www.sciencemag.org/cgi/content/full/science.aaa1348/DC1

# Supplementary Materials for

## Mutational landscape determines sensitivity to PD-1 blockade in non–small cell lung cancer

Naiyer A. Rizvi, Matthew D. Hellmann, Alexandra Snyder, Pia Kvistborg, Vladimir Makarov, Jonathan J. Havel, William Lee, Jianda Yuan, Phillip Wong, Teresa S. Ho, Martin L. Miller, Natasha Rekhtman, Andre L. Moreira, Fawzia Ibrahim, Cameron Bruggeman, Billel Gasmi, Roberta Zappasodi, Yuka Maeda, Chris Sander, Edward B. Garon, Taha Merghoub, Jedd D. Wolchok, Ton N. Schumacher, Timothy A. Chan

*Corresponding author. E-mail: chant@mskcc.org

Published 12 March 2015 on *Science* Express
DOI: 10.1126/science.aaa1348

**This PDF file includes:**



**Other Supplementary Material for this manuscript includes the following:**
(available at www.sciencemag.org/cgi/content/full/science.aaa1348/DC1)

Tables S3 to S6 as Excel files

## Materials and Methods

URLs

Memorial Sloan-Kettering Cancer Center cBioPortal for Cancer Genomics, http://www.cbioportal.org/; SnpEff, http://snpeff.sourceforge.net/; dbSNP, http://www.ncbi.nlm.nih.gov/projects/SNP/; ESP5400, http://evs.gs.washington.edu/EVS/; 1000 Genomes Project, http://www.1000genomes.org/; Ingenuity, http://www.ingenuity.com/products/ipa; SIFT, http://sift.jcvi.org/; PolyPhen-2 http://genetics.bwh.harvard.edu/pph2/; VarScan Copy Number Tool http://varscan.sourceforge.net/copy-number-calling.html; ATHLATES, http://www.broadinstitute.org/scientific-community/science/projects/viral-genomics/athlates; R package, e1071 http://CRAN.R-project.org/package=e1071.

Patients and clinical characteristics

All patients had stage IV non-small cell lung cancer (NSCLC) and were treated at Memorial Sloan Kettering Cancer Center (n=29) or the University of California at Los Angeles (n=5) on protocol NCT01295827 (Table S3). All patients initiated therapy in 2012-2013 and were treated at 10mg/kg every 2-3 weeks, except for 5 patients treated at 2mg/kg every 3 weeks. The overall response rate and progression-free survival are reported to be similar across dose and schedules (*8*). All patients had consented to Institutional Review Board-approved protocols permitting tissue collection and sequencing. PD-L1 expression was assessed prospectively by immunohistochemistry using a previously validated murine anti-human anti-PD-L1 antibody (clone 22C3, Merck & Co., Inc.). Analysis was performed on pre-treatment tumor tissue submitted as part of the clinical trial and scored by a central pathologist affiliated with the clinical trial. Membranous expression of PD-L1 on tumor cells and infiltrating immune cells were scored semi-quantitatively, as ≥50% membranous staining was considered strong, 1-49% was considered weak, and <1% was considered negative. PD-L1 scoring was available in 30 of 34 patients; four had unknown expression. Smoking status was evaluated using previously completed self-reported smoking questionnaires executed as standard of care at MSKCC or review of medical records at UCLA. Patients eligible for this analysis all received at least two doses of study therapy and were evaluable for response to pembrolizumab, and did not prematurely discontinue therapy due to toxicity or withdrawal of consent.

Tumor samples

All tumor tissue used for sequencing was obtained prior to dosing with pembrolizumab, except for one non-responder in whom post-treatment tissue was used (study ID DM123062). Tumor samples used for whole exome sequencing were either formalin-fixed paraffin-embedded (FFPE, n=26) or flash frozen material (n=8; Table S4). The site of procurement of tissue used and details of the tissue are listed in Table S4. The presence of tumor tissue in the sequenced samples was confirmed by examination of a representative hematoxylin and eosin-stained slide by thoracic pathologists (N.R. or A.M). Peripheral blood was collected and DNA isolated from all patients (Nucleospin Blood L, Machery-Nagel). DNA extraction was performed using the DNEasy kit (Qiagen).

Clinical efficacy analysis

Objective response to pembrolizumab was assessed by investigator-assessed immune-related response criteria (irRC) (*46*) by a study radiologist. Per protocol, CT scans were performed every nine weeks. Partial and complete responses were confirmed by a repeat imaging occurring at least 4 weeks after the initial identification of response; unconfirmed responses were considered stable or progressive disease dependent on results of the second CT scan. Durable clinical benefit (DCB) was defined as stable disease or partial response lasting longer than 6 months (week 27, the time of third protocol-scheduled response assessment). No durable benefit (NDB) was defined as progression of disease $\leq$ 6 months of beginning therapy. Patients who were still ongoing study therapy at the time of the data lock (October 10$^{th}$, 2014) but who had not yet reached 6 months of follow up were considered "Not reached" (NR). These patients were not included in the analysis of DCB/NDB, but were included in assessments of objective response and progression-free survival. For patients with ongoing response to study therapy, progression-free survival was censored at the date of the most recent imaging evaluation. For alive patients, overall survival was censored at the date of last known contact.

Whole-exome capture and sequencing

Whole-exome capture libraries were constructed via the Agilent Sure-Select Human All Exon v2.0, 44Mb baited target with the Broad in-solution hybrid selection process. Enriched exome libraries were sequenced on the HiSeq 2000 platform (Illumina) to generate paired-end reads (2x76bp) to a goal of 150X mean target coverage (Broad Institute, Cambridge, MA) (Fig. S12, Table S4). Targeted resequencing using a custom panel of 376 loci to a target coverage of 500X was performed using Ampliseq (Life Technologies).

HLA typing

HLA typing was performed at the MSKCC HLA typing lab New York Blood Center by high-resolution SeCore HLA sequence-based typing method (HLA-SBT) (Invitrogen). ATHLATES (*47*) was also used for HLA typing.

Exome analysis pipeline

Raw sequencing data were aligned to the hg37 genome build using the Burrows-Wheeler Aligner (BWA) version 0.7.10 (*48*) (Fig. S1). Further indel realignment, base-quality score recalibration and duplicate-read removal were performed using the Genome Analysis Toolkit (GATK) version 3.2.2 (*49*). Mutations were annotated using SnpEffect version 3.5d (build 2014-03-05) (*50*). Somatic Sniper version 1.0.0 (*51*), VarScan version 2.2.3 (*52*), Strelka version 1.0.13 (*53*) and MuTect version 1.4 (*54*) were used to generate single nucleotide variant (SNV) calls using default parameters. VarScan and Strelka were used to generate indel calls.

For variants called by one caller, filters included: coverage depth of 7X or greater and >10% variant nucleotide allelic fraction in tumor DNA, and >97% normal allelic fraction in normal DNA. All variants called by one caller were also manually reviewed using

Integrative Genomics Viewer (IGV) (*55*). For SNVs called by two or more callers, variants were accepted if they passed the same filters: coverage depth of 7X or greater and >10% variant nucleotide allelic fraction in tumor DNA, and >97% normal allelic fraction in normal DNA. However, two additional filtering steps were added. First, for those variants called by two, three, or four callers, manual review using IGV was conducted for variants that had a tumor DNA coverage of <7X, tumor variant allelic fraction of≤10%, or normal allelic fraction of ≤97%. Second, single nucleotide polymorphisms (SNPs) that were rare in the SNP data bases, present in tumor DNA, and had an allelic fraction of zero in normal DNA were also manually reviewed using IGV and included as somatic SNVs. Common SNPs were eliminated by comparison to 1000 Genomes Project, ESP6500 (National Heart, Lung and Blood Institute [NHLBI] GO Exome Sequencing Project) and dbSNP132 (*56-58*).

Called SNVs (Table S5) were evaluated as deleterious if denoted as such by snpEff (high), SIFT (*59*) (score <0.05) or PolyPhen-2 ("D" or "P") (*60*). SIFT and Polyphen2 prediction scores and GERP++ conservation scores were parsed from dbNSFP version 2.2 (*61*). Validation resequencing of detected mutations was determined to be 95%.

Molecular signature analysis

The mutation spectrum in each sample was calculated by analyzing nonsynonymous and synonymous exonic single nucleotide substitutions within their trinucleotide sequence context. That is, for each sample the percentage of each of the six possible single nucleotide changes (C>A:G>T, C>G:G>C, C>T:G>A, T>A:A>T, T>C:A>G, T>G:A>C, with the pyrimidine of the Watson-Crick base pair referenced first) within each of the 16 possible combinations of flanking nucleotides was calculated to generate a 96-feature vector that is used to represent the mutation spectrum for that sample. We utilized a Support Vector Machine (R package e1071) to generate a binary classifier to distinguish transversion low (TL) and transversion high (TH) tumors. Similar to a previously published analysis (*16*), the classifier was trained using tumors from lifelong never-smokers and patients with≥60 pack -years of smoking history as the respective controls. The training set was derived from publicly available exome sequencing and smoking history data from TCGA and previously published results (*62*). This classifier was applied to all sequenced patients in order to classify all samples as belonging to either the TL or TH categories.

*In silico* neoantigen prediction pipeline

All nonsynonymous point mutations identified were translated into strings of 17 amino acids with the mutant amino acid situated centrally using a bioinformatic tool called NAseek (*24*). A sliding window method was used to identify the 9 amino acid substrings within the mutant 17mer that had a predicted MHC Class I binding affinity of≤500nM to one (or more) of the patient-specific HLA alleles. Binding affinity for the mutant and corresponding wild type nonamer were analyzed using NetMHCv3.4 software (*25, 26, 63-65*) (Table S6).

Combinatorial Coding and Multimer Screening

HLA-A restricted candidate neoantigens were synthesized in-house (Netherlands Cancer Institute) and HLA-multimers containing these peptides were produced by micro-scale parallel UV-induced peptide exchange reactions as previously described (*28, 66*). Briefly, peptide-MHC complexes loaded with UV-sensitive peptide were subjected to 366-nm UV light (CAMAG) for one hour at 4$^{o}$C in the presence of candidate neoantigen peptide in a 384-well plate. pMHC multimers were generated using a total of 11 different fluorescent streptavidin (SA) conjugates (Invitrogen). For each pMHC monomer, conjugation was performed with two of these fluorochromes. $NaN_3$ (0.02% w/v) and an excess of D-biotin (26.4 mM, Sigma) were added to block residual binding sites.

