Exome Sequencing to Predict Neoantigens in Melanoma
Antonia L. Pritchard¹, Julie G. Burel², Michelle A. Neller³, Nicholas K. Hayward¹, J. Alejandro Lopez⁴,⁵, Martina Fatho⁵, Volker Lennerz⁵, Thomas Wölfl⁵, and Christopher W. Schmidt³

Abstract
The ability to use circulating peripheral blood cells and matched tumor sequencing data as a basis for neoantigen prediction has exciting possibilities for application in the personalized treatment of cancer patients. We have used a high-throughput screening approach, combining whole-exome sequence data, mRNA microarrays, and publicly available epitope prediction algorithm output to identify mutated proteins processed and displayed by patient tumors and recognized by circulating immune cells. Matched autologous melanoma cell lines and peripheral blood mononuclear cells were used to create mixed lymphocyte tumor cell cultures, resulting in an expansion of tumor-reactive T cells to use for mutated peptide screening. Five patients were investigated, three of whom had a durable complete response (CR; 15+ years) in an autologous melanoma-pulsed dendritic cell clinical trial. We identified seven mutated antigens in total that stimulated T-effector memory cells in two of the five patients. While the procedure did not result in clinically applicable neoantigens for all patients, those identified were likely important in tumor clearance, leading to durable CR. The nature of the screening process allows results to be obtained rapidly and is easily applicable to a wide variety of different tumor types.

Introduction
Although significant advances in therapeutic options for malignant melanoma have been made in recent years, when it is metastasized, it still kills >85% of patients within 5 years (1). Common tumor antigen–based immunotherapies can potentially selectively target cancer cells; however, this form of immunotherapy has yet to yield consistent meaningful clinical outcomes (2). It has been hypothesized that patient-specific neoepitopes, derived from protein-altering mutational events, result in stronger immune responses against the expressing tumor cells than common tumor antigens (3). Given the person-specific nature of these neoepitopes, they have been traditionally hard to identify using labor-intensive “forward” immunology approaches. In addition, these methods are unsuitable for timely epitope identification for patients in a clinical setting, in which the median survival time of stage IV melanoma patients is only 6 to 9 months.

Application of a rapid “reverse” methodology has been used to successfully predict immunogenic neoepitopes from tumor exome sequencing data recognized by tumor-infiltrating lymphocytes (TIL; refs. 4, 5). It is not always possible to isolate/obtain TILs from tumors, so we have established a methodology for identification of neoepitopes using circulating peripheral blood mononuclear cells (PBMC). Here, we demonstrate the application of a “reverse” immunology approach using tumor cell lines established from metastases of stage IV melanoma patients (6, 7) and recognized by T cells derived from PBMCs and expanded using mixed lymphocyte tumor culture (MLTC).

Materials and Methods
Tumor-specific mutations were identified using whole-exome sequencing, as previously described (8). Briefly, the matched blood and tumor pairs for samples A02, A06, D05, and D14 were sequenced using Illumina technology, whereas D41 was sequenced using the SOLiD platform. The samples were aligned to the hg19 reference genome, and all variants were called using ANNOVAR (9) and those unique to the tumor samples were identified. The minimum read depth was ≥10 reads, and the minimum mutation call had to be ≥20% of all reads; using Illumina OmniQuad arrays, the false-negative rate was determined to be 0.45%, and through Sanger sequencing of variants, the false-positive rate was determined to be 2%. The cell lines used in this work were created as part of the clinical trial protocols for a dendritic cell (DC) vaccine; they were created from metastatic lesions. The cell lines were tested against germline DNA from the blood to ensure that they were correctly matched, by using STR
profiling of 10 markers (QIMR Berghofer Scientific Services department); cell lines were additionally ensured they matched each other by whole-exome sequencing.

