

BCB 731:

Defense Against the Dark Arts



Optimist: A neoantigen fitness model
predicts tumour response to
checkpoint blockade immunotherapy

October 30th, 2023



A neoantigen fitness model predicts tumour response to checkpoint blockade immunotherapy

Marta Luksza¹, Nadeem Riaz^{2,3}, Vladimir Makarov^{3,4}, Vinod P. Balachandran^{5,6,7}, Matthew D. Hellmann^{7,8,9}, Alexander Solovyov^{10,11,12,13}, Naiyer A. Rizvi¹⁴, Taha Merghoub^{7,15,16}, Arnold J. Levine¹, Timothy A. Chan^{2,3,4,7}, Jedd D. Wolchok^{7,8,15,16} & Benjamin D. Greenbaum^{10,11,12,13}

THE NEW ENGLAND JOURNAL of MEDICINE

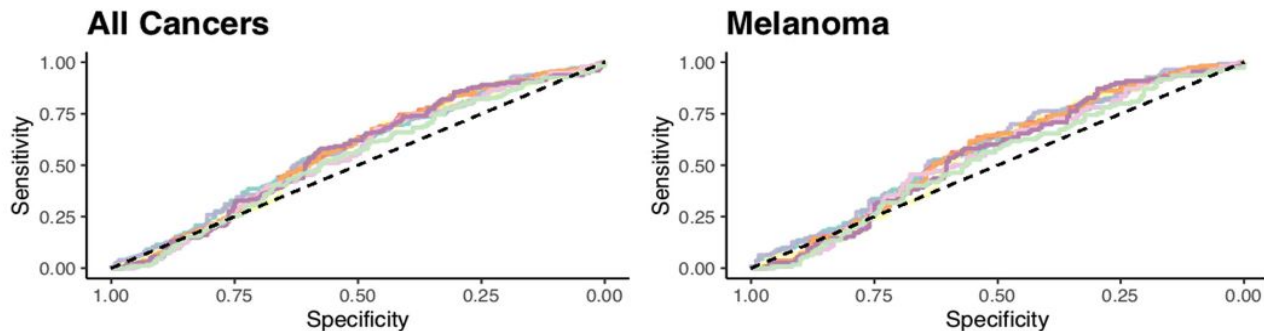
ORIGINAL ARTICLE

Genetic Basis for Clinical Response to CTLA-4 Blockade in Melanoma

Alexandra Snyder, M.D., Vladimir Makarov, M.D., Taha Merghoub, Ph.D., Jianda Yuan, M.D., Ph.D., Jesse M. Zaretsky, B.S., Alexis Desrichard, Ph.D., Logan A. Walsh, Ph.D., Michael A. Postow, M.D., Phillip Wong, Ph.D., Teresa S. Ho, B.S., Travis J. Hollmann, M.D., Ph.D., Cameron Bruggeman, M.A., Kasthuri Kannan, Ph.D., Yanyun Li, M.D., Ph.D., Ceyhan Elipenahli, B.S., Cailian Liu, M.D., Christopher T. Harbison, Ph.D., Lisu Wang, M.D., Antoni Ribas, M.D., Ph.D., Jedd D. Wolchok, M.D., Ph.D., and Timothy A. Chan, M.D., Ph.D.

Background

TMB predicts response to checkpoint blockade (Wood et al. 2020)

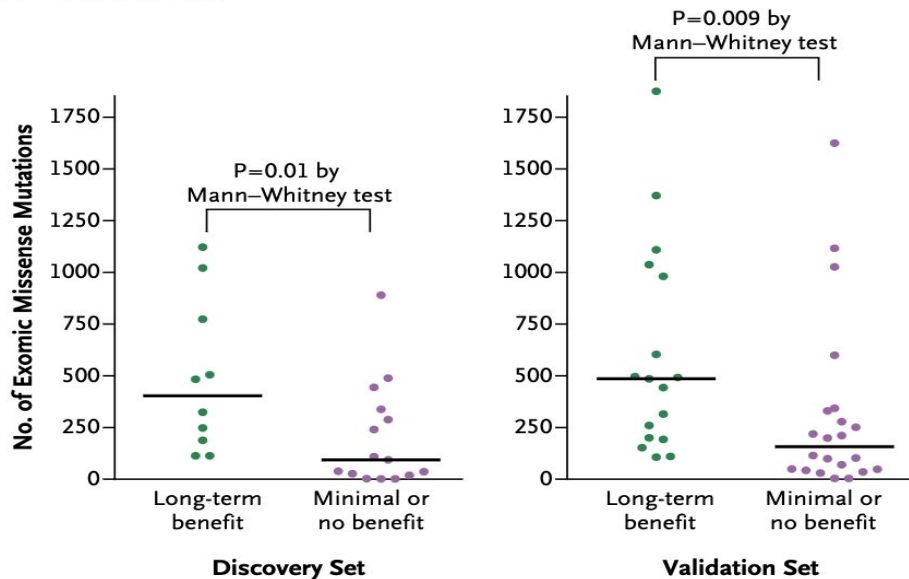


AUC by Burden and Cancer Type

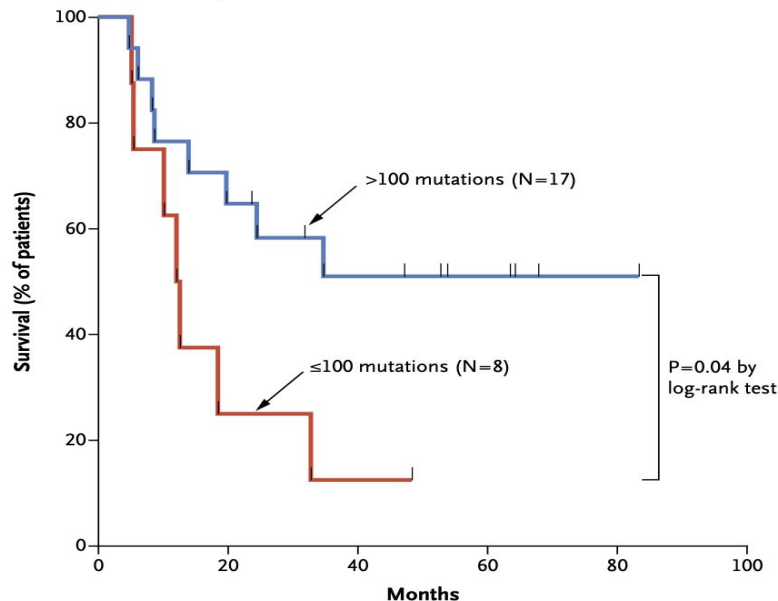
Cancer type	N	TMB1	TMB2	TMB1*HLA	NB	M	A	T	M*A	M*T	A*T	M*A*T
All	431	0.572	0.555	0.571	0.568	0.547	0.567	0.563	0.547	0.534	0.563	0.534
Melanoma	302	0.582	0.560	0.587	0.572	0.550	0.572	0.558	0.550	0.529	0.558	0.529
RCC	57	0.477	0.533	0.509	0.549	0.563	0.556	0.552	0.563	0.459	0.552	0.459
NSCLC	34	0.726	0.722	0.736	0.760	0.740	0.760	0.677	0.740	0.694	0.677	0.694

Snyder et al. (2014) shows TMB predicts outlier response to aCTLA-4

A Mutational Load



B Survival in Discovery Set



Data: Snyder et al. 2014

Table 1.

