# Deconvolving clinically relevant cellular immune crosstalk from bulk gene expression using CODEFACS and LIRICS

Kun Wang[**,1], Sushant Patkar[**,1,2], Joo Sang Lee[1,3], E. Michael Gertz[1], Welles Robinson[1,2], Fiorella Schischlik[1], David R. Crawford[1,4], Alejandro A. Schäffer[1], Eytan Ruppin[#,1]

[1] Cancer Data Science Laboratory, National Cancer Institute, NIH, Bethesda, MD

[2] Department of Computer Science, University of Maryland, College Park, MD

[3] Department of Artificial Intelligence & Department of Precision Medicine, School of Medicine, Sungkyunkwan University, Suwon, Republic of Korea

[4] Department of Cell Biology and Molecular Genetics, University of Maryland, College Park, MD

[**] These authors contributed equally

[#] Corresponding author:

eytan.ruppin@nih.gov

Key words: deconvolution, immunotherapy, tumor microenvironment, ligand-receptor interaction, cell-cell interaction, microsatellite instability

**Abstract**

The tumor microenvironment (TME) is a complex mixture of cell-types that interact with each other to affect tumor growth and clinical outcomes. To accelerate the discovery of such interactions, we developed CODEFACS (COnfident DEconvolution For All Cell Subsets), a deconvolution tool inferring cell-type-specific gene expression in each sample from bulk expression measurements, and LIRICS (LIgand Receptor Interactions between Cell Subsets), a supporting pipeline that analyzes the deconvolved gene expression from CODEFACS to identify clinically relevant ligand-receptor interactions between cell-types. Using 15 benchmark test datasets, we first demonstrate that CODEFACS substantially improves the ability to reconstruct cell-type-specific transcriptomes from individual bulk samples, compared to the state-of-the-art method, CIBERSORTx. Second, analyzing the TCGA, we uncover cell-cell interactions that specifically occur in TME of mismatch-repair-deficient tumors and are associated with their high response rates to anti-PD1 treatment. These results point to specific T-cell co-stimulating interactions that enhance immunotherapy responses in tumors independently of their mutation burden levels. Finally, using machine learning, we identify a subset of cell-cell interactions that predict patient response to anti-PD1 therapy in melanoma better than recently published bulk transcriptomics-based signatures. CODEFACS offers a way to study bulk cancer and normal transcriptomes at a cell type-specific resolution, complementing single-cell transcriptomics.

## Introduction

The importance of the tumor microenvironment (TME) in cancer has been recognized since the late 1800s[1]. The recent success of immune checkpoint blockade has further sparked interest in studying TME interactions that shape clinical outcomes following immunotherapy, aiming to find biomarkers of treatment response and new treatment opportunities[2]. One key step in studying these interactions is the characterization of the molecular profiles of different cell types in a patient's tumor sample. Fluorescence-activated cell sorting (FACS) and single-cell RNA sequencing have emerged as effective tools to address this challenge[3]. However, due to the cost of these procedures and scarcity of fresh tumor biopsies, the application of these approaches has remained limited. Given that bulk tumor gene expression from preserved biopsies is far more abundant, computational methods that can effectively extract cell-type-specific expression from such data, termed *deconvolution* algorithms, could be very helpful. If successful, such deconvolution methods can markedly advance our knowledge of the TME across many tumor types and different contexts, and beyond that, they may be readily applied to interrogate other large bulk expression datasets.

Several previous studies have developed a variety of expression deconvolution algorithms. DeMixT[4] was designed to estimate individual-specific expression for three cell components provided prior reference samples of two of these cell components. ISOpure[5] has aimed to derive sample-specific cancer cell expression with the assumption that the observed bulk gene expression profile is a mixture of predefined stromal and immune cell expression profiles that are shared across all the samples. Building on this work, Fox et al extended ISOpure to predict individual-specific non-tumor cell expression by subtracting cancer cell profiles from the bulk mixtures in a two-cell type model[6]. More recently, Newman et al[7] developed CIBERSORTx, the first approach that aims to predict the *sample-specific* gene expression of all cell types composing it by employing a set of novel deconvolution heuristics. As a proof of concept, Newman et al showed that CIBERSORTx can accurately reconstruct the cell-type-specific expression of genes in each input sample under certain modelling assumptions. This groundbreaking work has, however, some notable limitations: (1) The number of genes whose cell-type-specific

expression can be reconstructed in each sample is relatively small, especially for low-abundance cell types, and (2) their approach does not provide confidence estimations of the predictions made, while such estimations could be potentially useful in most deconvolution applications in the absence of ground truth data.

Here, we introduce a new deconvolution algorithm and software, CODEFACS (COnfident DEconvolution For All Cell Subsets), which markedly advances the ability to successfully deconvolve bulk gene expression data. CODEFACS receives as input bulk gene expression profiles of tumor samples and either pre-computed estimates of abundance of expected tumor, immunological and stromal cell types in each sample, or their prototypical molecular signatures, which serve as seeds for estimating the abundance of each cell type in each sample. CODEFACS then predicts the cell type-specific gene expression profiles of each sample. It is a heuristic approach aimed at maximizing the number of genes in each cell type whose expression across the samples can be confidently predicted via a heuristic method to estimate confidence. Using 15 benchmark datasets where the ground-truth is known, we show that CODEFACS robustly improves over CIBERSORTx, both in terms of gene coverage and the individual gene expression estimation accuracy. We additionally developed LIRICS (LIgand Receptor Interactions between Cell Subsets), a pipeline that integrates the output of CODEFACS with a database of prior immunological knowledge that we curated to infer the active cell-cell interaction landscape in each sample. These data can then be analyzed in conjunction with any sample-associated clinical annotations (e.g., response to treatment) to infer the most important clinically relevant immune interactions between the cell types in a given patient's cancer cohort.

Building on its enhanced coverage and accuracy, we next applied CODEFACS to reconstruct the cell-type-specific transcriptomes of ~8000 tumor samples from 21 cancer types in TCGA. Analyzing these fully deconvolved TCGA expression datasets using LIRICS we find a shared repertoire of intercellular interactions enriched in the TME of mismatch repair deficient tumors of different tissues of origin, which is associated with improved overall patient survival and high response rates to anti-PD1 treatment,

102 independently of their mutation burden levels. Finally, using machine learning techniques,
103 we identify a subset of intercellular TME interactions that are predictive of response to
104 immune checkpoint blockade treatment in melanoma patients.
105
106 In summary, CODEFACS and LIRICS present a new way to analyze large bulk
107 RNA-seq datasets to study cellular crosstalk in the TME of each patient, and to learn
108 more about the association of different tumor-immune interactions with different clinical
109 measures. The potential scope of applications of both CODEFACS and LIRICS goes
110 beyond studying the TME, as these tools can be applied to study any disease of interest
111 given bulk gene expression data and relevant reference signatures of cell types involved.
112

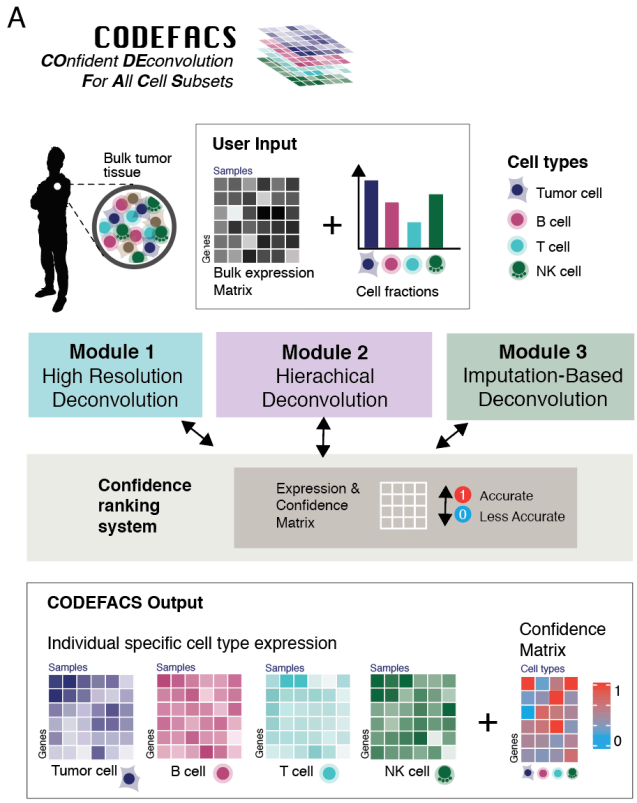
113 **Results**

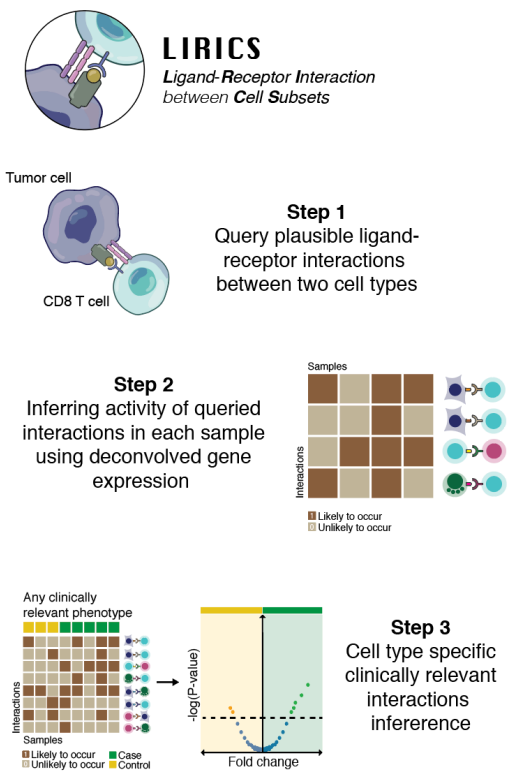
114 **Overview of CODEFACS and LIRICS**

115 CODEFACS is designed to characterize the tumor microenvironment by reconstructing
116 the cell-type-specific transcriptomes of each sample from bulk expression. It takes as
117 input the bulk RNA-seq expression values of a cohort of tumor samples and either the
118 estimations of the cell fractions of a pre-defined set of cell types in each sample or their
119 cell-type-specific molecular signature profiles, derived based on reference datasets or
120 from the literature[7].
121
122 CODEFACS then employs a heuristic approach that sequentially executes three
123 modules: *(module 1) high resolution deconvolution, (module 2) hierarchical deconvolution*
124 *and (module 3) imputation*. Each module is designed to predict the cell-type-specific
125 expression of genes in each sample; the second and third modules aim to overcome the
126 shortcomings of the previous modules. For a more detailed description of each module
127 and its modeling assumptions, see **Supplementary Note**. A key component of
128 CODEFACS is its *confidence ranking system*, which receives cell-type-specific
129 expression predictions from the different modules and labels them as high or low-
130 confidence estimations. Genes whose expression is determined with high confidence in
131 a given module are added to the output set, while low confidence predictions are
132 continued to be processed in subsequent modules (See **Fig. 1A** and **Supplementary**

**Note**). The final output of CODEFACS consists of two items: (a) a three-dimensional gene expression matrix, where each entry represents the predicted gene expression a gene in a given cell-type in a specific sample, and (b) a two-dimensional matrix of confidence scores ranging from [0,1] representing which gene-cell-type pairs have most confident predictions (~1) and which pairs have least confident predictions (~0; See **Fig. 1A, Output**). These scores can be further investigated to assess the quality of predictions for a given dataset.