Once the nonsynonymous and frameshift mutations were identified (total n = 1,061; Supplementary Table S1), they were verified by Sanger sequencing. Whether the gene was expressed or not was assessed using mRNA microarray (HT12v4; Illumina) and those that were undetectable were removed from further assessment. For each confirmed mutation in expressed genes, the 11 amino acids either side of the mutation were obtained using a LINUX-based script (available on request) based on the cDNA positional information output from Annotation variants of annotations. Frameshift mutations were manually curated. Identification of 9- and 10-mer peptides predicted to bind to patient-specific HLA-A and -B subtypes was carried out using this 21-mer and five programs employing different algorithms: SYFPEITHI (support vector machine modeling; ref. 10), RANKPEP (matrix-based method; ref. 11), ANN (artificial neural network; ref. 12), SMM (stabilized matrix method; ref. 13), and NetMHCpan 2.8 (an ANN, based on primate, mouse, cattle, and pig training sets; ref. 14); a summary of the prediction programs (10-12) were assessed using this 21-mer and five programs employing different algorithms: SYFPEITHI (support vector machine modeling; ref. 10), RANKPEP (matrix-based method; ref. 11), ANN (artificial neural network; ref. 12), SMM (stabilized matrix method; ref. 13), and NetMHCpan 2.8 (an ANN, based on primate, mouse, cattle, and pig training sets; ref. 14); a summary of the total data is in Supplementary Fig. S1; raw data are in Supplementary Table S2. HLA-C predictions were not employed, due to current limitations in predictive power (15).

MLTC were created using a standard protocol (3), and 70% to 88% of the final cultures were CD8+ (Supplementary Table S3). IFNγ ELISpot was performed using peptides at 2 μg/mL to stimulate MLTC cultured in AIM-V media, 10 IU/mL IL2, and 10% human AB serum (Sigma), plated at 3 × 10⁴ cells/well in 96-well Multiscreen® plates (Millipore; protocol available on request; antibodies used are detailed in Supplementary Table S4). Intracellular cytokine staining (ICS) was performed (antibodies used are detailed in Supplementary Table S4) on an LSR-Fortessa (BD Biosciences) and analyzed using FlowJo (FlowJo LLC). Positive controls were autologous tumor and PHA (ELISpot) and PMA/ionomycin (ICS); unstimulated wells allowed assessment of background MLTC stimulation.

Wild-type and mutant DNMT1 ORF (position 2086 to 2707) was amplified in a nested PCR from the D14 melanoma cell line (forward GCCATGAGCAGGCATGACCC; reverse CCTCAGCTCCAGCAGGCG), and purified PCR products were cloned into pcDNA3.1/V5-his TOPO TA (Life Technologies) and sequenced to identify a wild-type and mutant construct. COS-7 cells (2 × 10⁴) were transiently cotransfected (Lipofectamine 2000; Life Technologies) with HLA-A*02:04-cDNA or HLA-B*07:02 (100 ng) and either mutant or wt-DNMT1 plasmids (300 ng; with a control of HLA-A*02:04-cDNA or HLA-B*07:02 only) on Multiscreen-IP HTS plates (Merck Millipore) coated with anti-IFNγ monoclonal antibody (Mab1-D1K; 10 μg/mL; Mabtech). After 24 hours, 5,000 or 1,000 cells of D14-CTL clone 18B/85 were added and incubated for 20 hours before IFNγ ELISpot assays were performed. Additional controls were unstimulated CTls and CTls stimulated with D14-MEL#7 and D14-LCL (both at 50,000 cells/well).

**Results and Discussion**

One of the major obstacles to using patient-specific neoepitopes in tumor immunotherapy is the timely identification of immunogenic peptide sequences. Recent technologic advances allow the coding mutations in tumors to be relatively easily and rapidly identified and have been used to successfully predict the neoepitopes recognized by TILs (4, 5); unfortunately, it is not always possible to isolate or obtain TILs. We have therefore developed a method to identify immunogenic neoepitopes using peripheral circulating lymphocytes, which does not rely on a strong ex vivo response for reactivity to be detected. The use of MLTC increases the likelihood of detection of rare reactive T cells, due to proliferation of tumor-reactive CD8+ T cells, and peripheral blood is easy to obtain from all patients.

Three of the patients assessed, D05, D14, and D41, experienced long-lasting durable complete response (CR) to DC immunotherapy of greater than 15 years, to date (6, 7). PBMCs from all 5 patients had CD3+CD8+IFNγ+ T-cell expansion in response to coculture with autologous tumor (Supplementary Table S3).