Cohort summary

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Genetic Basis for Clinical Response to CTLA-4 Blockade in Melanoma

Alexandra Snyder, M.D., Vladimir Makarov, M.D., Taha Merghoub, Ph.D., Jianda Yuan, M.D., Ph.D., Jesse M. Zaretsky, B.S., Alexis Desrichard, Ph.D., Logan A. Walsh, Ph.D., Michael A. Postow, M.D., Phillip Wong, Ph.D., Teresa S. Ho, B.S., Travis J. Hollmann, M.D., Ph.D., Cameron Bruggeman, M.A., Kasthuri Kannan, Ph.D., Yanyun Li, M.D., Ph.D., Ceyhan Elipenahli, B.S., Cailian Liu, M.D., Christopher T. Harbison, Ph.D., Lisu Wang, M.D., Antoni Ribas, M.D., Ph.D., Jedd D. Wolchok, M.D., Ph.D., and Timothy A. Chan, M.D., Ph.D.

Patient samples

All analyzed samples were collected in accordance with local Internal Review Board policies as described in ref. 8 and summarized in Table 1. Thirty-four patients had tumor samples collected prior to initiating CTLA-4 blockade, and 30 patients had samples collected after initiating CTLA-4 blockade. Clinical benefit was defined as progression-free survival lasting for greater than 24 weeks after initiation of therapy (Online Data File 1). Nine **discordant** lesions were present, where overall patient benefit did not match individual tumor progression. See Table 1 for details about this patient cohort.

Group	Benefit	No benefit	Discordant
<i>N</i>	27	28	9
% Cutaneous	20/27	19/28	5/9
OS	3.7 (1.6–7.3)	0.8 (0.2–2.7)	4 (1.7–7.9)
Age	65 (33–81)	58.5 (18–79)	68 (40–90)
Mutations	611 (165–3,394)	321 (6–1,816)	549 (93–1,336)
Neoantigens	1,388 (209–6,502)	714.5 (3–4,510)	1,048 (197–2,584)

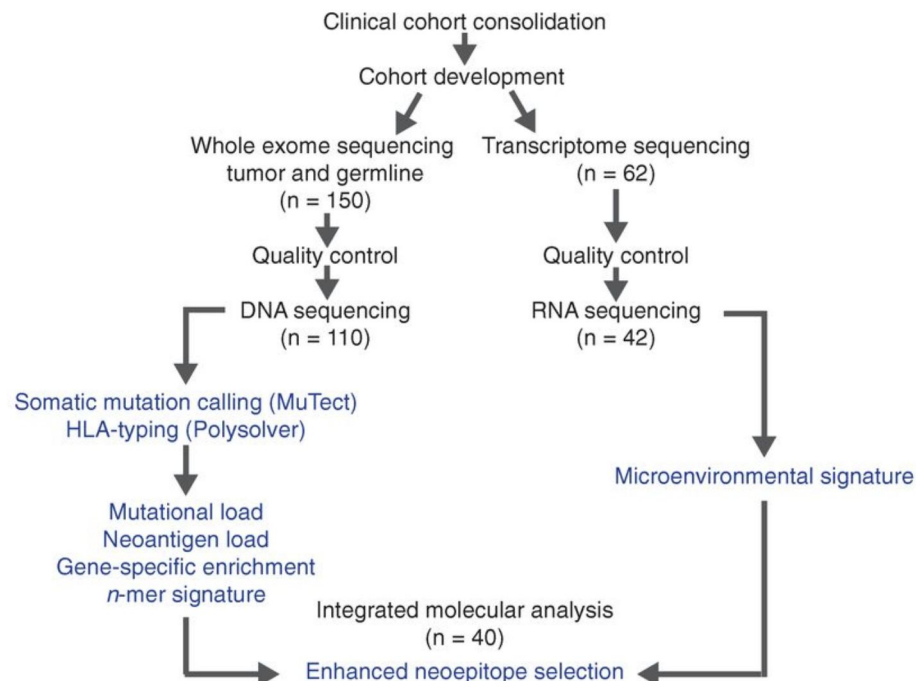
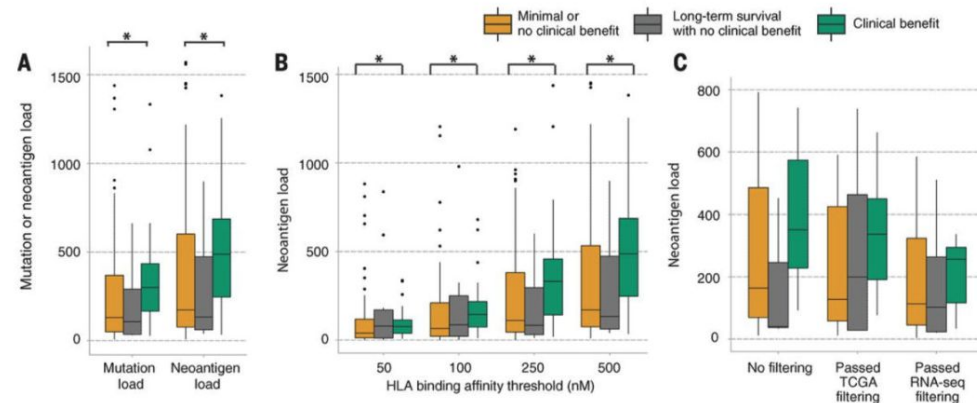
NOTE: Features of tumors from patients with clinical benefit, no benefit, or in which a discordant lesion was resected.

Abbreviation: OS, overall survival.

Data: Van Allen et al. (2015)

Genomic correlates of response to CTLA-4 blockade in metastatic melanoma

Eliezer M. Van Allen,^{1,2,3*} Diana Miao,^{1,2*} Bastian Schilling,^{4,5*} Sachet A. Shukla,^{1,2} Christian Blank,⁶ Lisa Zimmer,^{4,5} Antje Sucker,^{4,5} Uwe Hillen,^{4,5} Marnix H. Geukes Foppen,⁶ Simone M. Goldinger,⁷ Jochen Utikal,^{5,8,9} Jessica C. Hassel,¹⁰ Benjamin Weide,¹¹ Katharina C. Kaehler,¹² Carmen Loquai,¹³ Peter Mohr,¹⁴ Ralf Gutzmer,¹⁵ Reinhard Dummer,⁷ Stacey Gabriel,² Catherine J. Wu,^{1,2} Dirk Schadendorf,^{4,5,†} Levi A. Garraway^{1,2,3,†}



