

1

1 **LncRNA antigens- a novel resource to improve immunotherapy efficacy**

2 **predictions in Melanoma**

3 **Authors**

4 Sumaira Malik, School of Mathematical and Statistical Sciences, University of

5 Galway, Ireland, s.malik3@nuigalway.ie

6 Aaron Golden, School of Natural Sciences, University of Galway, Ireland,

7 aaron.golden@universityofgalway.ie

8 Corresponding author: Sumaira Malik, s.malik3@nuigalway.ie

9 **Key Words**

10 Immunotherapy efficacy, Biomarker, LncRNA sORFs, Prognostic value, Antigen

11 load, Melanoma, Tumor immune response

12 **Abstract**

13 **Background:**

14 ICI (immune checkpoint inhibitor) therapy is one of the most promising treatments for
15 melanoma. ICI response however varies among patients, emphasizing the importance
16 of identifying genomic biomarkers to predict likely therapeutic efficacy in advance of
17 treatment. We hypothesised that a lncRNA based immunogenicity (lnc-IM) score
18 could be used to predict individual response to ICI treatment, and that this could
19 complement the existing criterion for ICI selection based on tumor mutation burden
20 (TMB).

21 **Methodology:**

22 The TCGA-SKCM (n=101) and the ICI treated UCLA (n=25), MSKCC (n=16) and
23 DFCI (n=40) melanoma cohorts were used in this study, involving both clinical and
24 transcriptomic data. Each patient was assigned an lnc-IM score based on the number
25 of lncRNA sORF derived peptides predicted to be presented by their tumor's MHC-I
26 genotype. For the ICI treated cohorts, a combined antigen score was defined as a sum
27 of neo-antigen load (derived from TMB) and lnc-IM score. A logistic regression-
28 based classifier was used to predict ICI responses based on these combined antigen
29 scores.

30 **Results:**

31 Survival analysis showed improved overall survival among patients with low lnc-IM
32 scores (HR= 0.39, p=0.009) in the TCGA-SKCM cohort. We also observed a negative
33 association between tumor immune cell concentration and lnc-IM scores, with low
34 lnc-IM groups showing higher anti-tumor immune cell concentrations . Using the ICI
35 treated cohorts, we demonstrated that a classifier based on combined antigen scoring
36 improved the prediction of immunotherapy outcomes as compared to using TMB
37 alone, yielding an area under the curve (AUC) of 0.71 with an accuracy of 0.54 and
38 recall of 1. We also demonstrated a reduced rate of false negatives (14%) by using a
39 combined antigen score as compared to the use of TMB alone (33%) in ICI treated
40 cohorts.

41 **Conclusion:** Our findings suggest that the use of combined antigen scores (using lnc-
42 IM scores along with TMB derived neoantigen load) have potential in improving
43 immunotherapy efficacy predictions. Prospective validation in larger cohort sizes is
44 warranted.

45 **KEY MESSAGES:**

46 **What is already known on this topic**

47 Previous studies have established actionable associations between TMB neoantigen
48 load and immunotherapy responses.

49 **What this study adds**

50 This study introduces lnc-IM scores as a novel metric that predicts patients antigen
51 load based on translatable lncRNAs expression. These lnc-IM scores when combined
52 with TMB associated neoantigen load indicate an improvement in immunotherapy
53 efficacy predictions.

54 **How this study might affect research, practice or policy**

55 Future research is needed to further validate lnc-IM scores as a predictive biomarker
56 for immunotherapy response in various cancer types. The use of lnc-IM scores can
57 empower clinicians to make more informed decisions about administering
58 immunotherapy treatments, improving patient outcomes.

59 INTRODUCTION

60 Self versus non-self discrimination by T cells is a hallmark of cancer evasion by the
61 immune response [1]. T cells eliminate cancer cells by recognizing tumor-specific
62 antigens presented on the tumor cell's surface by MHC-I molecules [2]. Immune and
63 tumor cells possess "checkpoint proteins" such as PD-1, CTLA4 and PDL1 that keep
64 such immune responses in check with the binding of such checkpoint proteins
65 restricting T cells from killing tumor cells. Immune checkpoint inhibitor therapy (ICI)
66 is a promising immunotherapy that restores the T cells' capacity to attack and
67 eliminate tumor cells by blocking such checkpoint proteins [3].

68 The use of ICI therapy has been a significant achievement in the last decade for
69 cancer treatment and has demonstrated clear improvements in the survival rate of
70 cancer patients. Its use to date has been approved for multiple cancer types, including
71 melanoma - one of the first cancers to be treated with this therapy [4]. Despite some
72 remarkable successes, response to ICI therapy varies widely among individual
73 patients, emphasizing the importance of identifying genomic biomarkers to try and
74 predict an individual's likely response in advance of treatment. Tumor mutation
75 burden (TMB), programmed cell death ligand 1 (PDL1), and mismatch repair defect
76 (dMMR)/microsatellites Instability (MSI) are some of the currently FDA-approved
77 predictive biomarkers of ICI efficacy, of which TMB is the most widely used [5,6].
78 High TMB has been shown to predict improved therapeutic efficacy to ICI therapy in
79 various cancer types [7]. High TMB is associated with a high neoantigen burden [7];
80 hence, ICI therapy for patients whose cancers have a high TMB would be expected to
81 elicit significant T cell responses against those antigens, resulting in enhanced tumor
82 cell attrition. Although many trials have demonstrated the usability of TMB in clinical

83 practices, a recent study has questioned the concept of the universal usage of TMB as
84 a predictor of ICI efficacy [8], arguing that TMB prediction power is only accurate to
85 a subset of patients, and using TMB as a sole predictor of ICI response might deprive
86 potential patients who might otherwise respond to ICI therapy .

87 TMB-associated antigens (neoantigens) are derived from non-synonymous missense
88 mutations and originate from the coding region of the genome that encodes proteins
89 [9]. However, TMB is not the only source of antigens. The Encyclopedia of DNA
90 Elements (ENCODE) revealed that the human genome contains almost 80% of non-
91 coding RNAs in contrast to the 1.5% that encodes for proteins [10]. Among non-
92 coding RNAs, those longer than 200 nucleotides in length - collectively known as
93 long non-coding RNAs or lncRNAs - are of great interest on account of their evident
94 wide functional diversity [11]. LncRNAs have been studied for their ability to control
95 regulatory and cellular processes. LncRNAs regulate gene expression both as miRNA
96 sponges and mRNA sponges [12]. LncRNAs act as transcription regulators by
97 modifying the chromatin complexes and can activate or repress gene expression.
98 Moreover, they also control the binding of transcription regulatory factors leading to
99 the activation of nearby genes [13].