After Sanger sequencing verification of nonsynonymous and frameshift mutations identified by whole-exome sequencing (Supplementary Table S1), followed by mRNA microarray confirmation of gene expression and acquisition of the 11 amino acids on either side of the mutation, the 21-mer amino acid sequences were screened through five MHC-class I epitope prediction programs (10-14). Comparison between the algorithms revealed substantial differences in the number of epitopes predicted by individual mutation; SMM consistently predicted epitopes for a higher number of mutations, whereas SYFPEITHI and RANKPEP predicted epitopes for a fewer number (Supplementary Fig. S1). ANN, SMM, and NetMHCpan have an analogous scoring system allowing comparison between the resultant predicted epitopes. SMM predicted weaker binding affinities both when individual cell lines and the HLA types in common between multiple cell lines (HLA-A*02, HLA-A*03, and HLA-B*27) were assessed (Fig. 1). These data therefore support the recent hypothesis that these algorithms have distinct binding affinities for different HLA subtypes, with HLA-A*02 having a greater proportion of traditionally classified "strong binders" predicted than HLA-A*03 and HLA-B*27 (16). In addition, the recent work by Duan and colleagues (17) also suggested caution in using the traditional binding categories to assess results from the epitope prediction programs. This work demonstrated that mutations in an anchor residue for class-I MHC binding that resulted in a more rigid peptide were able to elicit a CD8+ T-cell response, regardless of the predicted binding affinity Kd score. It is therefore clear that while these epitope prediction algorithms are useful tools, there are important caveats to their use that must be considered, and future improvements are necessary before they can reliably predict the majority of immunogenic peptides present.

Results from the epitope prediction analyses were ranked based on binding scores from each model, with ANN, SMM, or NetMHCpan scores ≤500 nM plus support from at least one other method selected for peptide synthesis (Pepsets; Mimotopes; total dataset: Supplementary Table S2). The individual total number of mutations and complexity of HLA genotypes reflected the number of potential peptides that ranked within the cutoff criteria for testing (Supplementary Table S1).

In total, 642 peptides were tested in triplicate by ELISpot (raw data: Supplementary Fig. S2). Peptides stimulating positive ELISpot response (>30 spots once background was subtracted) were retested by ELISpot and by ICS. Five peptides consistently elicited IFNγ responses from the MLTC, as assessed by both ELISpot and ICS and two responses were...
detectable by ELISpot alone (Table 1). Comparison of the
algorithms predicting epitope binding for these seven peptides
revealed that only NetMHCpan predicted all seven would
bind (Table 1). Responses were particularly strong against
mutated peptides derived from CCT6A (D05) and DMNT1
(D14). CCT6A and WASL neoepitopes were also identi-

fied
via a "forward" cDNA expression and cloning approach
[(3) and Lennerz et al., manuscript in preparation]. In addi-
tion, clear populations of IFN
\(\gamma\)

high cells were observed (Table
1; FACS plots in Supplementary Fig. S3). Given the strength of
response, the DMNT1 neoepitope from patient D14 (IYKAP-

C

ENW) was selected to con-
firm the predicted HLA-A

/C3

24:02

restriction and that internal protein processing resulted in the
predicted neoepitope (Fig. 2). CCT6A (patient D05) was
previously con-

firmed to be restricted through HLA-B

/C3

27:05,
but WASL (patient D14) was found to be restricted through
HLA-B

/C3

51:01 (Lennerz and colleagues, manuscript in prepa-
ration), where it was predicted by the algorithms to bind HLA-
B

/C3

07:02 (e.g., by ANN: B

/C3

51:01 = 1,379 nM vs. B

/C3

07:02 = 66

nM).

Taken together (Table 1; Fig. 1; Supplementary Fig. S1), these
data indicate that although this approach was able to
successfully identify and assign HLA restriction of neoepitopes
stimulating a strong immune response, there is also a serious
potential for false-negative results, depending on the choice of
algorithm and interpretation of results.

Intriguingly, upon further scrutiny of PABPC3 (patient
D14), it was revealed that the dinucleotide mutation (GA/AG)
resulted in a 9-mer peptide sequence that is 100% homologous to
gene family member PABPC1 (whole-exome sequencing align-
ment: Supplementary Fig. S4). While PABPC1 is ubiquitously
expressed, PABPC3 has been reported as being testis-speci-
cific (18); mRNA microarray analysis showed that both genes were
expressed by the D14 melanoma cell line. The reason that toler-
ance to this peptide was not induced is therefore intriguing
and suggests that mutations resulting in endogenous sequences
may still have immunogenic potential.

Mid-expansion MLTC (day 18) were assessed for responses
to tumor/peptides. Despite fewer total CD8

\(^+\) cells, the pro-
portion of IFN
\(\gamma\)

\(+\) cells within this population was higher at
day 18 than at day 34 (Fig. 3), and responses were detectable
by ELISpot and ICS, suggesting MLTC generation could be
expedited for clinical application, while still retaining advan-
tages over the use of

ex vivo

cells. The CD3

\(+\) CD8

\(+\) IFN
\(\gamma\)

\(+\) cells

were phenotyped using memory T-cell markers (CCR7, CD27,
and CD45RA), revealing that the majority were of an effector
memory phenotype (Fig. 3; gating strategy: Supplementary
Fig. S5). The populations of na
c\(\divideontimes/C16\)ve antigen-speci-

fi
c T cells de-
clined between days 18 and 34, suggesting differentiation
occurred during MLTC expansion (Fig. 3).