For T-cell staining, a combinatorial encoding strategy was employed to be able to analyze for reactivity against up to 47 different peptides in parallel (*29*). PBL samples were thawed, treated with DNAse for 1h and stained with pMHC multimer panels for 15 min at 37C. Subsequently, anti–CD8–AF700 (Invitrogen), anti–CD4-FITC (Invitrogen), anti–CD14-FITC (Invitrogen), anti–CD16-FITC (Invitrogen), anti–CD19-FITC (Invitrogen), and LIVE/DEAD Fixable IR Dead Cell Stain Kit (Invitrogen) were added for an additional 20 min on ice. Data acquisition was performed on an LSR II flow cytometer (Becton Dickinson) with FACSDiva 6 software. Cutoff values for the definition of positive responses were ≥0.005% of total $CD8^+$ cells and ≥10 events.

For immunophenotypical analysis, day 44 PMLs were stained with HERC1 P3278S MHC multimers in two colors (qdot 625 (Invitrogen) and PerCPeFluor710 (ebioscience)) plus anti-CD45RA Ab (Invitrogen), anti-CCR7 Ab (BD Bioscience), anti-HLA-DR Ab (BD Bioscience), and anti-LAG-3 Ab (R&D systems). The immunophenotype of HERC1 P3278S reactive and bulk CD8+ T-cells were analyzed. Data were acquired using an LSR II flow cytometer (Becton Dickson) with FASCDiva 6 software.

Intracellular cytokine staining

HERC1 P3278S mutant and wild type peptides of 9 amino acids in length (Mutant: ASNA**S**SAAK, Wild type: ASNA**P**SAAK) were synthesized (GenScript Piscataway, NJ). 1.5 x $10^6$ patient PBLs were cultured with 1.5 x $10^6$ autologous PBLs pulsed with HERC1 P3278S mutant peptide in RPMI media containing 10% pooled human serum (PHS), 10 mM HEPES, 2 mM L-Glutamine, and 50 μM β-mercaptoethanol supplemented with IL-15 (10 ng/ml) and IL-2 (10 IU/ml), using methods previously described (*67*). Cells were harvested at day 12, stained with 3 μL PE-Cy5-CD107a (BD Pharmingen), and either left unstimulated, or stimulated by the addition of (a) mutant peptide or (b) wild type peptide for 2 hours. Cells were then treated with 1x Brefeldin A and monensin (BioLegend) for 4 hours, and subsequently stained with 1 μL Alexa Fluor 405-CD3 (Invitrogen), 3 μL APC-H7-CD8 (BD Bioscience), and 1 μL ECD-CD4 (Beckman Coulter). Upon subsequent washing and permeabilization, the cells were stained with the following antibodies to intracellular cytokines: 3 μL Alexa Fluor 647-IFN-γ (Biolegend), 3 μL PE-MIP-1β, and 1 μL PE-Cy7-TNF-α (BD Pharmingen). Data were acquired by flow cytometry (using CYAN flow cytometer, Summit software, Dako Cytomation California Inc., Carpinteria, CA). Analysis was done by FlowJo version 10.1, TreeStar, Inc. CD3+ single cell lymphocytes were gated for analysis (SS vs. FS [low, mid], FS vs. Pulse Width [all, low], and CD3 vs. "dump" channel [high, low]).

Statistics

Mann-Whitney test was used to compare mutation burdens and differences in the frequency of nucleotide changes. The log-rank test was used compare Kaplan-Meier survival curves and the Mantel-Haenszel method was used to determine hazard ratios. The proportion of objective responders/non-responders or DCB/NDB were compared using Fisher's exact test. The receiver operator characteristic (ROC) curve was generated by plotting the rate of DCB at various threshold settings of mutation burden. That is, the proportion of all DCB patients with mutation burden above any given cut point (sensitivity) is plotted against the proportion of the NDB patients that would also exceed the same cut point (1 – specificity). The area under the curve and exact 95% confidence intervals are reported. Correlations between nonsynonymous mutation burden and neoantigen burden, neoantigen burden and best overall response, and frequency of neoantigen burden/nonsynonymous mutation and best overall response were calculated using Spearman correlation formula. Statistical analyses were performed using GraphPad Prism v.6 (Graphpad Prism Software, San Diego, CA).

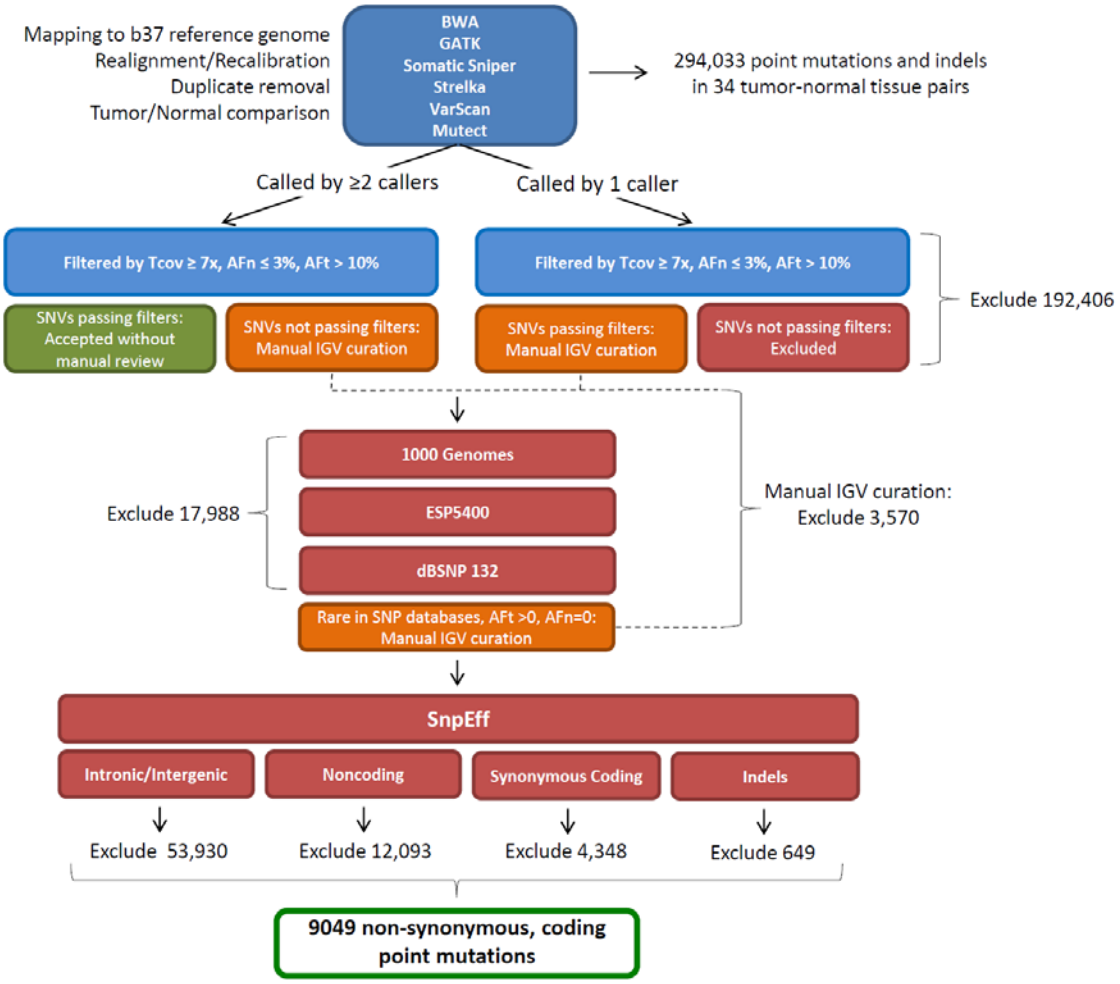



**Fig. S1.**
**Exome analysis pipeline.**

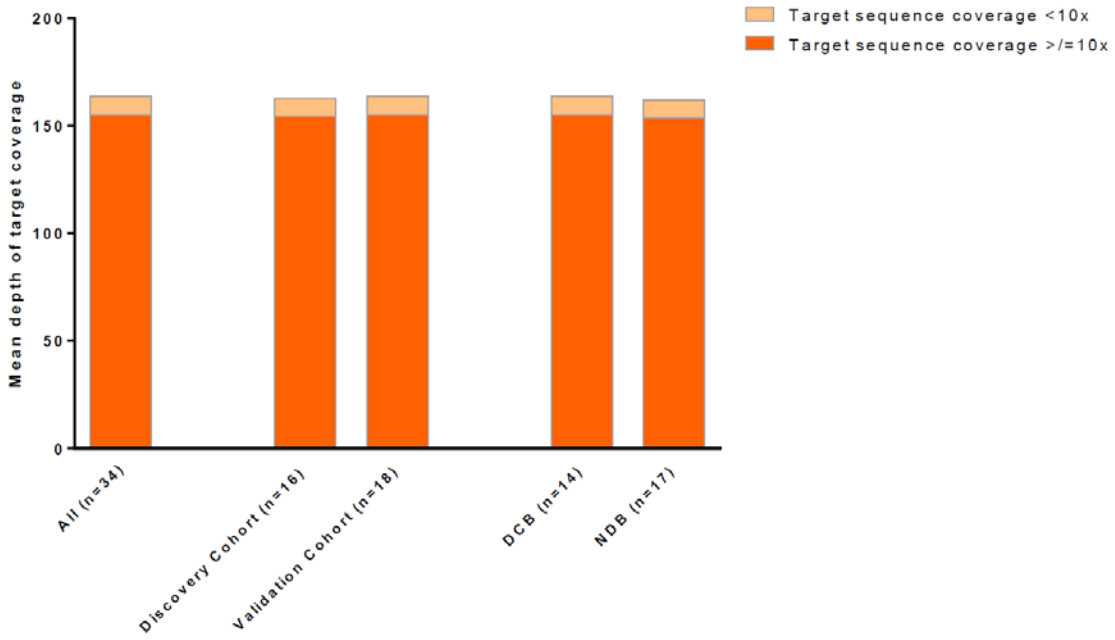
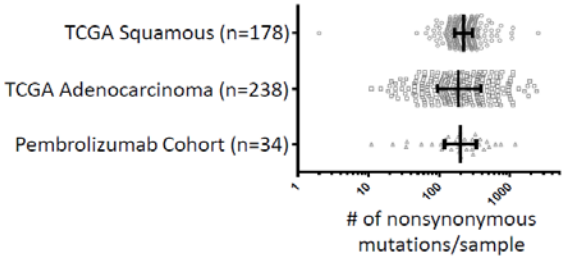



**Fig. S2**

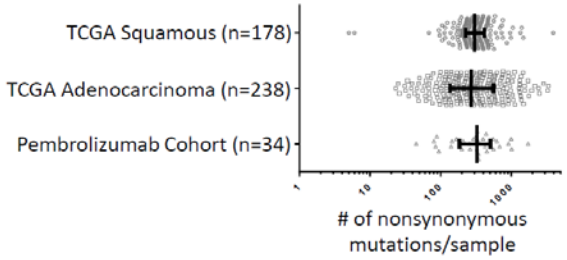
**Coverage and depth of target exome sequence.** Coverage and depth of sequenced exomes in all sequenced tumors, discovery and validation cohorts, and those with durable clinical benefit (DCB) or no durable benefit (NDB).

