Commentary
See related article by Fritsch et al., p. 522

Can We Predict Mutant Neoepitopes in Human Cancers for Patient-Specific Vaccine Therapy?

Eric R. Lutz and Elizabeth M. Jaffee

Most human cancers acquire tens to hundreds of somatic mutations (termed the "tumor mutome") during their development (1). Each of these mutations has the potential to generate one or more novel T-cell antigens (termed "neoepitopes") uniquely specific to each individual patient's tumor. Because these neoepitopes are not present in the germline, and are not encountered until after the onset of oncogenesis, repertoires of high-avidity T cells capable of recognizing them may avoid central tolerance and escape deletion in the thymus. For these reasons, numerous investigators have proposed that the tumor mutome provides an attractive source of antigenic targets for developing patient-specific tumor vaccines (2–5).

Because it is already possible to rapidly and comprehensively identify tumor mutations using next-generation DNA- and RNA-sequencing technologies (1), the first technical hurdle for the development of this approach has been overcome. However, it may not be practical to target the full repertoire of mutations expressed by a patient's tumor, especially in tumor types associated with high mutation rates, such as melanomas and lung cancers in cigarette smokers (1). Furthermore, regardless of the total number, only a fraction of mutations are expected to generate HLA-binding (known as MHC in mice) epitopes capable of serving as relevant vaccine targets, and it is possible that attempting to target all possible mutant neoepitopes may drown out the relevant targets and reduce efficacy. Even if it is possible and equally effective to target all possible neoepitopes, being selective would at least be advantageous from an economic and feasibility perspective. In addition, depending on the vector chosen (another variable that needs to be evaluated), there will be limits to the number of candidate neoepitopes that can be packaged into the vaccine. Thus, a critical challenge facing the development of patient-specific tumor vaccines is establishing guidelines for selecting which mutations should be included as vaccine targets, and which should be left out.

At a minimum, it will be necessary to choose (or at least enrich for) vaccine targets that are actually processed and presented by antigen-presenting cells and presented on HLA by the tumor to activate the T cells that can recognize these epitopes and mediate tumor lysis. Only considering CD8$^+$ T cells that recognize peptide epitopes typically 8 to 10 amino acids long, and occasionally 11 amino acids long, each mutation could generate 38 different peptides that could potentially bind to an HLA class I molecule. For any of these peptides to produce a targetable neoepitope, the peptide must be proteolytically exposed, but not destroyed, be chaperoned into the endoplasmic reticulum, and if capable (most are expected to be incapable), bind to MHC class I to be delivered to the cell surface for T-cell recognition. CD4$^+$ T-cell epitopes are longer and are processed differently, but also must be exposed and not destroyed, and they must have affinity for HLA class II molecules instead of HLA class I.

Because only peptides that can bind to HLA class I or II provide eligible T-cell targets, one possible strategy for selecting vaccine targets is to choose candidate neoepitopes based on their predicted affinities for the HLA molecules expressed by the patient, determined using HLA-binding affinity prediction algorithms (6–8). This strategy is often referred to as "reverse immunology." Although the algorithms have been updated and improved over time, this approach was initially used almost 15 years ago to identify an HLA-B7–restricted T-cell epitope derived from the tumor-associated antigen carcinoembryonic antigen (9). More recently, use of this approach has facilitated the identification of several mutant tumor neoepitopes recognized by cultured CD8$^+$ tumor-infiltrating lymphocytes used for adoptive immunotherapy in patients with melanoma that were associated with the development of clinical antitumor responses (10). In this issue of *Cancer Immunology Research*, Fritsch and colleagues provide new evidence supporting the use of this approach for selecting candidate targets for patient-specific tumor vaccines (11).

The authors initially pool together a comprehensive list of 40 previously identified mutant tumor neoepitopes recognized by patient CD8$^+$ T cells in association with improved clinical responses. The neoepitopes consisted of 35 missense mutations and five frameshift mutations representing seven different human cancer types, including both solid and hematologic tumors. Approximately 80% of the neoepitopes were tumor-specific somatic mutations, whereas the remaining 20% were polymorphic minor histocompatibility antigens identified following hematopoietic stem cell transplantations. Importantly, for most of the neoeptiopes, T-cell reactivity was more potent against the mutant peptide compared with the corresponding nonmutated native peptide. Because T-cell responses to these neoepitopes were associated with biologic antitumor responses, the authors suggest that they provide an ideal set of neoeptiopes for evaluating the potential for using reverse

Authors' Affiliation: Sidney Kimmel Cancer Center, Johns Hopkins University School of Medicine, Baltimore, Maryland

Corresponding Author: Elizabeth M. Jaffee, Sidney Kimmel Cancer Center, Johns Hopkins University School of Medicine, 1650 Orleans Street, Room 4MO7, Bunting-Blaustein Cancer Research Building, Baltimore, MD 21287. Phone: 410-955-2957; Fax: 410-614-8216; E-mail: ejaffee@jhmi.edu

American Association for **Cancer Research**

immunology to predict targetable neoepitopes for vaccine development.