100 LncRNAs have also been shown to be implicated in the creation of MHC-I associated
101 peptides with additional work demonstrating that a considerable number of tumor-
102 specific antigens originate from non-coding regions of the genome [14]. LncRNAs'
103 contribution to the cancer immunopeptidome is a topic of active research with a
104 particular focus being an assessment of whether such lncRNA-associated antigens are
105 associated with elevated cytotoxic T lymphocyte (CTL) responses. High CTL
106 responses have been associated with better survival and treatment outcomes [15].
107 Many lncRNAs possess intact short open reading frames (sORF) that can result in the

108 translation of short peptides in the dysregulated cancer transcriptome. Whilst short
109 peptides have been detected in mass-spectrometry based proteomic studies [16], direct
110 association of lncRNA sORFs with tumor immune microenvironment (TIM) has yet
111 to be fully characterised.

112 In this study we explore how the expression of lncRNA sORFs is associated with a
113 patient's likely response to ICI therapy. Specifically, we derive a novel metric, the lnc-
114 IM score, that estimates the level of sORF derived peptide presentation based on a
115 tumour's specific MHC-I genotype. In the first part of our study, we establish the
116 association of lnc-IM scores with the tumor immune microenvironment – determined
117 from known cellular biomarkers - and survival predictions in a melanoma cohort
118 (TCGA-SKCM) without any ICI treatment. In the second part, we utilize three ICI-
119 treated melanoma cohorts to examine how these lnc-IM scores can be used to improve
120 ICI efficacy predictions by their integration with each patient's TMB-associated
121 antigen count.

122 **METHODS**

123 **Data acquisition**

124 For this study, three different publicly available melanoma cohorts were utilized.
125 Clinical information and transcriptomic profiles (RNA-seq) of the TCGA-SKCM
126 cohort were downloaded from The Cancer Genome Atlas (TCGA, <https://portal.gdc.cancer.gov>). HLA typing of the TCGA-SKCM cohort was acquired from a
127 previous study [17]. A total of 101 patients were included in our analysis based on
128 the availability of MHC-I genotype and transcriptomic data and having basic clinical
129 information of age, gender, and overall survival. For ICI efficacy predictions, three
130 ICI-treated cohorts involving metastatic melanoma (UCLA) [18], melanoma

132 (MSKCC) [19] and metastatic melanoma (DFCI) [20] patient groups were utilized.
133 Clinical information, including overall survival, treatment response, TMB, neoantigen
134 load, and transcriptomic data, was retrieved from cBioPortal
135 (<https://www.cbioportal.org/>) [21]. Based on the criteria mentioned above, a total of
136 16 patients from the MSKCC cohort, 25 from the UCLA cohort and 40 from DFCI
137 cohort were selected for subsequent analysis. Tumor immune microenvironment
138 (TIM) analysis for the TCGA-SKCM cohort was performed using the xCell algorithm
139 [22], which estimates the the abundance of each immune cell type using expression
140 profiles of specific gene signatures for each cell type.

141 **Defining Lnc-IM scores**

142 A lncRNA-immunogenicity score (lnc-IM) for a patient was defined as the total
143 number of presentable lncRNA associated sORFs for that patient's tumor. In the first
144 step, all lncRNAs associated with short open reading frames (sORFs) were retrieved
145 from sORFs.org [23]. We selected our desired dataset from sORFs.org based on filters
146 (*species: humans and biotype = lncRNA*). This initial search resulted in 425 lncRNAs
147 and will be referred to as translatable lncRNAs in this work. These lncRNAs are
148 associated with ~ 3000 sORFs, as experimentally proven by different Riboseq
149 experiments [23]. Among these translatable lncRNAs, only overexpressed lncRNAs
150 were considered for subsequent analysis. The raw RNA-seq expression of translatable
151 lncRNAs was acquired from GDC (TCGA, <https://portal.gdc.cancer.gov>) and
152 cBioPortal (<https://www.cbioportal.org/>). Raw counts were converted to counts per
153 million (cpm) and log normalized using the edgeR [24] package in R. A translatable
154 lncRNA was considered for immunogenicity scoring if it passed the criteria of (log
155 cpm > 6) (Supplemental Figure 1). For each translatable lncRNA, all of its sORFs
156 were considered for lnc-IM scoring. The Patient Harmonic-mean Best Rank (PHBR)

157 score [17] was then assigned to each sORF, which is an estimate of its derived
158 peptide's MHC-I presentation likelihood. The PHBR score represents the harmonic
159 mean of best-ranked peptides (across a given patient's HLA alleles). A sORF with
160 $PHBR < 0.5$ was used to define a "presentable sORF". The resulting Inc-IM score
161 (Supplemental Figure 2) represents the total number of presentable sORFs in a given
162 patient's tumor.

163 **Combined antigen score**

164 For the ICI-treated cohorts, a combined antigen score was derived as a sum of each
165 patient's TMB associated neoantigen burden (using the methodology previously
166 articulated by [18, 19, 20]) and Inc-IM score (as previously defined by us). Any
167 overlapping loci between sORFs and neoantigen loci were identified and removed to
168 avoid the repetition of potential antigens.

169 **Statistical Analysis**

170 All analyses were implemented using R (v.3.6.3). Patients were divided into high,
171 low Inc-IM and combined antigen groups using the *surv_cutpoint* and
172 *surv_categorize* functions from the *survminer* package in R [25]. Cutpoints for
173 dividing data into high/low immunogenic groups based on Inc-IM scores and
174 combined antigen scores were chosen for each cohort separately using maximally
175 ranked statistics (Supplemental Figure 2 and Supplemental Table 1). To determine the
176 prognostic value of Inc-IM scores, Kaplan Meier survival plots were generated using
177 the *survival* R package [26]. Spearman's correlation and Wilcoxon test were
178 conducted using the *ggpubr* [27] R package to determine the association of Inc-IM
179 scores with tumor immune cells. A significance level of 0.05 was used as the cutoff,
180 with $p < 0.05$ considered as the statistically significant difference for all tests.

