Integrative genomic analysis identifies unique immune environments associated with immunotherapy response in diffuse large B cell lymphoma

Sravya Tumuluru[1*], James K. Godfrey[2*], Alan Cooper[3], Jovian Yu[3], Xiufen Chen[3], Brendan W. MacNabb[4], Girish Venkataraman[5], Yuanyuan Zha[3], Benedikt Pelzer[6], Joo Song[7], Gerben Duns[8], Christopher Bolen[9], Elicia Penuel[9], Ekaterina Postovalova[10], Nikita Kotlov[10], Aleksander Bagaev[10], Nathan Fowler[10], Sonali M. Smith[3], Christian Steidl[8,11], Justin Kline[1,3 #]

Affiliations:

[1] Committee on Cancer Biology, University of Chicago, Chicago, IL

[2] Department of Hematology & Hematopoietic Cell Transplantation, City of Hope, Duarte, CA

[3] Department of Medicine, Section of Hematology/Oncology, University of Chicago, Chicago, IL

[4] Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA

[5] Department of Pathology, University of Chicago, Chicago, IL

[6] Division of Hematology and Oncology, Weill Cornell Medicine, New York, NY

[7] Department of Pathology, City of Hope, Duarte, CA

[8] Centre for Lymphoid Cancer, BC Cancer, Vancouver, BC, Canada

[9] Genentech Inc., South San Francisco, CA

[10] BostonGene Corporation, Waltham, MA

[11] Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, BC, Canada

[*] These authors contributed equally

[#] Corresponding author

Corresponding Author:

Justin Kline, M.D.

Department of Medicine, Section of Hematology/Oncology, University of Chicago

900 East 57th Street, Chicago, IL, 60637

Phone: (773) 702-5550; Fax: (773) 702-3163; Email: jkline@bsd.uchicago.edu

Word Count: 5,824

Number of Figures/Tables: 6

Abstract

Most relapsed or refractory diffuse large B cell lymphoma (DLBCL) patients show primary resistance, or relapse after receiving immunotherapies such as CAR T cells, bispecific antibodies, or checkpoint blockade therapy, highlighting the need for a deeper understanding of mechanisms that regulate the immune environment and response to immunotherapy. Therefore, we performed bulk transcriptomic analysis of 862 diagnostic DLBCL specimens and segregated DLBCLs into four unique quadrants defined by cell-of-origin and immune-related gene set expression scores, termed DLBCL-immune quadrants (DLBCL-IQ). These quadrants consisted of activated B cell-like (ABC) hot, ABC cold, germinal center B cell-like (GCB) hot, GCB cold DLBCLs. Unique genomic alterations were enriched in each immune cluster, suggesting that lymphoma-intrinsic alterations modulate the DLBCL immune environment. For instance, *SOCS1* loss-of-function mutations were enriched among GCB hot DLBCLs, identifying a putative subset of inflamed DLBCLs that may be inherently susceptible to immunotherapy. Finally, treatment with the CD20 x CD3 bispecific antibody, mosunetuzumab, was associated with significantly improved progression-free survival of GCB hot DLBCL patients, underscoring the importance of our DLBCL-IQs in predicting bispecific antibody efficacy in this disease.

Significance

This in-depth characterization of DLBCL immune landscapes and associated genomic alterations should catalyze development of rationally designed, combinatorial studies of targeted and immune therapies. Moreover, the association between the immune environment and clinical outcomes to a bispecific antibody is important, and similar analyses can be employed to address how particular

56 DLBCL immune environments influence the efficacy of checkpoint blockade and CAR T cell
57 therapy.

58 ## Introduction

59 Although diffuse large B cell lymphoma (DLBCL) is often cured with R-CHOP (rituximab,
60 cyclophosphamide, doxorubicin, vincristine, and prednisone) chemoimmunotherapy,
61 approximately 40% of DLBCL patients will develop relapsed or refractory (r/r) disease, which is
62 challenging to effectively treat (1,2). Recently, immunotherapies have revolutionized the treatment
63 landscape of r/r DLBCL. For example, CD19-directed chimeric antigen receptor (CAR) T cell
64 therapy induces durable responses in ~30-40% of patients with r/r DLBCL (3–6). Complete
65 responses are also observed in up to 40% of r/r DLBCL patients who receive CD3 x CD20
66 bispecific antibodies (BsAbs) (7–9). However, in many r/r DLBCL patients, BsAbs and CAR T
67 cell therapy either fail completely or confer only fleeting responses. Finally, checkpoint blockade
68 therapy (CBT) with anti-PD-1 antibodies elicits responses in only a small subset (10-15%) of r/r
69 DLBCL patients (10,11). Collectively, these clinical observations indicate that while
70 immunotherapies can be effective against DLBCL, additional research is needed to understand the
71 cellular and molecular features of the DLBCL environment that underlie immunotherapy
72 responsiveness or resistance to enrich for the subset of patients who will derive maximal benefit
73 and to identify strategies to overcome mechanisms driving intrinsic or acquired resistance.

74 The role of the tumor immune environment in mediating response to immunotherapy has
75 been extensively explored in solid cancers, where numerous studies have demonstrated that a "T
76 cell-inflamed" environment identifies cancers against which a spontaneous anti-tumor immune
77 response has been raised and serves as a useful biomarker of response to CBT (12–15).
78 Transcriptional and immune cell profiling of DLBCL has also revealed subsets enriched in
79 immune cell infiltration and in expression of immune-related genes, which indicates that some
80 DLBCLs exhibit evidence of having stimulated an endogenous anti-lymphoma immune response

81 (16–19). However, the critical factors that shape the DLBCL immune environment remain
82 unknown and the impact of the immune environment on promoting or preventing immunotherapy
83 response has yet to be fully defined (10,20).

84 While bulk transcriptional profiling and genetic sequencing have uncovered important
85 insights into the molecular heterogeneity and biology of DLBCL, less is understood regarding the
86 nature and complexity of the DLBCL immune environment. In this study, we utilized an
87 integrative approach to 1) broadly characterize the DLBCL immune landscape, 2) provide insight
88 into the contribution of genomic alterations in shaping the DLBCL immune environment, and 3)
89 define the extent to which transcriptionally based immune clustering of DLBCL informs
90 responsiveness to CD3 x CD20 BsAb immunotherapy. Through these analyses, we uncovered four
91 unique DLBCL immune environments, each enriched for several distinct oncogenic alterations,
92 suggesting that lymphoma cell-intrinsic factors play an important role in orchestrating the immune
93 environment in this disease. For example, we identified loss-of-function (LoF) mutations in
94 *SOCS1* as significantly enriched in DLBCLs with inflamed environments. Moreover, we
95 demonstrate that genetic ablation of *Socs1* renders B cells more sensitive to the effects of T-cell
96 effector cytokines such as interferon-gamma (IFNγ). Finally, we associate the composition of the
97 immune environment with clinical outcomes to BsAb therapy in a retrospective analysis of tumor
98 samples collected from a large phase I/II clinical trial of r/r DLBCL patients.

99 ## Results

100 **Transcriptomic analysis identifies four immune-related DLBCL clusters.**

101 In solid cancers, a "T cell inflamed" gene signature identifies tumors against which a
102 spontaneous anti-tumor immune response has been induced and enriches for a subset of cancers
103 that are particularly vulnerable to immunotherapies, such as CBT (12–15). To determine if
104 DLBCLs could be similarly segregated, we performed bulk transcriptional analyses focused
105 specifically on immunological aspects of the lymphoma environment. Our discovery cohort
106 consisted of bulk transcriptomic data from diagnostic DLBCL biopsies from the National Cancer
107 Institute (NCI, n = 481) (21,22) and an internal dataset from the University of Chicago Medical
108 Center (UCMC, n = 96) (23), comprising a total of 577 cases (*See Methods for details)*. Gene set
109 variation analysis (GSVA) (24) was used to calculate sample-wise relative enrichment scores
110 reflecting the collective expression of genes comprising 19 gene sets that fell into 2 functional
111 groups: 1) immune-related signatures, and 2) COO-related signatures **(Supplementary Table 1)**.
112 Twelve immune-related signatures that reflect the presence or activation state of immune cell
113 subsets were manually curated from an extensive literature search and contained gene sets related
114 to: IFNγ response (12), type I IFN response (15), $CD8^+$ T cell activation (15,25,26) and exhaustion
115 (15), T helper subsets (T_h1 (25), T_h2 (25), T follicular helper (T_{fh}) (25), and regulatory T cells
116 (T_{reg}) (15)), macrophages (15), and dendritic cells (27). COO-related signatures were also included
117 in the GSVA because COO classification segregates DLBCLs into two subtypes - germinal center
118 B-cell-like (GCB) and activated B-cell-like (ABC) - with distinct clinical, transcriptional, and
119 genetic features (1,17,18,28–33). Therefore, we hypothesized that COO would significantly
120 influence the immune environment of DLBCL. To control for potential heterogeneity in DLBCL
121 immune landscapes related to COO, the second category of gene sets included in the GSVA were

122 derived from a validated gene expression-based COO classifier (34), as well as gene sets regulated
123 by IRF4 (35) and BCL6 (35), two transcription factors critical for COO classification. Altogether,
124 577 unique genes were included in the 19 gene sets selected for GSVA (<15% overlap between
125 individual gene sets).

126 GSVA scores for the 577 DLBCLs in the discovery cohort were next subjected to principal
127 component analysis (PCA), which revealed that the majority (69.8%) of variance in gene set
128 expression scores was explained by PC1 (45.38%) and PC2 (24.43%). Unsupervised k-means
129 clustering was performed on the PCA-transformed dataset and the optimal number of clusters was
130 determined to be four **(Supplementary Figure 1A)**. DLBCLs in our discovery cohort were then
131 assigned to one of four specific clusters as described below. Samples from NCI and UCMC
132 datasets were equally represented in each cluster, indicating the absence of significant batch effects
133 **(Supplementary Figure 1B and C)**.

134 Individual immune-related gene set expression scores were highly correlated, trended in
135 the same direction along PC1 **(Figure 1A)**, and contributed minimally to the PC2 axis
136 **(Supplementary Figure 1D and E).** Conversely, COO-related gene sets offered the largest
137 contribution to PC2 and did not contribute significantly to PC1 **(Supplementary Figure 1D and**
138 **E)**. ABC COO-related gene sets (ABCDLBCL-1, ABCDLBCL-2, IRF4Up-7) were highly
139 correlated and positively associated with PC2, while GCB COO-related gene sets (GCBDLBCL-
140 1, GCBDLBCL-2, IRF4Dn-1) were negatively associated with PC2 **(Figure 1A)**. Comparative
141 analyses of the transcriptomes of DLBCLs assigned to the ABC (PC2 > 0) and GCB (PC2 < 0)
142 clusters by GSVA identified 310 differentially expressed genes (|log2FC| >1.0, adj p value < 0.05).
143 In line with previous reports, several genes (*BATF* and *IRF4)*, known to drive classification of
144 ABC DLBCLs using molecular clustering methods, were also upregulated in DLBCLs classified

145 as ABC by our GSVA platform (34). Differentially expressed genes in GSVA-classified GCB
146 DLBCLs (*MYBL1*, *SERPIN9A*, and *LMO2*) were also highly expressed in bona-fide GCB
147 DLBCLs **(Supplementary Figure 1F)** (34). Additionally, unclassified DLBCLs, with
148 transcriptional features intermediate between ABC and GCB subsets, were clustered around PC2
149 ~ 0. This analysis indicated that two clusters within our dataset contained ABC DLBCLs and the
150 other two contained GCB DLBCLs. As an orthogonal validation of our COO-related PC2 axis, the
151 overlap between our GSVA-based COO calls with COO annotations defined through gene
152 expression-based (NCI) and RNAseq-based (UCMC) classifiers was compared (19,34). Molecular
153 COO calls were concordant with GSVA-based COO classifications in >95% of cases, indicating
154 that GSVA accurately clusters DLBCLs based on their true COO (**Figure 1B and C**).

155 As mentioned above, our GSVA model also segregated DLBCLs by distinct immune-
156 related gene set expression scores along PC1. Differential gene expression (DGE) analysis was
157 performed on these putative "hot" (PC1 > 0) and "cold" (PC1 < 0) DLBCLs. "Hot" DLBCLs were
158 characterized by significant upregulation of canonical genes related to immune cell infiltration and
159 activation (e.g. *CD2, CD3D, CD3E, GZMK, IFNG, CD8A, PRF1*) when compared to "cold"
160 DLBCLs **(Supplementary Figure 1G)**. Therefore, we determined that ABC DLBCLs could be
161 transcriptionally defined as "ABC hot" (n = 182, 31.5%, PC1 > 0, PC2 > 0), characterized by high
162 enrichment scores for ABC COO-related and immune-related gene sets, or "ABC cold" (n = 184,
163 31.9%, PC1 < 0, PC2 > 0). Similarly, GCB DLBCLs could be subdivided into "GCB hot" (n =122,
164 21.1%, PC1 >0, PC2 <0) and "GCB cold" (n = 89, 15.4%, PC1 < 0, PC2 < 0) clusters **(Figure**
165 **1D)**. Immune-related and COO-related gene sets were orthogonal and segregated DLBCLs into a
166 quadrant-like structure that we termed DLBCL-immune quadrants (DLBCL-IQ). Interestingly,
167 DLBCLs assigned to the ABC hot cluster were, on average, associated with higher expression

168 scores of immune-related gene sets (PC1 scores) compared to DLBCLs assigned to the GCB hot
169 cluster, suggesting that ABC hot DLBCLs are overall slightly more inflamed than their GCB hot
170 counterparts **(Figure 1E)**. Collectively, these results demonstrate that bulk transcriptional analysis
171 can effectively segregate DLBCLs into four unique clusters, each characterized by a distinct
172 distribution of immune-related (PC1) and COO-related (PC2) scores.

173 When an identical GSVA was performed on an independent DLBCL dataset (British
174 Columbia Cancer, BCC, n = 285) (36,37), a similar clustering pattern emerged **(Supplementary**
175 **Figure 2A)**, with immune-related gene sets contributing to PC1 and COO-related gene sets
176 contributing to PC2 **(Supplementary Figure 2B)**. Therefore, the NCI/UCMC and BCC datasets
177 were combined to power downstream analyses. In the combined dataset (n = 862), DLBCLs from
178 different sources (NCI/UCMC versus BCC) were similarly distributed among each of the four IQs
179 **(Supplementary Figure 2C and D)**, and again, immune-related gene sets contributed to PC1 and
180 COO-related gene sets contributed to PC2 **(Supplementary Figure 2E)**. As expected, the
181 combined dataset again was distributed into four IQs, demonstrating the stability of GSVA-based
182 clustering in classifying DLBCLs by COO and their immune environment **(Supplementary**
183 **Figure 2F)**.