Intriguingly, no consistently positive results were identi-

fied
for A02, A06, or D41 by ELISpot or ICS, despite MLTC reactivity.
Table 1. Peptides eliciting IFNγ release from MLTC measured by ELISpot and intracellular FACS

<table>
<thead>
<tr>
<th>Patient ID</th>
<th>Antigen</th>
<th>Predicted HLA restriction</th>
<th>Algorithms predicting HLA binding</th>
<th>Positive detection method</th>
<th>ELISpot CD3+CD8-</th>
<th>ELISpot CD3+CD8+</th>
<th>ICS IFNγ high CD3+CD8-</th>
<th>ICS IFNγ high CD8+</th>
</tr>
</thead>
<tbody>
<tr>
<td>D05</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>85%</td>
<td>n/a</td>
<td>0.137%</td>
<td>0.01%</td>
</tr>
<tr>
<td>D05</td>
<td>Unstimulated</td>
<td>—</td>
<td>—</td>
<td>ELISpot and ICS</td>
<td>8.71%</td>
<td>1.16%</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>D05</td>
<td>Tumor</td>
<td>—</td>
<td>—</td>
<td>ELISpot and ICS</td>
<td>33.60%</td>
<td>5.57%</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>D05</td>
<td>CCT6A: LRTKVYAEL</td>
<td>B*27:05</td>
<td>ANN, SMM, NetPan</td>
<td>ELISpot and ICS</td>
<td>0.22%</td>
<td>0.01%</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>D05</td>
<td>TRRAP: LLYQELLPL</td>
<td>A*02:01</td>
<td>All</td>
<td>ELISpot</td>
<td>87%</td>
<td>n/a</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>D14</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>1.75%</td>
<td>0.18%</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>D14</td>
<td>Tumor</td>
<td>—</td>
<td>ELISpot and ICS</td>
<td>17.50%</td>
<td>1.55%</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>D14</td>
<td>DNMT1: IYKAPÇENW</td>
<td>A*24:02</td>
<td>Rankpep, SMM, Netpan</td>
<td>ELISpot and ICS</td>
<td>38.90%</td>
<td>6.36%</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>D14</td>
<td>PABPC3: YYFPSQIAL</td>
<td>A*24:02</td>
<td>SYFP, ANN, SMM, Netpan</td>
<td>ELISpot and ICS</td>
<td>10.60%</td>
<td>2.17%</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>D14</td>
<td>MAGEA10: LYNQMEHLI</td>
<td>A*24:02</td>
<td>Rankpep, ANN, SMM, Netpan</td>
<td>ELISpot and ICS</td>
<td>15.60%</td>
<td>3.95%</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>D14</td>
<td>FMN2: HSVSSAFK</td>
<td>A*03:01</td>
<td>ANN, SMM, Netpan</td>
<td>ELISpot and ICS</td>
<td>29.00%</td>
<td>7.45%</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>D14</td>
<td>WASL: YPPPPPALL</td>
<td>B*07:02</td>
<td>Rankpep, ANN, SMM, Netpan</td>
<td>ELISpot</td>
<td>0.72%</td>
<td>0.15%</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>D41</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>85%</td>
<td>n/a</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>D41</td>
<td>Unstimulated</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.04%</td>
<td>0.00%</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>D41</td>
<td>Tumor</td>
<td>—</td>
<td>ELISpot and ICS</td>
<td>8.85%</td>
<td>1.07%</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>A02</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>80%</td>
<td>n/a</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>A02</td>
<td>Unstimulated</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.47%</td>
<td>0.09%</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>A02</td>
<td>Tumor</td>
<td>—</td>
<td>ELISpot and ICS</td>
<td>18.40%</td>
<td>4.47%</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>A06a</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>20%</td>
<td>n/a</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>A06a</td>
<td>Unstimulated</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>2.49%</td>
<td>0.17%</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>A06a</td>
<td>Tumor</td>
<td>—</td>
<td>ELISpot and ICS</td>
<td>29.60%</td>
<td>9.79%</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

NOTE: The predicted HLA restriction, antigen used to stimulate response, ELISpot well and the percentage of IFNγ-producing CD8+, and CD8+ IFNγhigh cells present in MLTC are shown; the percentages detected in unstimulated samples are shown, and were not subtracted from stimulated values. The percentage of each lineage in the total MLTC population is indicated for each cell line in the shaded gray section.