181 Three ICI-treated cohorts were utilized to determine the predictive value of combined
182 antigen score. Patients were stratified into the high and low combined immunogenic
183 groups as described above, and a logistic regression-based classifier was utilized to
184 predict immunotherapy responses using *stats* [28] package in R. Patients were
185 randomly assigned to training 70% and test 30% groups. For performance assessment,
186 the evaluation metrics area under the curve (AUC), accuracy and recall were
187 calculated using R packages *caret* and *caTools* [29, 30]. We compared the prediction
188 power of combined antigen scores with TMB using overall response rates, true
189 positive rates and false negative rates.

190 **RESULTS**

191 **Lnc-IM scores are associated with anti-tumor immune responses**

192 The tumor immune microenvironment (TIM) is critical in understanding disease
193 progression and predicting treatment responses. Tumor infiltrating lymphocytes
194 (TILs) comprises of a complex set of cells in the TIM that play important roles in both
195 tumor progression and suppression [31]. Based on these roles, these cells can be
196 divided into anti-tumor and pro-tumor immune cell groups. Tumor-associated
197 antigens are known to enhance TILs associated immune responses against tumor cells
198 [31]. To this end, we investigated the association of lncRNA antigen load using lnc-
199 IM scores with abundance of aDC, B cells, macrophages (M1 and M2), CD8-T cells,
200 CD4 memory T-cells, T regulatory cells, Th1 and Th2 cells. Six of these cell types
201 were characterized as anti-tumor immune cells based on their known involvement in
202 tumor surveillance mechanisms, and three were characterized as pro-tumor immune
203 cells. Three of six anti-tumor immune cells showed significant differences between
204 the high and low lnc-IM groups. None of the pro-tumor immune cells showed a

205 significant association between the two groups (Figure 1). Taken together these data
206 show a significant association between lnc-IM scores and elevated anti-tumor immune
207 responses.

208 **Lnc-IM scores are associated with survival predictions**

209 High tumor infiltrating lymphocytes (TILs) levels have been associated with the
210 immune system's capacity to eliminate tumor cells [32]. Hence, patients with high
211 TILs show improved survival compared to those with low TILs. Knowing the value of
212 TILs in survival prediction [33] and the association of lnc-IM scores with TILs
213 (Figure 1), we next questioned if lnc-IM scores could be used to predict survival in
214 the TCGA-SKCM cohort. The results showed that patients in the low lnc-IM category
215 (n = 59) showed better overall survival (HR = 0.39, p = 0.009) than in the high lnc-IM
216 category (n = 42) (Figure 2). The overall median survival remained at 1070 days
217 among the low lnc-IM group, while it reduced to 721 days in the high-IM group. The
218 baseline characteristics of all patients are shown in Table 1. In order to check the
219 association of any confounding variables with survival predictions, a multivariate
220 analysis was performed using age, gender, and cancer stage and lnc-IM scores (Figure
221 3). None except lnc-IM scores, were significantly associated with survival predictions.

	TCGA-SKCM	UCLA	MSKCC	DFCI
Gender				
Female	42	11	25	14
Male	59	27	39	26
Age				
≤50	15	7	14	12
>50	86	31	47	28
Treatment				
Iplimumab	-	-	60	40
Pembrolizumab	-	36	-	-
Nivolumab	-	2	-	-
Tremelimumab	-	-	4	-
Responders	-	21	14	5
Cancer stage				
I/II	72	-	-	-
III/IV	29	-	-	-

Table 1: Baseline features of cohorts used in this study. For ICI treated cohorts MSKCC, UCLA only subset of patients was selected for downstream analysis (as explained in Methods: Data acquisition).

223 **ICI efficacy predictions based on combined antigen score**

224 Based on the association of lnc-IM scores with both TILs and survival, we evaluated
225 how such scoring can help improve immunotherapy outcomes prediction. We
226 hypothesized that using a combined antigen score that includes both lncRNA-
227 associated antigen scores and TMB-associated neoantigen scores can give a fuller
228 picture of the tumor's immunopeptidome and so enhance the prediction of ICI
229 efficacy. Such scoring could also help predict ICI outcomes for patients' whose
230 tumors are not hypermutated (i.e have a high TMB) and so would not be typically
231 considered for such treatment. Using both lnc-IM scores and TMB-derived neoantigen

232 load, a combined antigen score was assigned to each patient in the ICI-treated cohorts.

233 A logistic regression-based classifier was then adopted to build a prediction model for

234 immunotherapy outcomes in a combined cohort.

235 We evaluated the model's performance; our results showed an overall AUC of 0.71 in

236 discriminating between responders and non-responders (accuracy = 0.54 and recall =

237 1). We compared this model with TMB based model (FDA approved cutoff > 10

238 mutations/megabase (mut/Mb) was used for dividing patients into high and low TMB

239 groups). The results showed an improved performance of combined antigen scoring as

240 compared to TMB based model (Table 2).

241

Evaluation Metric	Combined antigen score	TMB
AUC	0.71	0.53
Recall	1	0.4
Accuracy	0.54	0.5

Table 2: Evaluation metrics for combined antigen scoring and TMB based models.

242 **Combined antigen score performs better than TMB alone as biomarker**

243 To further assess whether predictions based on combined antigen score performed
244 better than using TMB alone, we compared overall response rates (ORR) between the
245 high TMB group and high combined antigen score in all three ICI treated cohorts.
246 The ORR improved in high combined antigen score as compared to high TMB among
247 UCLA and MSKCC cohorts while remaining the same for DFCI cohort (Figure 4).
248 Among these ICI treated cohorts, 21 out of 81 patients responded to therapy. Using a
249 high combined antigen score (cutoff criterion showed in Supplemental Material Table
250 1) as a classifier, 18 out of these 21 were correctly identified as responders. In
251 contrast, using TMB alone using the same classifier formalism correctly categorizes
252 only 14 responders. An additional advantage of using a combined strategy is the goal
253 of minimizing the false negative rate to ensure a robust classification that would not
254 deprive potential responders of therapy. In Figure 5 we show that the false negative
255 rate decreased from 33% to 14% using combined antigen score than TMB alone.
256 These results demonstrate strong evidence that using combined antigen score can help
257 improve prediction efficacy for patients who might have a low mutation burden but
258 still can benefit from treatment.