3A




| | Pembrolizumab Cohort | TCGA Lung Adenocarcinoma | TCGA Lung Squamous Cell Carcinoma |
|---|---|---|---|
| Number of values | 34 | 236 | 178 |
| Minimum | 11 | 11 | 0 |
| 25% Percentile | 118.8 | 93.5 | 163.5 |
| Median | 199.5 | 188 | 222.5 |
| 75% Percentile | 335 | 387.8 | 298.5 |
| Maximum | 1192 | 2402 | 2508 |

3B




| | Pembrolizumab Cohort | TCGA Lung Adenocarcinoma | TCGA Lung Squamous Cell Carcinoma |
|---|---|---|---|
| Number of values | 34 | 236 | 178 |
| Minimum | 94 | 22 | 3 |
| 25% Percentile | 235.3 | 131 | 219.5 |
| Median | 420 | 264.5 | 296 |
| 75% Percentile | 602.3 | 546 | 409.8 |
| Maximum | 2219 | 3306 | 3889 |

**Fig. S3**

**Median and interquartile range of mutations in the current study and in published series of NSCLC** (*15*, *16*). (A) Somatic nonsynonymous mutation burden. (B) Total exonic mutation burden.

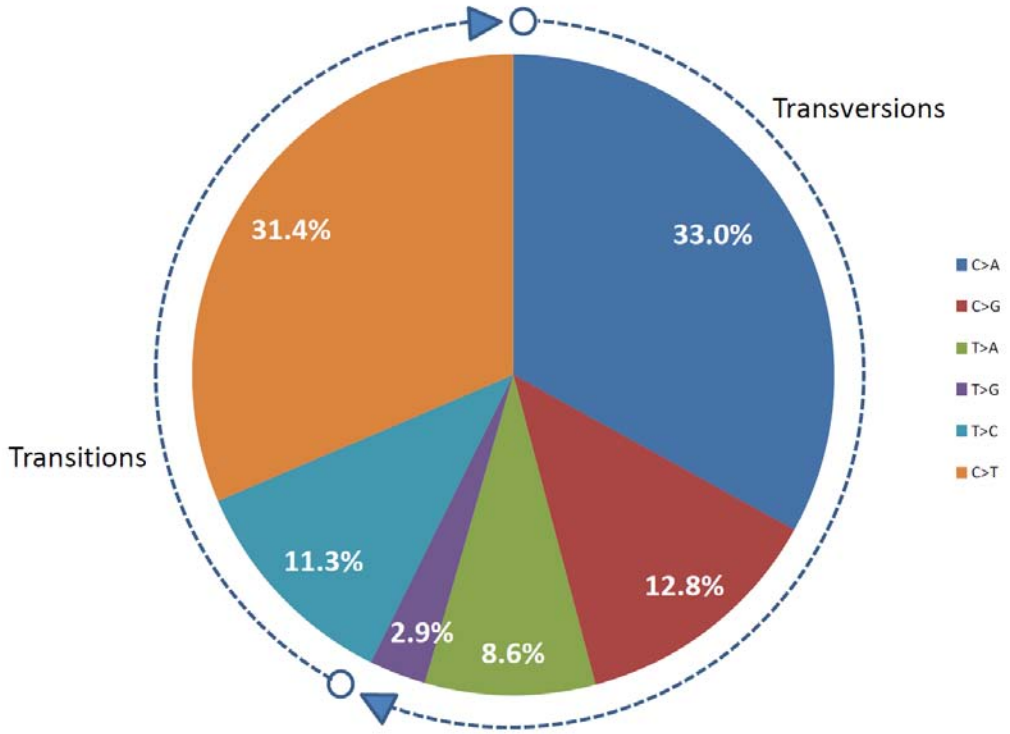



**Fig. S4.**

**Pattern of nucleotide changes in all tumors sequenced.** The spectrum and frequency of nucleotide changes in the pembrolizumab-treated NSCLCs is typical of non-small cell lung cancers.

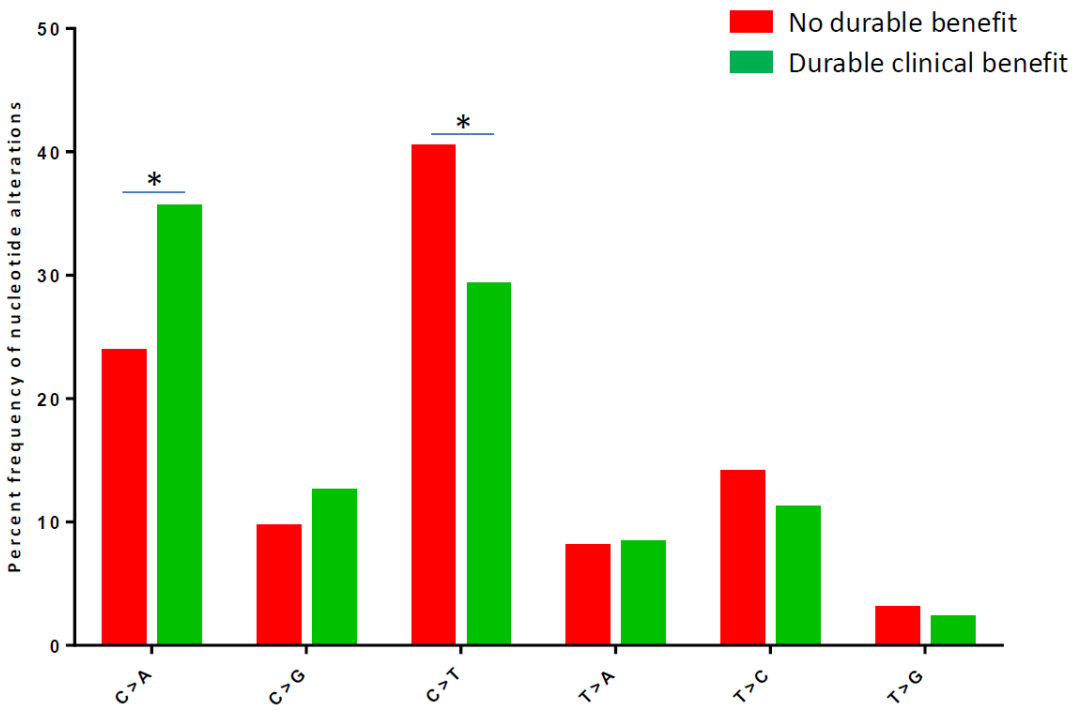



**Fig. S5.**

**Distribution of nucleotide alterations.** Histogram displays the frequency of nucleotide alterations found in nonsynonymous mutations of the tumors of those with durable clinical benefit (DCB) versus no durable benefit (NDB) to pembrolizumab (* denotes Mann-Whitney p=0.01).

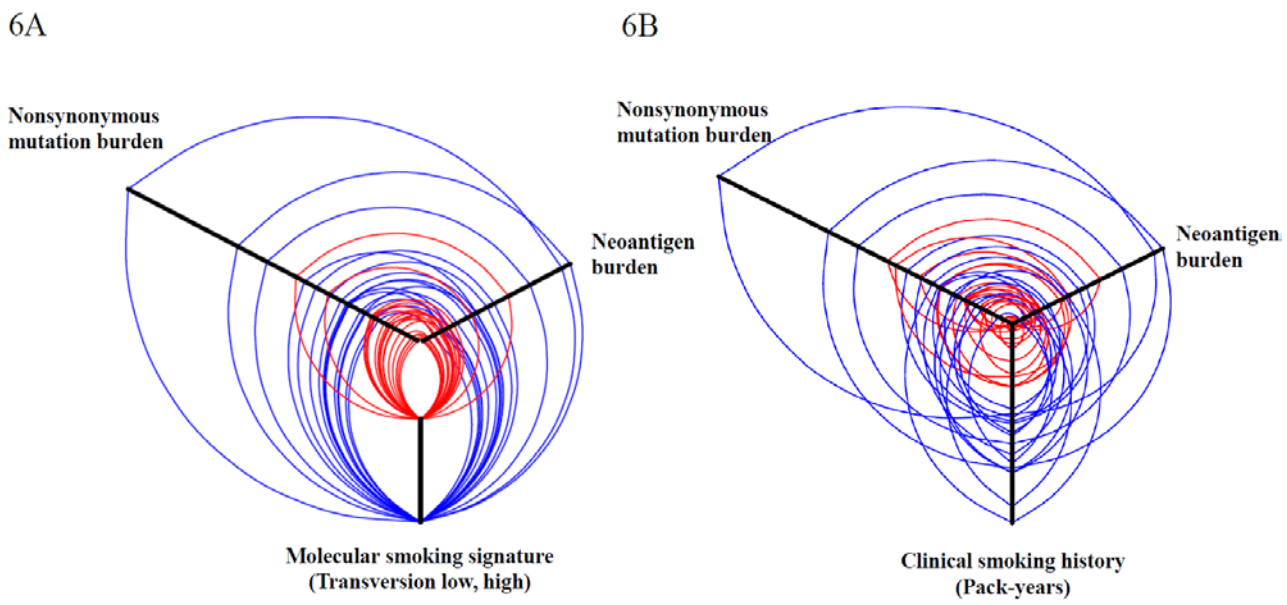



**Fig. S6.**

**Correlation of molecular smoking signature, nonsynonymous mutation burden, and neoantigen burden.** (A) This hive plot displays the relationship between molecular smoking signature, mutation and neoantigen burden for each tumor. Red lines depict transversion low tumors; blue lines depict transversion high tumors. Transversion low tumors have significantly lower mutation and neoantigen burden compared to transversion high tumors (Mann Whitney $p<0.0001$ for both). Nonsynonymous mutation burden correlates with neoantigen burden (Spearman $\rho$ 0.91, 95% CI 0.83-0.96, $p<0.0001$). (B) This hive plot displays the relationship between pack-years of tobacco consumption, mutation and neoantigen burden for each tumor. Red lines depict those who are light/never smokers (≤median pack-years of the cohort, 25); blue lines heavy smokers (>25 pack-years). Modest correlation is seen between pack-years and non-synonymous mutation burden (Spearman $\rho$ 0.31, 95% CI -0.05-0.59, $p=0.08$) as well as between pack-years and neoantigen burden (Spearman $\rho$ 0.35, 95% CI 0-0.62, $p=0.04$).