*ELISpot performed on cells selected on the basis of CD8 by autoMACS (Miltenyi) at day 28 and further cultured with irradiated tumor for 10 days, making a total of 44 days; insufficient cells remained for ICS, so this was performed on day 18 cells alone.
to autologous tumor (Table 1). Notably, D41 had durable CR resulting from DC immunotherapy (7), indicating that their immune response to autologous tumor was clinically effective. It is possible that these patients do not react to any neoepitopes and indeed D41 responds to three HLA-A restricted "traditional" cancer epitopes (data not shown); A02 and A06 are untested. The prediction algorithms or filtering process employed may have removed viable immunogenic neoepitopes, T-cell responses due to other mechanisms of creating neoepitopes (such as unidentified frameshifts, in/dels, or chromosomal translocations/gene fusions), or neoepitopes restricted through HLA-C/MHC-class II might be important in these patients. Having used sequencing data from tumor cell lines, it is possible that a nonendogenous mutational artifact developed in vitro may have been tested in this study or that a chromosomal aberration event may have occurred with culturing, masking the presence of an important endogenous mutation. It is also possible that immunogenic clonal mutations present in the tumor mass, but not grown out in the cell line, may not have been detected.

The patient samples used in this study are from clinical trials conducted over 15 years ago and as a result TILs are not available to perform a comparison with the reactivity identified in circulating CD8+ T cells. Given the recently published methods for identification of neoepitopes on TILs (4, 5), such investigations are now possible and would provide vital...
information on the distribution and diversity of neoepitope-reactive T cells within the body. In addition, combining assessments of the neoepitope landscape in TILs and circulating T-cell neoepitopes in patients with differing clinical responses to immunotherapies is likely to provide illuminating information, as recently suggested for clinical response to CTLA-4 (19).

In conclusion, while we have shown the current analysis procedures do not result in clinically applicable neoepitopes for all patients, for D05 and D14, the identified neoepitopes were potentially important in their durable CR due to DC immuno-therapy (6). This new methodology in cancer epitope prediction therefore has exciting possibilities for application in the personalized treatment of certain patients.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Authors’ Contributions
Conception and design: A.L. Pritchard, N.K. Hayward, M. Fatho, T. Wölfl, C.W. Schmidt
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): A.L. Pritchard, J.G. Burel, M.A. Neller, J.A. Lopez, M. Fatho, C.W. Schmidt
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): A.L. Pritchard, M.A. Neller, M. Fatho, V. Lennerz, T. Wölfl, C.W. Schmidt
Writing, review, and/or revision of the manuscript: A.L. Pritchard, J.G. Burel, M.A. Neller, N.K. Hayward, J.A. Lopez, M. Fatho, V. Lennerz, T. Wölfl, C.W. Schmidt
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): A.L. Pritchard, M. Fatho, V. Lennerz, C.W. Schmidt
Study supervision: A.L. Pritchard, N.K. Hayward, M. Fatho, C.W. Schmidt

Acknowledgments
The authors thank Mitchell Stark, Peter Johansson, and Lutz Krause for computational support. They are also grateful to the patients who donated samples, allowing this research to be carried out.

Grant Support
A.L. Pritchard is supported by Cure Cancer Australia (APP1051048). C.W. Schmidt is supported by the Rio Tinto Ride to Conquer Cancer (QIMR Berghofer Flagship award) and has ongoing support from the Alpha Delta chapter of Phi Beta Sigma. N.K. Hayward is supported by the NHMRC (APP1026112). This work was also funded by Buck Off Melanoma and A Day On The Farm.

Received April 2, 2015; revised May 21, 2015; accepted May 21, 2015; published OnlineFirst June 5, 2015.
References

Exome Sequencing to Predict Neoantigens in Melanoma
Antonia L. Pritchard, Julie G. Burel, Michelle A. Neller, et al.

Updated version
Access the most recent version of this article at:
doi:10.1158/2326-6066.CIR-15-0088

Supplementary Material
Access the most recent supplemental material at:
http://cancerimmunolres.aacrjournals.org/content/suppl/2015/06/05/2326-6066.CIR-15-0088.DC1

Cited articles
This article cites 19 articles, 6 of which you can access for free at:
http://cancerimmunolres.aacrjournals.org/content/3/9/992.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, contact the AACR Publications Department at permissions@aacr.org.