184

185 **Prognostic significance of DLBCL-IQs.**

186 Next, the extent to which our IQs correlated with clinical outcomes to initial treatment with
187 R-CHOP was assessed. Consistent with the known association of COO with survival to R-CHOP
188 chemotherapy, lymphomas assigned to GCB DLBCL-IQs showed improved progression-free
189 survival (PFS) and overall survival (OS) compared to those assigned to ABC DLBCL-IQs (18,30).
190 However, further stratifying within ABC or GCB DLBCLs according to IQ assignment did not

191 impact survival outcomes to R-CHOP therapy **(Supplementary Figure 3A and B)**, indicating that
192 the unique DLBCL immune environments were not associated with differential outcomes to R-
193 CHOP chemotherapy.
194
195 **Validation of transcriptionally defined DLBCL-IQs.**
196 Immune cell deconvolution analysis using CIBERSORTx was performed to determine the
197 degree to which GSVA-based IQs correlated with inferred proportions of various immune cell
198 populations in the DLBCL environment (38). As expected, higher estimated fractions of numerous
199 immune cell subsets were observed among ABC and GCB hot DLBCLs **(Figure 2A)**. Specifically,
200 GCB and ABC hot DLBCLs showed higher inferred proportions of $CD8^+$ T cells **(Figure 2B)**,
201 $CD4^+$ T cells **(Figure 2C)**, $FOXP3^+$ T_{regs} **(Supplementary Figure 4A)**, and macrophages
202 **(Supplementary Figure 4B)**, when compared to GCB and ABC cold DLBCLs.
203 Next, to confirm that DLBCL-IQs accurately reflect the actual immune cell composition
204 of specific lymphoma environments, multispectral immunofluorescence (mIF) was performed on
205 a subset of DLBCLs (n = 65, UCMC) for which paired RNA-seq data were available. Staining was
206 performed using fluorescently labeled antibodies against T cell markers (CD4 and CD8),
207 macrophage markers (CD68), DAPI, and B cells (PAX5). Macrophage ($CD68^+$) to DLBCL
208 ($PAX5^+$) cell ratios were similar across immune clusters **(Supplementary Figure 4C)** and were
209 not significantly correlated with sample-wise hot/cold (PC1) axis scores **(Supplementary Figure**
210 **4D)**. However, there were significant differences in T cell infiltration among the four DLBCL-
211 IQs. Representative images of a hot and cold DLBCL are shown in **Figure 2D**. DLBCLs
212 transcriptionally clustered by GSVA as ABC or GCB hot were characterized by significantly
213 higher ratios of $CD8^+$ T cells to DLBCL cells when compared to ABC and GCB cold counterparts

214 **(Figure 2E)**. Furthermore, $CD8^+$ T cell to DLBCL cell ratios were significantly correlated with
215 sample-wise hot/cold (PC1) axis scores **(Figure 2F)**. The same was true for $CD4^+$ T cell to DLBCL
216 cell ratios (**Figure 2G and H**). These data validate the accuracy of GSVA-based clustering in
217 classifying DLBCLs as harboring hot or cold immune environments.
218
219 **Immune environments of molecularly-defined DLBCL subtypes.**
220 Recently, several novel DLBCL subsets and their associated immune environments have
221 been described. For example, we previously demonstrated that *PD-L1* gene alterations (copy gains,
222 amplifications, translocations), identified by fluorescent *in-situ* hybridization (FISH) in ~20% of
223 DLBCLs, were associated with a non-GCB COO and a T cell inflamed environment (23).
224 Therefore, we anticipated our GSVA model would identify *PD-L1* gene-altered DLBCLs as being
225 enriched in the ABC hot cluster. Since *PD-L1* FISH was not available for all datasets,
226 amplifications in the *PD-L1* locus were identified through analysis of paired whole exome
227 sequencing (WES) (UCMC and BCC) or comparative genomic hybridization data (NCI). The
228 analysis was restricted to identification of overt *PD-L1* amplifications – defined as ≥ 4 copies – to
229 reduce the likelihood of false-positive and -negative calls. *PD-L1* amplifications were found in
230 4.7% of all DLBCLs, and were enriched among ABC hot cases, confirming our hypothesis.
231 However, given the low incidence of *PD-L1* gene-amplified DLBCLs, this result did not meet
232 statistical significance **(Figure 3A and B)**.
233 Dark zone signature positive ($DZsig^{pos}$) DLBCLs – formerly double hit signature positive
234 ($DHITsig^+$) DLBCLs – are a transcriptionally defined, high-risk GCB DLBCL subset associated
235 with a paucity of infiltrating immune cells (37,39). As expected, when the DZsig was applied to
236 our DLBCL immune clusters, $DZsig^{pos}$ DLBCLs were significantly enriched in the GCB cold

237 cluster. **(Figure 3C and D)**. Thus, these data suggest GSVA-based clustering accurately identifies
238 immune- and COO-associated features of established DLBCL subtypes, such as those harboring
239 *PD-L1* gene amplifications and those expressing the DZsig.

240 Kotlov and colleagues recently classified DLBCLs based on functional gene expression
241 signatures (F^{GES}) that reflected the relative abundance and functional features of cellular
242 constituents (19). This analysis also identified four DLBCL clusters, termed "lymphoma
243 microenvironments" (LMEs), including: 1) mesenchymal (LME-ME), characterized by high
244 abundance of stromal and extracellular matrix components; 2) germinal center-like (LME-GC),
245 characterized by presence of cells found in germinal centers; 3) inflammatory (LME-IN),
246 harboring an inflammatory milieu; 4) depleted (LME-DE), characterized by decreased expression
247 of F^{GES} scores related to immune cells in the environment. To determine the concordance between
248 LME subtypes and our GSVA-defined immune clusters, LME cluster annotations were applied to
249 DLBCLs in our combined NCI/UCMC/BCC dataset. We found that the ABC hot cluster
250 overlapped significantly with an inflammatory microenvironment (LME-IN) **(Figure 3E)**. ABC
251 hot DLBCLs exhibited high F^{GES} scores for several gene sets that reflect the presence and
252 activation of T cell subsets, including tumor infiltrating lymphocytes (TILs), T_{reg}, T cell traffic,
253 and JAK-STAT signaling **(Supplementary Figure 5)**. GCB and ABC cold DLBCLs were
254 characterized by low expression of T cell and myeloid cell-related F^{GES} scores but demonstrated
255 higher F^{GES} scores for B cell signaling when compared to hot DLBCLs **(Supplementary Figure**
256 **5)**, and strongly overlapped with LME-DE **(Figure 3E)**. LME-GC and LME-ME, however, did
257 not overlap significantly with any GSVA-defined immune clusters **(Figure 3E)**. Collectively,
258 these results suggest that LME and our GSVA clustering capture different aspects of the lymphoma
259 microenvironment.

260 **Genomic features associated with DLBCL-IQs.**

261 Emerging evidence indicates that specific oncogenic alterations and associated
262 transcriptional programs can significantly impact the composition of the tumor immune
263 environment and vulnerability to immunotherapies in solid tumors (40–48). Therefore, we
264 analyzed paired WES for DLBCLs in the combined cohort to identify genetic features associated
265 with each of our GSVA-based immune-related clusters.

266 In DLBCL, several landmark studies have developed genomic approaches to describe new
267 DLBCL subtypes based on the presence of specific co-occurring genetic alterations (21,29,49).
268 One such analysis classified DLBCLs into seven subtypes based on distinct genetic features using
269 a probabilistic classifier called LymphGen (21). However, the association of specific LymphGen
270 clusters with distinct immune environments has not been fully elucidated. LymphGen cluster
271 annotations were available for 754 DLBCLs in our combined dataset. Broadly speaking,
272 LymphGen classifications overlapped imperfectly with our DLBCL-IQs **(Figure 4A)**, with a few
273 notable exceptions. For example, DLBCLs characterized by activating *NOTCH1* mutations
274 (LymphGen N1 subtype), were significantly enriched in the ABC hot cluster, while DLBCLs with
275 *TP53* loss and aneuploidy (LymphGen A53 subtype), were almost exclusively found in the ABC
276 cold cluster. Lastly, DLBCLs with gain-of-function mutations in *EZH2*, *BCL2* translocations, and
277 *MYC* activation (LymphGen EZB-MYC^{+} subtype), were enriched in the GCB cold cluster.
278 Compared to other LymphGen subtypes, A53 and EZB-MYC^{+} DLBCLs have been shown to
279 exhibit low expression of signatures related to immune cell subsets, including T cells,
280 macrophages, and DCs, consistent with their assignment to ABC and GCB cold IQs, respectively
281 (21). The remaining four LymphGen subtypes showed little correlation with specific GSVA-based

282 immune-related clusters, and LymphGen unclassified DLBCLs (~40% of all DLBCLs) were
283 broadly distributed among all four IQs **(Figure 4A)**.

284 To examine whether individual mutations or copy number alterations (CNAs) are
285 associated with particular DLBCL immune environments, we compared the genetic landscapes of
286 DLBCLs assigned to each of the four GSVA-defined IQs (n = 842). The analysis was restricted to
287 essential driver genes to identify those most relevant to DLBCL biology **(Supplementary Table**
288 **2)** (28,29). Comparative analysis of hot DLBCLs versus cold DLBCLs independent of COO
289 identified several genetic lesions associated with cold immune environments, including *FOXO1*,
290 *MYD88,* and *TMEM30A*. Comparatively fewer genetic alterations were significantly enriched in
291 hot DLBCLs, but included mutations in *SOCS1*, *TNFRSF14,* and *CD274* **(Supplementary Figure**
292 **6A)**. As ABC and GCB DLBCLs arise in the context of distinct genetic alterations and oncogenic
293 pathways, we next compared oncogenic alterations specifically enriched in ABC hot versus ABC
294 cold DLBCLs, and GCB hot versus GCB cold DLBCLs. Alterations in genes such as *MYD88,*
295 *KLHL14, CARD11,* and *TMEM30A,* known to be potent drivers of chronic BCR signaling, were
296 strongly enriched in ABC cold DLBCLs (50–52) **(Figure 4B)**. Alterations in *CD274* (*PD-L1*) were
297 significantly enriched among ABC hot DLBCLs **(Figure 4B)**, concordant with our previous work
298 (23). Mutations in *RB1*, *FOXO1*, and *GNA13* were enriched in the GCB cold cluster, while those
299 in *TNFAIP3* and *SOCS1* were significantly associated with the GCB hot cluster **(Figure 4C)**.

300 To gain insight into the impact of oncogenic signaling pathways on DLBCL immune
301 environments, mutations and CNAs were grouped into functional pathways. Genetic alterations in
302 the BCR-dependent NF-κB signaling pathway were significantly enriched in ABC cold DLBCLs
303 compared to ABC hot DLBCLs **(Supplementary Figure 6B)** (52,53). ABC and GCB cold
304 DLBCLs were both significantly enriched for alterations in genes involved in p53 signaling and

305 cell cycle progression compared to ABC and GCB hot DLBCLs, respectively **(Supplementary**
306 **Figure 6C)** (54).

307 Finally, gene set enrichment analysis (GSEA) showed upregulation of immune-related
308 gene sets, including IFNγ targets, IFNα response genes, and IL2-STAT5 signaling pathway genes
309 in ABC hot and GCB hot clusters (**Figure 4D**). In contrast, G2M target gene sets were enriched in
310 ABC and GCB cold DLBCL clusters, suggesting dysregulated lymphoma cell cycling might play
311 a role in orchestrating cold DLBCL immune environments **(Figure 4D and E)**. GSEA also
312 demonstrated that ABC and GCB cold DLBCLs were significantly enriched for expression of *MYC*
313 target genes compared to ABC and GCB hot DLBCLs (**Figure 4D and F)**. Taken together, these
314 data show that several genetic lesions and oncogenic pathways are recurrently associated with
315 specific GSVA-defined DLBCL-IQs, suggesting that lymphoma cell-intrinsic alterations may
316 impact the immune environment in DLBCL.

317

318 **MYC activation is strongly associated with cold DLBCL immune environments.**

319 Because GSEA showed striking upregulation of MYC targets in cold DLBCLs, the role of
320 MYC in orchestrating a cold immune environment was further interrogated. We devised a MYC
321 GSVA score using genes from a previously published target gene signature (MycUp-4) to assign
322 DLBCLs into "MYC-High," "MYC-Low", or "MYC-Int" groups (55). IHC for MYC expression
323 was performed on DLBCLs in the BCC cohort with paired transcriptomic data (n = 266). DLBCLs
324 with high MYC protein expression – defined as ≥ 40% of lymphoma cells expressing MYC by
325 IHC staining – were consistently classified as MYC-high transcriptionally (MYC-High).
326 Conversely, DLBCLs identified as MYC-Low transcriptionally were characterized by low MYC

327 protein expression. Thus, expression of the transcriptional MYC signature accurately identifies
328 DLBCLs with high versus low MYC protein expression **(Supplementary Figure 7A)**.
329 ABC cold and GCB cold DLBCLs showed significantly higher expression of the MYC
330 activation signature compared to ABC hot or GCB hot DLBCLs **(Supplementary Figure 7B)** and
331 MYC-High DLBCLs were equally distributed between ABC cold (40%) and GCB cold (40%)
332 clusters **(Supplementary Figure 7C)**. Moreover, compared to their MYC-Low counterparts,
333 MYC-High DLBCLs in both ABC and GCB clusters contained significantly lower inferred
334 proportions of immune cell subsets when analyzed by CIBERSORTx **(Supplementary Figure**
335 **7D-F)**, suggesting a negative correlation between MYC activation and an inflamed environment.
336 DLBCLs in the MYC-High group were characterized by lower $CD8^+$ T cell to DLBCL cell ratios
337 compared to MYC-Low DLBCLs when assessed by mIF **(Supplementary Figure 7G)**. $CD8^+$ T
338 cell to DLBCL cell ratios were also negatively correlated with a sample wise MYC GSVA score
339 **(Supplementary Figure 7H)**. Similar findings were also observed for $CD4^+$ T cells
340 **(Supplementary Figure 7I and J)**.
341 The molecular mechanisms underlying MYC activation differ in ABC and GCB DLBCLs
342 (56–58). For example, *MYC* translocations to the immunoglobulin heavy chain (*IGH*) locus are
343 commonly found in GCB DLBCLs, while in ABC DLBCLs, MYC is more commonly activated
344 downstream of other oncogenic events such as BCR-dependent NF-κB signaling (56–58).
345 Interestingly, ABC cold DLBCLs had significantly higher MYC signature scores when compared
346 to GCB cold DLBCLs **(Supplementary Figure 7B)**, suggesting that ABC DLBCLs use distinct
347 mechanisms to drive more potent *MYC* activation. Comparative analysis of the genomes of GCB
348 MYC-High and GCB MYC-Low groups showed MYC-High GCB DLBCLs were enriched for
349 alterations in *MYC* and *DDX3X*, consistent with published reports of the oncogenic profiles of

350 DZsig^{pos} and EZB-MYC^{+} DLBCLs which share considerable overlap with our GSVA-defined
351 GCB MYC-High DLBCL group **(Supplementary Figure 7K)** (21,37,39). Conversely, ABC
352 MYC-High DLBCLs showed recurrent alterations in *MYD88^{L265P}, CD79B,* and *PRDM1* compared
353 to ABC MYC-Low DLBCLs **(Supplementary Fig 7L).** Interestingly, several genes enriched in
354 ABC *MYC*-high DLBCLs fell into the "BCR-dependent NF-κB" pathway, suggesting that
355 activation of this signaling pathway may drive MYC activity in ABC DLBCLs and promote a cold
356 immune environment (52,53).
357
358 ***SOCS1* mutations are enriched among GCB hot DLBCLs and enhance B-cell sensitivity to**
359 **IFNγ.**
360 We were particularly intrigued by the strong association between GCB hot DLBCLs and
361 putative LoF *SOCS1* mutations as SOCS1 is also frequently mutated in classical Hodgkin
362 lymphoma (cHL) (59–61) and primary mediastinal B cell lymphoma (PMBL) (59,62–64) – two
363 lymphoma subtypes associated with inflamed immune environments and sensitivity to CBT
364 (10,20,65–68). SOCS1 is a tumor suppressor and potent negative regulator of JAK/STAT
365 signaling, and LoF mutations in *SOCS1* have been implicated in enhancing oncogenic JAK/STAT
366 signaling in lymphoma downstream of cytokines such as IL-4 and IL-13 (69–73). However, as T
367 cell effector cytokines such as IFNγ also signal through the JAK/STAT pathway, we hypothesized
368 that LoF *SOCS1* mutations may render DLBCLs more sensitive to the effects of IFNγ (74–76).
369 *SOCS1* LoF alterations were identified in 16.2% of DLBCLs in our combined cohort and
370 were significantly associated with a GCB COO **(Figure 5A and B**). As expected, the incidence of
371 *SOCS1* mutations was significantly higher in the GCB hot cluster (30.3%) compared to other
372 clusters **(Figure 5A and B)**. *SOCS1* mutations were often missense or nonsense (82.3%) and

373 occurred with similar frequencies across the *SOCS1* gene **(Figure 5B),** consistent with the role of
374 SOCS1 as a tumor suppressor. Immune cell deconvolution revealed an increased inferred fraction
375 of several immune cell subsets, including $CD8^+$ T cells and $CD4^+$ T cells, in *SOCS1* mutant
376 compared to *SOCS1* wildtype GCB DLBCLs suggesting that GCB DLBCLs with *SOCS1*
377 mutations are more inflamed than their *SOCS1* wildtype counterparts **(Figure 5C)**. Moreover,
378 among GCB DLBCLs, GSEA demonstrated increased expression of IFNγ and IFNα signaling
379 pathway associated genes in *SOCS1* mutant DLBCLs compared to *SOCS1* WT DLBCLs **(Figure**
380 **5D)**.