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

```

graph TD
    A[Mapping to b37 reference genome  
Realignment/Recalibration  
Duplicate removal  
Tumor/Normal comparison] --> B[BWA  
GATK  
Somatic Sniper  
Strelka  
VarScan  
Mutect]
    B --> C[294,033 point mutations and indels  
in 34 tumor-normal tissue pairs]
    B -- "Called by ≥2 callers" --> D[Filtered by Tcov ≥ 7x, AFn ≤ 3%, Afi > 10%]
    B -- "Called by 1 caller" --> E[Filtered by Tcov ≥ 7x, AFn ≤ 3%, Afi > 10%]
    D --> F[SNVs passing filters:  
Accepted without manual review]
    D --> G[SNVs not passing filters:  
Manual IGV curation]
    E --> H[SNVs passing filters:  
Manual IGV curation]
    E --> I[SNVs not passing filters:  
Excluded]
    I --- J[Exclude 192,406]
    G --> K[1000 Genomes]
    G --> L[ESP5400]
    G --> M[dBSNP 132]
    G --> N[Rare in SNP databases, Afi > 0, AFn=0:  
Manual IGV curation]
    K --- O[Exclude 17,988]
    L --- O
    M --- O
    N --- P[Manual IGV curation:  
Exclude 3,570]
    O --> Q[SnpEff]
    N --> Q
    Q --> R[Intronic/Intergenic]
    Q --> S[Noncoding]
    Q --> T[Synonymous Coding]
    Q --> U[Indels]
    R --> V[Exclude 53,930]
    S --> W[Exclude 12,093]
    T --> X[Exclude 4,348]
    U --> Y[Exclude 649]
    V --- Z[9049 non-synonymous, coding  
point mutations]
    W --- Z
    X --- Z
    Y --- Z

```

The flowchart illustrates the variant calling pipeline for identifying non-synonymous coding point mutations. It begins with mapping to the b37 reference genome, followed by realignment/recalibration, duplicate removal, and tumor/normal comparison using tools like BWA, GATK, Somatic Sniper, Strelka, VarScan, and Mutect. This results in 294,033 point mutations and indels across 34 tumor-normal tissue pairs.

The variants are then filtered based on whether they were called by at least two callers or one caller. Both groups are filtered by Tcov $\geq 7\times$, AFn $\leq 3\%$, and Afi $> 10\%$.

For variants called by at least two callers, SNVs passing filters are accepted without manual review, while those not passing filters undergo manual IGV curation. For variants called by one caller, SNVs passing filters undergo manual IGV curation, while those not passing filters are excluded (192,406 total exclusions).

From the manual IGV curation step, variants are further evaluated against 1000 Genomes, ESP5400, dBSNP 132, and rare in SNP databases (Afi > 0 , AFn = 0). Variants failing these criteria are excluded (17,988 total exclusions), and those requiring manual IGV curation are also excluded (3,570 total exclusions).

The remaining variants are processed by SnpEff, which categorizes them as Intronic/Intergenic, Noncoding, Synonymous Coding, or Indels. These categories are then excluded: Intronic/Intergenic (53,930), Noncoding (12,093), Synonymous Coding (4,348), and Indels (649).

The final result is 9049 non-synonymous, coding point mutations.



$$\frac{\hat{N}_e(t+1)}{\hat{N}(t+1)} = \frac{1}{\hat{N}(t+1)} \sum_{i,t} N_i(t) \exp[f_i - \epsilon \mathcal{C}(\mathbf{a}_i, \mathbf{a}_v)]$$

$$H(\mathbf{Y}_{t+1} | \hat{\mathbf{Y}}_{t+1}) = \sum_{v \in K(t)} Y_v(t+1) \log[Y_v(t+1) / \hat{Y}_v(t+1)] \quad \frac{p}{p_0} \sim \exp\left[\sum_t \Delta \mathcal{H}(t)\right] \quad \Delta \mathcal{H}(t) = \tilde{m}(t) [-H(\mathbf{Y}_{t+1} | \hat{\mathbf{Y}}_{t+1}) + H(\mathbf{Y}_{t+1} | \mathbf{Y}_t)]$$

Physicists predict pathogen fitness

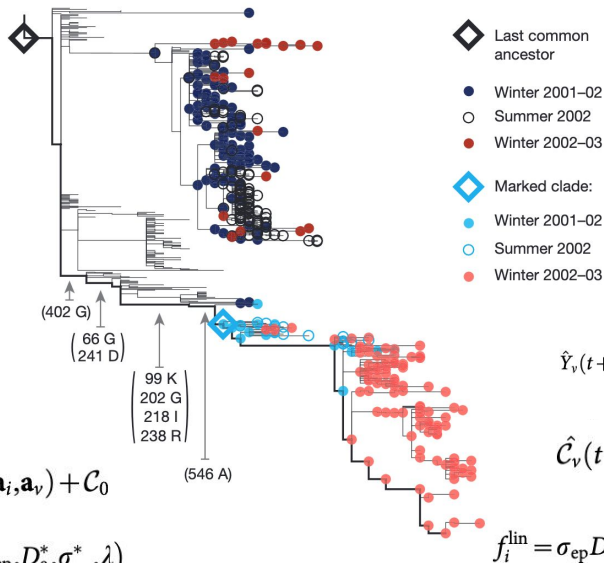
ARTICLE

doi:10.1038/nature13087

A predictive fitness model for influenza

Marta Luksza^{1,2} & Michael Lässig¹

The seasonal human influenza A/H3N2 virus undergoes rapid evolution, which produces significant year-to-year sequence turnover in the population of circulating strains. Adaptive mutations respond to human immune challenge and occur primarily in antigenic epitopes, the antibody-binding domains of the viral surface protein haemagglutinin. Here we develop a fitness model for haemagglutinin that predicts the evolution of the viral population from one year to the next. Two factors are shown to determine the fitness of a strain: adaptive epitope changes and deleterious mutations outside the epitopes. We infer both fitness components for the strains circulating in a given year, using population-genetic data of all previous strains. From fitness and frequency of each strain, we predict the frequency of its descendent strains in the following year. This fitness model maps the adaptive history of influenza A and suggests a principled method for vaccine selection. Our results call for a more comprehensive epidemiology of influenza and other fast-evolving pathogens that integrates antigenic phenotypes with other viral functions coupled by genetic linkage.



$$x_i = \frac{1}{z_\epsilon(t)} \left[\frac{m_i}{m(t)} + \frac{\epsilon}{m_{v(t,\kappa)}(t)} \right]$$

$$\hat{Y}_v(t+1) = \sum_{i \in \mathcal{S}(t)} \tilde{\rho}_{i,v} x_i \exp(f_i) \quad \text{for } v \in K(t)$$

$$\hat{\mathcal{C}}_v(t+1) = \sum_{i:t} x_i \exp(f_i) \mathcal{C}(\mathbf{a}_i, \mathbf{a}_v)$$

$$f_i^{\text{lin}} = \sigma_{\text{ep}} D_{\text{ep}}(\mathbf{a}_i, \mathbf{a}^*(t)) - \sigma_{\text{ne}} D_{\text{ne}}(\mathbf{a}_i, \mathbf{a}^*(t))$$

$$F_v(t) = \log \frac{\hat{Y}_v(t+1)}{Y_v(t)} \quad \tilde{\mathcal{S}}_v^{(\kappa)}(t) = \tilde{\mathcal{S}}_v^{(\kappa)} \cap \mathcal{S}(t) \quad \text{for } v \in K(t) \quad v = \frac{1}{\Delta t} \sum_t \sum_{v \in K(t)} Y_v(t) [F_v(t) - \bar{F}(t)]^2$$

$$\Delta \mathcal{H}(t) = \tilde{m}(t) \sum_{v \in K(t)} Y_v(t+1) F_v(t)$$

$$\Delta \mathcal{H}_{\text{tot}} = \sum_t \Delta \mathcal{H}(t; \sigma_{\text{ep}}^*(t), D_{00}^*, \sigma_{\text{ne}}^*, \lambda^*(t))$$