Given fully deconvolved gene expression data from CODEFACS, one can use LIRICS (LIgand Receptor Interactions between Cell Subsets) (**Fig. 1B**) to transform this data into a biologically interpretable feature space of active ligand-receptor interactions between cell types in each sample. Specifically, LIRICS takes the output of CODEFACS and processes it in three steps: (step 1) The first step queries a database of all plausible ligand-receptor interactions between any two cell types A and B, that we have systematically assembled and curated from the literature. This database is publicly available as part of LIRICS (see **Supplementary Note**, supplementary tables 1-3). (step 2) In the second step, given the deconvolved expression profiles of cell type A and cell type B in a given bulk tumor sample, LIRICS denotes as 'active' or 'likely to occur' ('1') the interactions where both the ligand and receptor are over-expressed in the relevant cell-types in that sample, or otherwise 'inactive' ('0'). A ligand or receptor is considered to be overexpressed in a given cell type if its expression exceeds the median expression in that cell type (**Supplementary Note**). (step 3) Finally, a Fisher's enrichment analysis is performed to test the association of the activity of specific ligand-receptor interactions with any relevant phenotypes of interest (e.g., treatment response, mutational subtype, etc.) (**Fig. 1B**). Furthermore, if required, one can collectively analyze the binary profiles returned by LIRICS from multiple independent datasets to augment sample size and increase statistical power. Finally, one can apply a false discovery rate (FDR) cut-off post multiple hypothesis correction and visualize significantly enriched interactions in a volcano plot as shown at the bottom of Fig 1B, or, alternatively, display their network structure as shown in subsequent biological applications.


A
CODEFACS
COnfident DEconvolution
For All Cell Subsets
Bulk tumor tissue
User Input
Samples
Genes
Bulk expression Matrix
Cell fractions
Cell types
Tumor cell
B cell
T cell
NK cell
Module 1
High Resolution Deconvolution
Module 2
Hierachical Deconvolution
Module 3
Imputation-Based Deconvolution
Confidence ranking system
Expression & Confidence Matrix
Accurate
Less Accurate
CODEFACS Output
Individual specific cell type expression
Confidence Matrix
Cell types
Tumor cell
B cell
T cell
NK cell


B
LIRICS
Ligand-Receptor Interaction between Cell Subsets
Tumor cell
CD8 T cell
Step 1
Query plausible ligand-receptor interactions between two cell types
Step 2
Inferring activity of queried interactions in each sample using deconvolved gene expression
Samples
Interactions
Likely to occur
Unlikely to occur
Any clinically relevant phenotype
Case
Control
-log(P-value)
Fold change
Step 3
Cell type specific clinically relevant interactions infererence

164

165

**Figure 1. Overview of CODEFACS and LIRICS. (A) CODEFACS** takes bulk gene expression profiles and prior knowledge of the cellular composition of each sample and executes a heuristic three step procedure to infer the deconvolved gene expression in each sample. In module 1, we perform a high-resolution deconvolution, which extends the CIBERSORTx algorithm. In module 2 (hierarchical deconvolution), bulk expression is modeled as a mixture of two components: a specific cell type of interest and all the remaining cell types. The expression for the cell type of interest is predicted by removing the estimated expression in the second component (using high-resolution deconvolution from module 1) from the bulk mixture. In module 3 – imputation-based deconvolution, we impute the cell-type-specific expression of a specific gene based on the predicted cell-type-specific expression of other high-confidence genes that are co-expressed with that gene in the bulk. Each module is designed to overcome the shortcomings of its predecessor based on their respective modeling assumptions. The confidence ranking system is responsible for classifying all the predictions at the end of each module into high-confidence or low-confidence predictions. Genes classified into the low-confidence class at the end of one module (e.g. module 1) are passed to the next module (e.g. module 2) for refinement. Finally, after all the three modules are executed, the prediction confidence levels are re-evaluated. The final output of CODEFACS consists of a 3-dimensional matrix with cell-type-specific gene expression predictions for each sample, along with estimated confidence scores of predictions for each gene in each cell type. For more details, see Supplementary note. **(B) LIRICS** takes the output of CODEFACS and processes it in three steps. In step 1, for each possible permutation of cell type pairs, LIRICS queries a literature-curated repository for enumerating all plausible ligand-receptor interactions between specified cell types. In step 2 this prior knowledge is integrated with the output from CODEFACS to infer which of the plausible cell-cell interactions are likely to occur or be “active” in each individual sample. The result is a binary matrix with rows representing each plausible cell-cell interaction and columns representing each patient’s tumor sample. Finally, in step 3, given any clinically relevant phenotype (e.g. response to therapy, driver mutation status, etc..), one can perform a Fisher’s enrichment analysis (shown at the bottom) to discover cell-grounded receptor-ligand interactions in the TME that are associated with the phenotype of interest. For more details, see Supplementary note.

166

167

168

169 **Evaluating the performance of CODEFACS**

170 To assess the accuracy of CODEFACS, we generated 15 benchmark datasets (see
171 Methods) by merging publicly available single cell RNA-seq[8,9] and FACS sorted purified
172 RNA-seq[10]. Thereafter, we applied CODEFACS to deconvolve these generated bulk
173 datasets and define the accuracy of its predictions by computing the Kendall correlation
174 between the predicted and ground truth expression in each cell type across individual
175 samples (the Kendall correlation provides a less inflated measure of accuracy by
176 accounting for ties in the data). In the main text, we show the results obtained on three
177 benchmark bulk datasets: one derived form a FACS-sorted lung cancer data, one from a
178 single cell melanoma RNA-seq data and from a single cell glioblastoma RNA-seq dataset,
179 respectively. Each bulk sample from these datasets represents a real biopsy from a
180 patient. We show that CODEFACS can predict the cell-type-specific expression of more
181 genes than CIBERSORTx (with Kendal's correlation ≥ 0.3) (**Fig. 2A-C**), and its predictions
182 are overall more accurate (**Fig. 2D-F**). The results for all the remaining 12 benchmark
183 datasets, created via artificial mixing of single cell profiles and single cell RNA-Seq
184 imputation (See Methods and **Supplementary Table 6** for more details), also show the
185 superiority of CODEFACS (**supplementary Fig. S1 and S2**). Overall, we observe that
186 the more abundant the cell type is, the better CODEFACS can predict its cell-type-specific
187 gene expression (**supplementary Fig. S3**).

188

189 Next, we quantified how well the confidence scores it returns align with the Kendall
190 scores that measure the true prediction accuracy with the ground truth for each (gene,
191 cell-type) pair to validate the claim that our confidence scores can help the user filter out
192 potentially noisy predictions in real bulk datasets. We quantified this using two metrics:
193 Spearman correlation and a classification AUC (Area Under the ROC Curve). Across all
194 the benchmark datasets analyzed, the Spearman correlation between confidence scores
195 of genes in each cell-type and their corresponding Kendall scores (quantifying the true
196 prediction accuracy) is strong and positive (**Fig. 2G** depicts the results for the FACS
197 sorted lung cancer benchmark dataset, and **supplementary Fig. S4** depicts the results
198 for the remaining benchmark datasets); To perform a classification-based quantification,
199 we grouped the genes in each cell-type into two classes based on the correlation between

200 their predicted and actual expression, informative (prediction accuracy ≥ 0.1 and p-value
201 ≤ 0.05) and uninformative (prediction accuracy < 0.1 or p-value > 0.05). We then tested
202 whether the confidence scores could be used to classify genes into these two classes for
203 each cell type. We find that the confidence score could effectively filter out uninformative
204 predictions (**Fig. 2H** depicts the results for the FACS sorted lung cancer benchmark
205 dataset, and **supplementary Fig. S5,** depicts for the remaining benchmark datasets).
206
207 Finally, to further evaluate CODEFACS on real bulk tumor data (where the ground
208 truth is unavailable), we applied it to deconvolve bulk expression data from 21 cancer
209 types (~8000 RNA-seq samples) in TCGA. To infer the cellular abundance of each cell
210 type in each sample which is required as input for CODEFACS, we made use of matched
211 bulk methylation data available for these samples and methylation-based reference
212 signature profiles of distinct cell types. These include 11 cell type signatures
213 (macrophages/dendritic cells--CD14+, B cells--CD19+, CD4+T cells, CD8+ T cells, T reg
214 cells, NK cells--CD56+, endothelial cells, fibroblasts, neutrophils, basophils, eosinophils
215 and tissue-specific tumor cells) obtained from MethylCIBERSORT [11]. Reassuringly, we
216 found strong Spearman correlations between the resulting predicted tumor cell fraction
217 and the tumor purity estimates derived from matched mutation and copy number data
218 (based on ABSOLUTE) for the same samples across 10 cancer types (Spearman
219 correlation: min=0.72, max=0.88, avg=0.8; **supplementary Fig. S6**). This testifies that
220 methylation-based cell fraction estimates indeed form a reliable basis for running
221 CODEFACS to deconvolve TCGA samples.
222
223 We then asked if CODEFACS can recover the expected cell-type-specific gene
224 expression signature of different cell types in a given cancer type. To this end, we
225 computed the Spearman correlation between (a) the mean deconvolved gene expression
226 of the top confidently deconvolved genes in a given cell type (confidence score ≥ 0.95)
227 and (b) the mean expression of these genes, which we derived from completely
228 independent single cell expression data of the same cancer type (Methods). We find that
229 (a) and (b) are substantially correlated (**Fig. 2I** depicts results for the TCGA-LUAD (lung
230 adenocarcinoma) dataset as an example and **supplementary Fig. S7** for the remaining

