1 Research Article

Immunogenomic intertumor heterogeneity across primary and metastatic sites in a patient with lung adenocarcinoma

Authors: Runzhe Chen^{1,2,#}, Jun Li^{2,#}, Junya Fujimoto³, Xin Hu², Kelly Quek¹, Ming
Tang², Akash Mitra², Carmen Behrens¹, Chi-Wan Chow³, Peixin Jiang¹, Latasha D.
Little², Curtis Gumbs², Xingzhi Song², Jianhua Zhang², Dongfeng Tan⁴, John V.
Heymach¹, Ignacio Wistuba³, P. Andrew Futreal², Don L. Gibbons^{1,*}, Lauren A. Byers^{1,*},
Jianjun Zhang^{1,2*}, Alexandre Reuben^{1,*}

9 Affiliations:

- ¹Department of Thoracic/Head and Neck Medical Oncology, the University of Texas MD
- 11 Anderson Cancer Center, Houston, Texas 77030, USA
- ²Department of Genomic Medicine, the University of Texas MD Anderson Cancer
 Center, Houston, Texas 77030, USA
- ³Department of Translational Molecular Pathology, the University of Texas MD
 Anderson Cancer Center, Houston, Texas 77030, USA
- ¹⁶ ⁴Department of Pathology, the University of Texas MD Anderson Cancer Center,
- 17 Houston, Texas 77030, USA
- 18 # co-first authors
- 19 *Correspondence: <u>dlgibbon@mdanderson.org</u> (D.L.G.), <u>lbyers@mdanderson.org</u>
- 20 (L.A.B.), jzhang20@mdanderson.org (J.J.Z.), areuben@mdanderson.org (A.R.)
- 21 Running title: Immunogenomic intertumor heterogeneity in lung cancer
- 22 **Keywords:** lung adenocarcinoma, intertumor heterogeneity, genomic, T cell repertoire

23

24

25 Abstract

Background: Lung cancer is the leading cause of cancer death, partially owing to its 26 extensive heterogeneity. The analysis of intertumor heterogeneity has been limited by 27 an inability to concurrently obtain tissue from synchronous metastases unaltered by 28 multiple prior lines of therapy. Methods: In order to study the relationship between 29 genomic, epigenomic and T cell repertoire heterogeneity in a rare autopsy case from a 30 31 young female never-smoker with late-stage lung adenocarcinoma (LUAD), we did whole-exome sequencing (WES), DNA methylation and T-cell receptor (TCR) 32 33 sequencing to characterize the immunogenomic landscape of one primary and 19 synchronous metastatic tumors. Results: We observed heterogeneous mutation, 34 35 methylation, and T cell patterns across distinct metastases including a set of prevalent T cell clonotypes which were completely excluded from left-side thoracic tumors. Though 36 37 a limited number of predicted neoantigens were shared, these were associated with homology of the T cell repertoire across metastases. Lastly, ratio of methylated 38 39 neoantigen coding mutations was negatively associated with T-cell density, richness and clonality. suggesting neoantigen methylation may partially drive 40 immunosuppression. **Conclusions:** Our study demonstrates heterogeneous genomic 41 and T cell profiles across synchronous metastases and how restriction of unique T cell 42 clonotypes within an individual may differentially shape the genomic and epigenomic 43 landscapes of synchronous lung metastases. 44

45

46 Background

Lung cancer is the leading cause of cancer death, partially owing to its extensive heterogeneity [1, 2]. It has been proposed that this extensive heterogeneity results from successive clonal expansion and selection of the fittest clones influenced by genomic accumulation and somatic epigenetic alterations [3-6]. However, tumor evolution may also be shaped by pressure from the immune system, which can prune the most immunogenic branches of the tumor [7].

T cells play a crucial role in preventing cancer development through antigen-specific 53 detection and destruction of malignant cells, though evolving tumors can eventually 54 escape immune surveillance through a process termed immunoediting [4, 8-11]. Few 55 studies have addressed the impact of the T cell repertoire in shaping metastatic 56 heterogeneity [12-14] with most work to date evaluating longitudinal changes spanning 57 multiple timepoints and therapies. Though these studies offer crucial insights, they do 58 not allow the evaluation of intrinsic intertumor heterogeneity in absence of selective 59 pressure from therapy. Furthermore, to date, the analysis of intertumor heterogeneity 60 has been limited by an inability to concurrently obtain tissue from synchronous 61 metastases unaltered by multiple prior lines of therapy [15-17]. 62

63 Here, we sought to study the relationship between genomic, epigenomic and T cell repertoire heterogeneity in a rare autopsy case from a female never-smoker in her early 64 30s with late-stage lung adenocarcinoma (LUAD) with more than 20 synchronous 65 metastases. We observed heterogeneous mutation, methylation, and T cell patterns 66 67 across distinct metastases including a set of prevalent T cell clonotypes which were completely excluded from left-side thoracic tumors. Our work further highlights 68 neoantigen methylation as a potential mechanism driving immunosuppression and 69 some of the hurdles facing the treatment of late-stage lung cancer. 70

71

72 Methods

73 Human subject research

We collected 20 tumor samples and one normal GI sample at autopsy. Collection and
use of patient samples were approved by the Institutional Review Board of the
University of Texas MD Anderson Cancer Center. Clinical information is presented in
Supplementary Table 1.

78 Sample collection

DNA of collected samples was isolated from FFPE tissues using the AllPrep DNA/RNA FFPE Kit (Qiagen, Hilden, Germany). Hematoxylin and eosin (H&E) slides of each case were reviewed by experienced lung cancer pathologists under the microscope to assess the percentage of pre-/micro-invasive neoplastic lesions and tumor tissues versus normal tissues. Tumor cell viability was also assessed by examining the presence of necrosis in the tissues. Only samples with enough viable tumor cells were selected for WES, methylation and immunoSEQ.

86 **DNA preparation**

Unstained tissue sections (10µm thick) were deparaffinized in xylene and 100% ethanol 87 (twice in each for 10 minutes). The macrodissected tumor areas of the deparaffinized 88 tissues were placed into a 1.5 mL collection tubes for DNA extraction. The tissue was 89 90 next suspended with Buffer PKD and Proteinase K from the Allprep FFPE kit. After incubating at 56 °C for 15 min then on ice for 5 min, the mixed solution was centrifuged 91 92 for 15 minutes at 20,000 x g. Finally, the DNA samples were quantified by Nano Drop 1000 Spectrophotometer (Thermo Scientific, Wilmington, DE, USA). The fragmentation 93 94 sizes were evaluated by the Agilent 2200 Tape Station system using the Genomic DNA Screen Tape Assay (Agilent Technologies, Santa Clara, CA, USA). 95

96 Whole-exome sequencing

97 Exome capture was performed on 500ng of genomic DNA per sample based on KAPA 98 library prep (Kapa Biosystems) using the Agilent SureSelect Human All Exon V4 kit 99 according to the manufacturer's instructions and paired-end multiplex sequencing was 100 performed on the Illumina HiSeq 2000 sequencing platform. The average sequencing 101 depth was 178x (ranging from 63x to 225x, standard deviation +/- 31).