$$\hat{X}_v(t+1, \epsilon) = \sum_{i:v,t} x_i \exp[f_i(\epsilon=0) - \epsilon \mathcal{C}(\mathbf{a}_i, \mathbf{a}_v) + \mathcal{C}_0]$$

$$\bar{\Phi}(\tilde{\mathbf{Y}}, t+1) = \sum_{v \in K(t)} \Phi_v(t) \tilde{Y}_v$$

$$(\sigma_{\text{ep}}^*(t), \lambda^*(t)) = \arg \max_{\sigma_{\text{ep}}, \lambda} \sum_{t'=t-8}^t \Delta \mathcal{H}(t'; \sigma_{\text{ep}}, D_{00}^*, \sigma_{\text{ne}}^*, \lambda)$$

$$\text{Var } \Phi(t) = \sum_{v \in K(t)} Y_v(t) [\Phi_v(t) - \Phi(t)]^2 = \text{Var } F(t)$$

$$\bar{\Phi}(t+1) = \sum_{v \in K(t)} \Phi_v(t) Y_v(t+1)$$

$$\tilde{\rho}_{i,v} = \frac{1}{n} \sum_{\kappa=1}^n \tilde{\epsilon}_{i,v}^{(\kappa)} \quad \text{with } \tilde{\epsilon}_{i,v}^{(\kappa)} = \begin{cases} 1 & \text{if } i \in \tilde{\mathcal{S}}_v^{(\kappa)} \\ 0 & \text{otherwise} \end{cases}$$

$$f_i(\epsilon) = f_i(\epsilon=0) - \epsilon \mathcal{C}(\mathbf{a}_i, \mathbf{a}_v) + \mathcal{C}_0(\epsilon) \quad \mathcal{Y}(\Phi, t) = \int d\Phi Y_v(t) \delta(\Phi - \Phi_v(t))$$

Łuksza pathogen fitness model

Our prediction is based on frequency and fitness data that depend only on information actually available at a given point in time. Consider a clade v containing a set of strains i with frequencies x_i in a given season t . The observed frequency of that clade in season t , which is denoted as $X_v(t)$, is simply the sum of these strain frequencies, $X_v(t) = \sum_{i:v,t} x_i$. This sum is defined as an average over strain trees, as detailed in Methods. Each strain has a Malthusian fitness or growth rate f_i (measured in units of 1/year), which is to be specified by our model. Given these initial data, we predict the frequency of that clade in the season 1 year later,

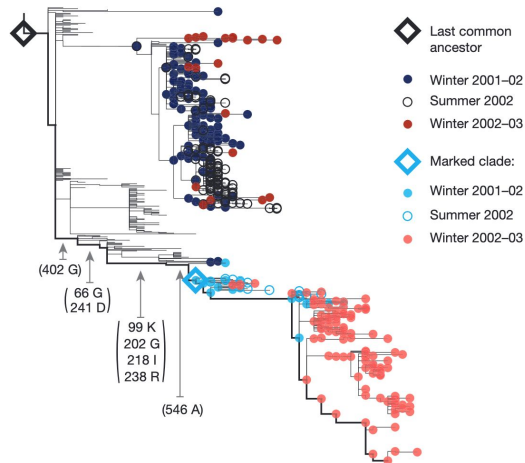
$$\hat{X}_v(t+1) = \sum_{i:v,t} x_i \exp(f_i) \quad (1)$$

Non-epitope mutations are predominantly under negative selection⁹, because they affect protein stability and other conserved molecular functions^{13,14}. Here we describe these effects by a simple mutational-load model: each strain incurs a fitness cost $\mathcal{L}(\mathbf{a}_i)$ that is the cumulative effect of recent non-epitope amino acid changes, which occur in its ancestral lineage in the current season (Methods).

Together we obtain a strain fitness of the form

$$f_i = f_0 - \mathcal{L}(\mathbf{a}_i) - \sum_{j: t_j < t_i} x_j \mathcal{C}(\mathbf{a}_i, \mathbf{a}_j) \quad (2)$$

with a constant f_0 ensuring the correct normalization of strain frequencies (Methods). Importantly, this strain-based model goes beyond a fitness model for individual mutations: it counts each new beneficial or



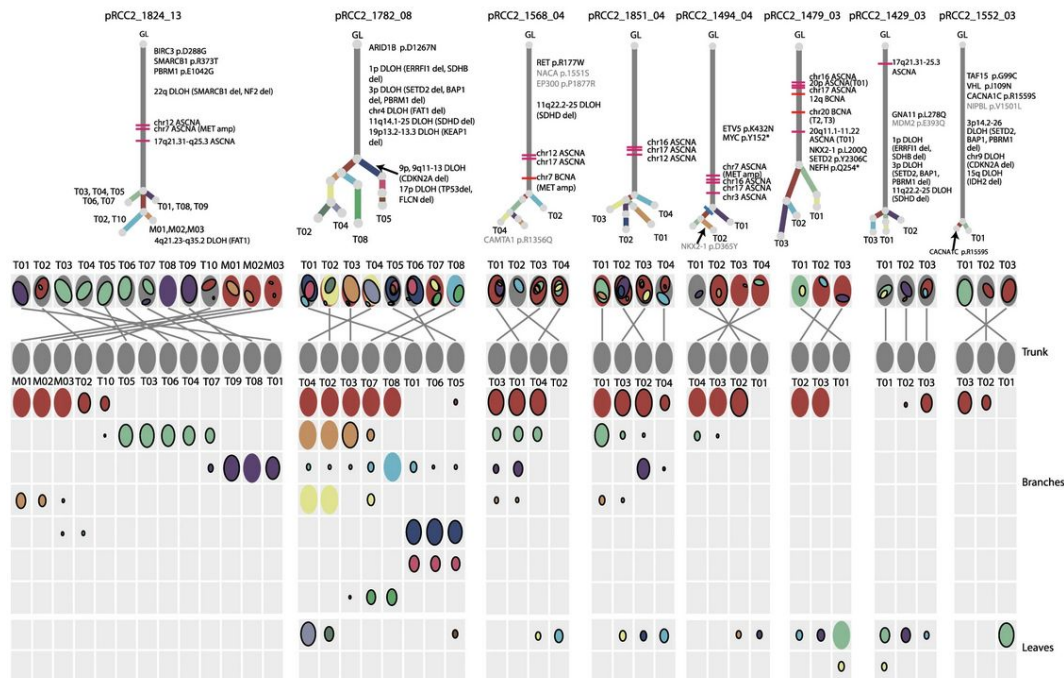
Prediction of cross-immunity. Another aggregate variable suitable for predictions is the average cross-immunity between a given strain v and the circulating strains in a given season,

$$\mathcal{C}_v(t) \equiv \sum_{i:t} x_i \mathcal{C}(\mathbf{a}_i, \mathbf{a}_v) \quad (24)$$

Our model predicts the expected average cross-immunity in season $t + 1$,

$$\hat{\mathcal{C}}_v(t+1) = \sum_{i:t} x_i \exp(f_i) \mathcal{C}(\mathbf{a}_i, \mathbf{a}_v) \quad (25)$$