Of the 40 neoepitopes chosen for this study, 31 were identified using an unbiased approach independent of HLA-binding affinity predictions. Using NetMHCpan version 2.4 (Center for Biological Sequence Analysis, Technical University of Denmark, Lingby, Denmark), HLA-binding affinities were predicted for each of the 38 possible peptides that could be generated by these 31 mutations. For all but one, the naturally recognized neoepitope was the peptide predicted to have the highest affinity. Furthermore, of the 31 neoepitopes analyzed, 20 were predicted to have strong affinity ($IC_{50} < 50$ nmol/L), three were predicted to have moderate affinity ($50 < IC_{50} < 150$), and four were predicted to have weak affinity ($150 < IC_{50} < 500$) for HLA. Only four (13%) of the neoepitopes were predicted not to bind ($IC_{50} > 500$). Thus, 27 (87%) of 31 neoepitopes would have been selected as HLA-binding peptides by NetMHCpan.

This retrospective analysis by Fritsch and colleagues suggests that NetMHCpan can potentially predict up to 85% of mutant neoepitopes. Although it may have been interesting to evaluate multiple algorithms, only NetMHCpan was used because it was found to be most accurate in a previous comparison of algorithms (6). Not mentioned by the authors, the accuracy of HLA-binding prediction algorithms depends on the number of epitopes available for training. Thus, predictions are more accurate for common HLA class I alleles for which more epitopes have been identified than for more rare alleles with fewer epitopes. This, of course, makes it a problem to predict neoepitopes for patients who do not have common HLA alleles, and could also lead to the exclusion of targets that would otherwise be presented by less predictable alleles. Only one of the neoepitopes analyzed was restricted to a rare HLA class I allele represented by less than 1,000 epitopes (the NFYC mutation restricted to HLA-B*52:01). Thus, the set of epitopes analyzed was not powered to assess whether this approach is better or worse for less common HLA alleles. However, it is interesting that neither the mutated nor wild-type NFYC peptide was predicted to bind to HLA-B52.

MHC-binding affinities were also predicted for nonmutated native peptides corresponding to the 35 missense neoepitopes analyzed. For 26 (74%) of them, both the wild-type and mutated peptides were predicted to have comparable affinities. The predicted affinity was higher for the mutated peptide for only six (17%) of the 35 neoepitopes, and neither of the peptides were predicted to bind for three (9%) of the 35 neoepitopes. This is an important finding and confirms that the potency of a T-cell response is determined by more than just the affinity of the neoepitope for its presenting HLA molecule. In fact, the affinity of the T-cell receptor for the HLA/peptide complex also contributes to the outcome of the T-cell response (12).

The predominance of neoepitopes arising from native peptides with similar predicted affinities may be explained by the increased probability of a mutation occurring within a nonanchor residue, which outnumber the two to three anchor positions within an MHC class I–binding peptide 3–4:1. How-

ever, as >98% of the human peptidome is not predicted to bind MHC by NetMHCpan (Fritsch and colleagues; unpublished data), these data suggest that neoepitopes more commonly result from mutations that alter a preexisting self-epitope than from mutations that generate new MHC-binding peptides.

These data support a reverse immunology approach for engineering patient-specific vaccines but also show that NetMHCpan will miss some neoepitopes capable of driving antitumor responses. In this study, this only accounted for 15% of the neoepitopes, but because most of the neoeptiopes evaluated were identified using cultured T cells, it is possible that neoepitopes of this type were unintentionally selected against. Furthermore, even for shared tumor antigens, for which reverse immunology has been being used to identify T-cell epitopes for nearly 15 years, very few studies have compared epitopes defined by reverse immunology with those defined by an unbiased approach for the same antigen. However, when the two approaches were compared for mesothelin, an antigen expressed by several human cancers, they did not identify the same CD8[+] T-cell epitopes (13, 14). In fact, none of the 10 mesothelin epitopes defined by an unbiased functional approach using patient T cells were predicted to bind HLA class I (13). It is possible that for shared antigens like mesothelin, T-cell repertoires specific for high-affinity HLA-binding epitopes are deleted or tolerized. This would explain why predicted high-affinity epitopes were not identified by the functional screen, and would be consistent with the low response rates associated with vaccines targeting shared tumor antigens. However, it is also possible that prediction algorithms really do miss more than 15% of HLA-binding peptides.

Beyond selecting neoepitopes that are actually presented by HLA, there are at least two additional minimal criteria that patient-specific vaccines must meet to have a chance of inducing an effective antitumor response. First, they must target neoepitopes that are in genes expressed by the tumor; and second, there must be a repertoire of T cells available to respond to them.