259 **DISCUSSION**

260 In this bioinformatics-based study, we introduced a novel metric “lnc-IM scores”
261 incorporating translatable lncRNA expression and patient MHC-I genotype and
262 combined them with TMB-associated neoantigen load to create a predictive
263 biomarker “combined antigen score” for ICI outcomes. TCGA SKCM data revealed
264 the prognostic value and association of lnc-IM scores with the tumor immune
265 microenvironment. Additionally, the value of combined antigen score as a predictive
266 biomarker was investigated using ICI-treated cohorts (UCLA n=27, MSKCC n=21,
267 DFCI n=40), previously used to investigate TMB-associated ICI response alone.

268 Among the TCGA SKCM cohort, lncRNA immunogenicity scores were significantly
269 associated with anti-tumor immune responses. TILs associated with anti-tumor
270 immune responses showed significant differences between high and low lnc-IM
271 groups. Interestingly, all of these cells were upregulated in the low lnc-IM group.
272 These findings appear to be in accordance with recent pan-cancer studies where high
273 TMB/high antigen groups were associated with depressed immune cell infiltration in
274 different cancer types [34]. One of the possible explanations for such trends could be
275 that the quality/immunogenicity of antigen being presented is responsible for a better
276 immune response than the quantity of antigens [35]. A study previously showed that
277 certain antigens belonging to ERVs were associated with high antigen specific CD8-
278 Tcell infiltrate as compared to other antigens that were associated with rare antigen
279 specific CD8-T cells population [15]. Other critical factors are the tumor immune
280 evasion mechanisms that are prevalent in high immunogenic groups. Cancers tend to
281 evade immune responses by downregulating MHC molecules and TAP3 proteins, so
282 these antigens are not efficiently presented on cancer cell surfaces [36]. Hence, we see
283 a decrease in immune cell infiltration in such cancers.

284 TILs have been linked to prognostic outcomes in various types of cancer. In general,
285 elevated levels of TILs within a tumor correspond to a more aggressive anti-tumor
286 immune microenvironment and are associated with better survival outcomes
287 compared to tumors with low TILs, as previously reported [37]. Our study (as
288 illustrated in Figure 1 and Figure 2) supports this finding by demonstrating that
289 patients with low Inc-IM tumors generally had greater anti-tumor TILs and better
290 survival outcomes than those with high Inc-IM tumors.

291 We hypothesized that using Inc-IM scores along with conventional TMB might help
292 improve efficacy predictions. We determined that the overall response rates improved
293 using high combined antigen score as compared to using TMB alone. An essential
294 factor is to ensure that the combined antigen score based classifier helps decrease the
295 false negative rate to avoid depriving actual responders of ICI therapy. We have
296 demonstrated that using the high combined antigen score compared to high TMB
297 alone cannot only improve the true positive rate but also help decrease the false
298 negative rate (Figure 5). In these cohorts, using high TMB alone as a biomarker could
299 deprive 33% of potential responders of potentially efficacious ICI therapy. However,
300 in our small study we demonstrated by using high combined antigen scores reduced
301 this percentage to 14%. These findings have specifically highlighted the predictive
302 power of combined antigen approach scores for responders with low TMB but high
303 Inc-IM scores.

304 The key limitations of this study are the small sample size, the retrospective nature of
305 cohorts used, and the variability of the type of ICI treatment administered. ICI
306 treatments vary based on the type of checkpoint protein targeted (PD-1 or CTLA4).
307 Both treatments show potential for melanoma treatment with 58% ORR for anti-PD-1
308 and 38% ORR for anti-CTLA4 therapy in melanoma [38]. Using both primary

309 (TCGA-SKCM) and metastatic cohorts (ICI treated) in our study may have
310 advantages and limitations. On the one hand, the inclusion of both sample types
311 increased the sample size of the study, enhancing our ability to identify more
312 inclusive and robust prognostic and predictive biomarkers. On the other hand, using
313 both cohorts may present challenges due to the potential differences in molecular
314 profiles between primary and metastatic tumors. We note however, a recent study
315 showing no significant difference between primary and metastatic melanoma TMB
316 [39]. We have demonstrated an association of the lnc-IM score with the TIM,
317 survival, and ICI outcomes prediction in three different cohorts providing confidence
318 as regards the added value of the lnc-IM scores as a biomarker. It is also worth noting
319 that the predictive power of combined antigen scoring was evaluated specifically in
320 metastatic tumors. Further validation in a larger cohort size is required to determine
321 the optimal cutoff for the lnc-IM and combined antigen scores.

322 Future research would ideally focus on determining the value of lnc-IM scores in
323 other cancer types that are not hypermutated yet still respond to ICI therapy, involving
324 larger cohort sizes.

325 **Acknowledgements:** This project was supported by a Hardiman Research
326 Scholarship, University of Galway, Ireland

327 **Competing interest declaration:** None declared

328 **Ethics approval:** All the data used in this study were derived from existing publicly
329 available databases and previously published studies. All the data were pre-
330 anonymized. Thus, ethical approval was not required.

331 **Data availability statement:** Data can be provided on request.