381 To directly test the hypothesis that loss of Socs1 function would render B cells more
382 sensitive to IFNγ, splenocytes from *Socs1*-deficient ($Cd19^{Cre/+}Socs1^{fl/fl}$) or *Socs1*-sufficient mice
383 ($Cd19^{Cre/+}$) were subjected to IFNγ stimulation. Representative gating of $CD19^+$ B cells and $CD3^+$
384 T cells is shown in **Supplementary Figure 8A**. As hypothesized, IFNγ response genes, such as
385 MHC-II, and PD-L1, were significantly more inducible by IFNγ in a dose-dependent manner in
386 splenic $CD19^+$ B cells from mice lacking *Socs1* compared to wildtype controls **(Figure 5E,**
387 **Supplementary Figure 8B)**. This effect was specific to $CD19^+$ cells in which *Socs1* had been
388 deleted, as MHC-II and PD-L1 were similarly inducible in $CD3^+$ T cells isolated from spleens of
389 conditional *Socs1*-sufficient and *Socs1*-deficient mice **(Supplementary Figure 8C)**. These data
390 demonstrate that DLBCLs with *SOCS1* mutations represent a novel subset of inflamed lymphomas
391 that may be particularly vulnerable to T cell-based immunotherapies due to enhanced IFNγ
392 sensitivity.

393

394 **DLBCL-IQs are associated with differential benefit to BsAb therapy.**

395 The presence of activated T cells in the tumor environment has been associated with
396 response to CBT in solid cancers (12–15,77,78). However, the extent to which the local immune
397 environment impacts the effectiveness of immunotherapies, such as BsAbs, has not been well
398 characterized in DLBCL. Therefore, we sought to define the degree to which pre-existing DLBCL
399 immune environments correlated with clinical benefit following treatment with mosunetuzumab,
400 a CD3 x CD20 BsAb.

401 First, using the immune- and COO-related gene sets described above, GSVA was
402 performed on transcriptomes of pre-treatment r/r DLBCL biopsies (n=74) from patients enrolled
403 on a phase I/II dose-escalation study of mosunetuzumab (7). Similar to our training and validation
404 datasets, PCA again revealed that PC1 reflected an immune-related axis and PC2 reflected a COO-
405 related axis **(Figure 6A)**. We also observed that r/r DLBCLs were segregated into the same
406 quadrant structure identified in our training and validation cohorts that were derived from
407 treatment-naïve DLBCLs. Each IQ was of a relatively similar size (ABC cold, n = 21, 28.3%; ABC
408 hot, n = 18, 24.3%; GCB cold, n = 19, 25.7%; GCB hot, n = 16, 21.6%), and was defined by the
409 distribution of PC1 and PC2 scores **(Figure 6B)**. The identification of comparable IQs in this r/r
410 dataset suggests that immune environments associated with r/r DLBCL may are similar to those
411 in the treatment naïve setting.

412 Next, we addressed whether GSVA-defined DLBCL-IQs were associated with differential
413 outcomes to mosunetuzumab treatment. Independent of COO, r/r DLBCLs characterized as "hot"
414 (n = 37) demonstrated significantly improved PFS with mosunetuzumab treatment (HR = 0.74
415 (0.59-0.94), p = 0.012) compared to those classified as "cold" (n = 37) **(Figure 6C)**. Outcomes to
416 mosunetuzumab treatment were next compared between ABC hot and ABC cold, as well as GCB
417 hot and GCB cold DLBCLs. Interestingly, patients with GCB hot DLBCLs demonstrated

418 significantly improved PFS compared to those with GCB cold DLBCLs (HR = 0.31 (0.14-0.67),
419 p = 0.0033) **(Figure 6D).** Patients with GCB cold DLBCLs demonstrated extremely poor outcomes
420 after treatment with mosunetuzumab, with a median PFS of only 2.3 months. Conversely, PFS
421 following mosunetuzumab treatment was similar among patients with ABC hot and ABC cold
422 DLBCLs **(Supplementary Figure 9A)**. Taken together, these data indicate that the pre-treatment
423 immune environment can significantly influence response to mosunetuzumab immunotherapy in
424 patients with r/r GCB DLBCL.
425
426
427
428
429
430
431
432
433
434
435
436
437

438 **Discussion**

439 Here, we curated immune- and COO-related gene sets and performed GSVA to analyze
440 bulk transcriptomes of 862 diagnostic DLBCL samples across three large datasets, which
441 effectively segregated DLBCLs into four immune-related clusters: GCB cold, GCB hot, ABC cold,
442 and ABC hot. The accuracy of transcriptional clustering was validated by mIF, which confirmed
443 ABC and GCB hot DLBCLs had significantly more tumor-infiltrating T cells compared to their
444 ABC and GCB cold DLBCL counterparts. In addition to characterizing the spectrum of immune
445 environments present in DLBCL, we identified strong associations between recurrent genetic
446 alterations and each GSVA-defined IQ, and showed how *SOCS1* LoF mutations, enriched in GCB
447 hot DLBCLs, functioned to increase B cell sensitivity to IFNγ. Finally, we uncovered a striking
448 relationship between the pre-existing immune landscape and response to CD3 x CD20 BsAb
449 immunotherapy in patients with r/r GCB DLBCL.

450 Previous transcriptional analyses have revealed significant upregulation of immune-related
451 genes in a fraction of DLBCLs. For example, using consensus clustering, Monti and colleagues
452 identified a "host response" subset of DLBCLs characterized by increased expression of immune-
453 related genes and robust T cell and dendritic cell infiltration (16). More recently, Kotlov and
454 colleagues employed Functional Gene Expression Signatures (F^{GES}) to analyze the transcriptomes
455 of over 4000 DLBCL specimens, revealing that roughly one quarter of DLBCLs harbored an
456 inflammatory microenvironment (LME-IN) (19). Thus, akin to solid cancers, there exists a
457 population of DLBCLs against which an endogenous anti-lymphoma immune response has been
458 activated. However, the mechanisms that orchestrate unique DLBCL immune environments, and
459 the potential association between the immune environment and immunotherapy response in

460 DLBCL, have not been well-explored. To address these important questions, we characterized
461 DLBCLs based on immune environmental features.

462 A central hypothesis underlying our analyses was that lymphoma cell-intrinsic alterations
463 significantly influence the composition of the DLBCL immune environment. Supporting this
464 notion, we found that each GSVA-defined immune cluster was significantly enriched for distinct
465 genetic lesions and oncogenic pathways. Far more recurrent genetic alterations were associated
466 with cold DLBCLs than with hot DLBCLs, suggesting that lymphomas may leverage numerous
467 oncogenic pathways to facilitate immune evasion. For example, among ABC DLBCLs, mutations
468 culminating in enhanced BCR activation and downstream NF-κB signaling, such as *MYD88*^{L265P},
469 *CARD11,* and *KLHL14*, were significantly associated with cold immune environments
470 (35,53,79,80). Additionally, LoF alterations in *TMEM30A* – a flippase enzyme that maintains
471 asymmetric phosphatidylserine distribution on the internal leaflet of the cell membrane – were also
472 significantly enriched among ABC cold DLBCLs, which is interesting since *TMEM30A* loss also
473 drives enhanced BCR signaling (50). The potential mechanism(s) by which oncogenic BCR
474 signaling in lymphoma cells leads to decreased T cell infiltration and/or activation remains
475 unknown but is an active area of investigation in our laboratory.

476 Among GCB DLBCLs, we identified several genes associated with "cold" environments.
477 These included *FOXO1* and *GNA13*, which are interestingly frequently mutated in Burkitt
478 lymphoma (BL) (81–83). While the immune environment of BL has not yet been fully
479 characterized, BL often exhibits low expression scores of immune-related gene sets when
480 compared to DLBCL (26,84,85). Thus, it is interesting to speculate that the presence of BL-like
481 mutations may also drive a cold immune environment in DLBCL.

482 Conversely, GCB hot DLBCLs were significantly enriched for *SOCS1* LoF mutations,
483 which are highly prevalent in cHL and PMBL, diseases with inflamed environments that are highly
484 susceptible to CBT. *SOCS1* is a potent negative regulator of IFNγ signaling – an effector cytokine
485 critical for anti-tumor immune responses (12,13,72–74,86). Supporting this notion, genetic
486 alterations that blunt IFNγ sensing in tumors, such as LoF mutations in *JAK1/2*, lead to resistance
487 to T cell-based immunotherapies (43,78,87–90). In contrast, loss of negative regulators of the IFNγ
488 signaling, for example *PTPN2, PBRM1,* and *ARID2*, sensitize tumor cells to IFNγ and T cell
489 mediated cytolysis, are associated with inflamed tumor environments, and have been linked with
490 CBT responsiveness in solid tumors (45,46,91,92). Although the extent to which LoF mutations
491 in *SOCS1* directly mediate sensitivity to CBT in lymphomas such as cHL and PMBL remains to
492 be elucidated, we hypothesize that *SOCS1* inactivation renders lymphoma cells more sensitive to
493 IFNγ signaling, as we find that genetic ablation of *Socs1* sensitizes B cells to IFNγ. Therefore,
494 these data suggest that *SOCS1*-altered DLBCLs may be inherently vulnerable to T cell-based
495 immunotherapies that rely on this key effector cytokine.

496 A shared feature of many cold DLBCLs was robust MYC pathway activation. These data
497 are supported by previous studies demonstrating strong associations between MYC activity and
498 decreased T cell infiltration in DLBCL (37,79,85). Interestingly, ABC cold DLBCLs showed
499 higher expression of MYC target genes compared to GCB cold DLBCLs, which was surprising
500 given that MYC activity in GCB DLBCL is frequently driven by the acquisition of *MYC*
501 translocations to the *IGH* locus that result in constitutive *MYC* expression (56–58). We hypothesize
502 that ABC DLBCL-associated oncogenic pathways, namely BCR-dependent NF-κB signaling, may
503 drive particularly potent MYC activity. Thus, while the mechanism(s) by which MYC activation
504 leads to a cold environment in DLBCL requires additional investigation, these data suggest that

505 therapies that indirectly modulate MYC activity by acting on upstream regulators, such as BCR
506 signaling, may restore T cell infiltration in a subset of DLBCLs and improve their sensitivity to T
507 cell-dependent immunotherapies.

508 Important from a translational perspective was our observation that GCB hot DLBCLs
509 were associated with improved PFS following treatment with the BsAb, mosunetuzumab, whereas
510 patients with GCB cold DLBCLs were less likely to benefit from this immunotherapy. Thus, for
511 r/r GCB DLBCLs, the pre-existing immune environment is clearly an important factor associated
512 with benefit from BsAb immunotherapy. Consistent with these findings, correlative analyses
513 performed on pre-treatment biopsies of r/r DLBCL patients treated with glofitamab, another CD3
514 x CD20 BsAb, demonstrated a trend toward higher effector $CD8^+$ T cell transcripts in DLBCL
515 patients who achieved a complete response (93). The finding that ABC hot and ABC cold DLBCLs
516 derive similar benefit from mosunetuzumab treatment, suggests there are likely qualitative COO-
517 related immune environmental differences not captured by our analysis and/or the existence of
518 other factors that could impact BsAb response.

519 While these results represent an important advance in our understanding of the DLBCL
520 immune environment, there are some limitations to our study. Our analysis utilized bulk
521 transcriptomic data from diagnostic DLBCL samples, and it is unclear how prior therapies impact
522 the immune landscape in this disease. We sought to address this caveat by applying our GSVA-
523 based clustering to the transcriptomes of a cohort of heavily pre-treated patients with r/r DLBCL.
524 Encouragingly, similar to untreated DLBCL samples, we observed four immune-related clusters
525 in this dataset, which preliminarily suggests a relatively limited effect of cytotoxic therapy on the
526 composition of DLBCL immune environments. Another limitation of the study was that the
527 association between our immune-related clusters and response to mosunetuzumab was

528 retrospective in nature. Finally, the utility of our GSVA-based clustering method in identifying
529 DLBCL immune environments sensitive to other immunotherapies such as CAR-T cell therapy or
530 CBT, was not investigated. However, a recent study of DLBCLs treated with CAR-T cell therapy
531 demonstrated that increased baseline T cell infiltration was associated with improved survival after
532 CAR-T cell treatment, implicating the pre-existing immune environment as an important factor in
533 CAR-T response in DLBCL (94).

534 In conclusion, this study advances our understanding of the spectrum of immune
535 landscapes present in DLBCL and is the first to identify a specific subset of inflamed DLBCLs
536 that exhibit improved outcomes with mosunetuzumab. We also identified several oncogenic
537 drivers of DLBCL that are associated with distinct immune landscapes in this disease, which
538 suggests that lymphoma cell-intrinsic alterations may significantly influence the composition of
539 the immune environment. Several genes and pathways that we identified as recurrently altered in
540 cold DLBCLs may represent novel therapeutic targets that could be modulated to restore T cell
541 infiltration into the DLBCL environment and potentially improve immunotherapy sensitivity.
542 Collectively, these results expand our knowledge of DLBCL biology, identify four unique
543 DLBCL-IQs, and implicate the pre-treatment lymphoma immune environment as a major factor
544 influencing survival of DLBCL patients treated with BsAb therapy.