102 Mutation calling

Tumor contents and major/minor copy number changes were estimated by Sequenza 103 (v2.1.2) [18]. To control those FFPE caused artifact contaminations, somatic single 104 nucleotide variants (SNVs) was first called using MuTect version 1.1.4 [19], VarScan 2 105 [20] and Strelka2 [21] with default setting, respectively. Then, the following filtering 106 criteria were applied to each callers: 1) sequencing depth $\geq 20 \times$ in tumor DNA and 107 108 \geq 10× in germline DNA; and 2) variant allele frequency (VAF) \geq 0.02 in tumor DNA and < 0.01 in germline DNA; and 3) the total number of reads supporting the variant calls is 109 \geq 4; and 4) variant frequency is < 0.01 in ESP6500, 1000 genome and EXAC databases; 110 Those mutations called by Mutect with a LOD score<10 was further filtered out. 111 112 Those mutations called by Strelka with a quality score below 35 was also filtered out. Finally, only those mutations were kept if called by any of the two tools and rescued if 113 any were rejected but shared by at least two tumors. Identified missense mutations 114 were manually reviewed using the Integrative Genomics Viewer version 2.3.61 [22, 23]. 115

116 **Phylogenetic analysis**

Ancestors were germ line DNA assuming with no mutations. The phylogenetic tree was generated as described [24]. A binary presence/absence matrix of all somatic mutations detected was used as input for R package phangorn version 2.0.2 [25].

120 Neoantigen predictions

Nonsynonymous mutations were identified from WES profiling and the binding affinity with patient-restricted MHC Class I molecules of all possible 9- and 10-mer peptides spanning the nonsynonymous mutations was evaluated with the NetMHC3.4 algorithm based on HLA-A, HLA-B, and HLA-C alleles of each patient [26-28]. Candidate peptides were considered as HLA binders when IC50<500 nM with high affinity binders presenting IC50<50 nM.

127 DNA methylation profiling and tumor–immune microenvironment deconvolution

Genomic DNA (approximately 500 ng) was bisulfite converted using EZ DNA 128 Methylation Kit (Zymo Research Corp. Irvine, CA, USA) following the manufacturer's 129 130 protocol. Bisulfite converted DNA materials were then processed and hybridized to the Infinium Human Methylation 450k arrays (Illumina, San Diego, CA, USA) according to 131 manufacturer's recommendation. Preprocessing and initial quality assessments of the 132 133 raw data were examined using the following R packages. Subset-quantile within-array normalization (SWAN) [29] was used to normalize raw methylation values. 134 IlluminaHumanMethylation450k.db annotation package was used to annotate the CpG 135 probes location. Before any genomics and statistical analyses were conducted, we 136 normalized and inspected the methylation data for the presence of substantial 137 confounding batch effects. Cellular deconvolution analyses were carried out using 138 139 estimated cellular fractions using MethylCIBERSORT [30].

140 TCRβ sequencing and comparison parameters

Immunosequencing of the CDR3 regions of human TCR^β chains was performed using 141 the protocol of immunoSEQ (Adaptive Biotechnologies, hsTCRβ Kit) [31-33]. Two sets 142 of PCRs were performed on DNA extracted from the tissues collected. The initial PCR 143 used a mix of multiplexed V- and J-gene primers which amplify all possible recombined 144 receptor sequences from the DNA sample, and then a second PCR designing to add 145 unique DNA barcodes to each PCR product was followed. After that, samples were 146 pooled together with a negative and a positive control. The pools were then sequenced 147 on an Illumina MiSeq platform using 150 cycle paired-end protocol and sequence-ready 148 primers. After finish the sequencing, the raw data were transferred to Adaptive 149 150 Biotechnologies and processed into a report including those passed quality-check samples and a normalized and annotated TCR^β profile repertoire accordingly. Profile of 151 TCR rearrangements is presented in **Supplementary Data**. 152

T-cell density in FFPE tissue samples was calculated by normalizing TCR-β template counts to the total amount of DNA usable for TCR sequencing, where the amount of usable DNA was determined by PCR-amplification and sequencing of housekeeping genes expected to be present in all nucleated cells. T-cell richness is a metric of T cell

diversity, and it is calculated by on the T-cell unique rearrangements. T-cell clonality is a 157 metric of T cell proliferation and reactivity, and it is defined as 1-Peilou's evenness. 158 159 Clonality ranges from 0 to 1: values approaching 0 indicate a very even distribution of frequency of different clones (polyclonal), whereas values approaching 1 indicate a 160 distinct asymmetric distribution in which a few activated clones are present at high 161 frequencies (monoclonal). Statistical analysis was performed in R version 3.2. JI is 162 conceptually a percentage of how many objects of two sets have in common out of how 163 many objects they have in total. JI = (number of rearrangements in common) / (total 164 number of rearrangements) 165

166 Statistical Analysis

Graphs were generated with GraphPad Prism 8.0. Spearman's rank correlations were calculated to assess the association between 2 continuous variables. Kruskal-Wallis tests were used for categorical variables with more than 2 levels. P-values less than 0.05 were considered to be statistically significant.

171

172 **Results**

173 **Patient information**

174 A female never-smoker in her early 30s presented to her primary care physician complaining of weakness in her upper right arm lasting for two weeks. Physical 175 176 examination was unremarkable, other than grade 3 weakness in her right upper limb. Shortly after, she was hospitalized due to acute venous thromboembolism of this arm. 177 178 She was started on anti-coagulants and underwent computed tomography (CT) scans of the chest, abdomen and pelvis and magnetic resonance imaging (MRI) of the brain 179 180 as part of the work up. Numerous nodules were detected suggestive of extensive metastasis and a 1.9 cm left lung mass was consistent with a lung primary (Fig. 1a-b). 181 A liver biopsy revealed the diagnosis of poorly differentiated LUAD. The patient 182 underwent palliative radiation therapy for C5-C7 spine metastases with 10x30 cGy and 183 184 one dose of chemotherapy with carboplatin and paclitaxel while awaiting molecular profiling results. Her condition deteriorated rapidly and she expired 13 days following
 her sole dose of chemotherapy. An autopsy was performed and widely disseminated
 metastatic carcinoma involving multiple systems and organs was observed.