332 **List of Abbreviations**

333 Immune checkpoint inhibitor therapy (ICI)

334 LncRNA based immunogenicity score (lnc-IM score)

335 Tumor mutation burden (TMB)

336 Area under the curve (AUC)

337 Programmed cell death ligand 1 (PDL1)

338 Mismatch repair defect (dMMR)

339 Microsatellites Instability (MSI)

340 Encyclopedia of DNA Elements (ENCODE)

341 Cytotoxic T lymphocyte (CTL)

342 Short open reading frames (sORFs)

343 Tumor immune microenvironment (TIM)

344 The Cancer Genome Atlas (TCGA)

345 Counts per million (cpm)

346 Patient Harmonic-mean Best Rank (PHBR)

347 Tumor infiltrating lymphocytes (TILs)

348 Overall response rates (ORR)

349 Hazard ratio (HR)

350 **REFERENCES**

351

- 352 1. Hanahan D, Weinberg RA. Hallmarks of cancer: The next generation. *Cell*.
353 2011;144(5):646–74.doi:[10.1016/j.cell.2011.02.013](https://doi.org/10.1016/j.cell.2011.02.013)
- 354 2. Kanaseki T, Torigoe T. Proteogenomics: Advances in cancer antigen research.
355 *Immunological Medicine*. 2019;42(2):65–70. doi:[10.1080/25785826.2019.1640500](https://doi.org/10.1080/25785826.2019.1640500)
- 356 3. He X, Xu C. Immune Checkpoint Signaling and cancer immunotherapy. *Cell*
357 *Research*. 2020;30(8):660-9.doi:[10.1038/s41422-020-0343-4](https://doi.org/10.1038/s41422-020-0343-4)
- 358 4. Chen Q, Li T, Yue W. Drug response to PD-1/PD-L1 blockade: Based on biomarkers.
359 *OncoTargets and Therapy*. 2018;11:4673–83.doi:[10.2147/OTT.S168313](https://doi.org/10.2147/OTT.S168313)
- 360 5. Havel JJ, Chowell D, Chan TA. The evolving landscape of biomarkers for checkpoint
361 inhibitor immunotherapy. *Nature Reviews Cancer*. 2019;19(3):133–
362 50.doi:[10.1038/s41568-019-0116-x](https://doi.org/10.1038/s41568-019-0116-x)
- 363 6. Yi M, Jiao D, Xu H, Liu Q, Zhao W, Han X, et al. Biomarkers for predicting efficacy
364 of PD-1/PD-L1 inhibitors. *Molecular Cancer*. 2018;17(1):129.doi:[10.1186/s12943-
365 018-0864-3](https://doi.org/10.1186/s12943-018-0864-3)
- 366 7. Kim JY, Kronbichler A, Eisenhut M, Hong SH, van der Vliet HJ, Kang J, et al.
367 Tumor mutational burden and efficacy of immune checkpoint inhibitors: A systematic
368 review and meta-analysis. *Cancers*. 2019;11(11):1798.doi:[10.3390/cancers11111798](https://doi.org/10.3390/cancers11111798)
- 369 8. Gurjao C, Tsukrov D, Imakaev M, Luquette LJ, Mirny LA. Limited evidence of
370 tumour mutational burden as a biomarker of response to immunotherapy. *BioRxiv*.
371 2020; 26(02):65.doi: <https://doi.org/10.1101/2020.09.03.260265>
- 372 9. Castle JC, Uduman M, Pabla S, Stein RB, Buell JS. Mutation-derived Neoantigens
373 for cancer immunotherapy. *Frontiers in Immunology*.
374 2019;10:1856.doi:[10.3389/fimmu.2019.01856](https://doi.org/10.3389/fimmu.2019.01856)
- 375 10. International Human Genome Sequencing Consortium. Initial sequencing and
376 analysis of the human genome. *Nature*. 2001;409: 860–
377 921.doi:<https://doi.org/10.1038/35057062>
- 378 11. Zhang F, Zhang L, Zhang C. Long noncoding RNAs and tumorigenesis: Genetic
379 Associations, molecular mechanisms, and therapeutic strategies. *Tumor Biology*.
380 2016;37(1):163–75.doi:[10.1007/s13277-015-4445-4](https://doi.org/10.1007/s13277-015-4445-4)
- 381 12. Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A cerna hypothesis: The Rosetta
382 Stone of a hidden RNA language? *Cell*. 2011;146(3):353–
383 8.doi:[10.1016/j.cell.2011.07.014](https://doi.org/10.1016/j.cell.2011.07.014)
- 384 13. Lee JT. Epigenetic regulation by long noncoding RNAs. *Science*.
385 2012;338(6113):1435–9.doi:[10.1126/science.1231776](https://doi.org/10.1126/science.1231776)
- 386 14. Boon T, Van Pel A. T cell-recognized antigenic peptides derived from the cellular
387 genome are not protein degradation products but can be generated directly by
388 transcription and translation of short subgenic regions. A hypothesis.
389 *Immunogenetics*. 1989;29(2):75–9. doi:[10.1007/BF00395854](https://doi.org/10.1007/BF00395854)