Although targeting neoepitopes in genes that are not expressed could indirectly contribute to the development of a postvaccination antitumor response, T cells specific to them would be incapable of mounting a direct cytolytic effect against the tumor. Thus, the ideal vaccine would include at least one target expressed by each and every tumor cell. Otherwise, in the absence of epitope spreading, some tumor cells could escape the vaccine response.

Identifying expressed mutations may be easily solved using RNA sequencing and/or gene expression analyses. However, the issue becomes more complex when the level of heterogeneity observed within tumors is considered (1). As tumors develop, they evolve numerous subclones, and gene expression between these subclones can vary. In addition, although the majority of subclones share a few driver mutations responsible for supporting tumor growth and survival, they have many more passenger mutations (mutations in genes not essential for tumor survival) that are more variable. The uneven ratio of driver to passenger mutations may pose a problem because it will likely force vaccines to target neoepitopes that are not

uniformly expressed across subclones and are not essential for survival. Tumor variation becomes more complex when considering variation between metastases, which derive from only a subset of subclones from the primary tumor. Fortunately, sequencing analysis can provide insight about the frequency of cells expressing each mutation within the tumor sample evaluated, but it is difficult to get this information at the clonal level, and it is not clear how accurately mutation profiles identified in a sampling of a tumor represent the tumor as a whole, as well as the unsampled metastatic lesions.

Ironically, the immune system has been shown to be partially responsible for shaping the clonal composition of tumors through a process termed immunoediting (15, 16). Although immunoediting demonstrates the potential of the immune system to mount an antitumor response, it may pose a problem for tumor mutome-targeted vaccines as immunoediting could eradicate tumor cells expressing the most immunogenic neoepitopes, and therefore best vaccine targets, leaving behind only less immunogenic and less optimal targets to be picked up by sequencing. Furthermore, the outgrowth of immunoedited tumors demonstrates the importance of targeting epitopes covering the entire repertoire of tumor subclones, and not necessarily just dominant neoepitopes.

Even if algorithms can predict neoepitopes most likely to be presented by HLA, what fraction of these will have an available T-cell repertoire capable of responding to them is not known, and will vary from tumor to tumor and between tumor subclones. Unfortunately, the data presented by Fritsch and colleagues cannot provide insight into this. However, there is one preclinical study (2) that may. In that study, 16 of 50 mutations identified in the murine B16F10 melanoma cell line that were predicted to generate MHC-binding T-cell epitopes could be recognized by, and induce T-cell responses, some of which were capable of mounting protective antitumor responses in tumor-bearing mice. This one available study suggests that approximately 33% of predicted neoepitopes will have an available T-cell repertoire. Because it is estimated that 10% of tumor mutations will generate peptides predicted to bind to each expressed HLA class I molecule (5), combining both of these estimates suggests that approximately 3% of mutations generate T-cell targets for any single HLA. Thus, depending on the number of different HLA class I alleles expressed by a patient, 9% to 18% of expressed mutations are estimated to provide valid $CD8^+$ T-cell vaccine targets; at least that can be predicted. Because *in vivo* immunogenicity screening has not been reported for mutations not predicted to generate MHC-binding peptides, as of yet, there are no data for estimating what fraction of these mutations provide valid targets. However, these data will be critical for assessing the full potential of predictive algorithms for developing this approach.

The antitumor responses observed in B16F10-challenged mice (2) provide encouragement for the development of patient-specific vaccines. However, there is at least one potential issue that cannot be easily addressed with transplantable tumor models—peripheral immune tolerance. Human tumors can take years to develop (1). Whether T cells recognizing mutant neoepitopes become tolerized during this time requires better models, or needs to be assessed in humans. Nonetheless, it is well known that most tumors possess immunosuppressive microenvironments (17). Therefore, vaccine-activated mutant neoepitope-specific T cells will almost certainly encounter some form of resistance at the tumor site, and ultimately, if patient-specific vaccines can be developed, they will likely need to be combined with immune modulation to achieve an optimal effect.

With all points considered, choosing neoepitope targets for the development of patient-specific vaccines is not so straightforward. Given that the data are currently limited, at least in early studies, it may be advantageous to design vaccines targeting the broadest repertoire of candidate neoepitopes possible. It may even be a good idea to use long peptides (17), or vectors encoding them, instead of predicted minimal epitopes, as long peptides have the potential to generate every possible neoepitope. This was the vaccination approach tested in the B16F10 model (2), and would provide a cushion in case the predictions are wrong, which was the case for one of the 31 neoepitopes analyzed by Fritsch and colleagues. At least for some tumor types with fewer mutations, it may even be feasible to target all expressed genes, regardless of binding predictions, which may require as few as 10 to 20 targets for several common tumors (1). An unbiased approach like this would enable a more formal evaluation of predictive algorithms and add to the data reported in this study.