To understand the genomic and T cell landscape of this extensively metastatic LUAD, 188 20 tumor samples (**Supplementary Table 1**) including the left lung (primary tumor, P), 189 thyroid gland (M01), left pleural cavity (M02), left hilar lymph node (M03), left-side 190 parietal pleura (M04), heart (M05), right lung (M06), right pleural cavity (M07), 12th 191 thoracic vertebra (M08), gastrointestinal (GI) tract (M09), liver (M10), 4 abdominal lymph 192 nodes (M11-14), left adrenal gland (M15), two metastases in the right kidney (M16 and 193 M17), left and right ovaries (M18 and M19) as well as one histologically normal sample 194 195 from the GI tract were collected and subjected to whole-exome sequencing (WES), DNA methylation array and T cell receptor (TCR) sequencing. 196

197 Distinct mutational profiles are seen across primary tumor and synchronous 198 metastases

Overall, 228 non-silent mutations were detected with an average of 76 per sample 199 (range=57-98). The number of non-silent mutations varied between tumors, with only 10 200 shared across all 20 samples (Fig. 1c). Of these non-silent mutations, 170 (75%) were 201 202 shared by at least two tumors while 58 (25%) were unique (Fig. 1d). When canonical cancer gene mutations were analyzed[34-36], commonly-mutated cancer genes 203 204 included TP53, CDKN2A, ASXL1 and MET in this patient (Supplementary Fig. 1a). Only TP53 mutation (chromosome 17 7578382, stop gain, spectrum G->C) was 205 206 detected in all tumors suggesting TP53 mutation was an early genomic event, while other cancer gene mutations were later events which may have followed subclonal 207 208 diversification. We also constructed a phylogenetic tree to depict the genomic heterogeneity and evolutionary trajectory of these metastatic tumors. As shown in 209 210 Supplementary Fig. 1b, the phylogenetic structure varied considerably between tumors highlighting profound genomic heterogeneity within this patient. We also utilized 211 the Jaccard index (JI), which takes into consideration the proportion of shared non-silent 212 mutations between any two samples. The JI ranged from 0.14 to 0.82 (average=0.49) 213

with more proximal tumors generally more genetically similar (**Supplementary Fig. 1c**). Homology between the primary and metastases ranged from 0.14 to 0.73 (average=0.33), with the thoracic lesions including the left pleural metastasis, left hilar lymph node and right pleural metastasis exhibiting the greatest similarity with the primary tumor. Taken together, these results reveal marked genomic heterogeneity across different metastases within the same patient.

220 The T cell infiltrate in distant metastases is more dense, diverse and reactive

The crucial role of T cells in immunoediting led us to study the T cell repertoire to further 221 investigate the spatial heterogeneity of T-cell responses[37, 38]. T-cell density, an 222 223 estimate of the fraction of T cells within a tumor, ranged from 3% to 38% (average=13%, Fig. 2a), while richness, a measure of T-cell diversity, ranged from 4,168 to 23,487 224 unique T-cell rearrangements (average=14,344 unique rearrangements, Fig. 2b). T-cell 225 clonality, a measure of T-cell reactivity, ranged from 0.02 to 0.05 (average=0.04, Fig. 226 2c). All TCR metrics were positively inter-correlated (Density vs. Richness: r=0.53, 227 p=0.02; Density vs. Clonality: r=0.54, p=0.02; Clonality vs. Richness: r=0.54, p=0.01, 228 Fig. 2d). Compared to prior work from our group in an early-stage LUAD cohort[39], 229 lower T-cell density (p<0.05) and clonality (p<0.0001) and higher richness (p<0.0001) 230 were observed in tumors from this patient (Supplementary Fig. 2a-e). 231

Tumors were then grouped anatomically. Non-thoracic tumors displayed higher T-cell 232 233 density (p<0.01), richness (p<0.0001) and clonality (p<0.01) than thoracic tumors 234 (Supplementary Fig. 3a-c), perhaps owing to their anatomical location far from the 235 primary tumor which are more likely to escape from immune surveillance[40, 41]. Lymph nodes serve as sites of T cell priming, activation, and modulation, leading us to 236 237 speculate that the interaction between metastatic cancer cells and T cells in lymph nodes may be distinct compared to other sites of metastases. However, no statistical 238 239 differences were observed in relation to lymph node involvement (Supplementary Fig. **3d-f**). Taken together, these data suggest differences in T cell response based on 240 anatomical site, that is, T cell exclusion, suppression and a more focused T cell 241 response in proximity to the primary tumor. 242

243 Distinct TCR repertoire profiles are associated with left-side thoracic tumors

To evaluate T-cell responses in the tumors, we next focused on the most prevalent TCR 244 clonotypes. Distinct clonotypes were detected in left-side thoracic tumors (left lung 245 tumor, left pleural cavity and rib, left pulmonary hilar lymph node and left side parietal 246 pleura) compared to others. Strikingly, the most prevalent clonotype in "other" tumors 247 (CACRPGNEAFF) was entirely undetectable in left-side thoracic tumors (P, M02, M03 248 249 and M04) (Fig. 2e). Similar trends were also observed among the top 5 and 10 TCR clonotypes with certain clonotypes completely excluded from left side thoracic tumors 250 251 (Supplementary Fig. 4a-b). These data illustrate spatial restriction even among the most prevalent T cell clonotypes across synchronous metastases. 252

T-cell repertoire heterogeneity is observed across differentially growing tumors

To gain deeper insights into TCR heterogeneity, we assessed the overlap between T 254 cell repertoires across different tumors. We first compared the proportion (JI) and 255 256 frequency of T cell clonotypes shared between the primary tumor and metastases. In accordance with the unique T cell clonotype pattern observed in tumors from the left 257 side thorax, proportions and frequencies of shared T cells were much more similar 258 259 between the three left thoracic metastases (M02, M03 and M04) and primary tumor (P) 260 (Fig. 3a-c). T cell repertoire heterogeneity was evident across all tumors, with an average JI value of 0.35 (ranging from 0.12 to 0.61) and more shared T cells between 261 262 proximal tumors (Fig. 4a), significantly higher than in a previously published cohort of 11 multi-region localized non-small cell lung cancers[42] (average 0.35 vs. 0.17, 263 264 p<0.0001) (Supplementary Fig. 5). Next, we studied the proportion and frequency of shared T cell clonotypes across all 20 tumors. In total, 599 shared T-cell clones were 265 266 shared across all tumors, with proportions ranging from 3.0% to 15.4% (average=5.39%) and frequencies accounting for 11.9% to 21.5% of the T cell repertoire 267 268 (average=15.96%) (Supplementary Fig. 6a-b). Of interest, both a greater proportion (p<0.01) and percentage (p<0.01) of shared T cell clones were observed in thoracic 269 tumors compared to non-thoracic tumors (Supplementary Fig. 6c-d). 270