231 cancer types that have publicly available scRNA-seq data). The concordance level is
232 higher for cell types that are abundant (e.g., tumor cells and fibroblasts) and decreases
233 for less abundant cell types. Additionally, we observed that tumor cells have the largest
234 fraction of genes whose expression is predicted with high confidence, with the highest in
235 thyroid cancer (THCA, ~67.4% of all genes). Furthermore, 7 KEGG pathways are
236 significantly enriched (adjusted p-value < 0.01) with highly confident genes in tumor cells
237 (confidence score ≥0.95) across the 21 cancer types (**supplementary Fig. S8**). Those
238 pathways mostly involve RNA transport, spliceosome, DNA replication, and mismatch
239 repair.
240

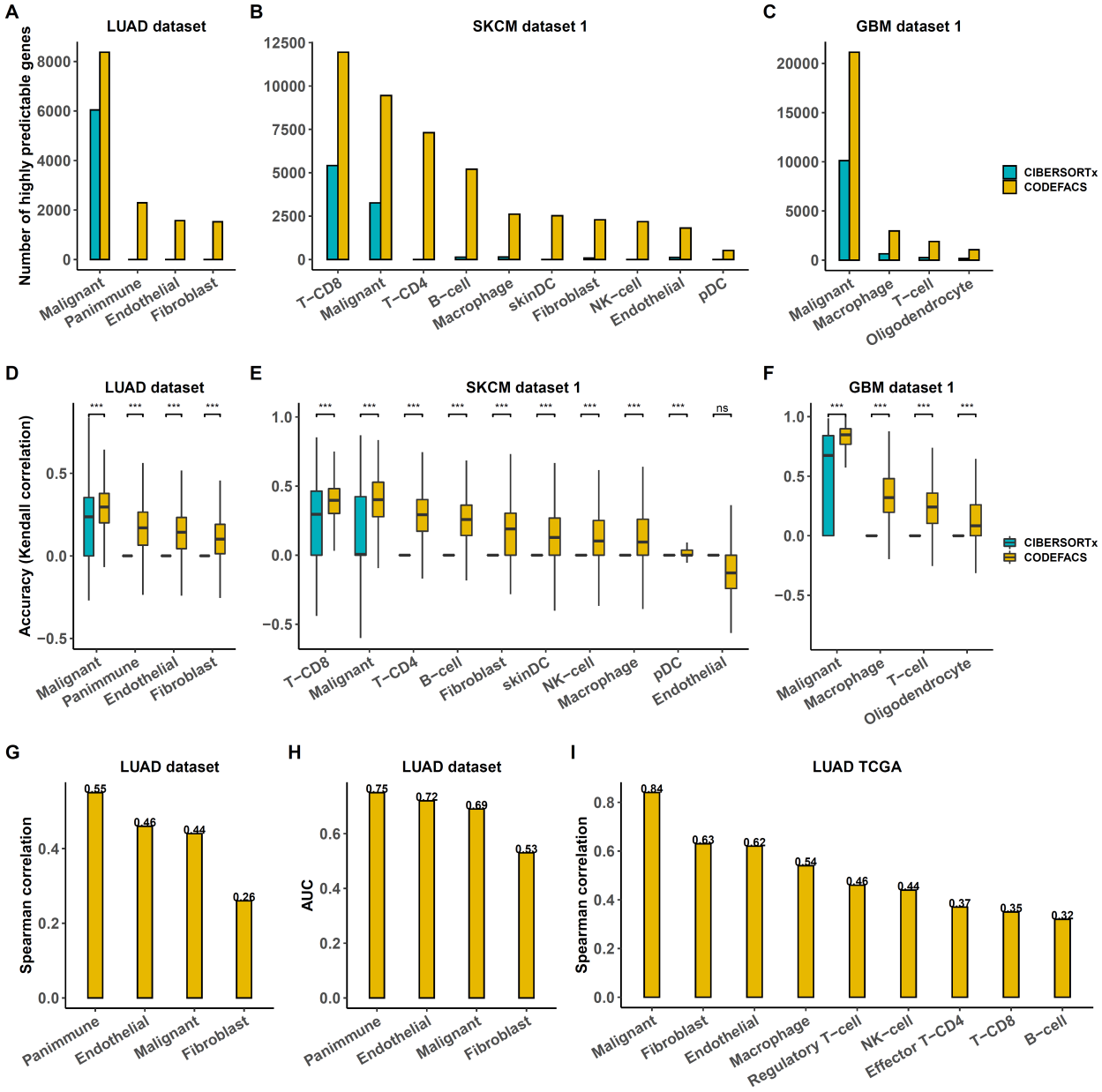

A
LUAD dataset
Number of highly predictable genes
Malignant
Panimmune
Endothelial
Fibroblast
B
SKCM dataset 1
T-CD8
Malignant
T-CD4
B-cell
Macrophage
skinDC
Fibroblast
NK-cell
Endothelial
pDC
C
GBM dataset 1
Malignant
Macrophage
T-cell
Oligodendrocyte
CIBERSORTx
CODEFACS
D
LUAD dataset
Accuracy (Kendall correlation)
E
SKCM dataset 1
F
GBM dataset 1
G
LUAD dataset
Spearman correlation
0.55
0.46
0.44
0.26
H
LUAD dataset
AUC
0.75
0.72
0.69
0.53
I
LUAD TCGA
0.84
0.63
0.62
0.54
0.46
0.44
0.37
0.35
0.32

**Figure 2. Evaluating the performance of CODEFACS. (A-C)** bar plots depicting the number of genes with a prediction accuracy (Kendall correlation) ≥ 0.3 with the ground truth for each cell type, estimated from bulk-generated samples of lung cancer (LUAD dataset; sample size = 26), melanoma (SKCM dataset 1; sample size = 28) and glioblastoma (GBM dataset 1; sample size = 24) benchmark datasets, as estimated by CODEFACS (yellow bars) and CIBERSORTx (blue bars). **(D-F)** boxplots depicting prediction accuracy distributions of all genes across different cell types in the lung cancer (LUAD with sample size 26), melanoma (SKCM with sample size 28) and glioblastoma (GBM with sample size 24) benchmark datasets, using CODEFACS (yellow) and CIBERSORTx (blue). A two-sided Wilcoxon signed rank test was performed to compare the prediction accuracies of CODEFACS and that of CIBERSORTx for each cell type in each dataset. *** denotes p-values < 2e-16. **(G)** Spearman correlations between prediction accuracies and confidence scores among cell types in the lung cancer benchmark dataset (LUAD dataset; sample size = 26). The y-axis indicates the spearman correlation coefficient value, while the x-axis indicates the cell type. **(H)** AUCs obtained in classifying informative and uninformative predictions among cell types in lung cancer benchmark dataset (LUAD dataset; sample size = 26). **(I)** bar plots depicting the Spearman correlations between mean deconvolved cell-type-specific expression in TCGA-LUAD and mean cell-type-specific expression derived from publicly available single cell datasets of LUAD. The y-axis indicates the Spearman correlation coefficient value, while the x-axis indicates the cell type. Details of all single cell RNA-Seq based pseudo bulk datasets can be found in supplementary table 6.

## Tumors with DNA mismatch repair deficiency have heightened T cell co-stimulation that is independent of their tumor mutation burden levels

In normal cells, DNA is constantly repaired in response to DNA damage or DNA replication errors[12]. However, defects in specific DNA repair pathways in cancer cells may result in the accumulation of many somatic mutations resulting in hypermutated tumors (TMB ≥ 10-20 mutation/Mb)[13–15]. One of the sources of hypermutability is a mismatch repair deficiency (MMRD), which leads to the accumulation of insertions and deletion mutations in microsatellite regions of the genome due to uncorrected DNA replication polymerase slippage events. This is known as microsatellite instability (MSI)[16,17]. Solid tumors with mismatch repair deficiency were shown to be sensitive to immune checkpoint

256 blockade (ICB) therapy irrespective of tumor type, leading the FDA to approve MSI as the
257 first cancer type agnostic biomarker for patients receiving anti-PD1 treatment[18]. The
258 reason behind this general sensitivity to anti-PD1 treatment is not completely understood.
259 Prior work has led to the prevailing hypothesis that elevated tumor mutation burden in
260 mismatch repair deficient tumors leads to more neoantigens, and thus is more likely to
261 activate a host immune response against tumor cells[16,19–21]. However, not all tumor types
262 with elevated tumor mutation burden have similar response rates to anti-PD1[22,23], and
263 recent studies have revealed that T cells recognize and respond to only a few neoantigens
264 per tumor[24–27]. More generally, when looking at non-synonymous tumor mutation burden,
265 MSI and survival data across the TCGA collection (**Fig 3A,B,** borrowed from[28,29]. See
266 Methods), we see a significant association between hypermutability and survival benefit
267 of patients in solid tumor types with a frequent underlying mismatch repair deficiency (**Fig.**
268 **3C.** log-rank test p-value = 0.00084), in contrast to other tumor types (**Fig 3D**, log-rank
269 test p-value = 0.4). These survival differences can be partially explained by the mutual
270 exclusivity of microsatellite instability and chromosomal instability, which has been
271 previously linked with a worse prognosis[28,30,31]. Taken together, these findings motivated
272 us to further study cellular immune crosstalk in the tumor microenvironment of mismatch
273 repair deficient tumors to gain additional cell-type-specific insights into their sensitivity to
274 anti-PD1.