545

546

547

548

549

550 ## Figure Legends

551 **Figure 1. Transcriptomic analysis identifies four unique DLBCL immune environments. A.**
552 Principal component analysis (PCA) biplot showing the contribution of immune-related gene sets
553 and COO-related gene sets to PC1 and PC2, respectively. **B**. Confusion matrix showing the
554 concordance between GSVA-based COO classification and molecularly-defined COO
555 designations for all DLBCLs in the discovery cohort (NCI/UCMC, n = 577). **C.** PCA plot showing
556 concordance between sample-wise GSVA COO scores and molecularly-defined COO calls for all
557 DLBCLs in the discovery cohort. Clusters are defined by the dashed line. **D.** PCA plot showing
558 sample-wise GSVA enrichment scores for DLBCLs in the discovery cohort, labeled by immune
559 quadrant (IQ) name (ABC Cold – dark blue, ABC Hot – red, GCB Cold – light blue, GCB Hot –
560 yellow). **E.** Heatmap showing sample-wise GSVA enrichment scores for all 19 gene sets.

561

562 **Figure 2. Validation of transcriptionally defined DLBCL-IQs**. **A.** Heatmap showing estimated
563 fraction of immune cell subtypes in each of the GSVA-based IQ as assessed by CIBERSORTx.
564 **B-C.** Violin plots showing absolute inferred proportions of $CD8^+$ T cells **(B)** and conventional
565 $CD4^+$ T cells **(C)** in each IQ. Statistical analysis by Kruskal-Wallis test followed by a post-hoc
566 Dunn's test with Benjamini-Hochberg (BH) adjusted p values **D.** Representative multispectral
567 immunofluorescence (mIF) images showing $Pax5^+$ lymphoma cells (green), $CD8^+$ T cells (red),
568 $CD4^+$ T cells (blue) for a cold (left) and hot (right) DLBCL. **E.** Box plot showing average $CD8^+$ T
569 cell : DLBCL cell ratios in immune-related clusters (n = 45). **F.** Scatter plot showing correlation
570 of hot/cold axis score (PC1) and $CD8^+$ T cell : DLBCL cell ratio (n = 45). **G.** Box plot showing
571 average $CD4^+$ T cell : DLBCL cell ratios in immune-related clusters (n = 45). **H.** Scatter plot
572 showing correlation of hot/cold axis score (PC1) and $CD4^+$ T cell : DLBCL cell ratio (n = 45).

573

574 **Figure 3. Immune environments of molecularly-defined DLBCL subtypes. A-B.** PCA plot **(A)**
575 and bar plot **(B)** showing frequency of *PD-L1* gene-amplified DLBCLs across GSVA-defined
576 immune clusters. **C-D.** PCA plot **(C)** and bar plot **(D)** showing the frequency of DZsig^{pos} DLBCLs
577 across GSVA-defined immune clusters. **E.** Alluvial plot demonstrating overlap between LME
578 clusters and GSVA-defined immune clusters. (DZsig^{pos} – dark zone signature positive; LME –
579 Lymphoma microenvironment; GC – germinal center; ME – mesenchymal; IN – inflamed; DE –
580 depleted). Fisher's exact test with BH-adjusted p values.

581

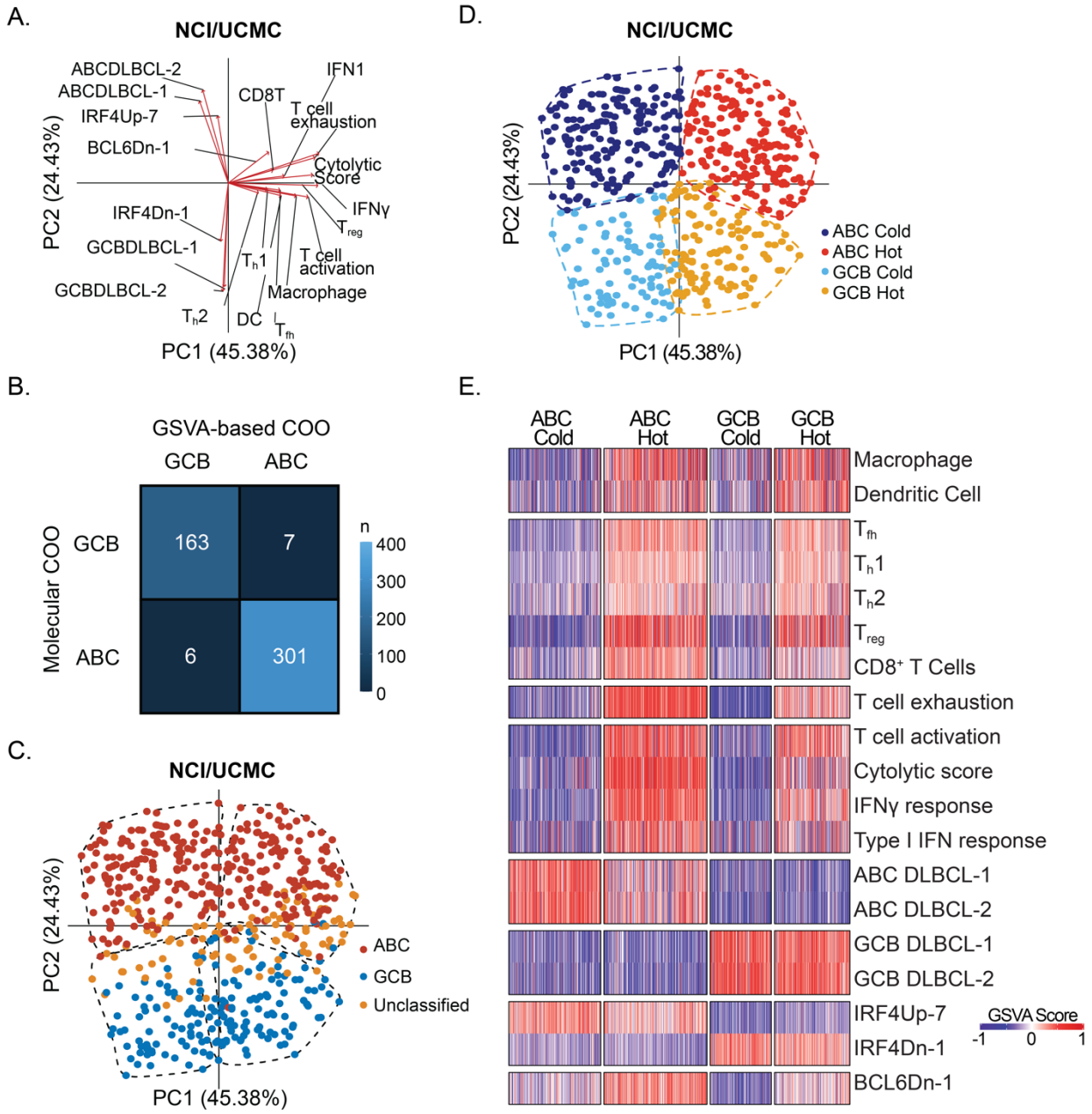
582 **Figure 4. Genomic features associated with DLBCL-IQs. A.** Alluvial plot showing overlap
583 between LymphGen subtypes and DLBC-IQs. **B.** Forest plot showing genetic alterations
584 recurrently associated with ABC Hot or ABC Cold DLBCLs. **C.** Forest plot showing genetic
585 alterations recurrently associated with GCB Hot or GCB Cold DLBCLs. **D.** Bar plot of gene set
586 enrichment analysis (GSEA) showing gene sets significantly enriched in each immune-related
587 cluster. **E.** GSEA plots showing upregulation of MYC target gene sets in ABC cold and GCB cold
588 DLBCLs. **F.** GSEA plots showing upregulation of G2M checkpoint related genes in ABC cold
589 and GCB cold DLBCLs. OR – odds ratio; ES – enrichment score; NES – normalized enrichment
590 score. Fisher's exact test with BH-adjusted p values displayed. (* adj. $p < 0.1$, ** adj. $p < 0.05$,
591 *** adj. $p < 0.01$).

592

593 **Figure 5. *SOCS1* mutations are enriched among GCB Hot DLBCLs and enhance B-cell**
594 **sensitivity to IFNγ signaling. A.** PCA plot (left) and bar plot (right) showing the frequency of
595 *SOCS1* loss of function (LoF) alterations in each immune-related cluster. **B.** Lollipop plots

596 showing frequency and distribution of mutations in *SOCS1* in GCB Hot (top) and GCB cold
597 (bottom) DLBCLs. **C.** Immune cell deconvolution (CIBERSORTx) showing absolute inferred
598 proportions of indicated immune cell subsets in *SOCS1* LoF DLBCLs compared to *SOCS1*
599 wildtype (WT) DLBCLs. **D.** GSEA plots showing upregulation of IFNγ response genes and IFNα
600 response genes in *SOCS1* LoF DLBCLs compared to *SOCS1* WT DLBCLs. Fisher's exact test
601 with BH-adjusted p values for categorical variables. Kruskal-Wallis test followed by post-hoc
602 Dunn's test with adjusted p values for continuous variables. **E.** Fold change in mean fluorescence
603 intensity (MFI) of H-2^b, I-A/I-E^b, and PD-L1 on $CD19^+$ splenocytes. $CD19^+$ splenocytes from
604 *Cd19*$^{Cre/+}$ (n = 6) or *Cd19*$^{Cre/+}$ *Socs1*$^{fl/fl}$ mice (n = 6) were cultured with media or the indicated
605 concentrations of IFNγ for 48 hours and expression levels of H-2^b, I-A/I-E^b, and PD-L1 were
606 measured. Mice were pooled from 3 independent biological replicates. Two-way ANOVA with
607 Bonferroni correction, adj. p values displayed (* adj. $p < 0.05$, ** adj. $p < 0.01$, *** adj. $p < 0.001$,
608 **** adj. $p < 0.0001$).
609
610 **Figure 6. R/R DLBCL-IQs are associated with differential benefit to BsAb therapy. A.** PCA
611 biplot showing the contribution of immune-related gene sets and COO-related gene sets to PC1
612 and PC2, respectively, for pre-treatment r/r DLBCL biopsies (n = 74) from patients treated with
613 mosunetuzumab. **B.** PCA plot showing sample-wise GSVA enrichment scores for DLBCL
614 biopsies from patients treated with mosunetuzumab. **C.** PFS for patients assigned to "hot" or
615 "cold" DLBCL immune clusters and treated with mosunetuzumab (top) and forest plot showing
616 hazard ratio of the association between PC1 score and PFS (bottom). **D.** PFS for patients assigned
617 to GCB Hot and GCB Cold DLBCL-IQs and treated with mosunetuzumab. Log-rank test with p
618 value displayed.

619 **FIGURES**



620

621 **Figure 1. Transcriptomic analysis identifies four unique DLBCL immune environments.**