271 Evolution of synchronous metastases may be shaped by the T cell repertoire

272 We next performed in silico prediction of HLA-A-, -B-, and -C-presented neoantigens using NetMHC3.4[26-28]. On average, 39 predicted neoantigens ($IC_{50} < 500$ nmol/L) 273 274 were detected per tumor, with the most (n=60) seen in the primary tumor and fewest (n=20) in the thyroid gland. Only 11 high binding affinity neoantigens were detected on 275 average ($IC_{50} < 50$ nmol/L) with the most (n=19) also in the primary tumor and least 276 (n=2) in the thyroid gland (Fig. 5a). This falls within range but below the average of 53 277 predicted neoantigens seen in non-smokers from TCGA (Fig. 5b). We then evaluated 278 the relationship between the T cell repertoire and predicted neoantigens. Predicted 279 neoantigen heterogeneity was also evident, with the average JI value of 0.44 (ranging 280 from 0.11 to 0.84, Fig. 4b), and a weak but statistically-significant positive correlation 281 between T-cell repertoire and neoantigen homology (r=0.12, p=0.0162, Fig. 4c), which 282 could suggest the distribution of T cells may be partially driven by their reactivity to 283 shared neoantigens. Interestingly, ratio of methylated neoantigen coding mutations was 284 negatively associated with T-cell density (r=-0.46, p=0.0549), richness (r=-0.55, 285 p=0.0152) and clonality (r=-0.61, p=0.0055) (Fig. 5c-e), suggesting neoantigen 286 287 methylation may contribute to immune suppression and potentially explaining the weak neoantigen associations with T cell repertoire homology. 288

In order to classify the T cell subsets infiltrating the tumor, we next performed cellular 289 290 deconvolution analyses using MethylCIBERSORT (Fig. 4d)[30]. Relative CD8⁺ T cell fraction but not CD4⁺ and other immune subsets was proportional to richness (r=0.75, 291 p=0.0002) and clonality (r=0.43, p=0.063) (Fig. 4e-f). Furthermore, CD8 to Treg ratio, 292 which correlates with a more favorable outcome in cancer[43, 44], was proportional to 293 294 T-cell richness (r=0.68, p=0.0012) and clonality (r=0.64, p=0.032) (Supplementary Fig. **7a-b**). A negative correlation between CD8⁺ T cell fraction and methylated neoantigen 295 coding mutations was also observed (r=-0.48, p=0.0391, Fig. 5e). These results 296 highlight the greater proliferative potential of CD8⁺ T cells and suggest T cell reactivity 297 and diversity may be mainly driven by the clonal expansion of CD8⁺ T cells at the 298 patient level, as previously suggested by our group and others in a large cohort of 299 300 early-stage non-small cell lung cancers [45]. Overall, our findings suggest that the

evolution of synchronous metastases may be shaped by the T cell response including inabsence of prior therapy.

303

304 **Discussion**

Metastasis is an evolutionary process shaped by the dynamic interactions between 305 tumor cells and host factors including immune surveillance [46]. T cells play a pivotal 306 role in mediating this process by recognizing antigens presented on MHC molecules at 307 the surface of tumors and carrying out cytotoxic responses against tumor cells 308 harboring their cognate antigens [47]. Accordingly, much importance has been 309 attributed to T cell infiltration in many solid tumors, with more T cells generally 310 associated with a better prognosis [37, 48, 49]. However, recent studies have 311 highlighted the impact of intratumor heterogeneity and bystander T cells [50-52], and 312 suggested that only ~10% of tumor-infiltrating lymphocytes are capable of recognizing 313 314 antigens presented by the tumor they have infiltrated [53], prompting deeper investigations into the T cell repertoire. Our understanding of the role of genomic and 315 immune heterogeneity in lung cancer has evolved in recent years, thanks to 316 investigations by our group and others into differences between regions of individual 317 318 tumors, synchronous metastases and between primary and metastatic tumors [54-58] highlighting potential spatial and temporal factors influencing clinical outcomes [59, 60]. 319 320 Here, we assess the characteristics of the T cell repertoire in a treatment-naïve non-321 smoking patient with synchronous lung metastases and depict the interplay between the 322 primary tumor and synchronous metastases [11, 55, 61], revealing extensive immunogenomic intertumor heterogeneity across primary and metastatic sites. 323

In our study, clonal *TP53* mutations were detectable in all tumors, suggestive of an early genomic event, in line with prior reports [24, 62]. Interestingly, a higher overlap in somatic mutations was observed across proximal tumors suggesting they are more genetically similar, potentially due to metastatic seeding from the primary tumor [63, 64]. Though our study focused on a single patient, the overlap in mutational burden observed between synchronous metastases is in line with previous reports in lung [58],

12

melanoma [65], kidney [66] and colon cancer [67]. Considering the role of somatic mutations in triggering T cell responses through the generation of neoantigens, this overlap suggests these somatic mutations may serve as potential therapeutic targets for vaccination or T cell engineering through targeting of unifying antigens present across all synchronous tumors. This is supported by the modest but significant correlation between shared mutations and shared TCRs though additional studies are needed to confirm these hypotheses.

We observed lower T cell repertoire heterogeneity across synchronous metastases in 337 338 our study than in our prior work assessing multi-region ITH in localized LUAD [55]. This difference could highlight the distinct resistance mechanisms at play in accelerated 339 340 progression in our study versus more gradual progression in early-stage LUAD which may have allowed for divergent genomic evolution and immune editing over decades. 341 This is reinforced by the absolute restriction of certain T cell clonotypes to metastases 342 surrounding the primary tumor, which could be reflective of the distinct antigenic 343 344 environments established in distal tumors. Unfortunately, our lack of deep immune phenotyping data precludes our ability to further investigate the role the distinct immune 345 microenvironments, including chemokine gradients and receptors, may have played in 346 establishing these vastly distinct T cell microenvironments. However, the presence of 347 shared T cell clonotypes could also be indicative of common responses against unifying 348 antigens displayed across synchronous metastases. 349

Aberrant methylation is involved in tumorigenesis in a variety of cancers [68, 69] and 350 could affect immune surveillance directly by regulating the expression of immune-351 352 related genes and/or potential neoantigens[70]. DNA methylation-induced silencing of genes that encode tumor associated antigens form a complex gene-regulatory network 353 354 for suppressing anti-tumor immune responses and therefore facilitating immune evasion [71]. Interestingly, in the current study, ratio of methylated neoantigen coding mutations 355 356 was negatively associated with T-cell density, richness and clonality, even at the individual tumor level. One could therefore hypothesize that therapeutic agents 357 358 modulating methylation could potentially reprogram the immune microenvironment and could exhibit some potential in treating these tumors. 359