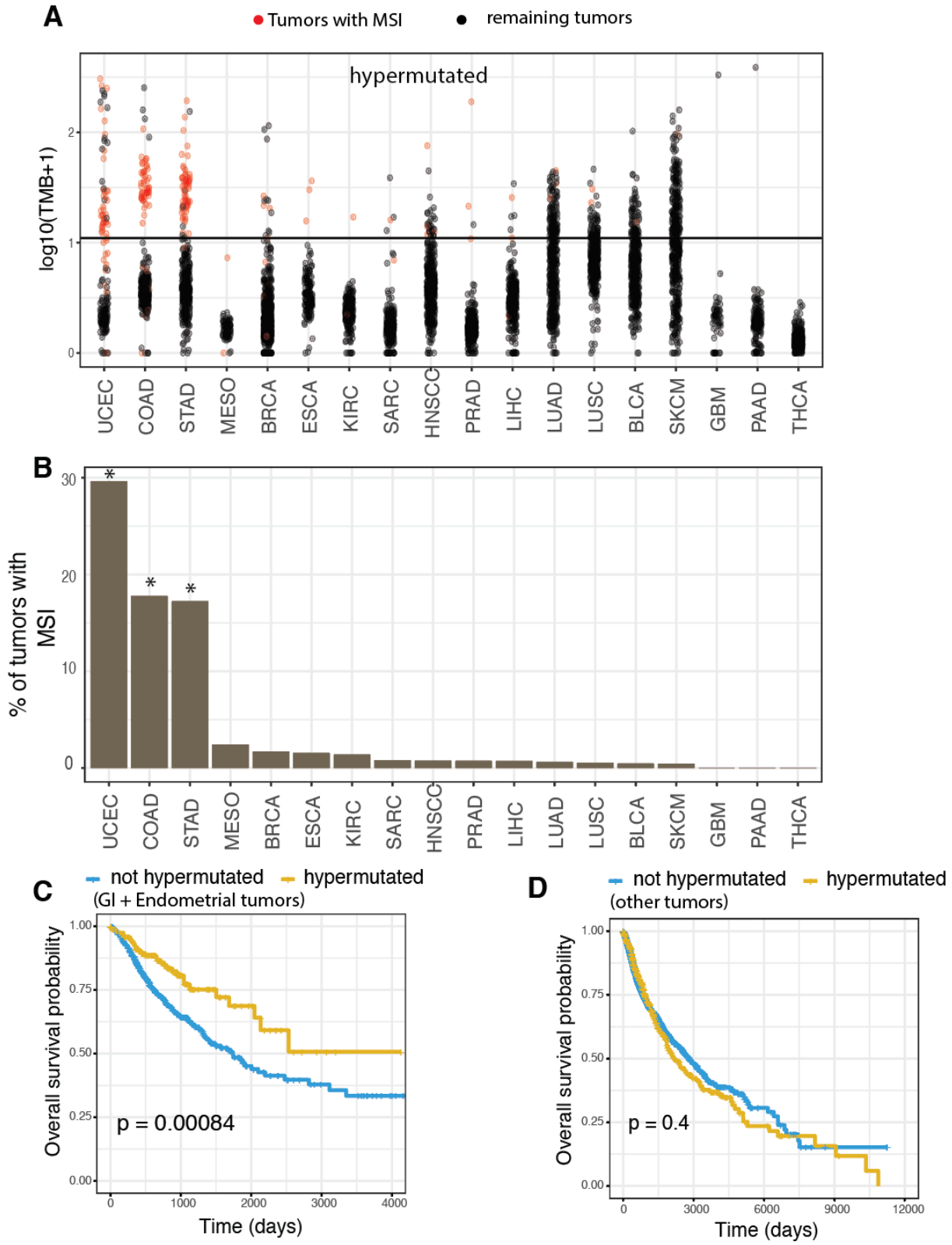
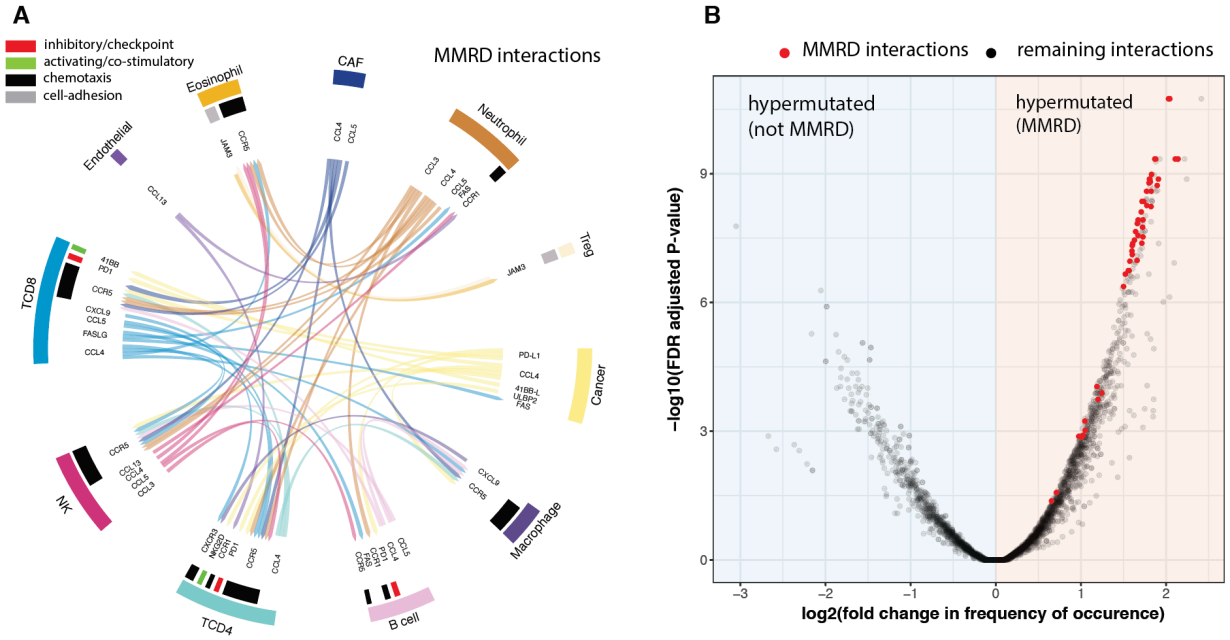



**Figure 3. (A)** The **landscape of tumor mutation burden and microsatellite instability across 18 different solid tumor types.** This panel plots the distribution of non-synonymous tumor

278 mutation burden on a logarithmic scale (Y-axis). All points above the horizontal line are typically
279 regarded as hyper-mutated tumors (> 10 mutations/Mb). All red points represent tumors with a
280 DNA mismatch repair deficiency detected via microsatellite instability (MSI). **(B) This panel**
281 **depicts the percentage of all tumor samples per cancer type with microsatellite instability**
282 **(Y-axis).** Tumor types marked with a * represent those where MSI is prevalent. **(C-D)**
283 **Comparison of overall survival of patients with tumors that are hypermutated vs not**
284 **hypermutated.** Left panel (**C**) In Gastro-Intestinal+Endometrial tumor types where MSI is
285 prevalent (marked with a * in panel B). Right panel (**D**) In other solid tumor types where tumors
286 rarely have an underlying mismatch repair deficiency. Statistical significance of differences in
287 survival was calculated using the log-rank test.
288
289 We hence aimed to identify cell-cell interactions that are differentially active
290 between microsatellite instable tumors (highlighted as red dots in Fig 3A) and
291 microsatellite stable tumors (highlighted as black dots in Fig 3A). To this end, we applied
292 CODEFACS to deconvolve the bulk gene expression of all solid tumors from TCGA and
293 integrated their predicted cell-type-specific gene expression levels with LIRICS. The top
294 50 interactions from this differential analysis (ordered by FDR adjusted p-value) are
295 shown in a network in **Fig. 4A**. These interactions are frequently active in mismatch repair
296 deficient tumors of distinct tumor types (**Supplementary Fig. S9**) and importantly, they
297 are more frequently active in hypermutated tumors with DNA mismatch repair deficiency
298 compared to other hypermutated tumors (**Fig. 4B, Fig. 3A**), testifying to their MSI
299 specificity. The top 50 MSI-specific interactions include the PDL1-PD1 checkpoint
300 interaction between tumor cells and CD8+ T-cells, but notably, T-cell activating/co-
301 stimulatory interactions such as the 41BBL-41BB interaction between Tumor cells and
302 CD8+ T cells, ULBP2-NKG2D between tumor cells and CD4+T cells, and chemotaxis
303 interactions involved in trafficking of lymphocytes in and around the tumor mass, such as
304 the CXCL9-CXCR3 chemokine interaction between macrophages and CD4+ T cells and
305 CCL3/4/5 – CCR5 interactions between various immune and stromal cell-types.
306
307 This shared heightened cellular crosstalk unique to the TME of mismatch repair
308 deficient tumors suggests that tumor infiltrating T cells can be activated by co-stimulatory
309 signals in the TME independent of overall tumor mutation burden, only to be kept in

310 balance by other immunoregulatory mechanisms such as the PD1-PDL1 checkpoint
311 interaction between CD8+ T cells and tumor cells. Our results indicate that when this
312 interaction is blocked by anti-PD1 treatment, the presence of other co-stimulatory
313 interactions can lead to the observed enhanced response of MMRD tumors to immune
314 checkpoint blockade therapy. This in turn raises the possibility that switching on specific
315 T cell co-stimulation signals in the TME may lead to better responses to anti-PD1
316 treatment independent of tumor mutation burden. Notably, recent pre-clinical studies have
317 shown that combination therapies aimed at enhancing such T-cell co-stimulating
318 interactions improve anti-tumor immune responses even in low TMB and highly immuno-
319 suppressive settings [32–37]. Currently, several clinical trials to assess the safety and
320 efficacy of these combinations are in progress [38–41].
321
322




323
324 **Figure 4. (A) Interaction network consisting of the top 50 interactions most highly activated**
325 **in TME of tumors with DNA mismatch repair deficiency.** Interactions highlighted in green
326 represent co-stimulatory interactions/having an activating effect on the target cell. Interactions
327 highlighted in red represent checkpoint interactions/having an inhibitory effect on the target cell.
328 Interactions highlighted in black represent pro-inflammatory/chemotaxis interactions involved in
329 inflammatory response and immune cell trafficking to tumor sites. Eos: Eosinophils, CAF: Cancer

330 associated fibroblasts. **(B) A volcano plot depicting on the x-axis the log2 fold change in the**
331 **frequency of occurrence of each cell-cell interaction in the TME of hypermutated tumors**
332 **with an underlying DNA mismatch repair deficiency vs other hypermutated tumors.** The y-
333 axis indicates the -log10 FDR adjusted p-value of the observed enrichment. Highlighted in red in
334 the scatter plot are the top 50 interactions that are most differentially active between all MSI vs
335 non-MSI tumors (shown in panel A).
336

337 **Machine learning guided discovery of cell-type-specific ligand-receptor**
338 **interactions that predict patient response to ICB therapy**