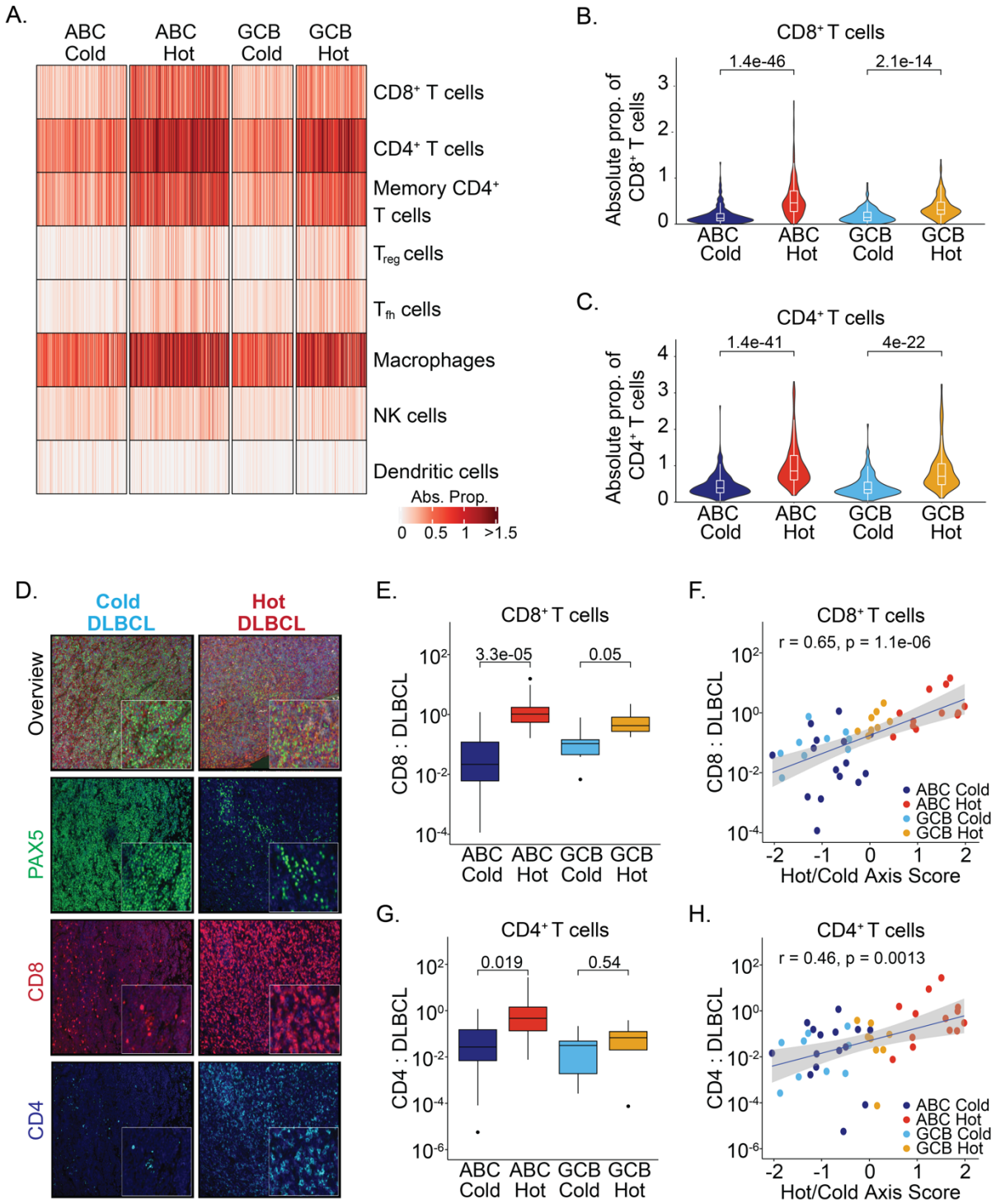
622

623

624

625

626

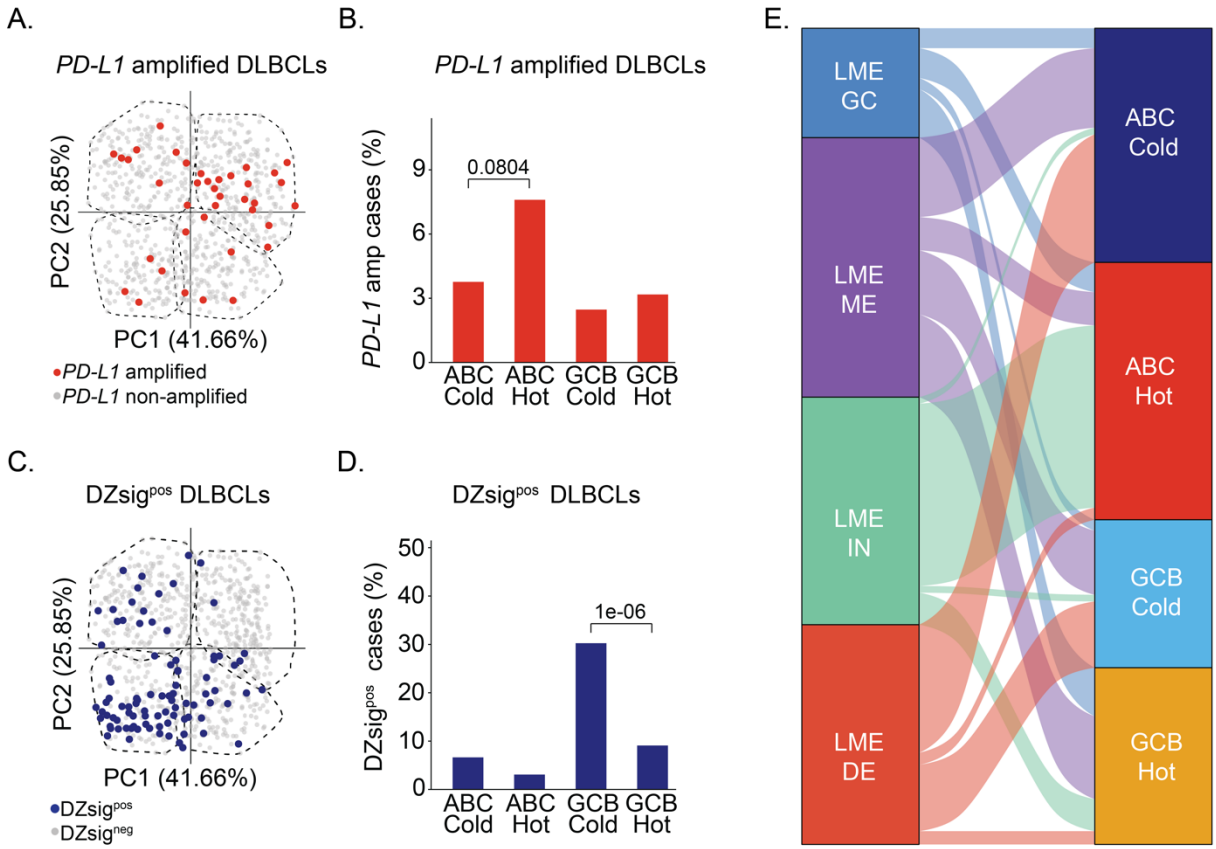


627

628 **Figure 2. Validation of transcriptionally defined DLBCL-IQs**.

629

630



631

632 **Figure 3. Immune environments of molecularly defined DLBCL subtypes.**

633

634

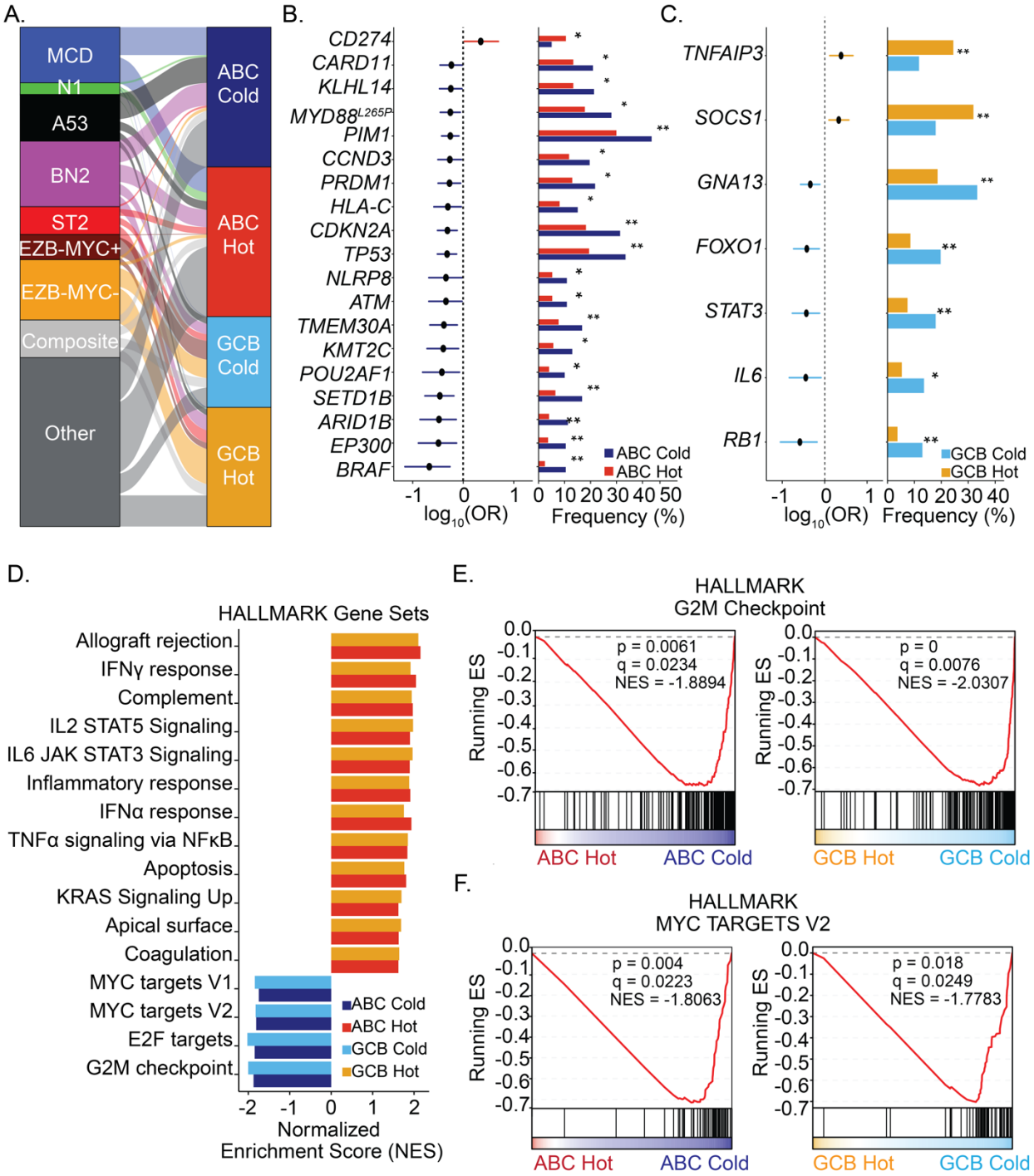


Figure 4. Genomic features associated with DLBCL-IQs.

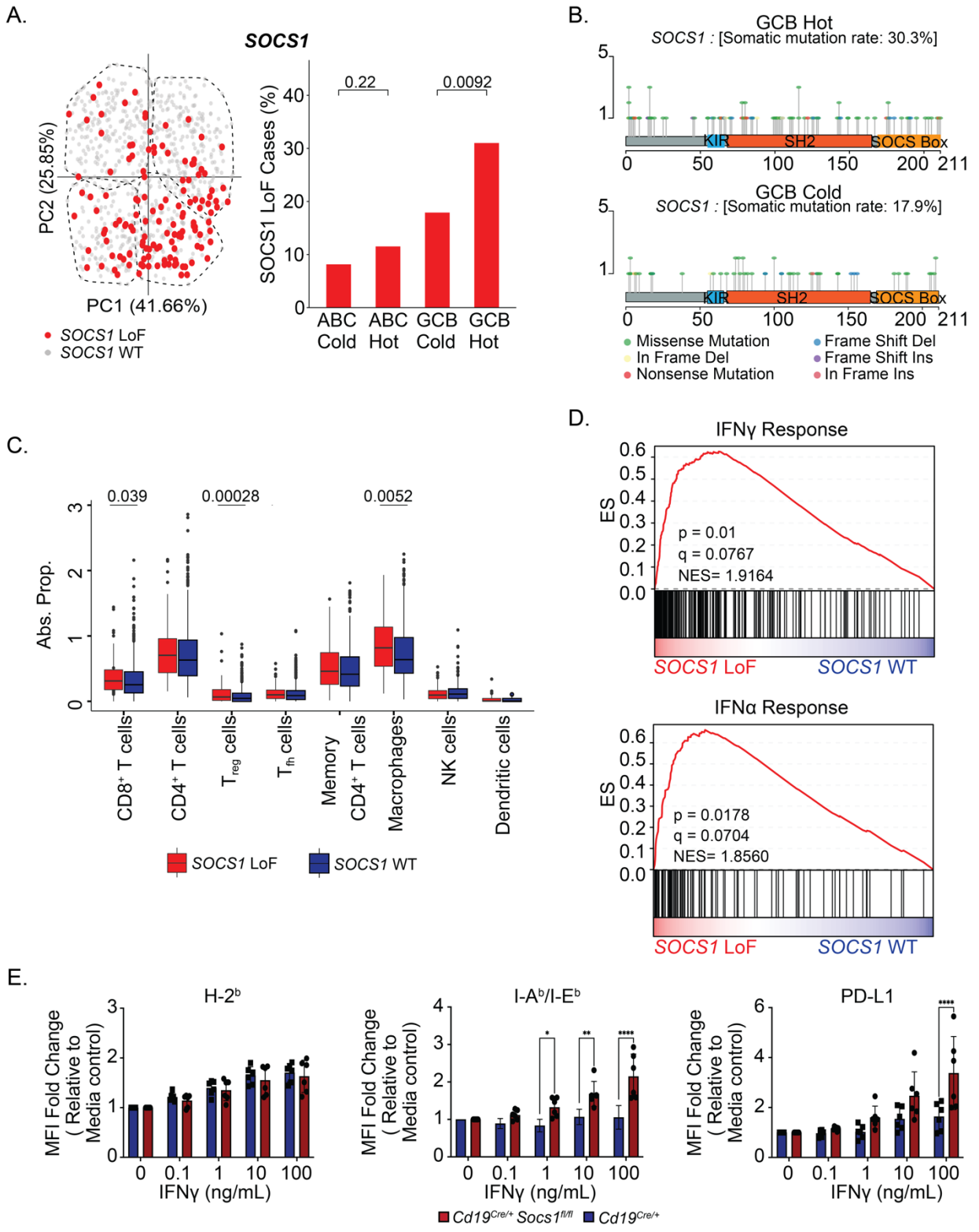


Figure 5. *SOCS1* mutations are enriched among GCB Hot DLBCLs and enhance B-cell sensitivity to IFNγ signaling.

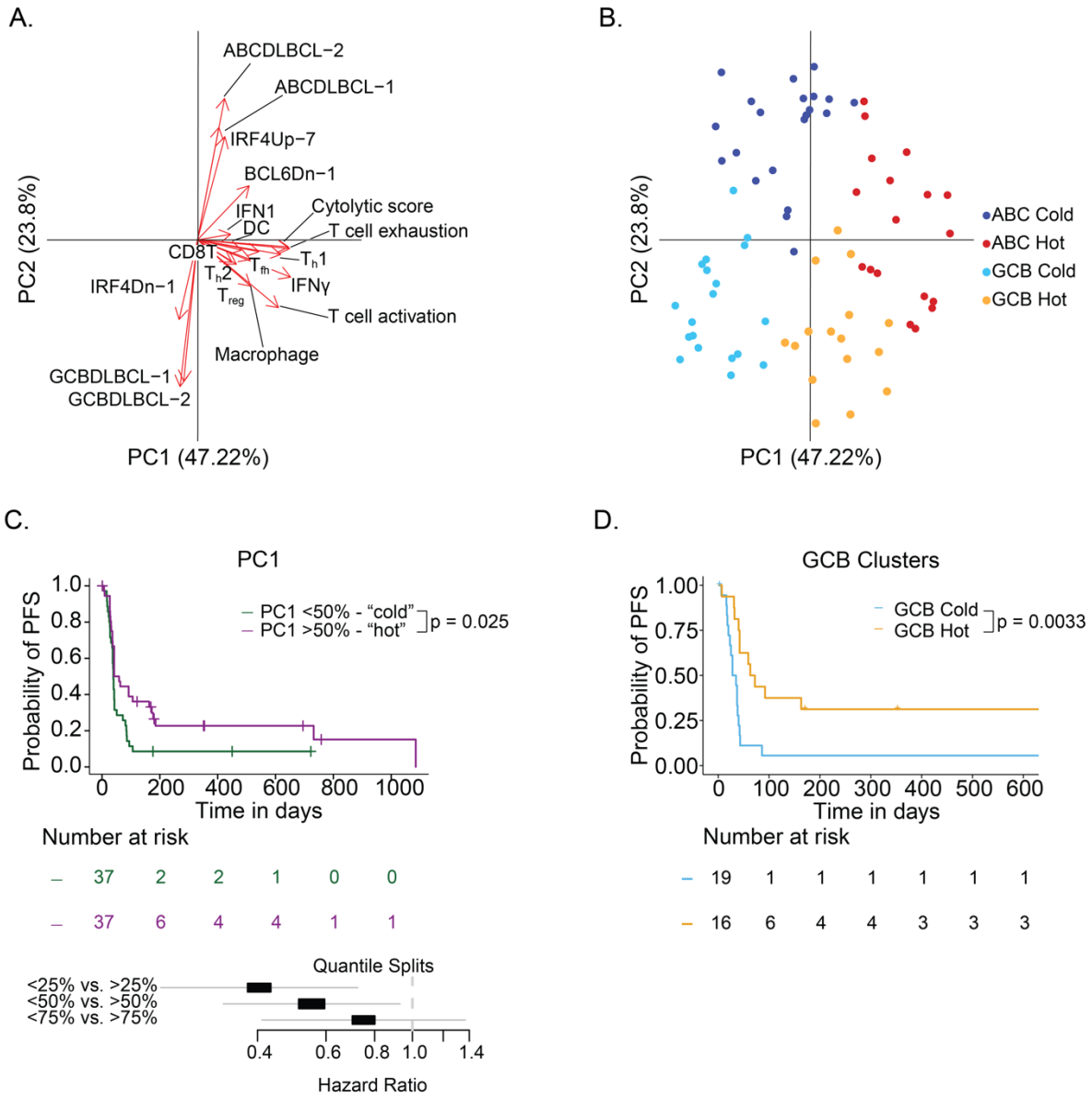


Figure 6. R/R DLBCL-IQs are associated with differential benefit to BsAb therapy.

Methods

644 **Data sets**

645 ***NCI cohort***: Pre-processed RNA-sequencing for 481 DLBCL biopsies were downloaded from
646 dbGaP (phs001444) (21). All patients had variants calls from whole exome sequencing and copy
647 number arrays.

648 ***UCMC cohort***: RNA-sequencing was performed on 96 DLBCL biopsies (84 treatment-naïve, 12-
649 relapsed/refractory, 2 unknown). Paired whole exome sequencing was successful for 76 cases.

650 ***BCC cohort***: Pre-processed RNA-sequencing for 285 treatment-naive DLBCL biopsies was
651 provided by BCC (36,37). Of these, 283 patients had available variant calls from targeted
652 sequencing and copy number arrays.