Tumor clonal structure



<https://doi.org/10.1038/s41467-020-16546-5>

OPEN

The genomic and epigenomic evolutionary history of papillary renal cell carcinomas

“Drivers”, clonality, allele frequency

Modeling the subclonal evolution of cancer cell populations

Diego Chowell^{1,2,‡}, James Napier³, Rohan Gupta³, Karen S. Anderson^{2,4}, Carlo C. Maley^{2,4,5,6,*}, and Melissa A. Wilson Sayres^{2,4,7,*}

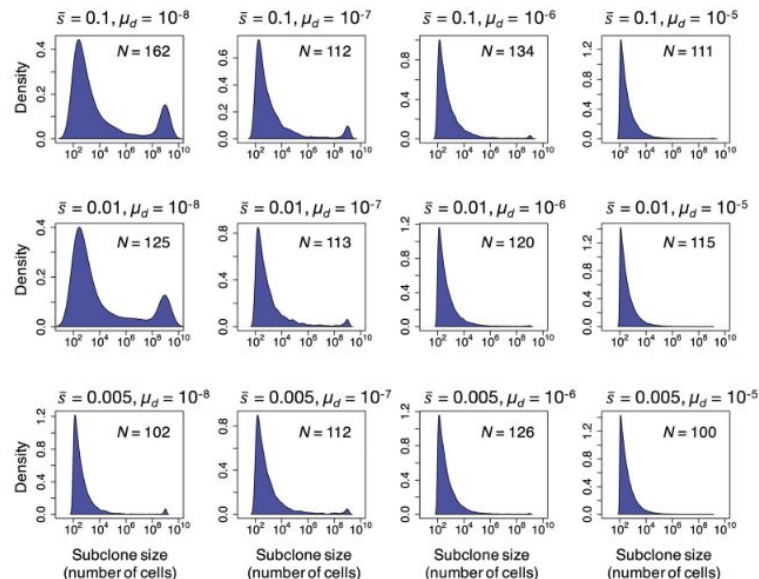
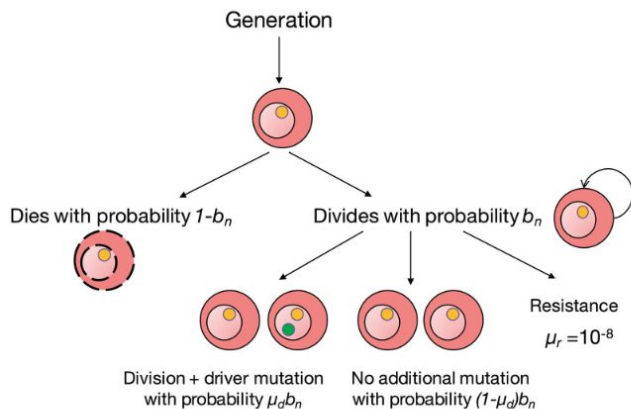


Figure 1.

Branching evolutionary process of cancer. Schematic representation of the process developed to simulate the subclonal evolution of cancer is presented below. For details of the process and assumptions, see main Quick Guide to Equations and Assumptions.

PhyloWGS: reconstructing cancer phylogenies from sequencing data

Deshwar *et al. Genome Biology* (2015) 16:35
DOI 10.1186/s13059-015-0602-8



METHOD

Open Access

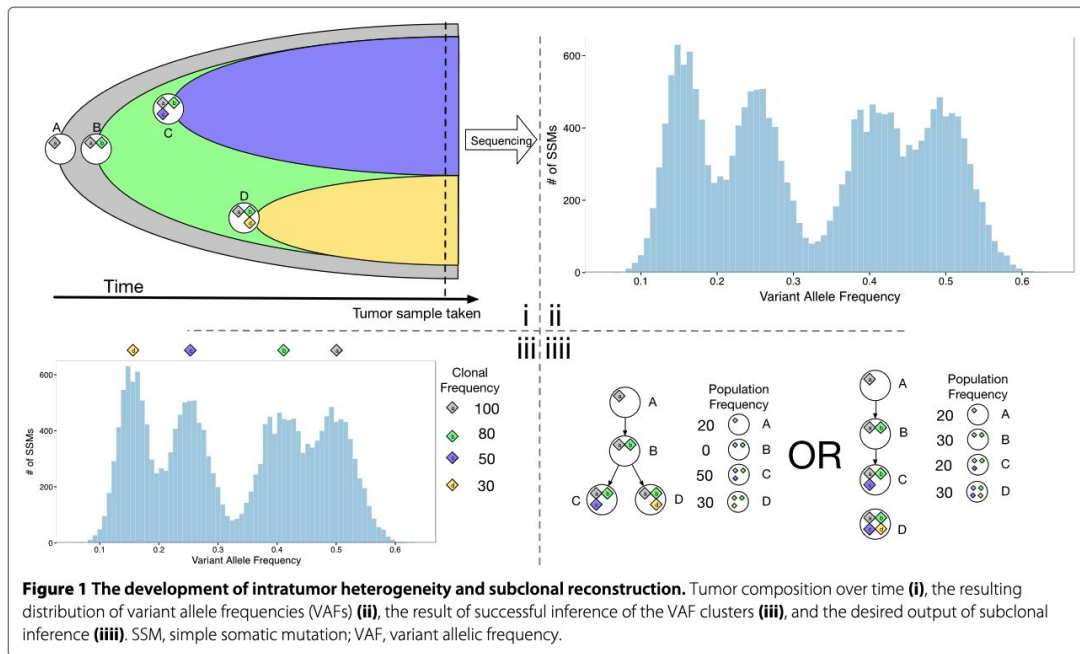
PhyloWGS: Reconstructing subclonal composition and evolution from whole-genome sequencing of tumors

Amit G Deshwar¹, Shankar Vembu², Christina K Yung³, Gun Ho Jang³, Lincoln Stein^{3,5}
and Qaid Morris^{1,2,4,5*}