- 390 15. Laumont CM, Vincent K, Hesnard L, Audemard É, Bonneil É, Laverdure J-P, et al.
391 Noncoding regions are the main source of targetable tumor-specific antigens. *Science*
392 *Translational Medicine*. 2018;10(470):eaau5516.doi:[10.1126/scitranslmed.aau5516](https://doi.org/10.1126/scitranslmed.aau5516)
- 393 16. Lv D, Chang Z, Cai Y, Li J, Wang L, Jiang Q, et al. TransLnc: A comprehensive
394 resource for translatable lncrnas extends immunopeptidome. *Nucleic Acids Research*.
395 2021;50(D1):D413–D420.doi:[10.1093/nar/gkab847](https://doi.org/10.1093/nar/gkab847).
- 396 17. Marty R, Kaabinejadian S, Rossell D, Slifker MJ, van de Haar J, Engin HB, et al.
397 MHC-I genotype restricts the oncogenic mutational landscape. *Cell*.
398 2017;171(6):1272–83.doi:[10.1016/j.cell.2017.09.050](https://doi.org/10.1016/j.cell.2017.09.050)
- 399 18. Hugo W, Zaretsky JM, Sun L, Song C, Moreno BH, Hu-Lieskovan S, et al. Genomic
400 and transcriptomic features of response to Anti-PD-1 therapy in metastatic melanoma.
401 *Cell*. 2016;165(1):35–44.doi:[10.1016/j.cell.2016.02.065](https://doi.org/10.1016/j.cell.2016.02.065)
- 402
- 403 19. Snyder A, Makarov V, Merghoub T, Yuan J, Zaretsky JM, Desrichard A, et al.
404 Genetic basis for clinical response to CTLA-4 blockade in melanoma. *New England*
405 *Journal of Medicine*. 2014;371(23):2189–99.doi:[10.1056/NEJMoa1406498](https://doi.org/10.1056/NEJMoa1406498)
- 406 20. Van Allen EM, Miao D, Schilling B, Shukla SA, Blank C, Zimmer L, Sucker A,
407 Hillen U, Geukes Foppen MH, Goldinger SM, Utikal J. Genomic correlates of
408 response to CTLA-4 blockade in metastatic melanoma. *Science*. 2015;350(6257):207-
409 1.doi:[10.1126/science.aad0095](https://doi.org/10.1126/science.aad0095)
- 410 21. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative
411 Analysis of Complex Cancer Genomics and clinical profiles using the cBioPortal.
412 *Science Signaling*. 2013;6(269):pl1.doi:[10.1126/scisignal.2004088](https://doi.org/10.1126/scisignal.2004088)
- 413 22. Aran D, Hu Z, Butte AJ. xCell: Digitally portraying the tissue cellular heterogeneity
414 landscape. *Genome Biology*. 2017;18(1):220.doi:[10.1186/s13059-017-1349-1](https://doi.org/10.1186/s13059-017-1349-1)
- 415 23. Olexiouk V, Crappé J, Verbruggen S, Verhegen K, Martens L, Menschaert G.
416 sORFs.org: A repository of small orfs identified by ribosome profiling. *Nucleic Acids*
417 *Research*. 2016;44(D1): D324–9.doi:[10.1093/nar/gkv1175](https://doi.org/10.1093/nar/gkv1175)
- 418 24. Robinson MD, McCarthy DJ, Smyth GK. edgeR: A bioconductor package for
419 differential expression analysis of digital gene expression data. *Bioinformatics*.
420 2010;26(1):139–40.doi:[10.1093/bioinformatics/btp616](https://doi.org/10.1093/bioinformatics/btp616)
- 421 25. Kassambara A, Kosinski M, Biecek P, Fabian S. Drawing survival curves using
422 'ggplot2' [R package survminer version 0.4.9]. The Comprehensive R Archive
423 Network. Comprehensive R Archive Network (CRAN); 2021. Available from:
424 <https://cran.r-project.org/web/packages/survminer/>
- 425 26. Therneau TM. Survival analysis [R package survival version 3.5-3]. The
426 Comprehensive R Archive Network. Comprehensive R Archive Network (CRAN);
427 2023. Available from: <https://cran.r-project.org/web/packages/survival/index.html>
- 428 27. Kassambara A. 'Ggplot2' based publication ready plots [R package ggpubr version
429 0.6.0]. The Comprehensive R Archive Network. Comprehensive R Archive Network
430 (CRAN); 2023. Available from:
431 <https://cran.r-project.org/web/packages/ggpubr/index.html>
- 432 28. R Core Team. R: A language and environment for statistical computing. R
433 Foundation for Statistical Computing, Vienna, Austria; 2013. Available from:
434 <http://www.R-project.org/>

- 435 29. Kuhn M. Classification and regression training [R package caret version 6.0-93]. The
436 Comprehensive R Archive Network. Comprehensive R Archive Network (CRAN);
437 2022. Available from: <https://cran.r-project.org/web/packages/caret/index.htm> l
- 438 30. Tuszynski J. Package catools. CRAN. Comprehensive R Archive Network (CRAN);
439 2021. Available from: <https://cran.r-project.org/web/packages/caTools/index.html>
- 440 31. He Z, Gu J, Luan T, Li H, Li C, Chen Z, et al. Comprehensive analyses of a tumor-
441 infiltrating lymphocytes-related gene signature regarding the prognosis and
442 immunologic features for immunotherapy in bladder cancer on the basis of WGCNA.
443 *Frontiers in Immunology*. 2022;13:973974.doi:10.3389/fimmu.2022.973974
- 444 32. Brummel K, Eerkens AL, de Bruyn M, Nijman HW. Tumour-infiltrating
445 lymphocytes: From prognosis to treatment selection. *British Journal of Cancer*.
446 2022;128(3):451–8.doi:10.1038/s41416-022-02119-4
- 447 33. Maibach F, Sadozai H, Seyed Jafari SM, Hunger RE, Schenk M. Tumor-infiltrating
448 lymphocytes and their prognostic value in cutaneous melanoma. *Frontiers in*
449 *Immunology*. 2020;11:2105.doi:10.3389/fimmu.2020.02105
- 450 34. Wang X, Li M. Correlate tumor mutation burden with immune signatures in human
451 cancers. *BMC Immunology*. 2019;20(1):4.doi:10.1186/s12865-018-0285-5
- 452 35. Laumont CM, Perreault C. Exploiting non-canonical translation to identify new
453 targets for T cell-based cancer immunotherapy. *Cellular and Molecular Life Sciences*.
454 2017;75(4):607–21.doi:10.1007/s00018-017-2628-4
- 455 36. Dhatchinamoorthy K, Colbert JD, Rock KL. Cancer immune evasion through loss of
456 MHC class I antigen presentation. *Frontiers in Immunology*.
457 2021;12:636568.doi:10.3389/fimmu.2021.636568
- 458 37. Gooden MJ, de Bock GH, Leffers N, Daemen T, Nijman HW. The prognostic
459 influence of tumour-infiltrating lymphocytes in cancer: A systematic review with
460 meta-analysis. *British Journal of Cancer*. 2011;105(1):93–
461 103.doi:10.1038/bjc.2011.189
- 462 38. Tie EN, Lai-Kwon J, Rtshiladze MA, Na L, Bozzi J, Read T, Atkinson V, Au-Yeung
463 G, Long GV, McArthur GA, Sandhu S. Efficacy of immune checkpoint inhibitors for
464 in-transit melanoma. *Journal for immunotherapy of cancer*.
465 2020;8(1):e000440.doi:10.1136/jitc-2019-000440
- 466 39. Gorris MA, van der Woude LL, Kroeze LI, Bol K, Verrijp K, Amir AL, Meek J,
467 Textor J, Figdor CG, de Vries IJ. Paired primary and metastatic lesions of patients
468 with ipilimumab-treated melanoma: high variation in lymphocyte infiltration and
469 HLA-ABC expression whereas tumor mutational load is similar and correlates with
470 clinical outcome. *Journal for Immunotherapy of Cancer*.
471 2022;10(5):e004329.doi:10.1136/jitc-2021-004329
472

473 **FIGURE LEGENDS**

Figure 1: Comparison of immune cells abundance between high and low immunogenic groups. A) Pro-tumor immune cells show no significant association

with lnc-IM scores. B) Association of anti-tumor immune cells with high and low lnc-IM scores.