653

654 **Molecular and genetic subtype classifications**

655 ***Cell of origin:*** COO classifications were compiled from previously described gene expression-
656 based classifiers. For the NCI cohort, all DLBCLs had COO calls from a DGE-based classifier
657 (34). For the BCC cohort, 283 DLBCLs had COO calls from Lymph3Cx (36). For the UCMC
658 cohort, all samples had COO calls from an RNAseq-based classifier from BostonGene (19).

659

660 ***LME classifications*****:** Lymphoma Microenvironment (LME) classifications for DLBCLs in the
661 NCI, BCC, and UCMC cohorts were provided by BostonGene (19).

662

663 ***Dark zone signature*****:** DZsig (formerly known as double hit signature) classifications for the BCC
664 and NCI cohorts were obtained from Ennishi et al. (37) and Wright et al. (21), respectively. DZsig

665 classifications for the UCMC cohort were provided by BostonGene from an RNAseq-based
666 classifier (21).
667
668 ***LymphGen* classifications:** LymphGen classifications for the BCC and NCI cohorts were obtained
669 from Wright et al. (21).
670
671 **Generation and processing of sequencing data**
672 ***Quality control:*** FastQC and FastQ Screen were used to assess read quality and other species
673 contamination (www.bioinformatics.babraham.ac.uk). SNP (single nucleotide polyphorism)-
674 calling using pileup in FACETS was used to assess sample mix-ups and contamination (95).
675
676 ***Mutation-calling and annotation:*** Whole exome sequencing data for the UCMC cohort was
677 generated as previously described (23). Mutation-calling for whole exome sequencing data for the
678 UCMC cohort was performed similarly to the analysis in Kotlov et al (19). Reads were trimmed
679 using fastp to eliminate low-quality reads and then were aligned to the GRCh38 reference genome
680 using BWA (v0.7.17) (96,97). Duplicate reads were removed using Picard's MarkDuplicates
681 (v2.6.0) ("Picard Toolkit", 2019. Broad Institute, GitHub Repository) and recalibrated by
682 BaseRecalibrator and ApplyBQSR from GATK4 (98). For samples with a paired normal reference,
683 short variants were called using Strelka (99). For tumor-only samples, Pisces (v5.2) was used
684 (100). Variants were annotated using Variant Effect Predictor (VEP) (101). For all data sets,
685 variants were restricted to these predicted to be deleterious.
686

687 ***Copy number analysis:*** Copy number variation (CNV) was assessed using Sequenza with
688 modifications from FACETS (95,102). CNV analysis was performed in the space of genes listed
689 in **Supplementary Table 2.** Sequenza input reference was modified to match FACETS, calculated
690 using facets-pileup. For tumor-only samples, calling versus average normal or tumor sample was
691 performed using a modified version of Sequenza. Average normal/tumor pileup files were
692 prepared to average normalized coverage and variant allele fraction (VAF) at each position. Some
693 samples were excluded due to low coverage and/or tumor content. For all data sets, heterozygous
694 deletions and low-level copy gains were excluded.
695
696 ***PD-L1 gene alterations*****:** *PD-L1* amplifications were defined by FISH when available but were
697 otherwise by sequencing. When assessed by sequencing, *PD-L1* amplifications were defined as ≥
698 4 copies.
699
700 ***RNA-seq processing and normalization:*** Paired-end and single-end RNA-seq data for the UCMC
701 cohort was generated as previously described (23). RNA-seq reads were aligned using kallisto
702 v0.42.4 to GENCODE v23 transcripts with default parameters (103). The protein-coding
703 transcripts, immunoglobulin heavy, kappa and lambda light chains, and TCR-related transcripts
704 were retained; noncoding RNA, histone, and mitochondria-related transcripts were removed.
705 Counts were processed using limma (v3.52.4) and edgeR (v3.38.4). Genes with low counts were
706 removed using a threshold of a 0.5 log2CPM average across all samples, as previously described
707 (24). Then, TMM (trimmed mean of M-values) normalization was performed, and counts were
708 quantified as log2CPM. Batch correction was performed using the removeBatchEffect function in
709 limma.

710 For DLBCLs in the mosunetuzumab cohort, robust library sizes were estimated using DESeq2,
711 and raw RNASeq counts were normalized via limma-voom.
712
713 **Differential gene expression (DGE) and gene set enrichment analysis (GSEA)**:
714 DGE was performed using voom, and differentially expressed genes were generally defined as
715 those with $|\log2FC| \geq 1$ and adjusted p value < 0.05 (104). GSEA was performed using the Desktop
716 client (v4.2.3) with default settings (105).
717
718 **Comparing variants between groups**
719 Variants were restricted to previously described DLBCL driver genes (28,29), listed in
720 **Supplementary Table 2.** Maftools (v2.12.0) was used to compare alterations between groups,
721 where additional filtering was performed to filter to genes altered in at least 10% of samples of the
722 smaller group (106).
723
724 **Mutational pathway analysis**
725 Genes in the BCR-dependent NFKB pathway (53,54) and cell cycle pathway-related genes (54)
726 were curated from previously published studies. The proportion of patients with an alteration in at
727 least one gene in the pathway was calculated and compared across clusters. Alterations for a given
728 gene were filtered to reflect their putative role as an oncogene or as a tumor suppressor gene.
729
730 **CIBERSORTx**
731 Immune cell deconvolution was performed using CIBERSORTx using the LM22 signature
732 matrix(38,107). The software was run on bulk-mode and absolute mode. For the input mixture file,

733 RNA-seq counts were first batch corrected using ComBat-Seq followed by quantification to
734 transcripts per million (TPM) (108). Cell types were condensed as such:
735 **T.cells** = "T cells CD8" + "T cells follicular helper (T_{fh})" + "T cells regulatory (Tregs)" + "T cells
736 CD4 naïve" + "T cells CD4 memory resting" + "T cells CD4 memory activated"+ "T cells gamma
737 delta"
738 **NK.cells** = "NK cells activated" + "NK cells resting"
739 **Memory.CD4.T.cells** = "T cells CD4 memory resting" + "T cells CD4 memory activated"
740 **Conv.CD4.T.cells** = "Memory.CD4.T.cells" + "T cells CD4 naïve" + "T cells follicular helper"
741 **CD4.T.cells** = "Memory.CD4.T.cells" + "T cells CD4 naïve" + "T cells follicular helper" + "T
742 cells regulatory (T_{regs})"
743 **Dendritic.cells** = "Dendritic cells activated" + "Dendritic cells resting"
744 **B.cells** = "B cells naïve" + "B cells memory" + "Plasma cells"
745 **Macrophages** = "Macrophages M0" + "Macrophages M1" + "Macrophages M2"
746
747 **Gene set compilation**
748 Gene sets were manually curated following an extensive literature search and focused on genes
749 related to 1) immune cell infiltration and activation, 2) cell-of-origin and 3) key transcription
750 factors related to cell-of-origin. Gene sets were excluded if they had fewer than 3 or more than 200
751 genes. Genes included in each gene set are listed in **Supplementary Table 1**.
752
753 ***Immune-related gene sets***: Twelve immune-related gene sets were included, known to be
754 associated with immune activity in other cancer types. Gene sets were related to activation state of
755 T cells (t.cell.activation (15), t.cell.exhaustion (15), IFNγ_ayers (12)), non-cellular mediators of

756 immune responses (ifn1_rooney (15), cytolytic.score (26)), and presence of immune cell subsets
757 (tfh_charoentong (25), th1_charoentong (25), th2_charoentong (25), treg_rooney (15),
758 cd8t_charoentong (25), macrophage_rooney (15), dc_xcell_total (27)). In general, immune-related
759 gene sets had less than 15% overlap based on pairwise comparisons.
760
761 ***Cell of origin gene sets*:** A total of four COO gene sets - 2 ABC DLBCL (ABCDLBC-1,
762 ABCDLBCL-2) specific gene sets and 2 GCB DLBCL specific gene sets (GCBDLBCL-1,
763 GCBDLBCL-2) - were included. Gene sets are derived from the seminal gene expression classifier
764 (34), found in SignatureDB (SigDB)(55). (https://lymphochip.nih.gov/signaturedb/index.html).
765
766 ***Transcription factor gene sets***: Three transcription factor related gene sets – related to IRF4
767 (IRF4Up-7 and IRF4Dn-1) and BCL6 (BCL6Dn-1) targets – were included from SigDB (55).
768
769 **Immune signature clustering model**
770 Gene set variation analysis (GSVA) using the 19 curated gene sets was performed on bulk
771 transcriptomes from 2 genomic datasets (NCI, n = 481; UCMC, n = 96) (21,23). The resulting
772 GSVA scores were resolved using principal component analysis (PCA). Unsupervised k-means
773 clustering was utilized to assign samples into unique clusters, where k = 4 clusters was optimally
774 selected using an elbow plot. Subsequent analyses using the third independent BCC dataset as well
775 as the 3-dataset combination were performed in the same structured steps, where equivalent
776 clustering was observed.
777
778 **Generation of MYCUp4 expression groups**

779 GSVA scores for the MYCUp-4 gene set (from SigDB (55)) were generated for all DLBCLs in
780 UCMC, NCI, and BCC datasets. DLBCLs were then split into ABC and GCB groups based on
781 their immune-related cluster membership, and a 25-75 percentile split within each COO group was
782 used to assign DLBCLs to "MYC-High" and "MYC-Low" groups. All other DLBCLs were
783 assigned to the "MYC-Int" group.
784
785 **Multispectral immunofluorescence**
786 Multi-spectral immunofluorescence (mIF) microscopy was performed on 65 DLBCLs (UCMC)
787 for which paired RNA-seq and GSVA data were available. mIF analysis was performed after
788 staining with fluorescence-labeled antibodies against a T cell panel and myeloid cell panel *(see*
789 *antibodies for mIF section*). Each slide was scanned using the Vectra Polaris (Akoya Biosciences)
790 imaging platform and the Phenochart software (PerkinElmer). Through Phenochart, at least 5
791 representative regions of interest per tissue section were acquired as multispectral images at 40x
792 magnification. Watershed segmentation was used to identify nuclei using DAPI staining and cell
793 borders of individual cells in each ROI. The supervised machine-learning algorithm in the inForm
794 software (v. 2.3) was used to classify each cell into specific phenotypes.
795 1. Slides were divided by panel and by specific features determined during the initial
796 watershed segmentation, including cell border detection, average cell size,
797 autofluorescence, and DAPI strength.
798 2. Samples were divided into 4 separate groups for each panel. 2 ROIs were chosen per
799 sample to train the machine-learning algorithm in inForm to identify the following
800 phenotypes with the associated markers:
801 1. **T cells** : $CD8^+$ T cell, $CD4^+$ T cell, $PAX5^+$ DLBCL cell

802 2. **Myeloid cells** : $CD68^+$ Macrophage, $CD11c^+$ dendritic cell

803 3. Once the initial training groups were processed, all ROIs in each group of samples were
804 classified using the matching training cohort.

805 4. Results of per-slide frequencies of each phenotype were tabulated in R using exported
806 values from inForm using the `phenoptr` package and the representative mIF images were
807 also exported through inForm. Given heterogeneity of tumor and microenvironment
808 composition in each ROI, comparisons were made between total number of each TME cell
809 population/phenotype normalized against the number of DLBCL cells present across all
810 ROIs per slide.

811

812 **Antibodies for mIF**

813

Panel 1 -- Myeloid	Location	Opal	Primary Ab Info	Clone	Ag retrieval	Primary Ab Diluent
PAX5	nucleus	620	BioCare Medical,CM 207	BC/24, mouse mAb	AR9	1:50 in Da Vinci Green
CD68	membrane+cytoplasmic	570	BioCare Medical, CM033A	KP1, mouse mAb	AR9	1:50 in Da Vinci Green
CD11c	membrane	540	BioCare Medical, ACI3122A	5D11, mouse mAB	AR9	1:100 in Renoir Red

814

815

Panel 2 – T cell	Location	Opal	Primary Ab Info	Clone	Ag retrieval	Primary Ab Diluent
CD4	membrane	520	BioCare Medical, ACI3148A	4B12, mouse mAb	AR9	1:50 in Van Gogh Yellow
PAX5	nucleus	620	BioCare Medical,CM 207	BC/24, mouse mAb	AR9	1:50 in Da Vinci Green
CD8	membrane	690	R&D, NBP232836B	C8/144B, mouse mAb	AR9	1:50

816 ***In vitro* IFNγ stimulation of splenocytes**

817 Spleens from *CD19*$^{Cre/wt}$ and *CD19*$^{Cre/wt}$; *Socs1*$^{fl/fl}$ mice were obtained from Ari Melnick (Weill
818 Cornell Medicine). Spleens were passed through a 70 µm cell strainer (Corning, 352350) and
819 ground using the plunger of a syringe with 10 mL of PBS to yield single cell suspensions.
820 Splenocytes were then washed twice in 10mL of PBS, followed by lysis of red blood cells in a
821 hypotonic solution and three washes in 10 mL of PBS. 5 x 10^5 splenocytes from single cell
822 suspensions were incubated with increasing concentrations of recombinant mouse IFNγ (0, 0.1, 1,
823 10, 100 ng/mL) for 48 hours (Peprotech, 315-05). Cells were then harvested and incubated with a
824 panel of antibodies containing anti-mouse CD19 (Biolegend,115504), anti-mouse CD3d
825 (Biolegend, 100244), anti-mouse H-2 (Biolegend, 125506), anti-mouse I-A/I-E (Biolegend,
826 107619), and anti-mouse PD-L1 (Biolegend, 124311). Splenocytes were analyzed by flow
827 cytometry for expression of H-2, I-A/I-E, and PD-L1 on $CD19^+$ and $CD3^+$ cell populations. Mean
828 fluorescence intensity (MFI) fold change for H-2, H-2, I-A/I-E, and PD-L1 was computed by
829 comparing treatment with IFNγ compared to media-only control. Data shown are the average of at
830 least 3 independent biological replicates. Statistical testing was conducted using a 2-way ANOVA
831 with Bonferroni correction for multiple comparisons.

832

833 **Statistical methods**

834 All statistical analyses were performed in R (v4.2.1). Statistical significance for categorical
835 variables was assessed with a Fisher's exact test for two-group comparisons; a multi-way Fisher's
836 exact test followed by post-hoc individual pairwise testing was used for comparisons of three or
837 more groups. Statistical significance of continuous variables was performed with a Mann-Whitney
838 U test for two-group comparisons and a Kruskal-Wallis test followed by a post-hoc Dunn's test

for comparisons across three or more groups. Survival analysis was performed using the survival package (v3.4.0), and the log-rank test was used to determine significance. A p value threshold of 0.05 was used to assess univariate statistical significance. Adjustment for multiple hypothesis testing was performed using the Benjamini-Hochberg method unless indicated otherwise.

Acknowledgements

The authors recognize the patients who participated in The University of Chicago Hoogland Lymphoma Biobank. We are grateful to the laboratory of Ari M. Melnick for the contribution of spleens from *Socs1*-sufficient and *Socs1*-deficient mice. We would also like to thank Genentech for their collaboration and for providing genomic data from patients enrolled in the GO29781 trial (NCT02500407). This work was completed using resources provided by the University of Chicago's Research Computing Center, the University of Chicago Human Immunologic Monitoring Facility (RRID: SCR_017916), and the University of Chicago Cytometry and Antibody Technology Facility (RRID: SCR_17760). This work was supported by grants from the Leukemia and Lymphoma Society, V Foundation for Cancer Research, and the National Cancer Institute (R01-CA273241-01A1).