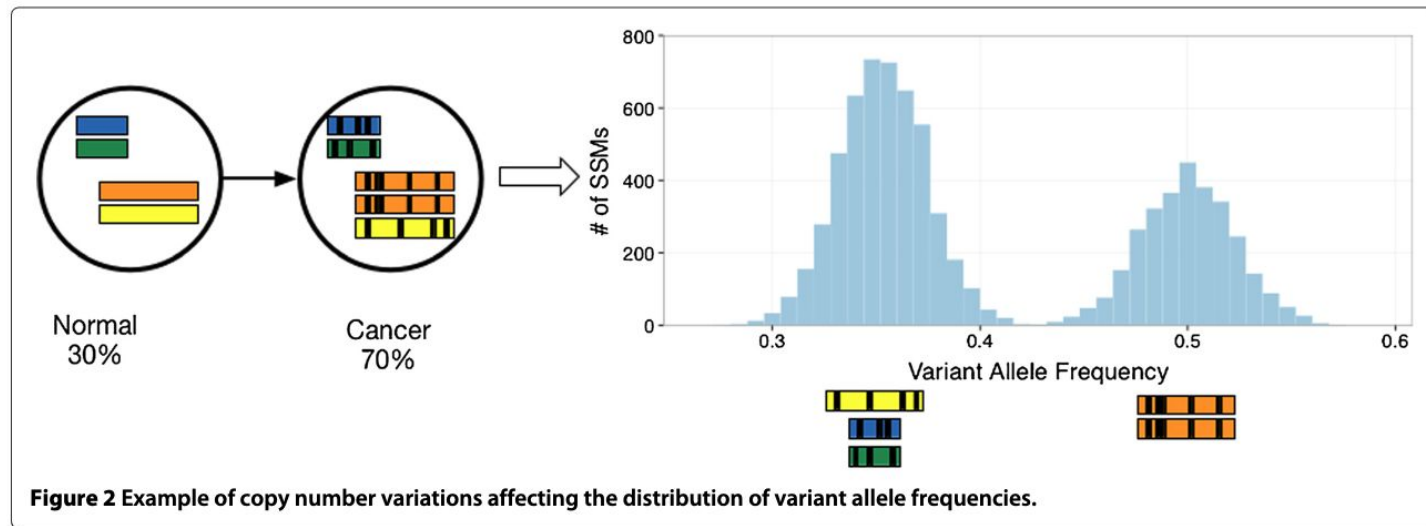
Abstract

Tumors often contain multiple subpopulations of cancerous cells defined by distinct somatic mutations. We describe a new method, PhyloWGS, which can be applied to whole-genome sequencing data from one or more tumor samples to reconstruct complete genotypes of these subpopulations based on variant allele frequencies (VAFs) of point mutations and population frequencies of structural variations. We introduce a principled phylogenetic correction for VAFs in loci affected by copy number alterations and we show that this correction greatly improves subclonal reconstruction compared to existing methods. PhyloWGS is free, open-source software, available at <https://github.com/morrislab/phylogws>.

PhyloWGS: identifying sub-clones by allele fraction clustering



PhyloWGS: adjusting for tumor “purity” / normal “contamination”



PhyloWGS: adjusting for copy number changes

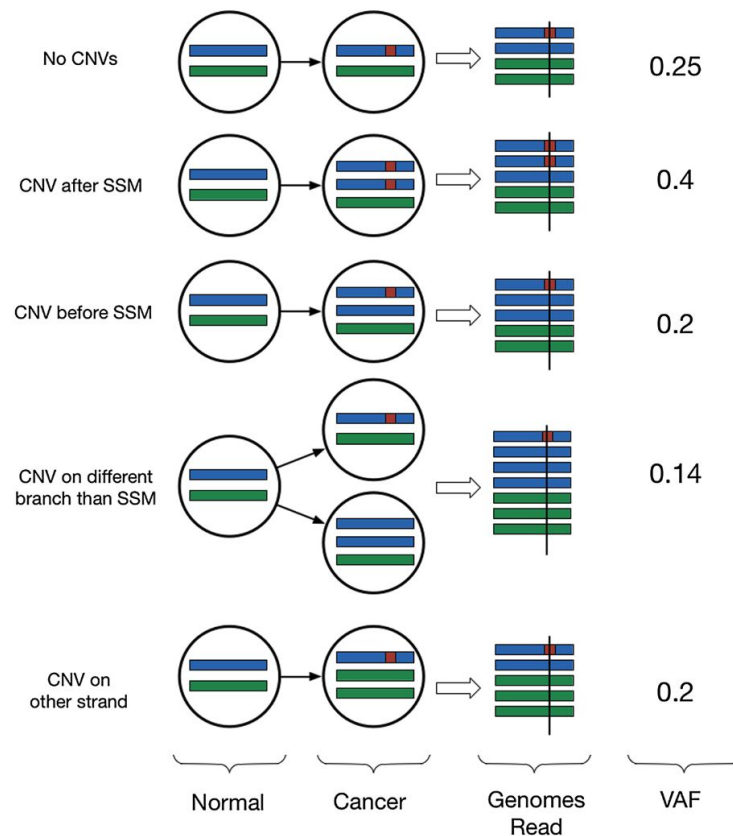


Figure 3 Changes to VAF caused by CNVs with different phylogenetic relationships. CNV, copy number variation; SSM, simple somatic mutation; VAF, variant allelic frequency.

The Paper

Let's apply the pathogen fitness model to tumor phylogenies

A neoantigen fitness model predicts tumour response to checkpoint blockade immunotherapy

Marta Luksza¹, Nadeem Riaz^{2,3}, Vladimir Makarov^{3,4}, Vinod P. Balachandran^{5,6,7}, Matthew D. Hellmann^{7,8,9}, Alexander Solovyyov^{10,11,12,13}, Naiyer A. Rizvi¹⁴, Taha Merghoub^{7,15,16}, Arnold J. Levine¹, Timothy A. Chan^{2,3,4,7}, Jedd D. Wolchok^{7,8,15,16} & Benjamin D. Greenbaum^{10,11,12,13}

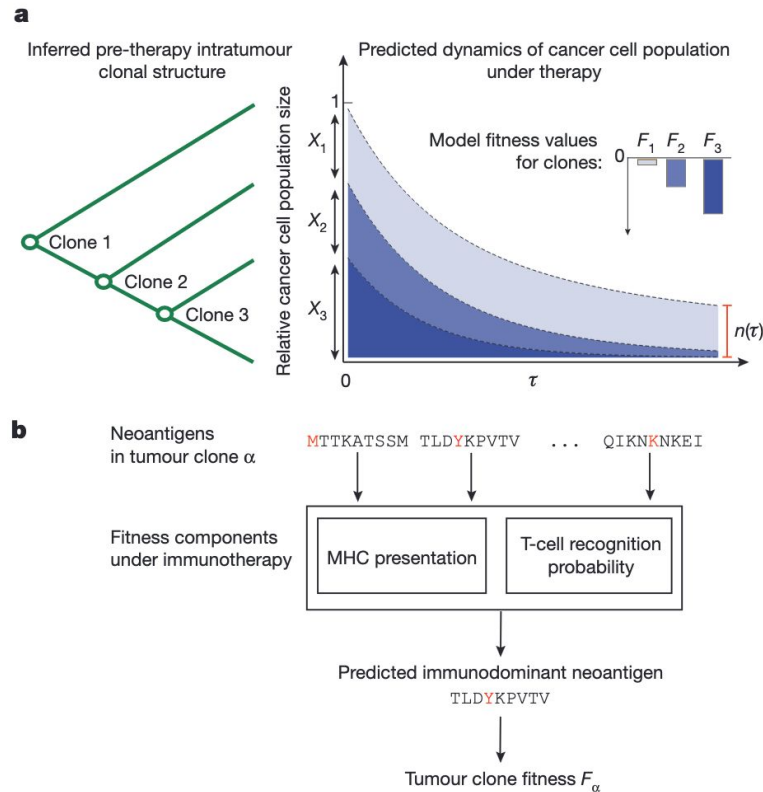
$$n(\tau) = \sum_{\alpha} X_{\alpha} \exp(F_{\alpha} \tau)$$

We propose a fitness model of tumour-immune interactions as a general mathematical framework to describe the evolutionary dynamics of cancer cell populations under checkpoint blockade immunotherapy and provide a proof of concept for its utility (Fig. 1). Analogous fitness models based on immune interactions have been successfully applied to human influenza⁷, HIV⁸ and chronic viral infections⁹. Checkpoint blockade exposes cancer cells to strong immune pressure on their neoantigens, reducing their reproductive success. Our model predicts the evolutionary dynamics of a cancer cell population after a finite time under such pressure. We compute $n(\tau)$, the predicted future effective size of a cancer cell population in a tumour relative to its effective size at the start of therapy. The size is a weighted sum over all of the genetic clones of a tumour (Fig. 1a and Methods),

$$n(\tau) = \sum_{\alpha} X_{\alpha} \exp(F_{\alpha} \tau) \quad (1)$$

Basic idea

- Use PhyloWGS to infer clones and their evolutionary relationship from tumor + normal DNA sequencing
- Define immune fitness of each sub-clone in terms of predicted neoantigens