Our study does exhibit certain limitations, including its focus on a single patient. 360 However, analysis of several synchronous tumor sites from a single patient with 361 362 advanced disease in absence of heavy pre-treatment is rarely possible, mainly due to the lack of clinical indication. Unfortunately, deeper analysis of underlying mechanisms, 363 immune cells and soluble factors influencing T cell trafficking and heterogeneity remain 364 unclear due to the archival nature of these samples and will require further investigation. 365 Despite these limitations, our study provides important evidence of differential tumor-366 immune responses co-existing in metastases within the same individual, related not 367 only to molecular alterations. As a result, our findings may also partially explain the 368 challenge of treating late-stage lung cancer due to the heterogeneity of metastases. 369 Additional genomic, transcriptomic and immune studies in patients with synchronous 370 371 metastases could help shed light on these and other mechanisms at play and provide therapeutic insights into late-stage non-small cell lung cancer. Lastly, our findings 372 emphasize the need to obtain multiple biopsies when feasible in patients with multiple 373 synchronous lung tumors as has been suggested in metastatic melanoma, even in 374 375 absence of selective pressure from therapy. Concurrent analysis of multiple tumors from the same patient could help identify unifying antigens and mechanisms capable of 376 377 prolonging survival in patients with metastatic lung cancer [65].

378

379 **Conclusion**

In summary, in this study, we present the immunogenomic landscape of one primary 380 381 tumor and 19 synchronous metastases from a minimally-pretreated young female never-smoker with late-stage LUAD. Fewer less diverse and reactive T cells infiltrated 382 383 the metastases nearest to the primary tumor, and a set of prevalent T cell clonotypes were excluded from left-side thoracic tumors further suggesting immune escape near 384 385 the primary site. Furthermore, shared predicted neoantigens were associated with homology of the T cell repertoire across metastases. Lastly, ratio of methylated 386 neoantigen coding mutations was negatively associated with T-cell density, richness 387 and clonality, suggesting neoantigen methylation may partially drive 388

immunosuppression. Our study demonstrates heterogeneous genomic and T cell
 profiles across synchronous metastases and how restriction of unique T cell clonotypes
 within an individual may differentially shape the genomic and epigenomic landscapes of
 synchronous lung metastases.

393

394 Abbreviations

- 395 CT: computed tomography
- 396 H&E: hematoxylin and eosin
- 397 JI: Jaccard index
- 398 LUAD: lung adenocarcinoma
- 399 MRI: magnetic resonance imaging
- 400 SNVs: single nucleotide variants
- 401 SWAN: subset-quantile within-array normalization
- 402 TCR: T cell receptor
- 403 WES: whole-exome sequencing
- 404 VAF: variant allele frequency
- 405

406 **References**

Tan WL, Jain A, Takano A, Newell EW, Iyer NG, Lim WT, et al. Novel therapeutic
 targets on the horizon for lung cancer. The Lancet Oncology. 2016; 17: e347-e62.

Negrao MV, Quek K, Zhang J, Sepesi B. TRACERx: Tracking tumor evolution to
impact the course of lung cancer. The Journal of thoracic and cardiovascular surgery.
2018; 155: 1199-202.

412 3. Nowell PC. The clonal evolution of tumor cell populations. Science. 1976; 194: 413 23-8.

414 4. Valastyan S, Weinberg RA. Tumor metastasis: molecular insights and evolving 415 paradigms. Cell. 2011; 147: 275-92.

416 5. Yates LR, Campbell PJ. Evolution of the cancer genome. Nat Rev Genet. 2012; 417 13: 795-806.

- 6. Quek K, Li J, Estecio M, Zhang J, Fujimoto J, Roarty E, et al. DNA methylation intratumor heterogeneity in localized lung adenocarcinomas. Oncotarget. 2017; 8: 21994-2002.
- 421 7. Galon J, Bruni D. Tumor Immunology and Tumor Evolution: Intertwined Histories.
 422 Immunity. 2020; 52: 55-81.
- 423 8. De Visser KE, Eichten A, Coussens LM. Paradoxical roles of the immune system 424 during cancer development. Nature reviews cancer. 2006; 6: 24-37.
- 9. Schreiber RD, Old LJ, Smyth MJ. Cancer immunoediting: integrating immunity's
 roles in cancer suppression and promotion. Science (New York, NY). 2011; 331: 156570.
- 428 10. Dunn GP, Old LJ, Schreiber RD. The three Es of cancer immunoediting. Annual
 429 review of immunology. 2004; 22: 329-60.
- 430 11. McGranahan N, Furness AJ, Rosenthal R, Ramskov S, Lyngaa R, Saini SK, et al.
 431 Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint
 432 blockade. Science. 2016; 351: 1463-9.
- 433 12. Angelova M, Mlecnik B, Vasaturo A, Bindea G, Fredriksen T, Lafontaine L, et al.
 434 Evolution of Metastases in Space and Time under Immune Selection. Cell. 2018; 175:
 435 751-65.e16.
- 436 13. Zhang AW, McPherson A, Milne K, Kroeger DR, Hamilton PT, Miranda A, et al.
 437 Interfaces of Malignant and Immunologic Clonal Dynamics in Ovarian Cancer. Cell.
 438 2018; 173: 1755-69.e22.
- 439 14. Jimenez-Sanchez A, Memon D, Pourpe S, Veeraraghavan H, Li Y, Vargas HA, et
 440 al. Heterogeneous Tumor-Immune Microenvironments among Differentially Growing
 441 Metastases in an Ovarian Cancer Patient. Cell. 2017; 170: 927-38.e20.
- 442 15. Shain AH, Bagger MM, Yu R, Chang D, Liu S, Vemula S, et al. The genetic
 443 evolution of metastatic uveal melanoma. Nature genetics. 2019; 51: 1123-30.
- Menzies AM, Yeh I, Botton T, Bastian BC, Scolyer RA, Long GV. Clinical activity
 of the MEK inhibitor trametinib in metastatic melanoma containing BRAF kinase fusion.
 Pigment cell & melanoma research. 2015; 28: 607.
- 447 17. Sanborn JZ, Chung J, Purdom E, Wang NJ, Kakavand H, Wilmott JS, et al.
 448 Phylogenetic analyses of melanoma reveal complex patterns of metastatic
 449 dissemination. Proceedings of the National Academy of Sciences. 2015; 112: 10995450 1000.
- 18. Favero F, Joshi T, Marquard AM, Birkbak NJ, Krzystanek M, Li Q, et al.
 Sequenza: allele-specific copy number and mutation profiles from tumor sequencing
 data. Ann Oncol. 2015; 26: 64-70.
- 454 19. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, et al.
 455 Sensitive detection of somatic point mutations in impure and heterogeneous cancer
 456 samples. Nature biotechnology. 2013; 31: 213-9.
- 457 20. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, et al. VarScan 2:
 458 somatic mutation and copy number alteration discovery in cancer by exome sequencing.
 459 Genome research. 2012; 22: 568-76.
- 460 21. Kim S, Scheffler K, Halpern AL, Bekritsky MA, Noh E, Kallberg M, et al. Strelka2:
 461 fast and accurate calling of germline and somatic variants. Nature methods. 2018; 15:
 462 591-4.