Author Contributions

J.K. designed and supervised the study, obtained funding, and wrote the manuscript. S.T., J.G., A.C., and J.Y., designed the study, wrote the manuscript, and performed all data analyses. Y.Z. performed and assisted with the analysis of the multispectral immunofluorescence imaging data. X.C., B.W.M., J.S., B.P., G.D., and C.S., assisted with data collection and provision of study materials. G.V. assisted with the collection of tumor samples from the University of Chicago Hoogland Lymphoma Biobank. G.D., E.P., N.K., A.B., and N.F. assisted with genomic analyses. C.B. and E.P. assisted with the analysis and interpretation of the mosunetuzumab data. All authors reviewed and edited the manuscript.

Disclosure of Conflicts of Interest

S.T., J.G., A.C., X.C., B.W.M., B.P., Y.Z., G.V., J.S., and G.D. have no conflicts of interest to disclose. J.Y. is an employee of AbbVie. E.P., N.K., A.B., and N.F. are employees of BostonGene. E.P. and C.B. are employees of Genentech Inc. S.M.S. has served as a consultant for Morphosys/Incyte, Janssen, BMS, Karyopharm, TGTX, and Celgene, and has research funding from FortySeven, TG Therapeutics, Pharmacyclics, Acerta, Karyopharm, Portola, Celgene, Novartis, Genentech/Roche, and Epizyme. C.S. received honoraria from Seattle Genetics, AbbVie, and Bayer, and holds research funding from Epizyme and Trillium Therapeutics Inc. J.K. receives research support from Merck, Secura Bio, and iTeos, has served on a Speakers Bureau for Kite/Gilead, and has served on advisory boards for Verastem, Seattle Genetics, MorphoSys, and Karyopharm.

Data and Code Availability

Raw sequencing data from UCMC will be available at gene expression omnibus (GEO). This paper does not report original code. Code to reproduce the GSVA clustering pipeline is available at https://github.com/kline-lab-group/immune-quadrant.

908

References

909 1. Sehn LH, Salles G. Diffuse Large B-Cell Lymphoma. N Engl J Med. Massachusetts Medical
910 Society; 2021;384:842–58.

911 2. Crump M, Neelapu SS, Farooq U, Van E, Neste D, Kuruvilla J, et al. Outcomes in refractory
912 diffuse large B-cell lymphoma: results from the international SCHOLAR-1 study Key
913 Points. 2017;

914 3. Neelapu SS, Locke FL, Bartlett NL, Lekakis LJ, Miklos DB, Jacobson CA, et al.
915 Axicabtagene Ciloleucel CAR T-Cell Therapy in Refractory Large B-Cell Lymphoma. N
916 Engl J Med. 2017;377:2531–44.

917 4. Locke FL, Miklos DB, Jacobson CA, Perales M-A, Kersten M-J, Oluwole OO, et al.
918 Axicabtagene Ciloleucel as Second-Line Therapy for Large B-Cell Lymphoma. N Engl J
919 Med. Massachusetts Medical Society; 2022;386:640–54.

920 5. Schuster SJ, Bishop MR, Tam CS, Waller EK, Borchmann P, McGuirk JP, et al.
921 Tisagenlecleucel in Adult Relapsed or Refractory Diffuse Large B-Cell Lymphoma.
922 https://doi.org/101056/NEJMoa1804980. Massachusetts Medical Society; 2018;380:45–56.

923 6. Abramson JS, Palomba ML, Gordon LI, Lunning MA, Wang M, Arnason J, et al.
924 Lisocabtagene maraleucel for patients with relapsed or refractory large B-cell lymphomas
925 (TRANSCEND NHL 001): a multicentre seamless design study. The Lancet. Elsevier;
926 2020;396:839–52.

927 7. Budde LE, Assouline S, Sehn LH, Schuster SJ, Yoon S-S, Yoon DH, et al. Single-Agent
928 Mosunetuzumab Shows Durable Complete Responses in Patients With Relapsed or
929 Refractory B-Cell Lymphomas: Phase I Dose-Escalation Study. J Clin Oncol. Wolters
930 Kluwer; 2022;40:481–91.

931 8. Dickinson MJ, Carlo-Stella C, Morschhauser F, Bachy E, Corradini P, Iacoboni G, et al.
932 Glofitamab for Relapsed or Refractory Diffuse Large B-Cell Lymphoma. N Engl J Med.
933 Massachusetts Medical Society; 2022;387:2220–31.

934 9. Thieblemont C, Phillips T, Ghesquieres H, Cheah CY, Clausen MR, Cunningham D, et al.
935 Epcoritamab, a Novel, Subcutaneous CD3xCD20 Bispecific T-Cell–Engaging Antibody, in
936 Relapsed or Refractory Large B-Cell Lymphoma: Dose Expansion in a Phase I/II Trial. J
937 Clin Oncol. Wolters Kluwer; 2023;41:2238–47.

938 10. Ansell SM. Checkpoint Blockade in Lymphoma. J Clin Oncol [Internet]. American Society
939 of Clinical Oncology; 2021 [cited 2023 Feb 23]; Available from:
940 https://ascopubs.org/doi/pdf/10.1200/JCO.20.01522?role=tab

941 11. Ansell SM, Minnema MC, Johnson P, Timmerman JM, Armand P, Shipp MA, et al.
942 Nivolumab for Relapsed/Refractory Diffuse Large B-Cell Lymphoma in Patients Ineligible
943 for or Having Failed Autologous Transplantation: A Single-Arm, Phase II Study.
944 https://doi.org/101200/JCO1800766. American Society of Clinical Oncology;
945 2019;37:481–9.

946 12. Ayers M, Lunceford J, Nebozhyn M, Murphy E, Loboda A, Kaufman DR, et al. IFN-γ-
947 related mRNA profile predicts clinical response to PD-1 blockade. J Clin Invest. American
948 Society for Clinical Investigation; 2017;127:2930–40.

949 13. Ott PA, Bang Y-J, Piha-Paul SA, Razak ARA, Bennouna J, Soria J-C, et al. T-Cell–Inflamed
950 Gene-Expression Profile, Programmed Death Ligand 1 Expression, and Tumor Mutational
951 Burden Predict Efficacy in Patients Treated With Pembrolizumab Across 20 Cancers:
952 KEYNOTE-028. J Clin Oncol. Wolters Kluwer; 2019;37:318–27.

953 14. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang T-H, et al. The Immune
954 Landscape of Cancer. Immunity. 2018;48:812-830.e14.

955 15. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and Genetic Properties of
956 Tumors Associated with Local Immune Cytolytic Activity. 2015; Available from:
957 http://dx.doi.org/10.1016/j.cell.2014.12.033

958 16. Monti S, Savage KJ, Kutok JL, Feuerhake F, Kurtin P, Mihm M, et al. Molecular profiling of
959 diffuse large B-cell lymphoma identifies robust subtypes including one characterized by
960 host inflammatory response. Blood. 2005;105:1851–61.

961 17. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, et al. Distinct types of
962 diffuse large B-cell lymphoma identified by gene expression profiling. Nature.
963 2000;403:503–11.

964 18. Lenz G, Wright G, Dave SS, Xiao W, Powell J, Zhao H, et al. Stromal gene signatures in
965 large-B-cell lymphomas. N Engl J Med. 2008;359:2313–23.

966 19. Kotlov N, Bagaev A, Revuelta MV, Phillip JM, Cacciapuoti MT, Antysheva Z, et al. Clinical
967 and Biological Subtypes of B-cell Lymphoma Revealed by Microenvironmental Signatures.
968 Cancer Discov. 2021;11:1468–89.

969 20. Kline J, Godfrey J, Ansell SM. The immune landscape and response to immune checkpoint
970 blockade therapy in lymphoma. Blood. 2020;135:523–33.

971 21. Wright GW, Huang DW, Phelan JD, Coulibaly ZA, Roulland S, Young RM, et al. A
972 Probabilistic Classification Tool for Genetic Subtypes of Diffuse Large B Cell Lymphoma
973 with Therapeutic Implications. Cancer Cell. Cell Press; 2020;37:551-568.e14.

974 22. Schmitz R, Wright GW, Huang DW, Johnson CA, Phelan JD, Wang JQ, et al. Genetics and
975 Pathogenesis of Diffuse Large B-Cell Lymphoma. N Engl J Med. Massachusetts Medical
976 Society; 2018;378:1396–407.

977 23. Godfrey J, Tumuluru S, Bao R, Leukam M, Venkataraman G, Phillip J, et al. PD-L1 gene
978 alterations identify a subset of diffuse large B-cell lymphoma harboring a T-cell–inflamed
979 phenotype. Blood. American Society of Hematology; 2019;133:2279–90.

980 24. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and
981 RNA-Seq data. BMC Bioinformatics. 2013;14:7.

982 25. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-cancer
983 Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and
984 Predictors of Response to Checkpoint Blockade. Cell Rep. 2017;18:248–62.

985 26. Dufva O, Pölönen P, Brück O, Keränen MAI, Klievink J, Mehtonen J, et al. Immunogenomic
986 Landscape of Hematological Malignancies. Cancer Cell. 2020;38:380-399.e13.

987 27. Aran D, Hu Z, Butte AJ. xCell: Digitally portraying the tissue cellular heterogeneity
988 landscape. Genome Biol. 2017;18.

989 28. Reddy A, Zhang J, Davis NS, Moffitt AB, Love CL, Waldrop A, et al. Genetic and
990 Functional Drivers of Diffuse Large B Cell Lymphoma. Cell. Cell Press; 2017;171:481-
991 494.e15.

992 29. Chapuy B, Stewart C, Dunford AJ, Kim J, Kamburov A, Redd RA, et al. Molecular subtypes
993 of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and
994 outcomes. Nat Med. Nature Publishing Group; 2018;24:679–90.

995 30. Scott DW, Mottok A, Ennishi D, Wright GW, Farinha P, Ben-Neriah S, et al. Prognostic
996 Significance of Diffuse Large B-Cell Lymphoma Cell of Origin Determined by Digital
997 Gene Expression in Formalin-Fixed Paraffin-Embedded Tissue Biopsies. J Clin Oncol Off J
998 Am Soc Clin Oncol. American Society of Clinical Oncology; 2015;33:2848–56.

999 31. Davis RE, Brown KD, Siebenlist U, Staudt LM. Constitutive Nuclear Factor κB Activity Is
1000 Required for Survival of Activated B Cell–like Diffuse Large B Cell Lymphoma Cells. J
1001 Exp Med. 2001;194:1861–74.

1002 32. Morin RD, Mendez-Lago M, Mungall AJ, Goya R, Mungall KL, Corbett RD, et al. Frequent
1003 mutation of histone-modifying genes in non-Hodgkin lymphoma. Nature. Nature
1004 Publishing Group; 2011;476:298–303.

1005 33. Pasqualucci L, Dalla-Favera R. Genetics of diffuse large B-cell lymphoma. Blood. American
1006 Society of Hematology; 2018;131:2307–19.

1007 34. Wright G, Tan B, Rosenwald A, Hurt EH, Wiestner A, Staudt LM. A gene expression-based
1008 method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma. Proc
1009 Natl Acad Sci U S A. 2003;100:9991–6.

1010 35. Yang Y, Shaffer AL, Emre NCT, Ceribelli M, Zhang M, Wright G, et al. Exploiting
1011 Synthetic Lethality for the Therapy of ABC Diffuse Large B Cell Lymphoma. Cancer Cell.
1012 2012;21:723–37.

1013 36. Duns G, Viganò E, Ennishi D, Sarkozy C, Hung SS, Chavez E, et al. Characterization of
1014 DLBCL with a PMBL gene expression signature. Blood. American Society of Hematology;
1015 2021;138:136–48.

1016 37. Ennishi D, Jiang A, Boyle M, Collinge B, Grande BM, Ben-Neriah S, et al. Double-Hit Gene
1017 Expression Signature Defines a Distinct Subgroup of Germinal Center B-Cell-Like Diffuse
1018 Large B-Cell Lymphoma. J Clin Oncol. 2019;37:190–201.

1019 38. Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, et al. Determining
1020 cell type abundance and expression from bulk tissues with digital cytometry. Nat
1021 Biotechnol. Nature Publishing Group; 2019;37:773–82.

1022 39. Alduaij W, Collinge B, Ben-Neriah S, Jiang A, Hilton LK, Boyle M, et al. Molecular
1023 determinants of clinical outcomes in a real-world diffuse large B-cell lymphoma
1024 population. Blood. 2023;141:2493–507.

1025 40. Spranger S, Bao R, Gajewski TF. Melanoma-intrinsic β-catenin signalling prevents anti-
1026 tumour immunity. Nature. Nature Publishing Group; 2015;523:231–5.

1027 41. Ennishi D, Takata K, Béguelin W, Duns G, Mottok A, Farinha P, et al. Molecular and
1028 genetic characterization of MHC deficiency identifies ezh2 as therapeutic target for
1029 enhancing immune recognition. Cancer Discov. American Association for Cancer Research
1030 Inc.; 2019;9:546–63.

1031 42. Shouval R, Alarcon Tomas A, Fein JA, Flynn JR, Markovits E, Mayer S, et al. Impact of
1032 TP53 Genomic Alterations in Large B-Cell Lymphoma Treated With CD19-Chimeric
1033 Antigen Receptor T-Cell Therapy. J Clin Oncol. Wolters Kluwer; 2022;40:369–81.

1034 43. Peng W, Chen JQ, Liu C, Malu S, Creasy C, Tetzlaff MT, et al. Loss of PTEN Promotes
1035 Resistance to T Cell-Mediated Immunotherapy. Cancer Discov. 2016;6:202–16.

1036 44. Peng D, Kryczek I, Nagarsheth N, Zhao L, Wei S, Wang W, et al. Epigenetic silencing of
1037 TH1-type chemokines shapes tumour immunity and immunotherapy. Nature. Nature
1038 Publishing Group; 2015;527:249–53.

1039 45. Manguso RT, Pope HW, Zimmer MD, Brown FD, Yates KB, Miller BC, et al. In vivo
1040 CRISPR screening identifies Ptpn2 as a cancer immunotherapy target. Nature. Nature
1041 Publishing Group; 2017;547:413–8.

1042 46. Pan D, Kobayashi A, Jiang P, Ferrari de Andrade L, Tay RE, Luoma AM, et al. A major
1043 chromatin regulator determines resistance of tumor cells to T cell–mediated killing.
1044 Science. 2018;359:770–5.

1045 47. Wellenstein MD, de Visser KE. Cancer-Cell-Intrinsic Mechanisms Shaping the Tumor
1046 Immune Landscape. Immunity. 2018;48:399–416.

1047 48. Spranger S, Gajewski TF. Mechanisms of Tumor Cell–Intrinsic Immune Evasion. Annu Rev
1048 Cancer Biol. 2018;2:213–28.