Figure 2: Association of lnc-IM scores with survival predictions in TCGA-SKCM cohort (n=101). Low lnc-IM group (n=59) is associated with better survival outcomes as compared to high lnc-IM group (n=42).

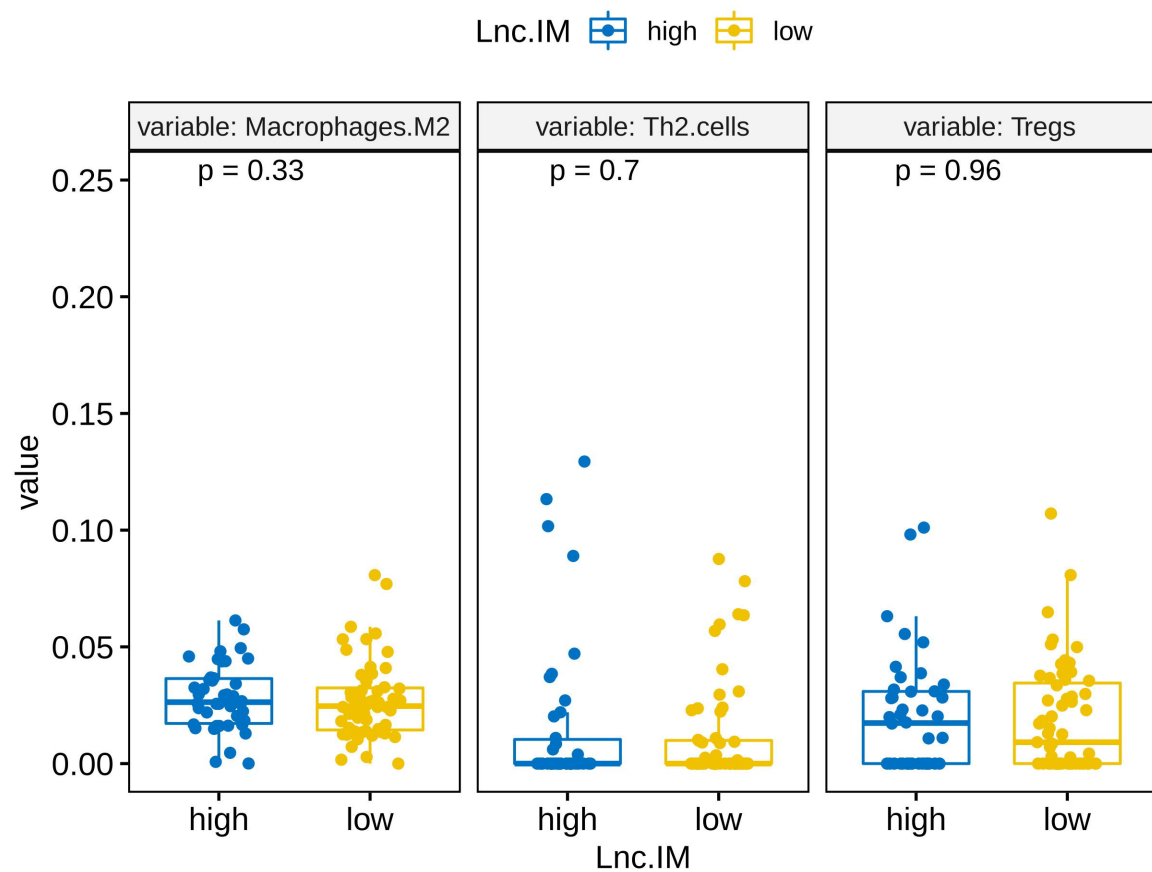
Figure 3: Multivariate analysis to identify any confounding variables. Only immunity count (lnc-IM scores) showed significance with survival predictions (TCGA-SKCM).

Figure 4: Comparison of overall response rate between high TMB category and high combined antigen score among all ICI treated cohorts.

Figure 5: Comparing true positive rate (TPR) and false negative rate (FNR) between high TMB and high combined antigen scores. High combined antigen scoring showed improved TPR of 85% as compared to high TMB (66%). FNR also showed improvement (14%) in high combined antigen classifier.