339 Given the small number of large publicly available transcriptomics datasets of patients
340 receiving immune checkpoint blockade therapy, we asked if we can effectively utilize this
341 large resource of deconvolved TCGA data we generated to transfer-learn cell-cell
342 interactions robustly predictive of response to immune checkpoint blockade therapy.
343 Specifically, since some mutations during tumor evolution can be immunogenic, we
344 hypothesized that one could potentially discover cell-type-specific ligand-receptor
345 interactions predictive of response to ICB therapy by a joint analysis of mutation and
346 deconvolved expression data from the TCGA. We focus on melanoma, currently the
347 tumor type best responding to ICB, where there are many independent publicly available
348 bulk expression datasets of patient's receiving anti-PD1 treatment and single-cell derived
349 cell-type-specific signatures, which serve as priors for the deconvolution of these bulk
350 datasets. Starting from the deconvolved TCGA-SKCM dataset as our training set (N=469),
351 we employed a genetic algorithm to find cell-type specific ligand-receptor interactions,
352 whose activation state best separates hypermutated melanoma tumors from non-
353 hypermutated tumors (**Fig 3A)**, assuming that the former group captures more samples
354 with some immunogenic mutations than the latter (see Methods, **Fig 5A**). We term the
355 interactions identified in this process *melanoma mutation specific functional interactions*
356 (*MSFI*)*,* and the network formed by these interactions is displayed in **Fig. 5B***.*
357

358 Having identified the MSFI interactions, we applied CODEFACS to deconvolve
359 the bulk expression data of pre-treatment samples from the three largest publicly
360 available melanoma datasets where patients received anti-PD1 treatment (either
361 monotherapy or in combination with anti-CTLA4; Methods) [42–44]. We then employed

362 LIRICS to the respective deconvolved expression of each of these checkpoint datasets,
363 *without any additional training*, and simply quantified the number of MSFI interactions that
364 are active in each of these patients' tumor samples, which we denote as the tumor's *MSFI*
365 *score*. Remarkably, we find that the MSFI score of each sample can robustly stratify
366 patients into those that are likely to respond to ICB vs those that are unlikely to respond
367 (**Fig. 5C**, progression free survival log rank test p-value: 0.00057, **Fig. 5D**, overall survival
368 log rank test p-value: 0.0031). **Supplementary Fig. S10** depicts the survival differences
369 for the two treatment groups separately (anti-PD1 monotherapy and anti-CTLA4 + anti-
370 PD1 combination). Additionally, **Supplementary Fig. S11** and **S12** depict the survival
371 differences for each ICB dataset separately. As evident, our results improve over recent
372 bulk gene expression based predictors of melanoma ICB therapy response (IMPRES [45],
373 TIDE [46] and the melanocytic plasticity signature (MPS) scores [47]). We note that the
374 performance levels of the latter on bulk expression datasets, where their original RNAseq
375 reads have been uniformly aligned and normalized as described in the Methods, is lower
376 than that reported in the original publications, pointing to the potential sensitivity of
377 expression-based predictors to the processing used and the need to do that in a uniform,
378 generally accepted manner (see Discussion).
379
380 To further evaluate the predictive performance of the MSFI score, we tested its
381 ability to predict partial or complete responders vs stable or progressive disease patients
382 in these datasets. To this end we plotted the receiver operator area under the curve (AUC)
383 obtained using it for classifying the patients to partial or complete responders vs stable or
384 progressive disease, and compared its performance to that obtained with the three other
385 published predictors for the different treatment groups (**Fig. 5E**). On average, the MSFI
386 score achieves an AUC of 0.63 (for anti-PD1 monotherapy the AUCs obtained are 0.6,
387 0.77 and 0.52 for the three individual ICB datasets, and for the combination ICB treatment
388 its 0.77, 0.49 and 0.63). A similar performance could not be achieved if the placement of
389 the ligand and receptor between interacting cell-types in the MSFI network was swapped
390 (average AUC 0.58) or by randomly shuffling the interaction activity profiles (average AUC
391 ~ 0.52), testifying that the selected cell-cell interactions are best predictive of response to
392 ICB. Comparing MSFI predictive performance in this response classification task to that

393 of the recent melanoma bulk expression-based predictors, TIDE achieves an average
394 AUC of 0.6 (for anti-PD1 monotherapy the AUCS are 0.46, 0.66 and 0.45 for the three
395 ICB datasets and 0.89, 0.55 and 0.61 for the combination), IMPRES achieves an average
396 AUC of 0.55 (for anti-PD1 monotherapy the AUCS are 0.59, 0.64 and 0.6 respectively for
397 the three ICB datasets and 0.61, 0.44 and 0.42 for the combination) and MPS achieves
398 an average AUC of 0.51 (for anti-PD1 monotherapy the AUCS are 0.61, 0.49 and 0.63
399 respectively for the three ICB datasets and 0.41, 0.59 and 0.38 for the combination).
400
401 Examining the MSFI network leads to interesting insights **(Fig. 5B)**. First, we find
402 an over-representation of cell-type-specific co-stimulatory/immune cell activating
403 interactions known from prior immunological literature (hypergeometric test p-value <
404 0.05)[48–58] (See Supplementary Table 1). Second, the MSFI network additionally includes
405 cytokine/chemokine interactions involved in pro-inflammatory response and the trafficking
406 of NK, T and B cells to the tumor site (responsible for better lymphocyte infiltration into
407 the tumor mass). Importantly, on fitting a multivariate Cox-proportional hazards model
408 with MSFI scores and TMB of each patient receiving anti-PD1 (wherever TMB data was
409 available), we see that a high MSFI score is significantly associated with improved
410 progression free survival (p-value: 0.013) and overall survival (p-value: 0.0258), whereas
411 high TMB is not (PFS p-value: 0.224, OS p-value: 0.477). These results further support
412 the findings from the TCGA analysis that heightened co-stimulatory and pro-inflammatory
413 signals in the TME can mobilize tumor infiltrating lymphocytes to generate an effective
414 host immune response upon immune checkpoint blockade independent of TMB.
415

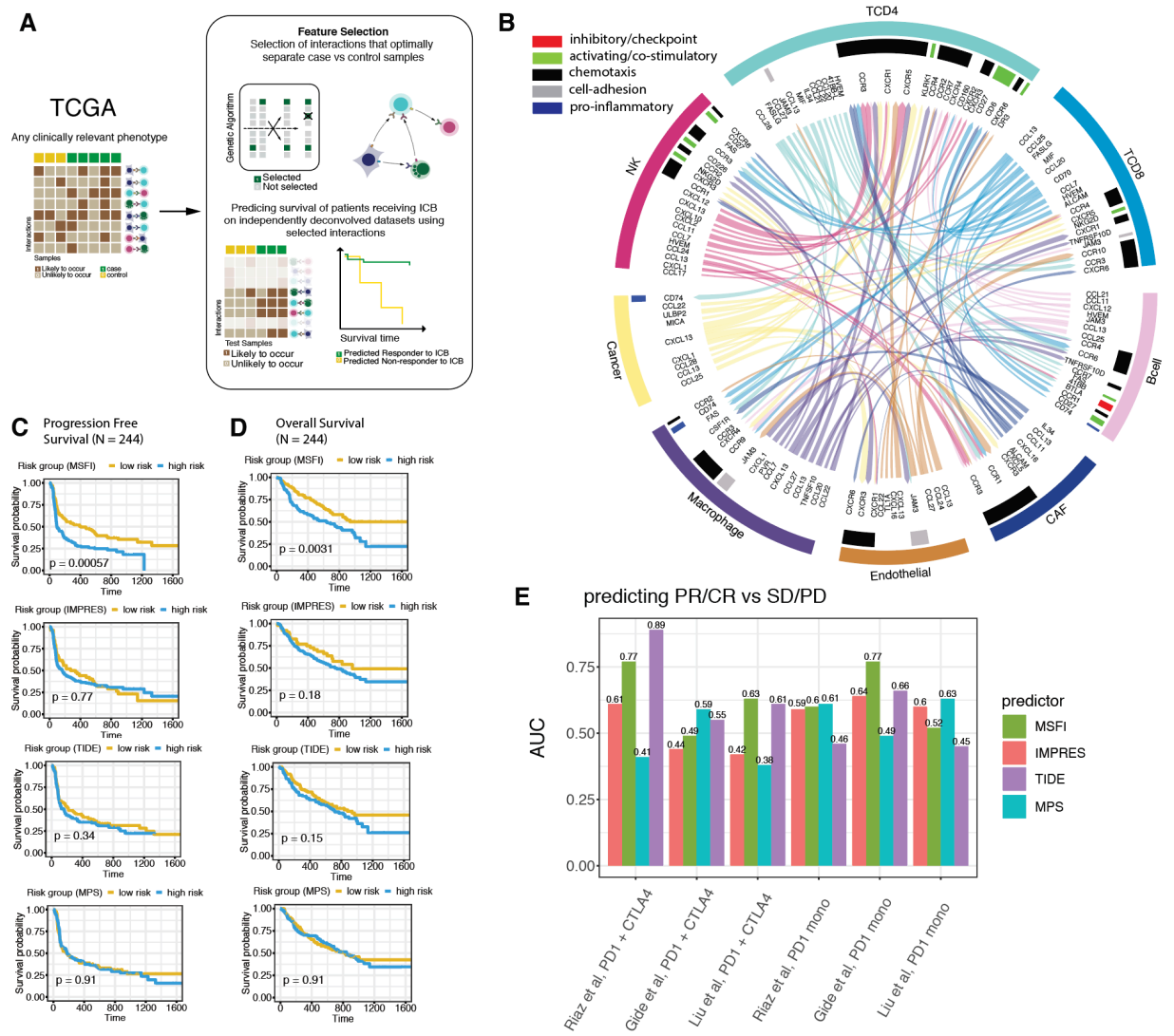



**Figure 5. (A) Overview of the machine learning analysis employed to identify cell type specific interactions that are predictive of response to immune checkpoint blockade therapy. (B) A chord diagram of the resulting MSFI network.** Each individual interaction is represented by a link from the source cell type (ligand expressing cell type) to the target cell type (receptor expressing cell type) and the color of the link represents the color of the source cell type. For interactions that are activating/co-stimulatory, the sector in the corresponding target cell type is highlighted in green. For inhibitory/checkpoint interactions, the sector in the target cell type is in highlighted red. Interactions involved in chemotaxis are highlighted in black and those I