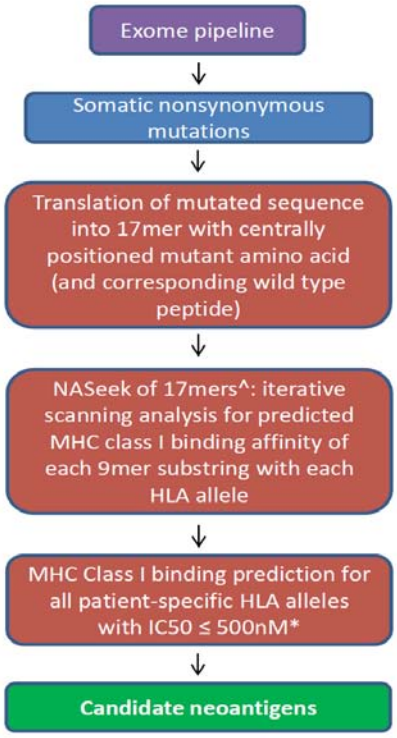



**Fig. S7.**

**Neoantigen analysis pipeline.** ^All steps are executed for predicted wild type and mutant. *MHC Class I prediction by NetMHCv3.4

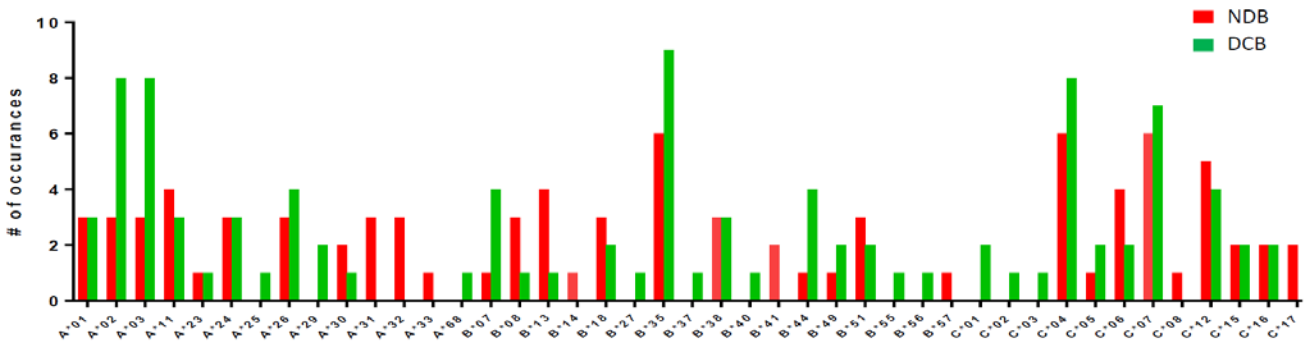



**Fig. S8.**

**HLA type and benefit to pembrolizumab.** Occurrence of durable clinical benefit (DCB) or no durable benefit (NDB) to pembrolizumab by specific HLA allele type.

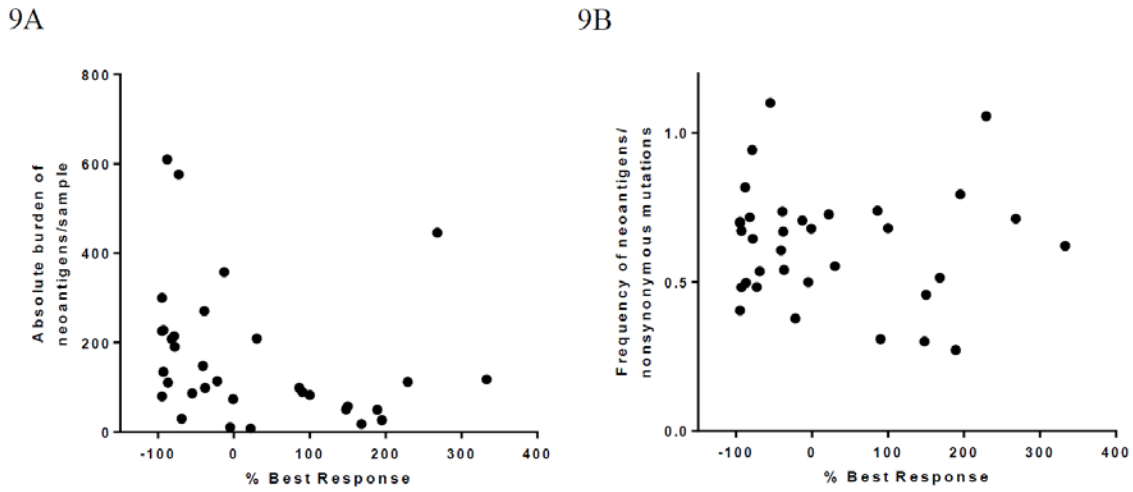



**Fig. S9.**

**Neoantigens and best objective response.** Correlation between best overall radiographic response to pembrolizumab with (A) the absolute quantity of candidate neoantigens (Spearman ρ -0.43, 95% CI -0.68- -0.10, p=0.01) or (B) the frequency of candidate neoantigens/nonsynonymous mutation (Spearman ρ -0.04, 95% CI -0.39-0.30, p=0.78).

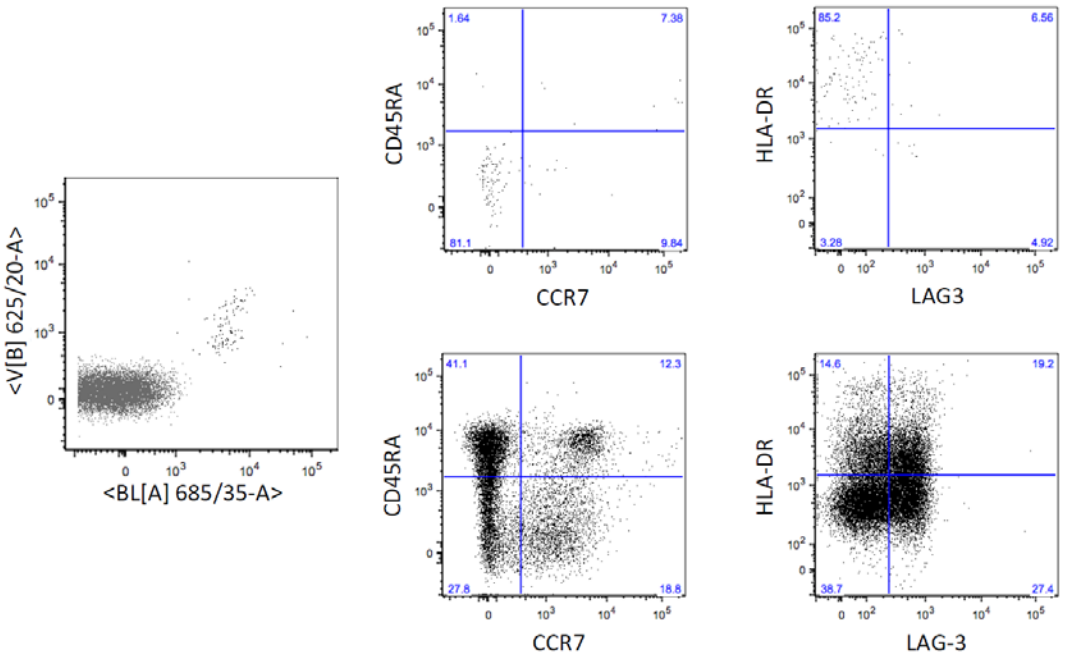



**Fig. S10.**

**Immunophenotype of neoantigen-specific T-cells.** In the left panel, peripheral blood lymphocytes (PBLs) from day 44 were used to identify HERC1 P3278S neoantigen (ASNA**S**SAAK) reactive T-cells using two-color MHC multimer staining, as described. Neoantigen-specific T-cells are represented by the events in the double positive position. Flow cytometry dot plots of staining of HERC1 P3278S neoantigen-specific T-cells (Top panels) and bulk CD8+ T-cells (Bottom panels) show expression of indicated phenotypic markers.