Fitness \sim immunogenicity of dominant neoantigen in the clone

(Fig. 1b and Methods). Here we model the fitness of a given clone α by the recognition potential of its dominant neoantigen,

$$F_{\alpha} = - \max_{i \in \text{clone } \alpha} (A_i \times R_i) \quad (2)$$

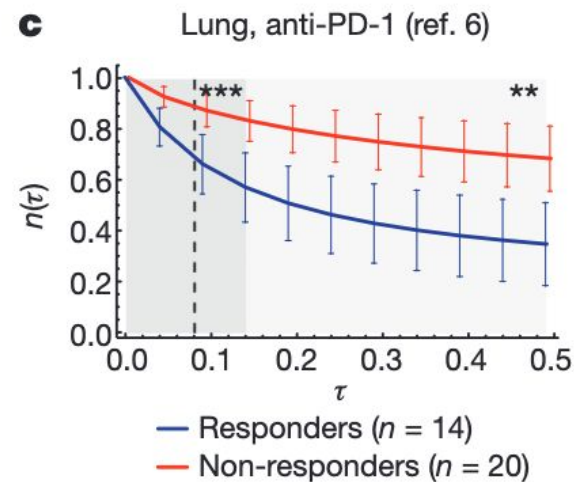
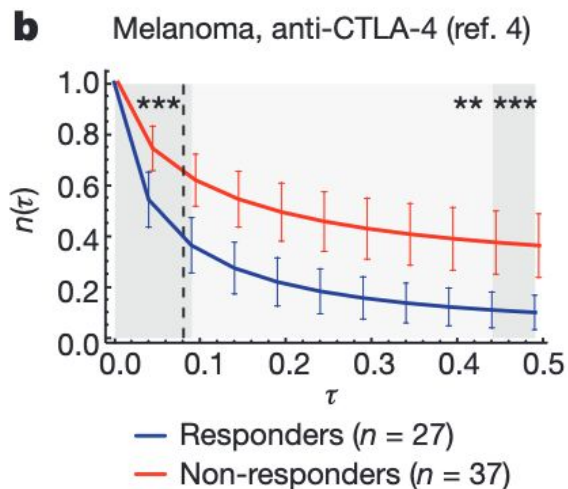
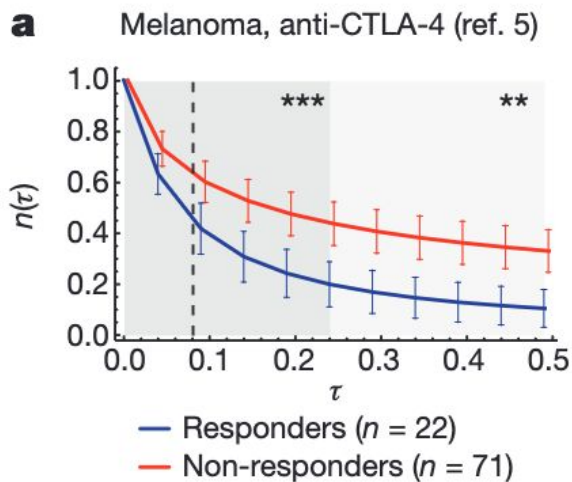
Recognition potential: “R” and “A”

$$n(\tau) = \sum_{\alpha} X_{\alpha} \exp \left(- \max_{i \in \text{clone } \alpha} (A_i \times R_i) \tau \right) \quad (5)$$

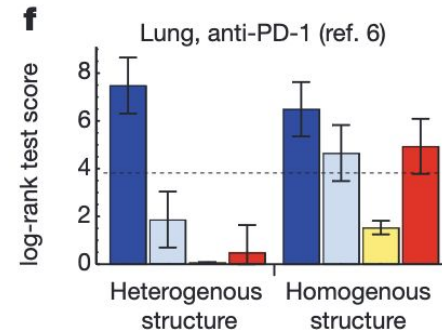
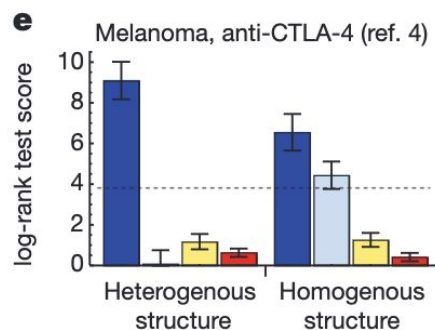
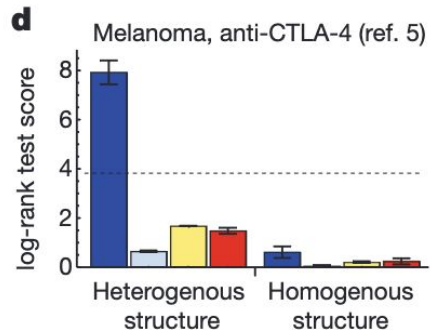
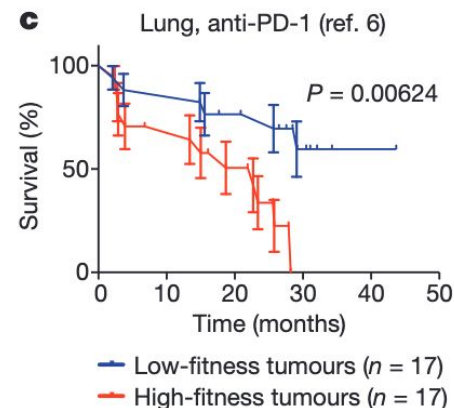
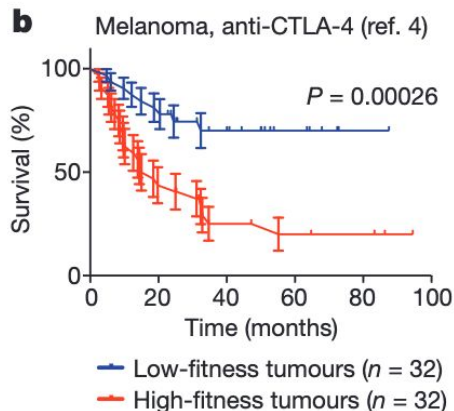
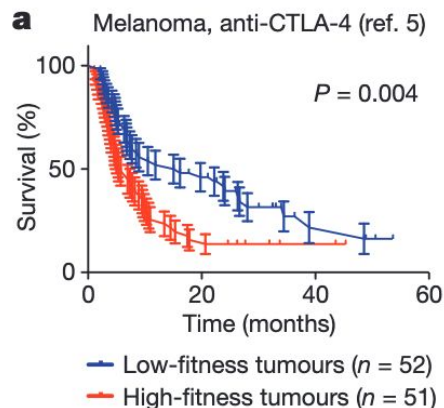
$$A = K_d^{\text{WT}} / K_d^{\text{MT}} \quad (7)$$

$$R = Z(k)^{-1} \sum_{e \in \text{IEDB}} \exp(-k(a - |\mathbf{s}, \mathbf{e}|)) \quad (10)$$