- 463 22. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, et 464 al. Integrative genomics viewer. Nature biotechnology. 2011; 29: 24-6.
- 465 23. Thorvaldsdottir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): 466 high-performance genomics data visualization and exploration. Briefings in 467 bioinformatics. 2013; 14: 178-92.
- 468 24. Jamal-Hanjani M, Wilson GA, McGranahan N, Birkbak NJ, Watkins TBK, Veeriah
 469 S, et al. Tracking the Evolution of Non-Small-Cell Lung Cancer. The New England
 470 journal of medicine. 2017; 376: 2109-21.
- 471 25. Schliep KP. phangorn: phylogenetic analysis in R. Bioinformatics (Oxford, 472 England). 2011; 27: 592-3.
- 473 26. Lundegaard C, Lamberth K, Harndahl M, Buus S, Lund O, Nielsen M. NetMHC474 3.0: accurate web accessible predictions of human, mouse and monkey MHC class I
 475 affinities for peptides of length 8-11. Nucleic acids research. 2008; 36: W509-12.
- 476 27. Lundegaard C, Lund O, Nielsen M. Accurate approximation method for prediction
 477 of class I MHC affinities for peptides of length 8, 10 and 11 using prediction tools trained
 478 on 9mers. Bioinformatics (Oxford, England). 2008; 24: 1397-8.
- A79 28. Nielsen M, Lundegaard C, Blicher T, Lamberth K, Harndahl M, Justesen S, et al.
 A80 NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and
 B locus protein of known sequence. PloS one. 2007; 2: e796.
- 482 29. Turker MS. The establishment and maintenance of DNA methylation patterns in 483 mouse somatic cells. Seminars in cancer biology. 1999; 9: 329-37.
- 484 30. Chakravarthy A, Furness A, Joshi K, Ghorani E, Ford K, Ward MJ, et al. Pan-485 cancer deconvolution of tumour composition using DNA methylation. Nature 486 communications. 2018; 9: 3220.
- 487 31. Robins HS, Campregher PV, Srivastava SK, Wacher A, Turtle CJ, Kahsai O, et
 488 al. Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T
 489 cells. Blood. 2009; 114: 4099-107.
- 32. Carlson CS, Emerson RO, Sherwood AM, Desmarais C, Chung MW, Parsons
 JM, et al. Using synthetic templates to design an unbiased multiplex PCR assay. Nature
 communications. 2013; 4: 2680.
- 33. Robins H, Desmarais C, Matthis J, Livingston R, Andriesen J, Reijonen H, et al.
 Ultra-sensitive detection of rare T cell clones. Journal of immunological methods. 2012;
 375: 14-9.
- 496 34. Ding L, Getz G, Wheeler DA, Mardis ER, McLellan MD, Cibulskis K, et al.
 497 Somatic mutations affect key pathways in lung adenocarcinoma. Nature. 2008; 455:
 498 1069-75.
- 499 35. Imielinski M, Berger AH, Hammerman PS, Hernandez B, Pugh TJ, Hodis E, et al.
 500 Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. Cell.
 501 2012; 150: 1107-20.
- 36. Govindan R, Ding L, Griffith M, Subramanian J, Dees ND, Kanchi KL, et al.
 Genomic landscape of non-small cell lung cancer in smokers and never-smokers. Cell.
 2012; 150: 1121-34.
- 505 37. Tumeh PC, Harview CL, Yearley JH, Shintaku IP, Taylor EJ, Robert L, et al. PD-506 1 blockade induces responses by inhibiting adaptive immune resistance. Nature. 2014; 507 515: 568-71.

38. Cha E, Klinger M, Hou Y, Cummings C, Ribas A, Faham M, et al. Improved
survival with T cell clonotype stability after anti-CTLA-4 treatment in cancer patients.
Science translational medicine. 2014; 6: 238ra70.

39. Reuben A, Gittelman R, Zhang J, Chen R, Quek K, Vence L, et al. OA 13.05
Immune, Molecular and T Cell Repertoire Landscape of 235 Resected Non-Small Cell
Lung Cancers and Paired Normal Lung Tissues. Journal of Thoracic Oncology. 2017;
12: S1780.

40. Gil Del Alcazar CR, Alečković M, Polyak K. Immune Escape during Breast Tumor Progression. Cancer immunology research. 2020; 8: 422-7.