1049 49. Lacy SE, Barrans SL, Beer PA, Painter D, Smith AG, Roman E, et al. Targeted sequencing
1050 in DLBCL, molecular subtypes, and outcomes: a Haematological Malignancy Research
1051 Network report. Blood. 2020;135:1759–71.

1052 50. Ennishi D, Healy S, Bashashati A, Saberi S, Hother C, Mottok A, et al. TMEM30A loss-of-
1053 function mutations drive lymphomagenesis and confer therapeutically exploitable
1054 vulnerability in B-cell lymphoma. Nat Med. Nature Research; 2020;26:577–88.

1055 51. Venturutti L, Teater M, Zhai A, Chadburn A, Babiker L, Kim D, et al. TBL1XR1 Mutations
1056 Drive Extranodal Lymphoma by Inducing a Pro-tumorigenic Memory Fate. Cell.
1057 2020;182:297-316.e27.

1058 52. Young RM, Shaffer AL, Phelan JD, Staudt LM. B cell Receptor Signaling in Diffuse Large
1059 B cell Lymphoma. Semin Hematol. 2015;52:77–85.

1060 53. Phelan JD, Young RM, Webster DE, Roulland S, Wright GW, Kasbekar M, et al. A
1061 multiprotein supercomplex controlling oncogenic signalling in lymphoma. Nature. Nature
1062 Publishing Group; 2018;560:387–91.

1063 54. Miao Y, Medeiros LJ, Li Y, Li J, Young KH. Genetic alterations and their clinical
1064 implications in DLBCL. Nat Rev Clin Oncol. Nature Publishing Group; 2019;16:634–52.

1065 55. Shaffer AL, Wright G, Yang L, Powell J, Ngo V, Lamy L, et al. A library of gene expression
1066 signatures to illuminate normal and pathological lymphoid biology. Immunol Rev.
1067 2006;210:67–85.

1068 56. Ott G, Rosenwald A, Campo E. Understanding MYC-driven aggressive B-cell lymphomas:
1069 pathogenesis and classification. Blood. 2013;122:3884–91.

1070 57. Friedberg JW. Double-Hit Diffuse Large B-Cell Lymphoma. J Clin Oncol. Wolters Kluwer;
1071 2012;30:3439–43.

1072 58. Copie-Bergman C, Cuillière-Dartigues P, Baia M, Briere J, Delarue R, Canioni D, et al.
1073 MYC-IG rearrangements are negative predictors of survival in DLBCL patients treated
1074 with immunochemotherapy: a GELA/LYSA study. Blood. 2015;126:2466–74.

1075 59. Green MR, Monti S, Rodig SJ, Juszczynski P, Currie T, O'Donnell E, et al. Integrative
1076 analysis reveals selective 9p24.1 amplification, increased PD-1 ligand expression, and
1077 further induction via JAK2 in nodular sclerosing Hodgkin lymphoma and primary
1078 mediastinal large B-cell lymphoma. Blood. 2010;116:3268–77.

1079 60. Wienand K, Chapuy B, Stewart C, Dunford AJ, Wu D, Kim J, et al. Genomic analyses of
1080 flow-sorted Hodgkin Reed-Sternberg cells reveal complementary mechanisms of immune
1081 evasion. Blood Adv. American Society of Hematology; 2019;3:4065–80.

1082 61. Reichel J, Chadburn A, Rubinstein PG, Giulino-Roth L, Tam W, Liu Y, et al. Flow sorting
1083 and exome sequencing reveal the oncogenome of primary Hodgkin and Reed-Sternberg
1084 cells. Blood. American Society of Hematology; 2015;125:1061–72.

1085 62. Chapuy B, Stewart C, Dunford AJ, Kim J, Wienand K, Kamburov A, et al. Genomic
1086 analyses of PMBL reveal new drivers and mechanisms of sensitivity to PD-1 blockade.
1087 Blood. American Society of Hematology; 2019;134:2369–82.

1088 63. Mottok A, Hung SS, Chavez EA, Woolcock B, Telenius A, Chong LC, et al. Integrative
1089 genomic analysis identifies key pathogenic mechanisms in primary mediastinal large B-cell
1090 lymphoma. Blood. American Society of Hematology; 2019;134:802–13.

1091 64. Noerenberg D, Briest F, Hennch C, Yoshida K, Hablesreiter R, Takeuchi Y, et al. Genetic
1092 Characterization of Primary Mediastinal B-Cell Lymphoma: Pathogenesis and Patient
1093 Outcomes. J Clin Oncol. Wolters Kluwer; 2023;JCO.23.01053.

1094 65. Ansell SM, Lesokhin AM, Borrello I, Halwani A, Scott EC, Gutierrez M, et al. PD-1
1095 Blockade with Nivolumab in Relapsed or Refractory Hodgkin's Lymphoma. N Engl J Med.
1096 2015;372:311–9.

1097 66. Armand P, Rodig S, Melnichenko V, Thieblemont C, Bouabdallah K, Tumyan G, et al.
1098 Pembrolizumab in Relapsed or Refractory Primary Mediastinal Large B-Cell Lymphoma.
1099 https://doi.org/101200/JCO1901389. American Society of Clinical Oncology;
1100 2019;37:3291–9.

1101 67. Armand P, Engert A, Younes A, Fanale M, Santoro A, Zinzani PL, et al. Nivolumab for
1102 Relapsed/Refractory Classic Hodgkin Lymphoma After Failure of Autologous
1103 Hematopoietic Cell Transplantation: Extended Follow-Up of the Multicohort Single-Arm
1104 Phase II CheckMate 205 Trial. J Clin Oncol. 2018;36:1428–39.

1105 68. Aoki T, Chong LC, Takata K, Milne K, Hav M, Colombo A, et al. Single-cell transcriptome
1106 analysis reveals disease-defining t-cell subsets in the tumor microenvironment of classic
1107 hodgkin lymphoma. Cancer Discov. American Association for Cancer Research Inc.;
1108 2020;10:406–21.

1109 69. Melzner I, Bucur AJ, Brüderlein S, Dorsch K, Hasel C, Barth TFE, et al. Biallelic mutation
1110 of SOCS-1 impairs JAK2 degradation and sustains phospho-JAK2 action in the MedB-1
1111 mediastinal lymphoma line. Blood. 2005;105:2535–42.

1112 70. Savage KJ, Monti S, Kutok JL, Cattoretti G, Neuberg D, De Leval L, et al. The molecular
1113 signature of mediastinal large B-cell lymphoma differs from that of other diffuse large B-
1114 cell lymphomas and shares features with classical Hodgkin lymphoma. Blood.
1115 2003;102:3871–9.

1116 71. Viganò E, Gunawardana J, Mottok A, Van Tol T, Mak K, Chan FC, et al. Somatic IL4R
1117 mutations in primary mediastinal large B-cell lymphoma lead to constitutive JAK-STAT
1118 signaling activation. Blood. 2018;131:2036–46.

1119 72. Alexander WS, Starr R, Fenner JE, Scott CL, Handman E, Sprigg NS, et al. SOCS1 is a
1120 critical inhibitor of interferon γ signaling and prevents the potentially fatal neonatal actions
1121 of this cytokine. Cell. Elsevier B.V.; 1999;98:597–608.

1122 73. Liau NPD, Laktyushin A, Lucet IS, Murphy JM, Yao S, Whitlock E, et al. The molecular
1123 basis of JAK/STAT inhibition by SOCS1. Nat Commun 2018 91. Nature Publishing Group;
1124 2018;9:1–14.

1125 74. Horvath CM. The Jak-STAT Pathway Stimulated by Interferon γ. Sci STKE. American
1126 Association for the Advancement of Science; 2004;2004:tr8–tr8.

1127 75. Villarino AV, Kanno Y, O'Shea JJ. Mechanisms and consequences of Jak–STAT signaling
1128 in the immune system. Nat Immunol. Nature Publishing Group; 2017;18:374–84.

1129 76. Ivashkiv LB. IFNγ: signalling, epigenetics and roles in immunity, metabolism, disease and
1130 cancer immunotherapy. Nat Rev Immunol. 2018;18:545–58.

1131 77. Hugo W, Zaretsky JM, Sun L, Song C, Moreno BH, Hu-Lieskovan S, et al. Genomic and
1132 Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma. Cell.
1133 2016;165:35–44.

1134 78. Kalbasi A, Ribas A. Tumour-intrinsic resistance to immune checkpoint blockade. Nat Rev
1135 Immunol. 2020;20:25–39.

1136 79. Choi J, Phelan JD, Wright GW, Häupl B, Huang DW, Shaffer AL, et al. Regulation of B cell
1137 receptor-dependent NF-κB signaling by the tumor suppressor KLHL14. Proc Natl Acad
1138 Sci. Proceedings of the National Academy of Sciences; 2020;117:6092–102.

1139 80. Staudt LM. Oncogenic Activation of NF-κB. Cold Spring Harb Perspect Biol.
1140 2010;2:a000109.

1141 81. Love C, Sun Z, Jima D, Li G, Zhang J, Miles R, et al. The genetic landscape of mutations in
1142 Burkitt lymphoma. Nat Genet. Nature Publishing Group; 2012;44:1321–5.

1143 82. Zhou P, Blain AE, Newman AM, Zaka M, Chagaluka G, Adlar FR, et al. Sporadic and
1144 endemic Burkitt lymphoma have frequent FOXO1 mutations but distinct hotspots in the
1145 AKT recognition motif. Blood Adv. 2019;3:2118–27.

1146 83. Trinh DL, Scott DW, Morin RD, Mendez-Lago M, An J, Jones SJM, et al. Analysis of
1147 FOXO1 mutations in diffuse large B-cell lymphoma. Blood. 2013;121:3666–74.

1148 84. Scott DW, Gascoyne RD. The tumour microenvironment in B cell lymphomas. Nat Rev
1149 Cancer. 2014;14:517–34.

1150 85. Sha C, Barrans S, Cucco F, Bentley MA, Care MA, Cummin T, et al. Molecular High-Grade
1151 B-Cell Lymphoma: Defining a Poor-Risk Group That Requires Different Approaches to
1152 Therapy. J Clin Oncol Off J Am Soc Clin Oncol. American Society of Clinical Oncology;
1153 2019;37:202–12.

1154 86. Yoshimura A, Naka T, Kubo M. SOCS proteins, cytokine signalling and immune regulation.
1155 Nat Rev Immunol 2007 76. Nature Publishing Group; 2007;7:454–65.

1156 87. Patel SJ, Sanjana NE, Kishton RJ, Eidizadeh A, Vodnala SK, Cam M, et al. Identification of
1157 essential genes for cancer immunotherapy. Nature. Nature Publishing Group;
1158 2017;548:537–42.

1159 88. Wee ZN, Li Z, Lee PL, Lee ST, Lim YP, Yu Q. EZH2-Mediated Inactivation of IFN-γ-JAK-
1160 STAT1 Signaling Is an Effective Therapeutic Target in MYC-Driven Prostate Cancer. Cell
1161 Rep. Elsevier; 2014;8:204–16.

1162 89. Zaretsky JM, Garcia-Diaz A, Shin DS, Escuin-Ordinas H, Hugo W, Hu-Lieskovan S, et al.
1163 Mutations Associated with Acquired Resistance to PD-1 Blockade in Melanoma.
1164 https://doi.org/101056/NEJMoa1604958. Massachusetts Medical Society; 2016;375:819–
1165 29.

1166 90. Gao J, Shi LZ, Zhao H, Chen J, Xiong L, He Q, et al. Loss of IFN-γ Pathway Genes in
1167 Tumor Cells as a Mechanism of Resistance to Anti-CTLA-4 Therapy. Cell. 2016;167:397-
1168 404.e9.

1169 91. Baumgartner CK, Ebrahimi-Nik H, Iracheta-Vellve A, Hamel KM, Olander KE, Davis TGR,
1170 et al. The PTPN2/PTPN1 inhibitor ABBV-CLS-484 unleashes potent anti-tumour
1171 immunity. Nature. Nature Publishing Group; 2023;622:850–62.

1172 92. Miao D, Margolis CA, Gao W, Voss MH, Li W, Martini DJ, et al. Genomic correlates of
1173 response to immune checkpoint therapies in clear cell renal cell carcinoma. Science.
1174 2018;806:801–6.

1175 93. Bröske A-ME, Korfi K, Belousov A, Wilson S, Ooi C-H, Bolen CR, et al.
1176 Pharmacodynamics and molecular correlates of response to glofitamab in
1177 relapsed/refractory non-Hodgkin lymphoma. Blood Adv. 2022;6:1025–37.

1178 94. Scholler N, Perbost R, Locke FL, Jain MD, Turcan S, Danan C, et al. Tumor immune
1179 contexture is a determinant of anti-CD19 CAR T cell efficacy in large B cell lymphoma.
1180 Nat Med. Nature Publishing Group; 2022;28:1872–82.

1181 95. Shen R, Seshan VE. FACETS: allele-specific copy number and clonal heterogeneity analysis
1182 tool for high-throughput DNA sequencing. Nucleic Acids Res. 2016;44:e131.

1183 96. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor.
1184 Bioinformatics. 2018;34:i884–90.

1185 97. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform.
1186 Bioinformatics. 2009;25:1754–60.

1187 98. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, del Angel G, Levy-Moonshine A, et
1188 al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best
1189 practices pipeline. Curr Protoc Bioinforma Ed Board Andreas Baxevanis Al.
1190 2013;11:11.10.1-11.10.33.

1191 99. Saunders CT, Wong WSW, Swamy S, Becq J, Murray LJ, Cheetham RK. Strelka: accurate
1192 somatic small-variant calling from sequenced tumor–normal sample pairs. Bioinformatics.
1193 2012;28:1811–7.

1194 100. Dunn T, Berry G, Emig-Agius D, Jiang Y, Lei S, Iyer A, et al. Pisces: an accurate and
1195 versatile variant caller for somatic and germline next-generation sequencing data.
1196 Bioinformatics. 2019;35:1579–81.

1197 101. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, et al. The Ensembl
1198 Variant Effect Predictor. Genome Biol. 2016;17:122.

1199 102. Favero F, Joshi T, Marquard AM, Birkbak NJ, Krzystanek M, Li Q, et al. Sequenza: allele-
1200 specific copy number and mutation profiles from tumor sequencing data. Ann Oncol.
1201 Elsevier; 2015;26:64–70.

1202 103. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq
1203 quantification. Nat Biotechnol. Nature Publishing Group; 2016;34:525–7.

1204 104. Law CW, Chen Y, Shi W, Smyth GK. voom: precision weights unlock linear model
1205 analysis tools for RNA-seq read counts. Genome Biol. 2014;15:R29.

1206 105. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene
1207 set enrichment analysis: a knowledge-based approach for interpreting genome-wide
1208 expression profiles. Proc Natl Acad Sci U S A. National Academy of Sciences;
1209 2005;102:15545–50.

1210 106. Mayakonda A, Lin D-C, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and
1211 comprehensive analysis of somatic variants in cancer. Genome Res. 2018;28:1747–56.

1212 107. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of
1213 cell subsets from tissue expression profiles. Nat Methods. Nature Publishing Group;
1214 2015;12:453–7.

1215 108. Zhang Y, Parmigiani G, Johnson WE. ComBat-seq: batch effect adjustment for RNA-seq
1216 count data. NAR Genomics Bioinforma. 2020;2:lqaa078.