods attempt to capture the behavior of the immune system as a whole. The resulting descriptions tend to have a multi-scale (hierarchical) character in both space and time \[229,230\]. The wide variety of available modeling formalisms and applications has been surveyed by Narang et al. \[231\].

Mathematical modeling has begun to impact the clinic through efforts to optimize dosage and timing (“schedule optimization”), which have gained a foothold in chemotherapy \[232,233\] and radiotherapy \[234,235\]. There is now a rich literature on the modeling of immunotherapy \[236,237\]. As an example, modeling the kinetics of the immune response \[239\] reveals the possibility that a proper choice of schedule can summon a robust T cell
response, overcoming what appears to be tumor resistance. To increase their impact, such models may need to integrate into Bayesian adaptive trials. One theme borrowed from the physics community is to develop simplified abstract models. Such models can generate powerful predictions, derived from the concept of universality. The observation of unexpected dynamical patterns in the immune system such as oscillations over several days suggests that phenomenological models have an important role to play.

Detailed mechanistic models employ knowledge at the molecular or cell level to explain and predict phenomena. A recent attempt to model the cancer-immune system interaction using 12 immune cell types and 13 cytokines plus cancer cells finds steady state “basins” corresponding to escape, elimination, and equilibrium phases in immunoediting, while also finding oscillatory states. The interested reader is referred to volumes focused on cancer and the immune system. A textbook on computational immunology has recently been released.

From an artificial intelligence perspective, therapy can be viewed as planning in the presence of uncertainty. The idea that the immune system can be “steered” has been demonstrated by proof-of-concept in silico work. Cancer cells can be treated as adversaries in a game theory context. In the clinical trials arena, reinforcement learning approaches promise a model-free approach to sequential treatment selection.

CONCLUSION
After a long history of doubt and failure, checkpoint blockade therapy has opened the door for cancer immunotherapy. With this key modality now accepted, the full weight of technological progress can be brought to bear. New tools provide windows through which the process of disease and treatment can be viewed. Their integration will allow increasingly sophisticated descriptions of immune system and tumor state. Neoantigen cancer vaccines in particular are beneficiaries of this new environment and are poised to lead the way to more precise and effective therapies.

While neoantigen vaccines can now be created with workflows that are increasingly standardized and almost routine, many challenges lie ahead to elicit their true potential. Foremost is gaining a better understanding of primary and secondary resistance. This can be viewed in the light of theories in which cancer cells and the immune system train each other, for better or worse, over decades.

Combination therapies are now pursued as the next step forwards. The examination of all possible protocols may however become infeasible or at least inefficient. Principled methods will need to be developed to systematically identify promising approaches and learn from both successes and failures. This complexity is also an opportunity to formalize therapy as a strategy and not simply an application of magic bullets. Over the longer term, this promises growth in novel interventions that integrate technology, data, models, and algorithms as part of an interdisciplinary biomedical science.

DECLARATIONS
Acknowledgments
The authors wish to express their gratitude to Rong (Shirly) Li, who created Figures 1 and 2. They also wish to thank Qiubin Lin, Bin Yu, Gang Fang, Antoine Danchin, Chenli Liu, and Chung-I Wu for enlightening conversations.

Authors’ contributions
Conception: Huang JD
Design: Tokuyasu TA, Huang JD
Financial support and sponsorship
This work was supported by a National Basic Research Program of China (973 Program, No. 2014CB745202); the Shenzhen Peacock Team Project (No. KQTD2015033117210153); the Shenzhen Science and Technology Innovation Committee Basic Science Research Grants (No. JCYJ20170413154523577; No. JCYJ20150629151046896); the Guangdong Natural Science Funds for Distinguished Young Scholar Grant (No. S2013050016987).

Conflicts of interest
There are no conflicts of interest.

Patient consent
Not applicable.

Ethics approval
Not applicable.

Copyright
© The Author(s) 2018.

REFERENCES
12. Manz BN, Jackson BL, Petit RS, Dustin ML, Groves J. T-cell triggering thresholds are modulated by the number of antigen within individual T-cell receptor clusters. Proc Natl Acad Sci U S A 2011;108:9089-94.


139. Young MR. Cancer immunology with a focus on understudied cancers as targets for immunotherapy. *Int J Mol Sci* 2017;18:E127.