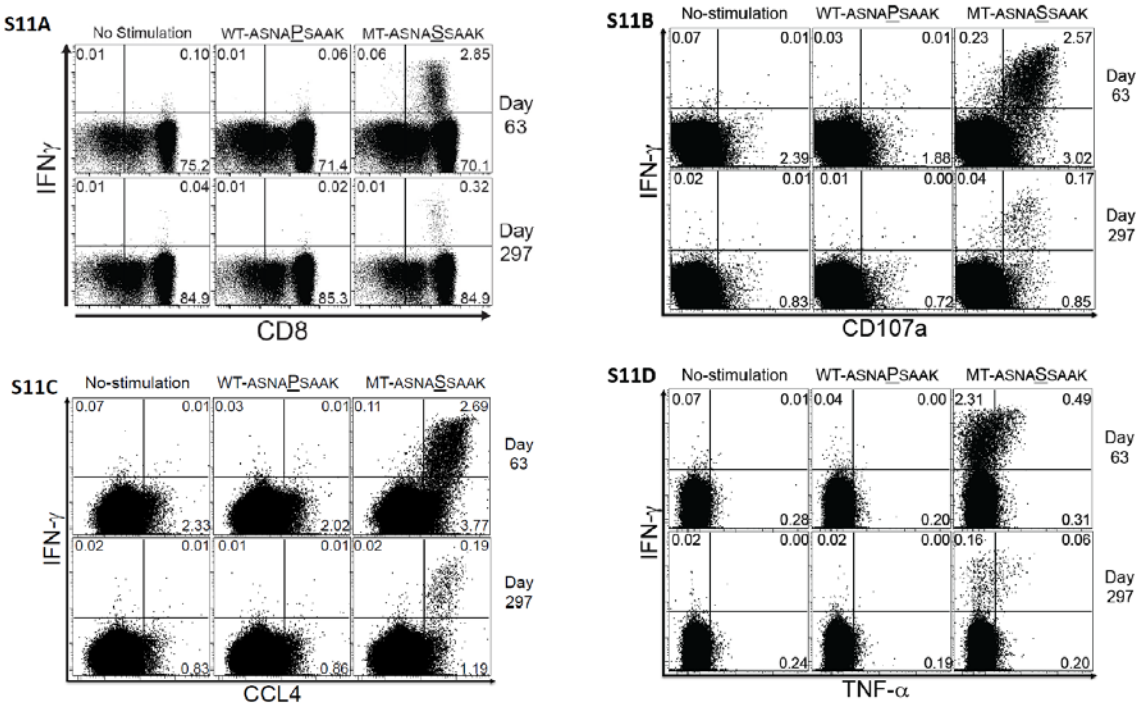



**Fig. S11.**

**PBLs pulsed with mutant peptide for 12 days followed by stimulation with mutant peptide,wild type peptide or no stimulation control show a polyfunctional CD8+ T-cell response to the mutant peptide only. (A)** Neoantigen-induced IFNγ production by CD3+CD8+ T-cells at day 63 and day 297 after initiation of therapy. **(B)** Co-staining of CD3+CD8+ cells for CD107a and IFNγ after no stimulation, stimulation with wild type peptide, or stimulation with mutant peptide. **(C)** Co-staining of CD3+CD8+ cells for CCL4 and IFNγ after no stimulation, simulation with wild type peptide, or stimulation with mutant peptide. **(D)** Co-staining of CD3+CD8+ cells for TNF-α and IFNγ after no stimulation, simulation with wild type peptide, or stimulation with mutant peptide.

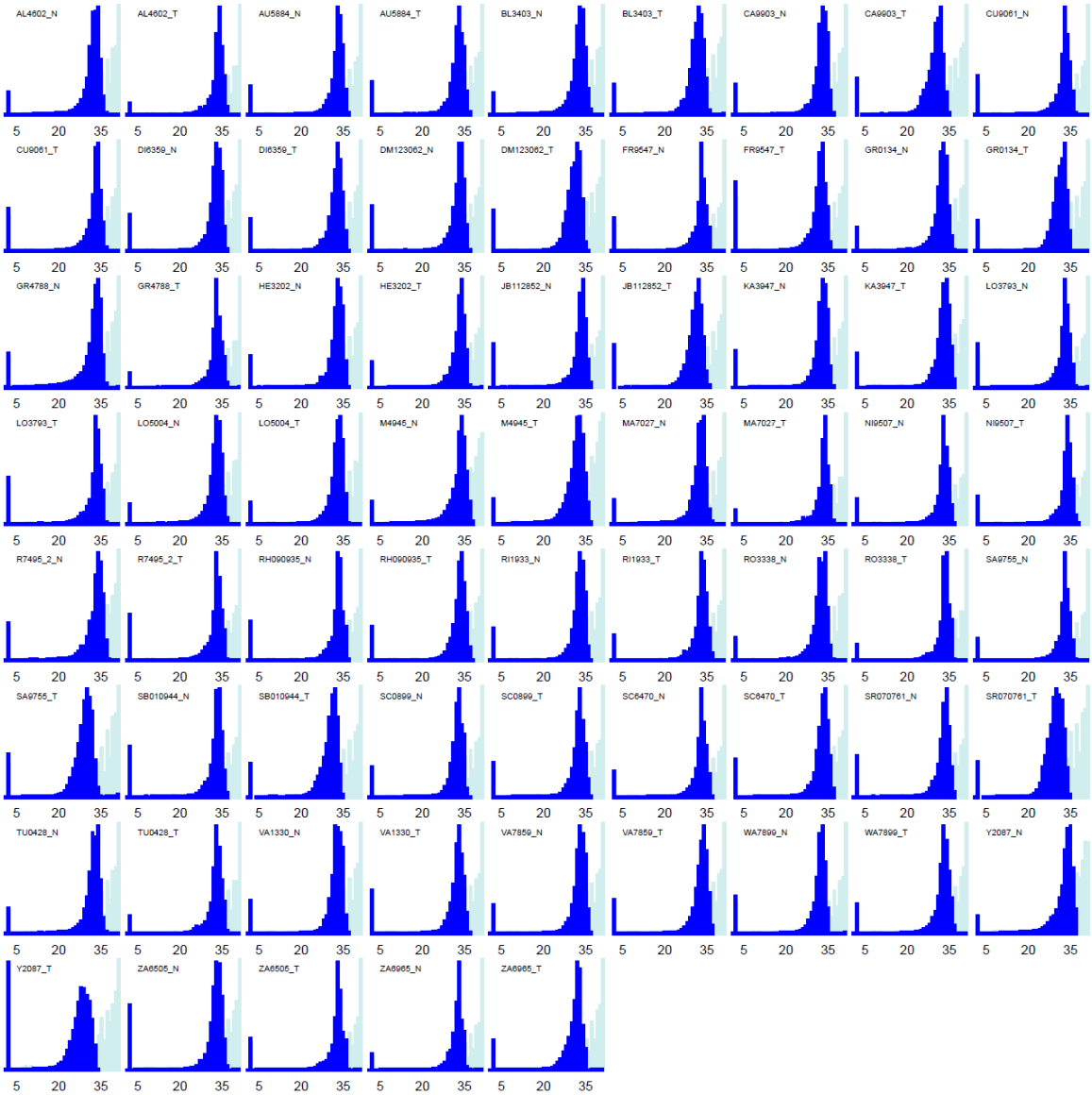



**Fig. S12.**
**DNA quality metrics.**

| | All patients | Discovery Cohort | Validation Cohort |
|---|---|---|---|
| *n* | 34 | 16 | 18 |
| Age (median, range) | 63 years (41-80) | 62 (56-73) | 63 (41-80) |
| Gender (n, %) | | | |
| Male | 16 (47) | 8 (50) | 8 (44) |
| Female | 18 (53) | 8 (50) | 10 (56) |
| Smoking status (n, %) | | | |
| Current | 7 (21) | 5 (31) | 2 (11) |
| Former | 21 (62) | 10 (63) | 11 (61) |
| Never | 6 (17) | 1 (6) | 5 (28) |
| Histology (n, %) | | | |
| Adenocarcinoma | 29 (85) | 15 (94) | 14 (78) |
| Squamous | 4 (12) | 1 (6) | 3 (17) |
| NSCLC NOS | 1 (3) | 0 (0) | 1 (6) |
| Driver oncogenes (n, %) | | | |
| KRAS mutant | 8 (24) | 5 (31) | 3 (17) |
| EGFR mutant | 2 (6) | 1 (6) | 1 (6) |
| ALK re-arranged | 1 (3) | 0 | 1 (6) |
| MK3475 as *n* line of therapy (n, %) | | | |
| 1st | 9 (26) | 5 (31) | 4 (22) |
| 2nd | 11 (33) | 6 (38) | 5 (28) |
| 3rd | 5 (15) | 1 (6) | 4 (22) |
| 4th and beyond | 9 (26) | 4 (25) | 5 (28) |
| PDL1 expression (n, %) | | | |
| Strong (≥50% membranous staining) | 10 (29) | 5 (31) | 5 (28) |
| Weak (1-49%) | 14 (41) | 6 (38) | 8 (44) |
| Negative (<1%) | 6 (18) | 3 (19) | 3 (17) |
| Unknown | 4 (12) | 2 (12) | 2 (11) |
| Dose of pembrolizumab (n, %) | | | |
| 2mg/kg | 5 (15) | 0 (0) | 5 (28) |
| 10mg/kg | 29 (85) | 16 (100) | 13 (72) |
| Confirmed objective response by irRC | | | |
| Partial response | 12 (35) | 5 (31) | 7 (39) |
| Stable disease | 9 (26) | 5 (31) | 5 (28) |
| Progressive disease | 13 (39) | 6 (38) | 6 (33) |
| Durable clinical benefit (PR/SD) > 6 months | | | |
| Yes | 14 (41) | 7 (44) | 7 (39) |
| No | 17 (50) | 9 (56) | 8 (44) |
| Not yet reached 6 month follow up | 3 (9) | 0 (0) | 3 (17) |
| Somatic non-synonymous mutation burden (median, range) | 200 (11-1192) | 207 (11-746) | 200 (35-1192) |
| Total somatic exonic mutation burden (median, range) | 327 (45-1732) | 309 (45-1011) | 322 (85-1732) |

**Table S1.**

**Summary of clinical and genomic characteristics.**

| Mutations quantified | Patient population | *n* | Median # mutations/ sample (range) | ORR (n, %) | Fisher's Exact *P* | DCB (n of x patients evaluable at 6 months, %) | Fisher's Exact *P* | Median PFS (months) | Log-rank *P* | HR for PFS (95% CI) |
|---|---|---|---|---|---|---|---|---|---|---|
| Nonsynonymous mutations | Disc. Cohort: high burden | 8 | 313 (228-746) | 5 (63) | 0.03 | 6 of 8 (75) | 0.04 | 14.5 | 0.01 | 0.19 (0.05-0.70) |
| | Disc. Cohort: low burden | 8 | 127 (11-190) | 0 (0) | | 1 of 8 (13) | | 3.7 | | |
| | Valid. Cohort: high burden | 9 | 368 (201-1192) | 5 (56) | 0.33 | 5 of 6^ (83) | 0.04 | Not reached | 0.006 | 0.15 (0.04-0.59) |
| | Valid. Cohort: low burden | 9 | 122 (35-198) | 2 (22) | | 2 of 9 (22) | | 3.4 | | |
| | | | | | | | | | | |
| Nonsynonymous mutations | All patients: high burden | 17 | 324 (201-1192) | 10 (59) | 0.01 | 11 of 14^ (79) | 0.0011 | Not reached | 0.0004 | 0.19 (0.08-0.47) |
| | All patients: low burden | 17 | 122 (11-198) | 2 (12) | | 3 of 17 (18) | | 3.4 | | |
| | | | | | | | | | | |
| Total exonic mutations | All patients: high burden | 17 | 494 (328-1732) | 8 | 0.28 | 9 of 14^ | 0.08 | 14.5 | 0.045 | 0.40 (0.17-0.98) |
| | All patients: low burden | 17 | 190 (45-326) | 4 | | 5 of 17 | | 4.1 | | |

**Table S2.**

**Nonsynonymous, total exonic mutation burden, and association with clinical efficacy to pembrolizumab.** Analyzed independently, nonsynonymous mutation burden significantly correlates with improved confirmed ORR, DCB, and PFS (with the exception of ORR for the validation cohort**,** p=0.33). Clinical efficacy strongly correlates with nonsynonymous mutation burden in the overall set of sequenced NSCLCs. High total exonic mutation burden correlates less strongly with improved clinical efficacy. ^Denotes that three patients are currently undergoing therapy and have not yet reached 6 months of follow-up; as such, these patients are not included in the DCB/NDB calculations and are removed from the numerator and denominator.