- 517 41. Walker R, Poleszczuk J, Pilon-Thomas S, Kim S, Anderson A, Czerniecki BJ, et 518 al. Immune interconnectivity of anatomically distant tumors as a potential mediator of 519 systemic responses to local therapy. Scientific reports. 2018; 8: 9474.
- 42. Reuben A, Gittelman R, Gao J, Zhang J, Yusko EC, Wu C-J, et al. TCR repertoire intratumor heterogeneity in localized lung adenocarcinomas: an association with predicted neoantigen heterogeneity and postsurgical recurrence. Cancer discovery. 2017; 7: 1088-97.
- 43. Baras AS, Drake C, Liu J-J, Gandhi N, Kates M, Hoque MO, et al. The ratio of CD8 to Treg tumor-infiltrating lymphocytes is associated with response to cisplatinbased neoadjuvant chemotherapy in patients with muscle invasive urothelial carcinoma of the bladder. Oncoimmunology. 2016; 5: e1134412-e.
- 44. Jordanova ES, Gorter A, Ayachi O, Prins F, Durrant LG, Kenter GG, et al. Human leukocyte antigen class I, MHC class I chain-related molecule A, and CD8+/regulatory T-cell ratio: which variable determines survival of cervical cancer patients? Clinical Cancer Research. 2008; 14: 2028-35.
- 45. Qi Q, Liu Y, Cheng Y, Glanville J, Zhang D, Lee JY, et al. Diversity and clonal selection in the human T-cell repertoire. Proceedings of the National Academy of Sciences of the United States of America. 2014; 111: 13139-44.
- 535 46. Labelle M, Hynes RO. The initial hours of metastasis: the importance of 536 cooperative host-tumor cell interactions during hematogenous dissemination. Cancer 537 discovery. 2012; 2: 1091-9.
- 538 47. Kitamura T, Qian B-Z, Pollard JW. Immune cell promotion of metastasis. Nature 539 Reviews Immunology. 2015; 15: 73-86.
- 48. Robert L, Tsoi J, Wang X, Emerson R, Homet B, Chodon T, et al. CTLA4
 blockade broadens the peripheral T-cell receptor repertoire. Clinical cancer research :
 an official journal of the American Association for Cancer Research. 2014; 20: 2424-32.
- 49. Reuben A, Gittelman R, Zhang J, Quek K, Vence LM, Behrens C, et al.
 Association of the T-cell receptor landscape with survival in non-small cell lung cancer.
 Journal of Clinical Oncology. 2018; 36: 140-.
- 546 50. Chiou S-H, Tseng D, Reuben A, Mallajosyula V, Molina IS, Conley S, et al. 547 Global analysis of shared T cell specificities in human non-small cell lung cancer 548 enables HLA inference and antigen discovery. Immunity. 2021; 54: 586-602. e8.
- 549 51. Reuben A, Zhang J, Chiou SH, Gittelman RM, Li J, Lee WC, et al. 550 Comprehensive T cell repertoire characterization of non-small cell lung cancer. Nat 551 Commun. 2020; 11: 603.

552 52. Simoni Y, Becht E, Fehlings M, Loh CY, Koo S-L, Teng KWW, et al. Bystander 553 CD8+ T cells are abundant and phenotypically distinct in human tumour infiltrates. 554 Nature. 2018; 557: 575-9.

555 53. Scheper W, Kelderman S, Fanchi LF, Linnemann C, Bendle G, de Rooij MA, et al. 556 Low and variable tumor reactivity of the intratumoral TCR repertoire in human cancers. 557 Nature medicine. 2019; 25: 89-94.

558 54. Zhang J, Fujimoto J, Zhang J, Wedge DC, Song X, Zhang J, et al. Intratumor 559 heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. 560 Science. 2014; 346: 256-9.

- 561 55. Reuben A, Gittelman R, Gao J, Zhang J, Yusko EC, Wu CJ, et al. TCR 562 Repertoire Intratumor Heterogeneity in Localized Lung Adenocarcinomas: An 563 Association with Predicted Neoantigen Heterogeneity and Postsurgical Recurrence. 564 Cancer Discov. 2017; 7: 1088-97.
- 565 56. Lee WC, Reuben A, Hu X, McGranahan N, Chen R, Jalali A, et al. Multiomics 566 profiling of primary lung cancers and distant metastases reveals immunosuppression as 567 a common characteristic of tumor cells with metastatic plasticity. Genome Biol. 2020; 21: 568 271.
- 569 57. Hong L, Negrao MV, Dibaj SS, Chen R, Reuben A, Bohac JM, et al. 570 Programmed Death-Ligand 1 Heterogeneity and Its Impact on Benefit From Immune 571 Checkpoint Inhibitors in NSCLC. Journal of thoracic oncology : official publication of the 572 International Association for the Study of Lung Cancer. 2020; 15: 1449-59.
- 573 58. Liu Y, Zhang J, Li L, Yin G, Zhang J, Zheng S, et al. Genomic heterogeneity of 574 multiple synchronous lung cancer. Nature communications. 2016; 7: 1-8.
- 575 59. Kim R, Keam B, Kim S, Kim M, Kim SH, Kim JW, et al. Differences in tumor 576 microenvironments between primary lung tumors and brain metastases in lung cancer 577 patients: therapeutic implications for immune checkpoint inhibitors. BMC cancer. 2019; 578 19: 1-10.
- 579 60. Vignot S, Frampton GM, Soria JC, Yelensky R, Commo F, Brambilla C, et al.
 580 Next-generation sequencing reveals high concordance of recurrent somatic alterations
 581 between primary tumor and metastases from patients with non-small-cell lung cancer.
 582 Journal of clinical oncology : official journal of the American Society of Clinical Oncology.
 583 2013; 31: 2167-72.
- 584 61. Zhang J, Fujimoto J, Zhang J, Wedge DC, Song X, Zhang J, et al. Intratumor
 585 heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing.
 586 Science (New York, NY). 2014; 346: 256-9.
- 587 62. Leong TL, Gayevskiy V, Steinfort DP, De Massy MR, Gonzalez-Rajal A, Marini 588 KD, et al. Deep multi-region whole-genome sequencing reveals heterogeneity and 589 gene-by-environment interactions in treatment-naive, metastatic lung cancer. Oncogene. 590 2019; 38: 1661-75.
- 591 63. Cheung KJ, Ewald AJ. A collective route to metastasis: Seeding by tumor cell 592 clusters. Science. 2016; 352: 167-9.
- 593 64. Heyde A, Reiter JG, Naxerova K, Nowak MA. Consecutive seeding and transfer 594 of genetic diversity in metastasis. Proceedings of the National Academy of Sciences. 595 2019; 116: 14129-37.

596 65. Reuben A, Spencer CN, Prieto PA, Gopalakrishnan V, Reddy SM, Miller JP, et al. 597 Genomic and immune heterogeneity are associated with differential responses to 598 therapy in melanoma. NPJ Genom Med. 2017; 2.

66. Becerra MF, Reznik E, Redzematovic A, Tennenbaum DM, Kashan M, Ghanaat M, et al. Comparative genomic profiling of matched primary and metastatic tumors in renal cell carcinoma. European urology focus. 2018; 4: 986-94.

602 67. Lim B, Mun J, Kim JH, Kim CW, Roh SA, Cho DH, et al. Genome-wide mutation 603 profiles of colorectal tumors and associated liver metastases at the exome and 604 transcriptome levels. Oncotarget. 2015; 6: 22179-90.

605 68. Ehrlich M. DNA hypomethylation in cancer cells. Epigenomics. 2009; 1: 239-59.

606 69. Park C, Jeong K, Park JH, Jung S, Bae JM, Kim K, et al. Pan-cancer methylation 607 analysis reveals an inverse correlation of tumor immunogenicity with methylation 608 aberrancy. Cancer immunology, immunotherapy : CII. 2021; 70: 1605-17.

70. Rosenthal R, Cadieux EL, Salgado R, Al Bakir M, Moore DA, Hiley CT, et al.
Neoantigen-directed immune escape in lung cancer evolution. Nature. 2019; 567: 47985.