427 mediating a pro-inflammatory response are highlighted in blue, cell-adhesion interactions are
428 highlighted in grey. **(C) Kaplan-Meier plot depicting the progression free survival of the**
429 **combined set of melanoma patients receiving immune checkpoint blockade (N= 244).** On
430 the top, the patients are stratified into low-risk/high-risk groups based on the median value of
431 MSFI score from LIRICS. Second from top, patients stratified into low/high risk groups based on
432 median IMPRES score[45]. Third from top, patients stratified into low/high risk groups based on
433 median TIDE score[46], Bottom, patients stratified into low/high risk groups based on median MPS
434 score[47]. **(D) Kaplan-Meier plots depicting the overall survival of all melanoma patients**
435 **receiving immune checkpoint blockade (N= 244).** On the top, the patients are stratified into
436 low-risk/high-risk groups based on the median value of MSFI score from LIRICS. Second from
437 top, patients stratified into low/high risk groups based on median IMPRES score[45]. Third from top,
438 patients stratified into low/high risk groups based on median TIDE score[46], Bottom, patients
439 stratified into low/high risk groups based on median MPS score[47]. Survival differences among
440 patients that received anti-PD1 monotherapy vs anti-CTLA4 + anti-PD1 combination are shown
441 in supplementary figure 10 **(E) Area under the ROC curves in predicting Complete/Partial-**
442 **response (based on RECIST v1.1) to immune checkpoint blockade therapy for the different**
443 **scores.** X-axis marks patients grouped by dataset source and treatment regimen. PD1 mono
444 represents patients that received anti-PD1 monotherapy. PD1 + CTLA4 represents patients that
445 additionally received anti-CTLA4 besides anti-PD1.
446

447 **Discussion**

448 This study presents a computational tool, CODEFACS and a pipeline, LIRICS, that
449 enables an (averaged) 'virtual single cell' characterization of the TME from bulk tumor
450 expression data. Applying these tools, we identify cell type specific ligand-receptors
451 interactions that are active and functionally important within individual tumor
452 microenvironments, in modifying patients' survival and response to ICB. Applying
453 CODEFACS to ~8000 tumors from TCGA, we estimate the cell-type-specific gene
454 expression profiles of each individual tumor sample thus enabling the analysis of the
455 TCGA at a cell-type-specific resolution. Integrating these data with LIRICS, we
456 systematically characterized the immune cellular crosstalk of the tumor
457 microenvironments of different tumor types. We identified a shared core of intercellular
458 TME interactions in DNA mismatch repair deficient tumors, which are associated with
459 improved patient survival and high sensitivity to immune checkpoint blockade therapy.

460 One potentially interesting implication of these findings is that immunomodulators
461 enhancing T-cell co-stimulation (e.g, via the 41BB receptor) might improve patients'
462 response to ICB irrespective of their tumor mutation burden. Finally, focusing on
463 melanoma, we show that one can bootstrap on the large deconvolved data resource from
464 TCGA using machine learning techniques to discover cell-cell interactions within the TME
465 that successfully predict patients' response to immune checkpoint blockade.
466
467 While we have provided a toolkit to discover clinically relevant cell-cell interactions
468 from bulk tumor expression, there are limitations that should be noted and potentially
469 further improved upon in the future. First, several recent publications have reported
470 discrepancies between different RNA-seq expression quantification methods based on
471 the reference transcriptome version used and choice of method (alignment based vs
472 alignment free)[59–63]. This can potentially affect reproducibility of findings. Hence,
473 whenever possible, we recommend that all bulk RNA-seq datasets are homogeneously
474 pre-processed using the same RNA-seq quantification method and reference
475 transcriptome before the application of our tools (See Methods), and any clinically
476 relevant findings are further followed up. Indeed, following this notion, in this study we
477 have pre-processed all bulk RNA-seq datasets in a uniform manner, using STAR (v2.7.6a)
478 + RSEM (v1.3.3) and GENCODE v23 human genome annotation (See Methods for more
479 details).
480
481 Second, regarding limitations of CODEFACS itself (1) it requires prior information
482 about the cell type composition of the input tumors, or alternatively, knowledge of the
483 pertaining cell-types' gene expression or methylation signatures that can be used to infer
484 their abundances, and its accuracy depends on the accuracy of the latter. (2) its prediction
485 power is limited to subsets of the whole exome genes, and its performance deteriorates
486 for lowly-abundant cell types. However, the confidence scores provided partially alleviate
487 this limitation, allowing the user to rank genes in each cell-type by the quality of
488 predictions of the expression of each gene in a given cell-type. Third, regarding LIRICS,
489 it currently restricted to well defined ligand-receptor interactions between tumor, immune,
490 stromal and epithelial cell types and does not consider the spatial localization of cells in

the TME. The inclusion of the latter with the advent of forthcoming spatial transcriptomics data is likely to lead to considerably more informative interaction inference approaches.

Although this work focuses on studying the tumor microenvironment, the tools presented here can be applied to discover important cell-cell interactions in noncancerous tissues under a variety of normal and disease states. One interesting application that we envision is the characterization of clinically relevant intercellular interactions occurring at the maternal-fetal interface using corresponding bulk gene expression data and pregnancy outcome information, whose elucidation may help treat and mitigate preeclampsia and other pregnancy related complications. One can also use our tools to study bulk gene expression data from pre-malignant tissue samples and compare them against malignant samples to elucidate cell-cell interaction dynamics on the way malignancy. Finally, one can use our tools to deconvolve expression data from autoimmune disorders to learn more about the underlying immune interactions. In sum, the computational tools developed and presented in this study offer a cost-effective way to study immune responses at a cell type specific resolution from the widely abundant bulk gene expression in both health and disease, complementing first-line single cell technologies in situations where the latter are less applicable.

## Data availability

All deconvolved expression data generated in this study will be made publicly available upon publication.

## Code availability

The tools and all codes for reproducing the figures of this study will be made publicly available upon publication.

## Methods

### Data collection and pre-processing

**Single cell RNA-seq datasets:** To benchmark the performance of CODEFACS, we first set out to obtain publicly available single cell RNA-seq datasets where both tumor and

522 non-tumor cells were successfully isolated. This search led us to the identification of nine
523 such single cell RNA-seq datasets from the literature, each from a different cancer type.
524 Collection of additional single cell datasets was frozen after Dec 2019. Details of the
525 single cell datasets that we curated for this study in **Supplementary Table 4**. For each
526 dataset sequenced on the SmartSeq2 platform, the log normalized transcript counts for
527 each gene in each sequenced cell were made publicly available by the original authors.
528 For the application of deconvolution, these counts were transformed back to the
529 Transcripts Per Million (TPM) scale. For datasets sequenced on the 10x platform, UMI
530 counts for each gene were made publicly available and were scaled by the library size of
531 each cell and multiplied by a factor of 1 million to get expression values in TPM scale.
532
533 **Bulk RNA-seq datasets:** Gene expression and matching bulk tumor methylation data
534 from fresh frozen tumor biopsies in TCGA were downloaded from (http://xena.ucsc.edu/)
535 [64]. In addition, publicly available bulk expression data from formalin fixed paraffin
536 embedded tumor biopsies of melanoma patients receiving immune checkpoint blockade
537 treatment were downloaded from[42–44]. All bulk RNA-seq datasets were collected such that
538 they have a sufficiently large sample size to reliably perform complete deconvolution of
539 expression profiles (> 4 times the number of cell-types involved) [7]. Collection of datasets
540 was frozen after Dec 2019. Details on bulk RNA-seq datasets deconvolved in this study
541 can be found in **Supplementary Table 5**. To maintain consistency with the pipeline used
542 for preprocessing TCGA data, bulk gene expression levels in immune checkpoint
543 blockade datasets were re-quantified using STAR v2.7.6a and RSEM v1.3.3[65] with
544 GENCODE v23 human genome annotation[66]. Furthermore, to mitigate technical biases,
545 between-sample scaling factors were estimated using TMM method implemented in
546 edgeR[67] and TPM values in each sample were further rescaled by these scaling factors[68].
547
548 **Generation of simulated bulk RNA-seq datasets:** To evaluate the performance of
549 CODEFACS, we generated 14 different pseudo-bulk RNA-seq datasets from mixing
550 experiments with single cell data. Each sample in each benchmark dataset has matching
551 cell type specific gene expression profiles derived from averaging single cell RNA-seq
552 profiles of individual cells from the same sample and same cell type. These profiles serve

553 as the ground truth for the evaluation of deconvolution performance. To avoid any
554 circularity in our validations, for each of the single cell datasets involved, single cell data
555 from 4 randomly chosen patients were separated from the rest. These data were used to
556 derive reference gene expression signatures for each cell type. The mixing experiments
557 were then performed on single cell data of the remaining patients that were hidden from
558 the reference signature derivation process. In addition, we simulated technical replicates
559 for each pseudo-bulk sample, wherein we injected noise in the pseudo-bulk expression
560 of a few randomly chosen genes and then renormalized the expression data by the
561 sample library size. This procedure simulates mRNA composition noise that is commonly
562 observed in bulk RNA-seq datasets due to technical differences in sample preparation
563 [64-66]. The datasets and the mixing experiment used to generate the pseudo-bulk
564 samples are listed in **Supplementary Table 6**. In addition, we obtained a FACS sorted
565 lung cancer dataset which include purified RNA-seq for four cell types [10] and generated
566 a pseudo bulk correspondingly. In total, 15 benchmark datasets were generated.
567

568 **Deconvolving immune checkpoint blockade (ICB) and TCGA bulk RNA-seq**
569 **datasets using CODEFACS**

570 For the application of CODEFACS, molecular profiles of signature genes of each cell type
571 of interest are needed to estimate the relative cell fractions in the bulk. We used single
572 cell expression derived signatures as priors to deconvolve the melanoma ICB datasets.
573 To derive these signatures from single cell data, we first start out by obtaining the class
574 labels of each cell type of interest. These data are publicly available for each single cell
575 dataset we collected. Hence, we primarily use these labels in our study (unless further
576 refinement of labels into specific cell subtypes of interest is needed for a specific usage).
577 With a collection of single cell expression profiles and matching cell type labels as input,
578 we used CIBERSORT online tool to derive a cell-type-specific signature matrix.
579 Thereafter, we applied CODEFACS to ICB datasets with default parameters settings and
580 batch correction requirement specified. For TCGA deconvolution, we first estimated cell
581 fractions based on bulk methylation and then applied CODEFACS to corresponding bulk
582 gene expression for the 21 cancer types which have both types of data available. We
583 chose methylation signatures over expression-based signatures for TCGA analysis for