**Table S3.**

**Detailed clinical and genomic characteristics of individual patients.** Table is provided in Other Supplementary Material as an Excel file.

**Table S4.**

**Quality metrics for all samples.** Table is provided in Other Supplementary Material as an Excel file.

**Table S5.**
**Mutation list.** Table is provided in Other Supplementary Material as an Excel file.

**Table S6.**

**Immunogenic mutations, HLA Types, neoantigens, and predicted MHC binding.**
Table is provided in Other Supplementary Material as an Excel file.

## References


1. W. B. Coley, The treatment of malignant tumors by repeated inoculations of erysipelas. With a report of ten original cases. 1893. *Clin. Orthop. Relat. Res.* (262): 3–11 (1991). Medline

2. F. S. Hodi, S. J. O'Day, D. F. McDermott, R. W. Weber, J. A. Sosman, J. B. Haanen, R. Gonzalez, C. Robert, D. Schadendorf, J. C. Hassel, W. Akerley, A. J. van den Eertwegh, J. Lutzky, P. Lorigan, J. M. Vaubel, G. P. Linette, D. Hogg, C. H. Ottensmeier, C. Lebbé, C. Peschel, I. Quirt, J. I. Clark, J. D. Wolchok, J. S. Weber, J. Tian, M. J. Yellin, G. M. Nichol, A. Hoos, W. J. Urba, Improved survival with ipilimumab in patients with metastatic melanoma. *N. Engl. J. Med.* **363**, 711–723 (2010). Medline doi:10.1056/NEJMoa1003466

3. S. L. Topalian, F. S. Hodi, J. R. Brahmer, S. N. Gettinger, D. C. Smith, D. F. McDermott, J. D. Powderly, R. D. Carvajal, J. A. Sosman, M. B. Atkins, P. D. Leming, D. R. Spigel, S. J. Antonia, L. Horn, C. G. Drake, D. M. Pardoll, L. Chen, W. H. Sharfman, R. A. Anders, J. M. Taube, T. L. McMiller, H. Xu, A. J. Korman, M. Jure-Kunkel, S. Agrawal, D. McDonald, G. D. Kollia, A. Gupta, J. M. Wigginton, M. Sznol, Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N. Engl. J. Med.* **366**, 2443–2454 (2012). Medline doi:10.1056/NEJMoa1200690

4. J. D. Wolchok, H. Kluger, M. K. Callahan, M. A. Postow, N. A. Rizvi, A. M. Lesokhin, N. H. Segal, C. E. Ariyan, R. A. Gordon, K. Reed, M. M. Burke, A. Caldwell, S. A. Kronenberg, B. U. Agunwamba, X. Zhang, I. Lowy, H. D. Inzunza, W. Feely, C. E. Horak, Q. Hong, A. J. Korman, J. M. Wigginton, A. Gupta, M. Sznol, Nivolumab plus ipilimumab in advanced melanoma. *N. Engl. J. Med.* **369**, 122–133 (2013). Medline doi:10.1056/NEJMoa1302369

5. C. Robert, A. Ribas, J. D. Wolchok, F. S. Hodi, O. Hamid, R. Kefford, J. S. Weber, A. M. Joshua, W. J. Hwu, T. C. Gangadhar, A. Patnaik, R. Dronca, H. Zarour, R. W. Joseph, P. Boasberg, B. Chmielowski, C. Mateus, M. A. Postow, K. Gergich, J. Elassaiss-Schaap, X. N. Li, R. Iannone, S. W. Ebbinghaus, S. P. Kang, A. Daud, Anti-programmed-death-receptor-1 treatment with pembrolizumab in ipilimumab-refractory advanced melanoma: A randomised dose-comparison cohort of a phase 1 trial. *Lancet* **384**, 1109–1117 (2014). Medline doi:10.1016/S0140-6736(14)60958-2

6. T. Powles, J. P. Eder, G. D. Fine, F. S. Braiteh, Y. Loriot, C. Cruz, J. Bellmunt, H. A. Burris, D. P. Petrylak, S. L. Teng, X. Shen, Z. Boyd, P. S. Hegde, D. S. Chen, N. J. Vogelzang, MPDL3280A (anti-PD-L1) treatment leads to clinical activity in metastatic bladder cancer. *Nature* **515**, 558–562 (2014). Medline doi:10.1038/nature13904

7. S. M. Ansell, A. M. Lesokhin, I. Borrello, A. Halwani, E. C. Scott, M. Gutierrez, S. J. Schuster, M. M. Millenson, D. Cattry, G. J. Freeman, S. J. Rodig, B. Chapuy, A. H. Ligon, L. Zhu, J. F. Grosso, S. Y. Kim, J. M. Timmerman, M. A. Shipp, P. Armand, PD-1 blockade with nivolumab in relapsed or refractory Hodgkin's lymphoma. *N. Engl. J. Med.* **372**, 311–319 (2015). Medline doi:10.1056/NEJMoa1411087

8. E. B. Garon *et al.*, Anti-tumor activity of pembrolizumab (Pembro; MK-3475) and correlation with programmed death ligand 1 (PD-L1) expression in a pooled analysis of patients (pts) with advanced non-small cell lung carcinoma (NSCLC). *Ann. Oncol.* **25**, LBA43 (2014).

9. G. P. Pfeifer, Y. H. You, A. Besaratinia, Mutations induced by ultraviolet light. *Mutat. Res.* **571**, 19–31 (2005). Medline doi:10.1016/j.mrfmmm.2004.06.057

10. G. P. Pfeifer, M. F. Denissenko, M. Olivier, N. Tretyakova, S. S. Hecht, P. Hainaut, Tobacco smoke carcinogens, DNA damage and p53 mutations in smoking-associated cancers. *Oncogene* **21**, 7435–7451 (2002). Medline doi:10.1038/sj.onc.1205803

11. M. S. Lawrence, P. Stojanov, P. Polak, G. V. Kryukov, K. Cibulskis, A. Sivachenko, S. L. Carter, C. Stewart, C. H. Mermel, S. A. Roberts, A. Kiezun, P. S. Hammerman, A. McKenna, Y. Drier, L. Zou, A. H. Ramos, T. J. Pugh, N. Stransky, E. Helman, J. Kim, C. Sougnez, L. Ambrogio, E. Nickerson, E. Shefler, M. L. Cortés, D. Auclair, G. Saksena, D. Voet, M. Noble, D. DiCara, P. Lin, L. Lichtenstein, D. I. Heiman, T. Fennell, M. Imielinski, B. Hernandez, E. Hodis, S. Baca, A. M. Dulak, J. Lohr, D. A. Landau, C. J. Wu, J. Melendez-Zajgla, A. Hidalgo-Miranda, A. Koren, S. A. McCarroll, J. Mora, R. S. Lee, B. Crompton, R. Onofrio, M. Parkin, W. Winckler, K. Ardlie, S. B. Gabriel, C. W. Roberts, J. A. Biegel, K. Stegmaier, A. J. Bass, L. A. Garraway, M. Meyerson, T. R. Golub, D. A. Gordenin, S. Sunyaev, E. S. Lander, G. Getz, Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**, 214–218 (2013). Medline

12. L. B. Alexandrov, S. Nik-Zainal, D. C. Wedge, S. A. Aparicio, S. Behjati, A. V. Biankin, G. R. Bignell, N. Bolli, A. Borg, A. L. Børresen-Dale, S. Boyault, B. Burkhardt, A. P. Butler, C. Caldas, H. R. Davies, C. Desmedt, R. Eils, J. E. Eyfjörd, J. A. Foekens, M. Greaves, F. Hosoda, B. Hutter, T. Ilicic, S. Imbeaud, M. Imielinski, N. Jäger, D. T. Jones, D. Jones, S. Knappskog, M. Kool, S. R. Lakhani, C. López-Otín, S. Martin, N. C. Munshi, H. Nakamura, P. A. Northcott, M. Pajic, E. Papaemmanuil, A. Paradiso, J. V. Pearson, X. S. Puente, K. Raine, M. Ramakrishna, A. L. Richardson, J. Richter, P. Rosenstiel, M. Schlesner, T. N. Schumacher, P. N. Span, J. W. Teague, Y. Totoki, A. N. Tutt, R. Valdés-Mas, M. M. van Buuren, L. van 't Veer, A. Vincent-Salomon, N. Waddell, L. R. Yates, J. Zucman-Rossi, P. A. Futreal, U. McDermott, P. Lichter, M. Meyerson, S. M. Grimmond, R. Siebert, E. Campo, T. Shibata, S. M. Pfister, P. J. Campbell, M. R. Stratton; Australian Pancreatic Cancer Genome Initiative; ICGC Breast Cancer Consortium; ICGC MMML-Seq Consortium; ICGC PedBrain, Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013). Medline

13. B. Vogelstein, N. Papadopoulos, V. E. Velculescu, S. Zhou, L. A. Diaz Jr., K. W. Kinzler, Cancer genome landscapes. *Science* **339**, 1546–1558 (2013). Medline doi:10.1126/science.1235122

14. R. Govindan, L. Ding, M. Griffith, J. Subramanian, N. D. Dees, K. L. Kanchi, C. A. Maher, R. Fulton, L. Fulton, J. Wallis, K. Chen, J. Walker, S. McDonald, R. Bose, D. Ornitz, D. Xiong, M. You, D. J. Dooling, M. Watson, E. R. Mardis, R. K. Wilson, Genomic landscape of non-small cell lung cancer in smokers and never-smokers. *Cell* **150**, 1121–1134 (2012). Medline doi:10.1016/j.cell.2012.08.024