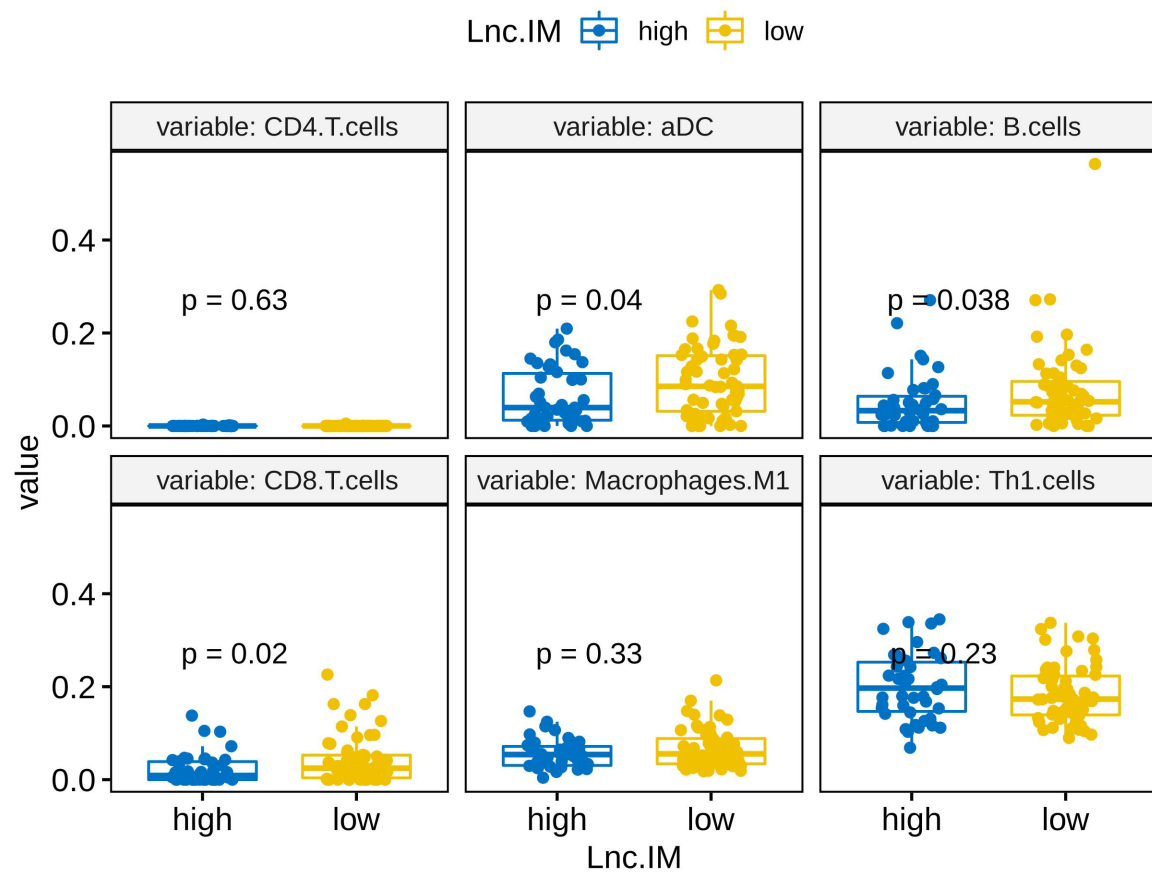
A

Pro-tumor immune cells

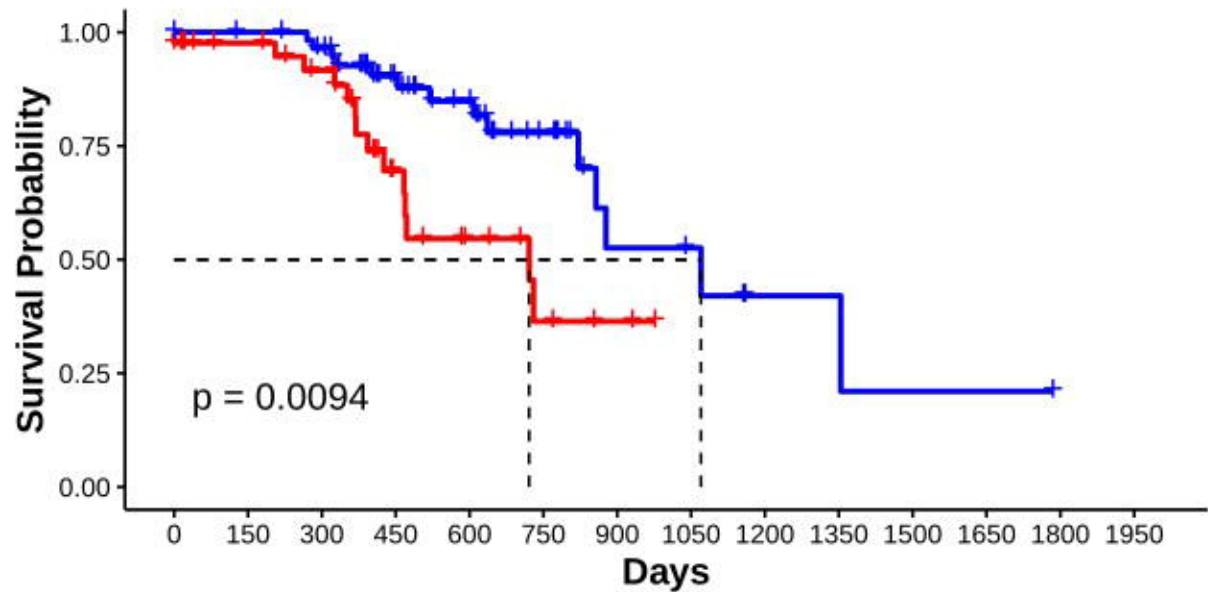


B

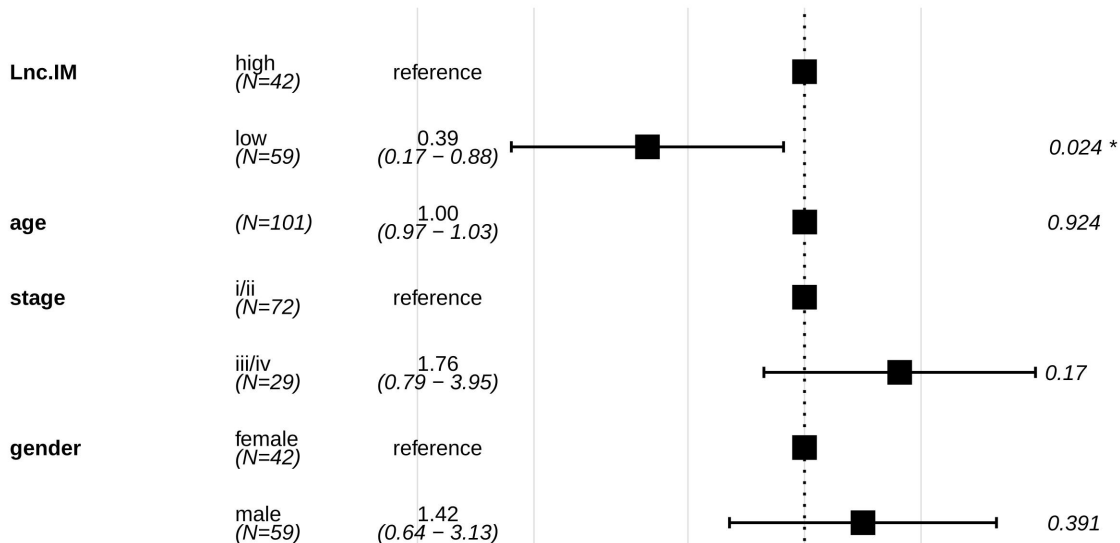
Anti-tumor immune cells



Strata Lnc.IM=high Lnc.IM=low



Hazard ratio



Events: 28; Global p-value (Log-Rank): 0.07153

AIC: 203.52; Concordance Index: 0.69