584 two reasons. First, single cell expression data with consistent cell types across 21 cancer
585 types are not available. Second, DNA methylation-based signatures are considered to be
586 more stable marks of cellular identity compared to dynamic RNA expression derived
587 signatures [69]. The methylation-based cell type signatures (**Supplementary Table 7**)
588 were obtained from MethylCIBERSORT [11]. We applied CODEFACS to TCGA datasets
589 with default parameters settings and without batch correction requirement specified. For
590 the details of CODEFACS algorithm and parameter settings, see Supplementary note 1.
591
592 **Machine learning for discovery of cell-cell interactions predictive of clinical**
593 **responses to immune checkpoint blockade in melanoma**
594 We used a genetic algorithm, which is a randomized heuristic search algorithm designed
595 to select optimal features for a prediction task given some user-defined fitness function
596 for training [70]. In this setting, the features are ligand-receptor interactions between cell-
597 types, the prediction task is predicting response to ICB treatment and the fitness function
598 is defined as the accuracy of predicting a user defined phenotype based on the total
599 number of interactions from those selected occurring in a given sample; accuracy is
600 quantified by the AUC. To aid in faster convergence of the genetic algorithm during
601 training, the size of the search space is reduced by first assessing the fold change of each
602 feature between the two classes specified in the training dataset and incorporated into
603 the fitness function (e.g., hypermutated vs non-hypermutated). After this step, only
604 features with a fold change >1 is passed to the genetic algorithm for further optimization.
605
606 The algorithm starts out by randomly generating sets of features. This is defined
607 as the seed population. These sets iteratively evolve via the phenomenon of natural
608 selection enforced by the user defined fitness function. Specifically, for each subsequent
609 iteration, features from the best performing sets, as determined by the user defined fitness
610 function, in the current iteration are mixed at random followed by random new feature
611 additions or dropouts (referred to as mutations) to build a new generation of feature sets
612 and the process repeats. Eventually, after a number of epochs, which we set to 100, the
613 fitness function converges to an optimum and the best set of features for the prediction
614 task is returned to the user. Since the fitness function landscape is often non-convex and

615 the training process is stochastic, we repeat the training process 500 times, each with a
616 randomly chosen seed population, and eventually choose frequently selected features
617 over all solutions to reach a solution we suspect is close to the global optimum solution.
618 For our plots, we set the threshold to frequency > 100 times. The probability of any feature
619 being selected more than 100 times by random chance based on this approach is
620 estimated to be < 0.01. Results are qualitatively similar for more stringent thresholds. The
621 genetic algorithm was implemented in R using the genalg package [71]
622

623 **Acknowledgements**
624 This research is supported in part by the Intramural Research Program of the National
625 Institutes of Health, National Cancer Institute, Center for Cancer Research. This work
626 utilized the computational resources of the NIH HPC Biowulf cluster. We would
627 additionally like to acknowledge Dr. Noam Auslander, Dr. Chi-Ping Day and other
628 members of the Cancer Data Science Laboratory for their helpful feedback on this work
629

630 **References:**

631 1. Paget S. the Distribution of Secondary Growths in Cancer of the Breast. *Lancet*.
632 1889;133(3421):571-573. doi:10.1016/S0140-6736(00)49915-0
633 2. Smyth MJ, Ngiow SF, Ribas A, Teng MWL. Combination cancer immunotherapies
634 tailored to the tumour microenvironment. *Nat Rev Clin Oncol*. 2016;13(3):143-158.
635 doi:10.1038/nrclinonc.2015.209
636 3. Wagner A, Regev A, Yosef N. Revealing the vectors of cellular identity with single-cell
637 genomics. *Nat Biotechnol*. 2016;34(11):1145-1160. doi:10.1038/nbt.3711
638 4. Wang Z, Cao S, Morris JS, et al. Transcriptome Deconvolution of Heterogeneous Tumor
639 Samples with Immune Infiltration. *iScience*. 2018;9:451-460.
640 doi:10.1016/j.isci.2018.10.028
641 5. Quon G, Haider S, Deshwar AG, Cui A, Boutros PC, Morris Q. Computational purification
642 of individual tumor gene expression profiles leads to significant improvements in
643 prognostic prediction. *Genome Med*. 2013;5(3). doi:10.1186/gm433
644 6. Fox NS, Haider S, Harris AL, Boutros PC. Landscape of transcriptomic interactions
645 between breast cancer and its microenvironment. *Nat Commun*. 2019;10(1):1-13.
646 doi:10.1038/s41467-019-10929-z
647 7. Newman AM, Steen CB, Liu CL, et al. Determining cell type abundance and expression

648 from bulk tissues with digital cytometry. *Nat Biotechnol*. 2019;37(7):773-782.
649 doi:10.1038/s41587-019-0114-2
650 8. Jerby-Arnon L, Shah P, Cuoco MS, et al. A Cancer Cell Program Promotes T Cell
651 Exclusion and Resistance to Checkpoint Blockade. *Cell*. 2018;175(4):984-997.e24.
652 doi:10.1016/j.cell.2018.09.006
653 9. Ghoshdastider U, Naeini MM, Rohatgi N, et al. Data-driven inference of crosstalk in the
654 tumor microenvironment. *bioRxiv*. Published online January 1, 2019:835512.
655 doi:10.1101/835512
656 10. Gentles AJ, Hui ABY, Feng W, et al. A human lung tumor microenvironment interactome
657 identifies clinically relevant cell-type cross-talk. *Genome Biol*. 2020;21(1):107.
658 doi:10.1186/s13059-020-02019-x
659 11. Chakravarthy A, Furness A, Joshi K, et al. Pan-cancer deconvolution of tumour
660 composition using DNA methylation. *Nat Commun*. 2018;9(1):1-13. doi:10.1038/s41467-
661 018-05570-1
662 12. Drake JW, Charlesworth B, Charlesworth D, Crow JF. Rates of spontaneous mutation.
663 *Genetics*. 1998;148(4):1667-1686. doi:10.1007/978-94-011-4385-1_17
664 13. Loeb LA. Mutator Phenotype May Be Required for Multistage Carcinogenesis. *Cancer*
665 *Res*. 1991;51(12):3075-3079.
666 14. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in
667 human cancer. *Nature*. 2013;500(7463):415-421. doi:10.1038/nature12477
668 15. Muzny DM, Bainbridge MN, Chang K, et al. Comprehensive molecular characterization of
669 human colon and rectal cancer. *Nature*. 2012;487(7407):330-337.
670 doi:10.1038/nature11252
671 16. Funkhouser WK, Lubin IM, Monzon FA, et al. Relevance, Pathogenesis, and testing
672 algorithm for mismatch repair-defective colorectal carcinomas: A report of the association
673 for molecular pathology. *J Mol Diagnostics*. 2012;14(2):91-103.
674 doi:10.1016/j.jmoldx.2011.11.001
675 17. Cortes-Ciriano I, Lee S, Park WY, Kim TM, Park PJ. A molecular portrait of microsatellite
676 instability across multiple cancers. *Nat Commun*. 2017;8. doi:10.1038/ncomms15180
677 18. Boyiadzis MM, Kirkwood JM, Marshall JL, Pritchard CC, Azad NS, Gulley JL. Significance
678 and implications of FDA approval of pembrolizumab for biomarker-defined disease. *J*
679 *Immunother Cancer*. 2018;6(1). doi:10.1186/s40425-018-0342-x
680 19. Galon J, Costes A, Sanchez-Cabo F, et al. Type, density, and location of immune cells
681 within human colorectal tumors predict clinical outcome. *Science (80- )*.

682 2006;313(5795):1960-1964. doi:10.1126/science.1129139
683 20. Llosa NJ, Cruise M, Tam A, et al. The vigorous immune microenvironment of
684 microsatellite instable colon cancer is balanced by multiple counter-inhibitory
685 checkpoints. *Cancer Discov*. 2015;5(1):43-51. doi:10.1158/2159-8290.CD-14-0863
686 21. Le DT, Uram JN, Wang H, et al. PD-1 blockade in tumors with mismatch-repair
687 deficiency. *N Engl J Med*. 2015;372(26):2509-2520. doi:10.1056/NEJMoa1500596
688 22. Yarchoan M, Hopkins A, Jaffee EM. Tumor Mutational Burden and Response Rate to PD-
689 1 Inhibition. *N Engl J Med*. Published online 2017. doi:10.1056/nejmc1713444
690 23. Imakaev M. Limited evidence of tumour mutational burden as a biomarker of response to
691 immunotherapy. *bioRxiv*. Published online January 1, 2020:2020.09.03.260265.
692 doi:10.1101/2020.09.03.260265
693 24. Robbins PF, Lu YC, El-Gamil M, et al. Mining exomic sequencing data to identify mutated
694 antigens recognized by adoptively transferred tumor-reactive T cells. *Nat Med*.
695 2013;19(6):747-752. doi:10.1038/nm.3161
696 25. Bassani-Sternberg M, Bräunlein E, Klar R, et al. Direct identification of clinically relevant
697 neoepitopes presented on native human melanoma tissue by mass spectrometry. *Nat*
698 *Commun*. 2016;7. doi:10.1038/ncomms13404
699 26. Roh W, Chen PL, Reuben A, et al. Integrated molecular analysis of tumor biopsies on
700 sequential CTLA-4 and PD-1 blockade reveals markers of response and resistance. *Sci*
701 *Transl Med*. 2017;9(379). doi:10.1126/scitranslmed.aah3560
702 27. Kalaora S, Wolf Y, Feferman T, et al. Combined analysis of antigen presentation and T-
703 cell recognition reveals restricted immune responses in melanoma. *Cancer Discov*.
704 2018;8(11):1366-1375. doi:10.1158/2159-8290.CD-17-1418
705 28. Taylor AM, Shih J, Ha G, et al. Genomic and Functional Approaches to Understanding
706 Cancer Aneuploidy. *Cancer Cell*. 2018;33(4):676-689.e3. doi:10.1016/j.ccell.2018.03.007
707 29. Bonneville R, Krook MA, Kautto EA, et al. Landscape of Microsatellite Instability Across
708 39 Cancer Types. *JCO Precis Oncol*. 2017;(1):1-15. doi:10.1200/po.17.00073
709 30. Davoli T, Uno H, Wooten EC, Elledge SJ. Tumor aneuploidy correlates with markers of
710 immune evasion and with reduced response to immunotherapy. *Science (80- )*. Published
711 online 2017. doi:10.1126/science.aaf8399
712 31. Auslander N, Wolf YI, Koonin E V. Interplay between DNA damage repair and apoptosis
713 shapes cancer evolution through aneuploidy and microsatellite instability. *Nat Commun*.
714 2020;11(1). doi:10.1038/s41467-020-15094-2
715 32. Chen S, Lee LF, Fisher TS, et al. Combination of 4-1BB agonist and PD-1 antagonist