15. See supplementary text available on *Science* Online.

16. P. S. Hammerman *et al.*; Cancer Genome Atlas Research Network, Comprehensive genomic characterization of squamous cell lung cancers. *Nature* **489**, 519–525 (2012). Medline

17. Cancer Genome Atlas Research Network, Comprehensive molecular profiling of lung adenocarcinoma. *Nature* **511**, 543–550 (2014). Medline

18. O. D. Abaan, E. C. Polley, S. R. Davis, Y. J. Zhu, S. Bilke, R. L. Walker, M. Pineda, Y. Gindin, Y. Jiang, W. C. Reinhold, S. L. Holbeck, R. M. Simon, J. H. Doroshow, Y. Pommier, P. S. Meltzer, The exomes of the NCI-60 panel: A genomic resource for cancer biology and

systems pharmacology. *Cancer Res.* **73**, 4372–4382 (2013). Medline doi:10.1158/0008-5472.CAN-12-3342

19. D. Hoffmann, I. Hoffmann, K. El-Bayoumy, The less harmful cigarette: A controversial issue. a tribute to Ernst L. Wynder. *Chem. Res. Toxicol.* **14**, 767–790 (2001). Medline doi:10.1021/tx000260u

20. R. Hindges, U. Hübscher, DNA polymerase delta, an essential enzyme for DNA transactions. *Biol. Chem.* **378**, 345–362 (1997). Medline

21. C. Palles, J. B. Cazier, K. M. Howarth, E. Domingo, A. M. Jones, P. Broderick, Z. Kemp, S. L. Spain, E. Guarino, I. Salguero, A. Sherborne, D. Chubb, L. G. Carvajal-Carmona, Y. Ma, K. Kaur, S. Dobbins, E. Barclay, M. Gorman, L. Martin, M. B. Kovac, S. Humphray, A. Lucassen, C. C. Holmes, D. Bentley, P. Donnelly, J. Taylor, C. Petridis, R. Roylance, E. J. Sawyer, D. J. Kerr, S. Clark, J. Grimes, S. E. Kearsey, H. J. Thomas, G. McVean, R. S. Houlston, I. Tomlinson, CORGI ConsortiumWGS500 Consortium, Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. *Nat. Genet.* **45**, 136–144 (2013). Medline doi:10.1038/ng.2503

22. J. F. Goodwin, K. E. Knudsen, Beyond DNA repair: DNA-PK function in cancer. *Cancer Discov* **4**, 1126–1139 (2014). Medline doi:10.1158/2159-8290.CD-14-0358

23. X. Wang, L. Zou, H. Zheng, Q. Wei, S. J. Elledge, L. Li, Genomic instability and endoreduplication triggered by RAD17 deletion. *Genes Dev.* **17**, 965–970 (2003). Medline doi:10.1101/gad.1065103

24. S. Dogan, R. Shen, D. C. Ang, M. L. Johnson, S. P. D'Angelo, P. K. Paik, E. B. Brzostowski, G. J. Riely, M. G. Kris, M. F. Zakowski, M. Ladanyi, Molecular epidemiology of EGFR and KRAS mutations in 3,026 lung adenocarcinomas: Higher susceptibility of women to smoking-related KRAS-mutant cancers. *Clin. Cancer Res.* **18**, 6169–6177 (2012). Medline doi:10.1158/1078-0432.CCR-11-3265

25. A. Snyder, V. Makarov, T. Merghoub, J. Yuan, J. M. Zaretsky, A. Desrichard, L. A. Walsh, M. A. Postow, P. Wong, T. S. Ho, T. J. Hollmann, C. Bruggeman, K. Kannan, Y. Li, C. Elipenahli, C. Liu, C. T. Harbison, L. Wang, A. Ribas, J. D. Wolchok, T. A. Chan, Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N. Engl. J. Med.* **371**, 2189–2199 (2014). Medline

26. M. Nielsen, C. Lundegaard, P. Worning, S. L. Lauemøller, K. Lamberth, S. Buus, S. Brunak, O. Lund, Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Sci.* **12**, 1007–1017 (2003). Medline doi:10.1110/ps.0239403

27. C. Lundegaard, K. Lamberth, M. Harndahl, S. Buus, O. Lund, M. Nielsen, NetMHC-3.0: Accurate web accessible predictions of human, mouse and monkey MHC class I affinities for peptides of length 8-11. *Nucleic Acids Res.* **36** (Web Server), W509–W512 (2008). Medline doi:10.1093/nar/gkn202

28. M. S. Rooney, S. A. Shukla, C. J. Wu, G. Getz, N. Hacohen, Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* **160**, 48–61 (2015). Medline doi:10.1016/j.cell.2014.12.033

29. B. Rodenko, M. Toebes, S. R. Hadrup, W. J. van Esch, A. M. Molenaar, T. N. Schumacher, H. Ovaa, Generation of peptide-MHC class I complexes through UV-mediated ligand exchange. *Nat. Protoc.* **1**, 1120–1132 (2006). Medline doi:10.1038/nprot.2006.121

30. R. S. Andersen, P. Kvistborg, T. M. Frøsig, N. W. Pedersen, R. Lyngaa, A. H. Bakker, C. J. Shu, P. Straten, T. N. Schumacher, S. R. Hadrup, Parallel detection of antigen-specific T cell responses by combinatorial encoding of MHC multimers. *Nat. Protoc.* **7**, 891–902 (2012). Medline doi:10.1038/nprot.2012.037

31. J. M. Taube, A. Klein, J. R. Brahmer, H. Xu, X. Pan, J. H. Kim, L. Chen, D. M. Pardoll, S. L. Topalian, R. A. Anders, Association of PD-1, PD-1 ligands, and other features of the tumor immune microenvironment with response to anti-PD-1 therapy. *Clin. Cancer Res.* **20**, 5064–5074 (2014). Medline doi:10.1158/1078-0432.CCR-13-3271

32. R. S. Herbst, J. C. Soria, M. Kowanetz, G. D. Fine, O. Hamid, M. S. Gordon, J. A. Sosman, D. F. McDermott, J. D. Powderly, S. N. Gettinger, H. E. Kohrt, L. Horn, D. P. Lawrence, S. Rost, M. Leabman, Y. Xiao, A. Mokatrin, H. Koeppen, P. S. Hegde, I. Mellman, D. S. Chen, F. S. Hodi, Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. *Nature* **515**, 563–567 (2014). Medline doi:10.1038/nature14011

33. P. C. Tumeh, C. L. Harview, J. H. Yearley, I. P. Shintaku, E. J. Taylor, L. Robert, B. Chmielowski, M. Spasic, G. Henry, V. Ciobanu, A. N. West, M. Carmona, C. Kivork, E. Seja, G. Cherry, A. J. Gutierrez, T. R. Grogan, C. Mateus, G. Tomasic, J. A. Glaspy, R. O. Emerson, H. Robins, R. H. Pierce, D. A. Elashoff, C. Robert, A. Ribas, PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* **515**, 568–571 (2014). Medline doi:10.1038/nature13954

34. R. D. Schreiber, L. J. Old, M. J. Smyth, Cancer immunoediting: Integrating immunity's roles in cancer suppression and promotion. *Science* **331**, 1565–1570 (2011). Medline doi:10.1126/science.1203486

35. T. Matsutake, P. K. Srivastava, The immunoprotective MHC II epitope of a chemically induced tumor harbors a unique mutation in a ribosomal protein. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3992–3997 (2001). Medline doi:10.1073/pnas.071523398

36. H. Matsushita, M. D. Vesely, D. C. Koboldt, C. G. Rickert, R. Uppaluri, V. J. Magrini, C. D. Arthur, J. M. White, Y. S. Chen, L. K. Shea, J. Hundal, M. C. Wendl, R. Demeter, T. Wylie, J. P. Allison, M. J. Smyth, L. J. Old, E. R. Mardis, R. D. Schreiber, Cancer exome analysis reveals a T-cell-dependent mechanism of cancer immunoediting. *Nature* **482**, 400–404 (2012). Medline doi:10.1038/nature10755

37. J. C. Castle, S. Kreiter, J. Diekmann, M. Löwer, N. van de Roemer, J. de Graaf, A. Selmi, M. Diken, S. Boegel, C. Paret, M. Koslowski, A. N. Kuhn, C. M. Britten, C. Huber, O. Türeci, U. Sahin, Exploiting the mutanome for tumor vaccination. *Cancer Res.* **72**, 1081–1091 (2012). Medline doi:10.1158/0008-5472.CAN-11-3722

38. T. Schumacher, L. Bunse, S. Pusch, F. Sahm, B. Wiestler, J. Quandt, O. Menn, M. Osswald, I. Oezen, M. Ott, M. Keil, J. Balß, K. Rauschenbach, A. K. Grabowska, I. Vogler, J. Diekmann, N. Trautwein, S. B. Eichmüller, J. Okun, S. Stevanović, A. B. Riemer, U. Sahin, M. A. Friese, P. Beckhove, A. von Deimling, W. Wick, M. Platten, A vaccine targeting mutant IDH1 induces antitumour immunity. *Nature* **512**, 324–327 (2014). Medline doi:10.1038/nature13387

39. M. M. Gubin, X. Zhang, H. Schuster, E. Caron, J. P. Ward, T. Noguchi, Y. Ivanova, J. Hundal, C. D. Arthur, W. J. Krebber, G. E. Mulder, M. Toebes, M. D. Vesely, S. S. Lam, A. J. Korman, J. P. Allison, G. J. Freeman, A. H. Sharpe, E. L. Pearce, T. N. Schumacher, R. Aebersold, H. G. Rammensee, C. J. Melief, E. R. Mardis, W. E. Gillanders, M. N.