Predicted tumor sizes



Survival stratified by fitness



Neoantigen fitness model: ■ $A \times R$ ■ A ■ R ■ Neoantigen load

Homework

Data

Sample	Months	Status					
Pat02	53.6547357	0					
Pat03	3.28766763	1					
Pat04	32.4492795	0					
Pat06	5.29314488	1					
Pat08	4.60273468	1					
Pat100	11.8356035	1					
Pat101	9.46848277	1					
Pat103	34.4547568	1					
Pat104	7.79177228	1					
Pat106	8.21916908	1					
Pat109	2.72876413	1					
Pat11	25.9725743	0					
Pat110	10.5205364	1					
Pat113	9.89587957	1					
Pat115	4.76711806	1					
Pat117	30.0492821	0					
Pat118	10.2903997	1					
Pat119	26.5643545	0					
Pat121	4.01095451	1					
Pat123	28.0438049	1					
Pat124	4.79999474	1					
Pat126	21.0739495	0					
Pat127	10.9150565	1					
Pat128	3.71506442	1					
Pat129	17.7205285	0					
Pat13	24.0657271	1					
Pat130	1.47945043	1					
Pat131	8.44930581	1					
Pat132	22.2246332	0					
Pat133	17.8191586	1					
Pat135	2.66301078	1					
Pat138	48.5917276	1					
Pat139	2.25479995	1					
►	Survival VanAllen et al.	Survival Snyder et al.	Survival Rizvi et al.				

A	B	C	D	E	F	G	H	I	J	K
ID	MUTATION_Sample	WT.Peptide	MT.Peptide	MT.Allele	WT.Score	MT.Score	HLA			
1	1_44084332:AL4602	RGTECTGLI	RVTETCGLI	C1502	1208	124	A0301,A3201,B0801,B5101,C0702,C1502			
2	1_56990069:AL4602	IGRLRPHFL	IVRLRPHFL	B0801	305	165	A0301,A3201,B0801,B5101,C0702,C1502			
3	1_15219203:AL4602	SSNGPHGSV	SSNVPHGSV	C1502	73	210	A0301,A3201,B0801,B5101,C0702,C1502			
4	1_15822729:AL4602	ALWFRKRCF	ALWFRKLCF	A3201	547	284	A0301,A3201,B0801,B5101,C0702,C1502			
5	1_15941037:AL4602	NTREHDQLI	NSREHDQLI	C1502	41	42	A0301,A3201,B0801,B5101,C0702,C1502			
6	1_15950577:AL4602	MKRKNFTEV	MKRKNFTEL	B0801	258	63	A0301,A3201,B0801,B5101,C0702,C1502			
7	1_23134964:AL4602	HTLTFFING	HTLTFFINW	A3201	11745	59	A0301,A3201,B0801,B5101,C0702,C1502			
8	1_24836727:AL4602	RAFMKILGK	RALMKILGK	A0301	36	89	A0301,A3201,B0801,B5101,C0702,C1502			
9	1_24836727:AL4602	EVTRAFMKI	EVTRALMKI	C1502	241	327	A0301,A3201,B0801,B5101,C0702,C1502			
10	10_7357937:AL4602	LTVHVTQPK	LTVHMTQPK	A0301	442	460	A0301,A3201,B0801,B5101,C0702,C1502			
11	11_2064826:AL4602	KVVYFTATF	KVVYFMATF	A3201	5	11	A0301,A3201,B0801,B5101,C0702,C1502			
12	11_2064826:AL4602	VYFTATFPY	VYFMATFPY	C0702	162	220	A0301,A3201,B0801,B5101,C0702,C1502			
13	11_2064826:AL4602	TSQKVVYFT	TSQKVVYFM	C1502	8765	430	A0301,A3201,B0801,B5101,C0702,C1502			
14	11_5532201:AL4602	NVQEIFVVF	NVHEIVFVF	C1502	341	222	A0301,A3201,B0801,B5101,C0702,C1502			
15	11_5611444:AL4602	ALKRTLNR	ALKRTFTNR	A0301	670	359	A0301,A3201,B0801,B5101,C0702,C1502			
16	11_5611444:AL4602	RTLNRFKI	RTFTNRFKI	A3201	36	14	A0301,A3201,B0801,B5101,C0702,C1502			
17	11_5611444:AL4602	LTNRFKIPI	FTNRFKIPI	A3201	104	377	A0301,A3201,B0801,B5101,C0702,C1502			
18	11_5611444:AL4602	RTLNRFKI	RTFTNRFKI	C1502	136	154	A0301,A3201,B0801,B5101,C0702,C1502			
19	11_5611444:AL4602	LTNRFKIPI	FTNRFKIPI	C1502	148	78	A0301,A3201,B0801,B5101,C0702,C1502			
20	11_6313776:AL4602	ISIPLDSNM	ISIQLDSNM	C1502	310	452	A0301,A3201,B0801,B5101,C0702,C1502			
21	11_8287819:AL4602	QVDEHSKPP	QVDEHSKPL	C1502	17298	178	A0301,A3201,B0801,B5101,C0702,C1502			
22	12_4662254:AL4602	KVKSPVEK	VMVKSPVEK	A0301	19358	89	A0301,A3201,B0801,B5101,C0702,C1502			
23	12_4662254:AL4602	KVKSPVEKK	MVKSPVEKK	A0301	99	155	A0301,A3201,B0801,B5101,C0702,C1502			
24	12_9334662:AL4602	IARMFIFAI	NARMFIFAI	B0801	667	145	A0301,A3201,B0801,B5101,C0702,C1502			
25	12_2100797:AL4602	ALSFYSIAK	VLSFSYIAK	A0301	27	24	A0301,A3201,B0801,B5101,C0702,C1502			
26	12_2100797:AL4602	KMFLAALSF	KMFLAVLSF	A3201	4	3	A0301,A3201,B0801,B5101,C0702,C1502			
27	12_2100797:AL4602	LAALSFYSI	LAVLSFSYI	B5101	295	282	A0301,A3201,B0801,B5101,C0702,C1502			
28	12_2105140:AL4602	AMGFQSMV	ATGFQSMVI	A3201	75	125	A0301,A3201,B0801,B5101,C0702,C1502			
29	12_2105140:AL4602	LAMGFQSM	LATGFQSMV	C1502	49	181	A0301,A3201,B0801,B5101,C0702,C1502			
30	12_2105140:AL4602	AMGFQSMV	ATGFQSMVI	C1502	2826	79	A0301,A3201,B0801,B5101,C0702,C1502			
31	12_2539828:AL4602	VVGAGGVGI	VVGASGVGI	A0301	410	217	A0301,A3201,B0801,B5101,C0702,C1502			
32	12_5282062:AL4602	LSGEGVSPV	LSGEGVSPV	C1502	771	238	A0301,A3201,B0801,B5101,C0702,C1502			
33	12_5282062:AL4602	GUSPVMISV	VUSPVMISV	C1502	2091	364	A0301,A3201,B0801,B5101,C0702,C1502			
►	VanAllen et al.	Snyder et al.	Rizvi et al.	+						

⬢ *Fin* ⬢