716 promotes antitumor effector/memory CD8 T cells in a Poorly Immunogenic Tumor Model.
717 *Cancer Immunol Res*. Published online 2015. doi:10.1158/2326-6066.CIR-14-0118
718 33. Ma HS, Poudel B, Torres ER, et al. A CD40 Agonist and PD-1 Antagonist Antibody
719 Reprogram the Microenvironment of Nonimmunogenic Tumors to Allow T-cell–Mediated
720 Anticancer Activity. *Cancer Immunol Res*. Published online 2019. doi:10.1158/2326-
721 6066.CIR-18-0061
722 34. Barber A. Costimulation of Effector CD8+ T Cells: Which Receptor is Optimal for
723 Immunotherapy? *MOJ Immunol*. Published online 2014.
724 doi:10.15406/moji.2014.01.00011
725 35. Melero I, Hirschhorn-Cymerman D, Morales-Kastresana A, Sanmamed MF, Wolchok JD.
726 Agonist antibodies to TNFR molecules that costimulate T and NK cells. *Clin Cancer Res*.
727 Published online 2013. doi:10.1158/1078-0432.CCR-12-2065
728 36. Chester C, Sanmamed MF, Wang J, Melero I. Immunotherapy targeting 4-1BB:
729 Mechanistic rationale, clinical results, and future strategies. *Blood*. Published online 2018.
730 doi:10.1182/blood-2017-06-741041
731 37. Philipson BI, O'Connor RS, May MJ, June CH, Albelda SM, Milone MC. 4-1BB
732 costimulation promotes CAR T cell survival through noncanonical NF-κB signaling. *Sci*
733 *Signal*. Published online 2020. doi:10.1126/scisignal.aay8248
734 38. Siu LL, Steeghs N, Meniawy T, et al. Preliminary results of a phase I/IIa study of BMS-
735 986156 (glucocorticoid-induced tumor necrosis factor receptor–related gene [GITR]
736 agonist), alone and in combination with nivolumab in pts with advanced solid tumors. *J*
737 *Clin Oncol*. Published online 2017. doi:10.1200/jco.2017.35.15_suppl.104
738 39. Tolcher AW, Sznol M, Hu-Lieskovan S, et al. Phase Ib study of utomilumab (PF-
739 05082566), a 4-1BB/CD137 agonist, in combination with pembrolizumab (MK-3475) in
740 patients with advanced solid tumors. *Clin Cancer Res*. Published online 2017.
741 doi:10.1158/1078-0432.CCR-17-1243
742 40. Cohen EEW, Pishvaian MJ, Shepard DR, et al. A phase Ib study of utomilumab (PF-
743 05082566) in combination with mogamulizumab in patients with advanced solid tumors. *J*
744 *Immunother Cancer*. Published online 2019. doi:10.1186/s40425-019-0815-6
745 41. Choi Y, Shi Y, Haymaker CL, Naing A, Ciliberto G, Hajjar J. T-cell agonists in cancer
746 immunotherapy. *J Immunother Cancer*. Published online 2020. doi:10.1136/jitc-2020-
747 000966
748 42. Riaz N, Havel JJ, Makarov V, et al. Tumor and Microenvironment Evolution during
749 Immunotherapy with Nivolumab. *Cell*. 2017;171(4):934-949.e15.

750 doi:10.1016/j.cell.2017.09.028
751 43. Gide TN, Quek C, Menzies AM, et al. Distinct Immune Cell Populations Define Response
752 to Anti-PD-1 Monotherapy and Anti-PD-1/Anti-CTLA-4 Combined Therapy. *Cancer Cell*.
753 2019;35(2):238-255.e6. doi:10.1016/j.ccell.2019.01.003
754 44. Liu D, Schilling B, Liu D, et al. Integrative molecular and clinical modeling of clinical
755 outcomes to PD1 blockade in patients with metastatic melanoma. *Nat Med*.
756 2019;25(12):1916-1927. doi:10.1038/s41591-019-0654-5
757 45. Auslander N, Zhang G, Lee JS, et al. Robust prediction of response to immune
758 checkpoint blockade therapy in metastatic melanoma. *Nat Med*. 2018;24(10):1545-1549.
759 doi:10.1038/s41591-018-0157-9
760 46. Jiang P, Gu S, Pan D, et al. Signatures of T cell dysfunction and exclusion predict cancer
761 immunotherapy response. *Nat Med*. 2018;24(10):1550-1558. doi:10.1038/s41591-018-
762 0136-1
763 47. Pérez-Guijarro E, Yang HH, Araya RE, et al. Multimodel preclinical platform predicts
764 clinical response of melanoma to immunotherapy. *Nat Med*. 2020;26(5):781-791.
765 doi:10.1038/s41591-020-0818-3
766 48. Billadeau DD, Leibson PJ. ITAMs versus ITIMs: striking a balance during cell regulation.
767 *J Clin Invest*. 2002;109(2):161-168. doi:10.1172/jci14843
768 49. Staub E, Rosenthal A, Hinzmann B. Systematic identification of immunoreceptor tyrosine-
769 based inhibitory motifs in the human proteome. *Cell Signal*. 2004;16(4):435-456.
770 doi:10.1016/j.cellsig.2003.08.013
771 50. Murphy K, Weaver C. *Janeway's Immunobiology. Ninth Edition*.; 2018.
772 doi:10.1086/696793
773 51. Chen L, Flies DB. Molecular mechanisms of T cell co-stimulation and co-inhibition. *Nat*
774 *Rev Immunol*. 2013;13(4):227-242. doi:10.1038/nri3405
775 52. Campbell KS, Purdy AK. Structure/function of human killer cell immunoglobulin-like
776 receptors: Lessons from polymorphisms, evolution, crystal structures and mutations.
777 *Immunology*. 2011;132(3):315-325. doi:10.1111/j.1365-2567.2010.03398.x
778 53. Pende D, Falco M, Vitale M, et al. Killer Ig-like receptors (KIRs): Their role in NK cell
779 modulation and developments leading to their clinical exploitation. *Front Immunol*.
780 2019;10(MAY). doi:10.3389/fimmu.2019.01179
781 54. Ward-Kavanagh LK, Lin WW, Šedý JR, Ware CF. The TNF Receptor Superfamily in Co-
782 stimulating and Co-inhibitory Responses. *Immunity*. 2016;44(5):1005-1019.
783 doi:10.1016/j.immuni.2016.04.019

784 55. Gonçalves CM, Henriques SN, Santos RF, Carmo AM. CD6, a rheostat-type signalosome
785 that tunes T cell activation. *Front Immunol*. 2018;9. doi:10.3389/fimmu.2018.02994
786 56. Steri M, Orru V, Idda ML, et al. Overexpression of the cytokine BAFF and autoimmunity
787 risk. *N Engl J Med*. 2017;376(17):1615-1626. doi:10.1056/NEJMoa1610528
788 57. Chen M, Lin X, Liu Y, et al. The function of BAFF on T helper cells in autoimmunity.
789 *Cytokine Growth Factor Rev*. 2014;25(3):301-305. doi:10.1016/j.cytogfr.2013.12.011
790 58. Varin MM, Le Pottier L, Youinou P, Saulep D, Mackay F, Pers JO. B-cell tolerance
791 breakdown in Sjögren's Syndrome: Focus on BAFF. *Autoimmun Rev*. 2010;9(9):604-608.
792 doi:10.1016/j.autrev.2010.05.006
793 59. Wu DC, Yao J, Ho KS, Lambowitz AM, Wilke CO. Limitations of alignment-free tools in
794 total RNA-seq quantification. *BMC Genomics*. 2018;19(1). doi:10.1186/s12864-018-4869-
795 5
796 60. Sahraeian SME, Mohiyuddin M, Sebra R, et al. Gaining comprehensive biological insight
797 into the transcriptome by performing a broad-spectrum RNA-seq analysis. *Nat Commun*.
798 2017;8(1). doi:10.1038/s41467-017-00050-4
799 61. Everaert C, Luypaert M, Maag JLV, et al. Benchmarking of RNA-sequencing analysis
800 workflows using whole-transcriptome RT-qPCR expression data. *Sci Rep*. 2017;7(1).
801 doi:10.1038/s41598-017-01617-3
802 62. Teng M, Love MI, Davis CA, et al. A benchmark for RNA-seq quantification pipelines.
803 *Genome Biol*. 2016;17(1). doi:10.1186/s13059-016-0940-1
804 63. Robert C, Watson M. Errors in RNA-Seq quantification affect genes of relevance to
805 human disease. *Genome Biol*. 2015;16(1). doi:10.1186/s13059-015-0734-x
806 64. Goldman MJ, Craft B, Hastie M, et al. Visualizing and interpreting cancer genomics data
807 via the Xena platform. *Nat Biotechnol*. 2020;38(6):675-678. doi:10.1038/s41587-020-
808 0546-8
809 65. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq
810 quantification. *Nat Biotechnol*. 2016;34(5):525-527. doi:10.1038/nbt.3519
811 66. Harrow J, Frankish A, Gonzalez JM, et al. GENCODE: The reference human genome
812 annotation for the ENCODE project. *Genome Res*. 2012;22(9):1760-1774.
813 doi:10.1101/gr.135350.111
814 67. Robinson MD, McCarthy DJ, Smyth GK. edgeR: A Bioconductor package for differential
815 expression analysis of digital gene expression data. *Bioinformatics*. 2009;26(1):139-140.
816 doi:10.1093/bioinformatics/btp616
817 68. Robinson MD, Oshlack A. A scaling normalization method for differential expression

818 analysis of RNA-seq data. *Genome Biol*. 2010;11(3):R25. doi:10.1186/gb-2010-11-3-r25
819 69. Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev*. 2002;16(1):6-21.
820 doi:10.1101/gad.947102
821 70. Melanie M. An introduction to genetic algorithms By Melanie Mitchell. MIT Press,
822 Cambridge, MA. (1996). 205 pages. $30.00. *Comput Math with Appl*. 1996;32(6):133.
823 doi:10.1016/S0898-1221(96)90227-8
824 71. Willighagen E, Ballings M, Ballings MM. Package 'genalg.' Published online 2015.
825