Liu M, Zhou J, Chen Z, Cheng AS. Understanding the epigenetic regulation of
 tumours and their microenvironments: opportunities and problems for epigenetic
 therapy. The Journal of pathology. 2017; 241: 10-24.

615	
616	
617	
618	
619	
620	
621	
622	
623	
624	Declarations
625	Ethics approval

- 626 Collection and use of patient samples were approved by the Institutional Review Board 627 of the University of Texas MD Anderson Cancer Center.
- 628 **Consent for publication**
- 629 Not applicable

630 **Competing interests**

L.A.B. serves on advisory committees for AstraZeneca, AbbVie, GenMab, BergenBio, 631 Pharma Mar SA, Sierra Oncology, Merck, Bristol Myers Squibb, Genentech, and Pfizer 632 and has research support from AbbVie, AstraZeneca, GenMab, Sierra Oncology and 633 634 Tolero Pharmaceuticals. I.W. reports grants and personal fees from Genentech/Roche, Bayer, Bristol-Myers Squibb, AstraZeneca/Medimmune, Pfizer, HTG Molecular, Merck, 635 and Guardant Health; personal fees from GlaxoSmithKline and MSD; grants from 636 Oncoplex, DepArray, Adaptive, Adaptimmune, EMD Serono, Takeda, Amgen, Karus, 637 638 Johnson & Johnson, Iovance, 4D, Novartis, Oncocyte and Akoya. J.J.Z. reports research funding from Merck, Johnson and Johnson, and consultant fees from BMS, 639 Johnson and Johnson, AstraZeneca, Geneplus, OrigMed and Innovent outside the 640 submitted work. J.V.H. reports honorariums from AstraZeneca, Boehringer-Ingelheim, 641 Catalyst, Genentech, GlaxoSmithKline, Guardant Health, Foundation medicine, Hengrui 642 Therapeutics, Eli Lilly, Novartis, Spectrum, EMD Serono, Sanofi, Takeda, Mirati 643 Therapeutics, BMS, BrightPath Biotherapeutics, Janssen Global Services, Nexus 644 Health Systems, EMD Serono, Pneuma Respiratory, Kairos Venture Investments, 645 Roche and Leads Biolabs. A.R. serves on the Scientific Advisory Board and has 646 received honoraria from Adaptive Biotechnologies. The other authors declare no 647 competing interests. 648

649 Author contributions

A.R., D.L.G., L.A.B. and J.J.Z. conceived the study. J.L. and R.C. led the data analysis.
J.F, D.T., and C.W.C. led the pathological assessment, multi-region sample preparation
and DNA extraction. R.C., K.Q. and M.T. collected resected specimens and clinical data.
L.L. and C.G. performed DNA preparation and whole-exome sequencing. X.S. and

J.H.Z. performed sequencing raw data processing. J.L., X.H., K.Q., M.T. and A.M.
performed downstream bioinformatics analyses. R.C., C.B, P.J., J.V.H., I.W., P.A.F.,
D.L.G, L.A.B., J.J.Z. and A.R. interpreted the data for clinical and pathological
correlation. R.C. and J.L. performed statistical analyses. R.C., J.J.Z and A.R. wrote the
paper. All authors edited the manuscript.

659 Acknowledgements

This work was supported by Conquer Cancer Foundation ASCO Young Investigator 660 Award, MD Anderson Physician Scientist Award, Cancer Prevention & Research 661 Institute of Texas (CPRIT) Multiple Investigator Award, TJ Martell Foundation, NIH/NCI 662 663 R01-CA207295, NIH/NCI U01-CA213273 and Department of Defense (LC170171). A.R. is supported by the Exon 20 Group, Rexanna's Foundation for Fighting Lung Cancer, 664 the Waun Ki Hong Lung Cancer Research Fund, MD Anderson's Lung 665 Cancer Moon Shot and the University Cancer Foundation via the Institutional Research 666 Grant program at the University of Texas MD Anderson Cancer Center. 667

668

669 Figure legends

Figure 1. Synchronous metastatic tumors exhibit heterogeneous growth and somatic mutation and neoantigen patterns. a) Anatomical map of representative biospecimen collection sites in the patient. b) Representative CT and MRI scans of different resected tumors: primary lung tumor, cervical lymph node enlargement, liver metastasis and C5 spine compression fracture by metastasis. c) Non-silent mutation counts in tumors. d) Fraction of shared and unique non-silent mutations across tumors.

Figure 2. Characterization of TCR repertoire metrics across tumors. T-cell a)
 density, b) richness and c) clonality. d) Correlation between T-cell fraction, richness and
 clonality. e) Distribution of most prevalent TCR clonotype.

Figure 3. T-cell repertoire heterogeneity is observed across differentially growing
 tumors. a) Number of T-cell clonotypes in the primary tumor (red), metastases (blue) or

shared (purple). b) Shared T-cell clonotype proportions and c) frequencies between the
 primary tumor and metastases.

Figure 4. Evolution of synchronous metastases may be shaped by the T cell
repertoire. a) Quantification of TCR repertoire heterogeneity across tumors by Jaccard
index (JI). b) Quantification of predicted neoantigen heterogeneity across tumors by JI.
c) Correlation between TCR repertoire JI and predicted neoantigen JI. d) Deconvolution
of immune components and T cell subpopulations by MethylCIBERSORT. e)
Correlation between T-cell richness and estimated CD8⁺ T-cell fraction. f) Correlation

Figure 5. Methylated neoantigen burden is inversely correlated with T cell repertoire metrics. a) Total number of predicted neoantigens across tumors. b) Comparison of predicted neoantigens between TCGA cohort and tumors in our study. Correlation between ratio of methylated neoantigen coding mutations and T-cell c) density, d) richness and e) clonality.

а



b







Figure 2



Ρ

P

Ρ

P

P 2936

a



b





4342

4383

4617

4995

5008

14340

3079 M03

3593 M04

M02

15578

a

d

Fraction

1.0-

0.8

0.4

0.2-

0.0

0.6

Shared TCR Jaccard index 0.6 F M01 M02 M03 M04 M05 M06 0.4 M07 M08 M09 M10 M11 M12 M13 M14 M15 M16 M17 M18 M19 M09 M11 M12 M15 M15 M16 M17 M18 M18 M18 M08 M01 M02 M07

Treg

Monocytes

M03 M19 M05 M08 M02 M07

CD8⁺ T cells Eosinophils

M11 M16 M15 M15 M17 M17 M17 M17 M18 M18

B cells

b



Relative CD8⁺ T cell fraction

Relative CD8⁺ T cell fraction