Artyomov, R. D. Schreiber, Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens. *Nature* **515**, 577–581 (2014). Medline doi:10.1038/nature13988

40. M. Yadav, S. Jhunjhunwala, Q. T. Phung, P. Lupardus, J. Tanguay, S. Bumbaca, C. Franci, T. K. Cheung, J. Fritsche, T. Weinschenk, Z. Modrusan, I. Mellman, J. R. Lill, L. Delamarre, Predicting immunogenic tumour mutations by combining mass spectrometry and exome sequencing. *Nature* **515**, 572–576 (2014). Medline doi:10.1038/nature14001

41. F. Duan, J. Duitama, S. Al Seesi, C. M. Ayres, S. A. Corcelli, A. P. Pawashe, T. Blanchard, D. McMahon, J. Sidney, A. Sette, B. M. Baker, I. I. Mandoiu, P. K. Srivastava, Genomic and bioinformatic profiling of mutational neoepitopes reveals new rules to predict anticancer immunogenicity. *J. Exp. Med.* **211**, 2231–2248 (2014). Medline doi:10.1084/jem.20141308

42. N. van Rooij, M. M. van Buuren, D. Philips, A. Velds, M. Toebes, B. Heemskerk, L. J. van Dijk, S. Behjati, H. Hilkmann, D. El Atmioui, M. Nieuwland, M. R. Stratton, R. M. Kerkhoven, C. Kesmir, J. B. Haanen, P. Kvistborg, T. N. Schumacher, Tumor exome analysis reveals neoantigen-specific T-cell reactivity in an ipilimumab-responsive melanoma. *J. Clin. Oncol.* **31**, e439–e442 (2013). Medline doi:10.1200/JCO.2012.47.7521

43. P. F. Robbins, Y. C. Lu, M. El-Gamil, Y. F. Li, C. Gross, J. Gartner, J. C. Lin, J. K. Teer, P. Cliften, E. Tycksen, Y. Samuels, S. A. Rosenberg, Mining exomic sequencing data to identify mutated antigens recognized by adoptively transferred tumor-reactive T cells. *Nat. Med.* **19**, 747–752 (2013). Medline doi:10.1038/nm.3161

44. M. Rajasagi, S. A. Shukla, E. F. Fritsch, D. B. Keskin, D. DeLuca, E. Carmona, W. Zhang, C. Sougnez, K. Cibulskis, J. Sidney, K. Stevenson, J. Ritz, D. Neuberg, V. Brusic, S. Gabriel, E. S. Lander, G. Getz, N. Hacohen, C. J. Wu, Systematic identification of personal tumor-specific neoantigens in chronic lymphocytic leukemia. *Blood* **124**, 453–462 (2014). Medline doi:10.1182/blood-2014-04-567933

45. C. Linnemann, M. M. van Buuren, L. Bies, E. M. Verdegaal, R. Schotte, J. J. Calis, S. Behjati, A. Velds, H. Hilkmann, D. E. Atmioui, M. Visser, M. R. Stratton, J. B. Haanen, H. Spits, S. H. van der Burg, T. N. Schumacher, High-throughput epitope discovery reveals frequent recognition of neo-antigens by CD4+ T cells in human melanoma. *Nat. Med.* **21**, 81–85 (2015). Medline doi:10.1038/nm.3773

46. E. Tran, S. Turcotte, A. Gros, P. F. Robbins, Y. C. Lu, M. E. Dudley, J. R. Wunderlich, R. P. Somerville, K. Hogan, C. S. Hinrichs, M. R. Parkhurst, J. C. Yang, S. A. Rosenberg, Cancer immunotherapy based on mutation-specific CD4+ T cells in a patient with epithelial cancer. *Science* **344**, 641–645 (2014). Medline

47. J. D. Wolchok, A. Hoos, S. O'Day, J. S. Weber, O. Hamid, C. Lebbé, M. Maio, M. Binder, O. Bohnsack, G. Nichol, R. Humphrey, F. S. Hodi, Guidelines for the evaluation of immune therapy activity in solid tumors: Immune-related response criteria. *Clin. Cancer Res.* **15**, 7412–7420 (2009). Medline doi:10.1158/1078-0432.CCR-09-1624

48. C. Liu, X. Yang, B. Duffy, T. Mohanakumar, R. D. Mitra, M. C. Zody, J. D. Pfeifer, ATHLATES: Accurate typing of human leukocyte antigen through exome sequencing. *Nucleic Acids Res.* **41**, e142 (2013). Medline doi:10.1093/nar/gkt481

49. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009). Medline doi:10.1093/bioinformatics/btp324

50. M. A. DePristo, E. Banks, R. Poplin, K. V. Garimella, J. R. Maguire, C. Hartl, A. A. Philippakis, G. del Angel, M. A. Rivas, M. Hanna, A. McKenna, T. J. Fennell, A. M. Kernytsky, A. Y. Sivachenko, K. Cibulskis, S. B. Gabriel, D. Altshuler, M. J. Daly, A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011). Medline doi:10.1038/ng.806

51. G. De Baets, J. Van Durme, J. Reumers, S. Maurer-Stroh, P. Vanhee, J. Dopazo, J. Schymkowitz, F. Rousseau, SNPeffect 4.0: On-line prediction of molecular and structural effects of protein-coding variants. *Nucleic Acids Res.* **40** (D1), D935–D939 (2012). Medline doi:10.1093/nar/gkr996

52. D. E. Larson, C. C. Harris, K. Chen, D. C. Koboldt, T. E. Abbott, D. J. Dooling, T. J. Ley, E. R. Mardis, R. K. Wilson, L. Ding, SomaticSniper: Identification of somatic point mutations in whole genome sequencing data. *Bioinformatics* **28**, 311–317 (2012). Medline doi:10.1093/bioinformatics/btr665

53. D. C. Koboldt, Q. Zhang, D. E. Larson, D. Shen, M. D. McLellan, L. Lin, C. A. Miller, E. R. Mardis, L. Ding, R. K. Wilson, VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* **22**, 568–576 (2012). Medline doi:10.1101/gr.129684.111

54. C. T. Saunders, W. S. Wong, S. Swamy, J. Becq, L. J. Murray, R. K. Cheetham, Strelka: Accurate somatic small-variant calling from sequenced tumor-normal sample pairs. *Bioinformatics* **28**, 1811–1817 (2012). Medline doi:10.1093/bioinformatics/bts271

55. K. Cibulskis, M. S. Lawrence, S. L. Carter, A. Sivachenko, D. Jaffe, C. Sougnez, S. Gabriel, M. Meyerson, E. S. Lander, G. Getz, Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* **31**, 213–219 (2013). Medline doi:10.1038/nbt.2514

56. J. T. Robinson, H. Thorvaldsdóttir, W. Winckler, M. Guttman, E. S. Lander, G. Getz, J. P. Mesirov, Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011). Medline doi:10.1038/nbt.1754

57. S. T. Sherry, M. Ward, K. Sirotkin, dbSNP-database for single nucleotide polymorphisms and other classes of minor genetic variation. *Genome Res.* **9**, 677–679 (1999). Medline

58. NHLBI GO Exome Sequencing Project (ESP), http://evs.gs.washington.edu/EVS (Exome Variant Server).

59. M. Via, C. Gignoux, E. G. Burchard, The 1000 Genomes Project: New opportunities for research and social challenges. *Genome Med* **2**, 3 (2010). Medline doi:10.1186/gm124

60. P. Kumar, S. Henikoff, P. C. Ng, Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* **4**, 1073–1081 (2009). Medline doi:10.1038/nprot.2009.86

61. I. A. Adzhubei, S. Schmidt, L. Peshkin, V. E. Ramensky, A. Gerasimova, P. Bork, A. S. Kondrashov, S. R. Sunyaev, A method and server for predicting damaging missense mutations. *Nat. Methods* **7**, 248–249 (2010). Medline doi:10.1038/nmeth0410-248

62. X. Liu, X. Jian, E. Boerwinkle, dbNSFP v2.0: A database of human non-synonymous SNVs and their functional predictions and annotations. *Hum. Mutat.* **34**, E2393–E2402 (2013). Medline doi:10.1002/humu.22376

63. M. Imielinski, A. H. Berger, P. S. Hammerman, B. Hernandez, T. J. Pugh, E. Hodis, J. Cho, J. Suh, M. Capelletti, A. Sivachenko, C. Sougnez, D. Auclair, M. S. Lawrence, P. Stojanov, K. Cibulskis, K. Choi, L. de Waal, T. Sharifnia, A. Brooks, H. Greulich, S. Banerji, T. Zander, D. Seidel, F. Leenders, S. Ansén, C. Ludwig, W. Engel-Riedel, E. Stoelben, J. Wolf, C. Goparju, K. Thompson, W. Winckler, D. Kwiatkowski, B. E. Johnson, P. A. Jänne, V. A. Miller, W. Pao, W. D. Travis, H. I. Pass, S. B. Gabriel, E. S. Lander, R. K. Thomas, L. A. Garraway, G. Getz, M. Meyerson, Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* **150**, 1107–1120 (2012). Medline doi:10.1016/j.cell.2012.08.029

64. M. Nielsen, C. Lundegaard, P. Worning, C. S. Hvid, K. Lamberth, S. Buus, S. Brunak, O. Lund, Improved prediction of MHC class I and class II epitopes using a novel Gibbs sampling approach. *Bioinformatics* **20**, 1388–1397 (2004). Medline doi:10.1093/bioinformatics/bth100

65. M. Nielsen, C. Lundegaard, O. Lund, C. Keşmir, The role of the proteasome in generating cytotoxic T-cell epitopes: Insights obtained from improved predictions of proteasomal cleavage. *Immunogenetics* **57**, 33–41 (2005). Medline doi:10.1007/s00251-005-0781-7

66. C. Lundegaard, O. Lund, M. Nielsen, Accurate approximation method for prediction of class I MHC affinities for peptides of length 8, 10 and 11 using prediction tools trained on 9mers. *Bioinformatics* **24**, 1397–1398 (2008). Medline doi:10.1093/bioinformatics/btn128

67. M. Toebes, M. Coccoris, A. Bins, B. Rodenko, R. Gomez, N. J. Nieuwkoop, W. van de Kasteele, G. F. Rimmelzwaan, J. B. Haanen, H. Ovaa, T. N. Schumacher, Design and use of conditional MHC class I ligands. *Nat. Med.* **12**, 246–251 (2006). Medline

68. J. Yuan, S. Gnjatic, H. Li, S. Powel, H. F. Gallardo, E. Ritter, G. Y. Ku, A. A. Jungbluth, N. H. Segal, T. S. Rasalan, G. Manukian, Y. Xu, R. A. Roman, S. L. Terzulli, M. Heywood, E. Pogoriler, G. Ritter, L. J. Old, J. P. Allison, J. D. Wolchok, CTLA-4 blockade enhances polyfunctional NY-ESO-1 specific T cell responses in metastatic melanoma patients with clinical benefit. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 20410–20415 (2008). Medline doi:10.1073/pnas.0810114105