Thus, the data provided by Fritsch and colleagues are encouraging. But with few studies available, it may still be too risky to rely only on prediction algorithms. Additional studies in patients are needed to determine the accuracy of these algorithms. Nonetheless, as studies are being designed on this basis, let us hope the conclusion made by Fritsch and colleagues turns out to be the right one.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

References

1. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA Jr, Kinzler KW. Cancer genome landscapes. Science 2013;339:1546–58.
2. Castle JC, Kreiter S, Diekmann J, Löwer M, van de Roemer N, de Graaf J, et al. Exploiting the mutanome for tumor vaccination. Cancer Res 2012;72:1081–91.
3. Hacohen N, Fritsch EF, Carter TA, Lander ES, Wu CJ. Getting personal with neoantigen-based therapeutic cancer vaccines. Cancer Immunol Res 2013;1:11–5.
4. Heemskerk B, Kvistborg P, Schumacher TN. The cancer antigenome. EMBO J 2013;32:194–203.

5. Segal NH, Parsons DW, Peggs KS, Velculescu V, Kinzler KW, Vogelstein B, et al. Epitope landscape in breast and colorectal cancer. Cancer Res 2008;68:889–92.

6. Lundegaard C, Lund O, Nielsen M. Prediction of epitopes using neural network based methods. J Immunol Methods 2011;374:26–34.

7. Parker KC, Bednarek MA, Coligan JE. Scheme for ranking potential HLA-A2 binding peptides based on independent binding of individual peptide side-chains. J Immunol 1994;152:163–75.

8. Rammensee H, Bachmann J, Emmerich NP, Bachor OA, Stevanovic S. SYFPEITHI: database for MHC ligands and peptide motifs. Immunogenetics 1999;50:213–9.

9. Lu J, Celis E. Use of two predictive algorithms of the world wide web for the identification of tumor-reactive T-cell epitopes. Cancer Res 2000;60:5223–7.

10. Robbins PF, Lu YC, El-Gamil M, Li YF, Gross C, Gartner J, et al. Mining exomic sequencing data to identify mutated antigens recognized by adoptively transferred tumor-reactive T cells. Nat Med 2013;19: 747–52.

11. Fritsch EF, Rajasagi M, Ott PA, Brusic V, Hacohen N, Wu CJ, et al. HLA-binding properties of tumor neoepitopes in humans. Cancer Immunol Res 2014;2:522–29.

12. Bowerman NA, Crofts TS, Chlewicki L, Do P, Baker BM, Christopher Garcia K, et al. Engineering the binding properties of the T cell receptor: peptide:MHC ternary complex that governs T cell activity. Mol Immunol 2009;46:3000–8.

13. Lutz E, Yeo CJ, Lillemoe KD, Biedrzycki B, Kobrin B, Herman J, et al. A lethally irradiated allogeneic granulocyte-macrophage colony stimulating factor-secreting tumor vaccine for pancreatic adenocarcinoma. A phase II trial of safety, efficacy, and immune activation. Ann Surg 2011;253:328–35.

14. Thomas AM, Santarsiero LM, Lutz ER, Armstrong TD, Chen YC, Huang LQ, et al. Mesothelin-specific CD8(+) T cell responses provide evidence of in vivo cross-priming by antigen-presenting cells in vaccinated pancreatic cancer patients. J Exp Med 2004; 200:297–306.

15. Matsushita H, Vesely MD, Koboldt DC, Rickert CG, Uppaluri R, Magrini VJ, et al. Cancer exome analysis reveals a T-cell–dependent mechanism of cancer immunoediting. Nature 2012; 482:400–4.

16. Schreiber RD, Old LJ, Smyth MJ. Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion. Science 2011;331:1565–70.

17. Gajewski TF, Meng Y, Blank C, et al. Immune resistance orchestrated by the tumor microenvironment. Immunol Rev 2006;213:131–45.

18. Melief CJ. Synthetic vaccine for the treatment of lesions caused by high risk human papilloma virus. Cancer J 2011;17:300–1.

Cancer Immunology Research

Can We Predict Mutant Neoepitopes in Human Cancers for Patient-Specific Vaccine Therapy?

Eric R. Lutz and Elizabeth M. Jaffee

Updated version	Access the most recent version of this article at: doi:10.1158/2326-6066.CIR-14-0041

Cited articles	This article cites 18 articles, 10 of which you can access for free at: http://cancerimmunolres.aacrjournals.org/content/2/6/518.full#ref-list-1

E-mail alerts	Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions	To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.
Permissions	To request permission to re-use all or part of this article, use this link http://cancerimmunolres.aacrjournals.org/content/2/6/518. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.