Multi-Omics Profiling Reveals Gene Signatures and Therapeutic Targets in HER2-Guided Gastric Cardia Adenocarcinoma Patients

Li-Dong Wang[1, *], Pengwei Xing[2], Xue-Ke Zhao[1], Xin Song[1], Meng-Xia Wei[1], Duo You[1, 3], Ling-Ling Lei[1], Rui-Hua Xu[1], Ran Wang[1], Lei Ma[4], Pei-Nan Chen[5], Xinmin Li[6], Minglu Xie[2], Miao Zhao[2], He Zhang[7], Hui-Fang Lv[4], Ai-Li Li[8], Xian-Zeng Wang[9], She-Gan Gao[10, *], Xingqi Chen[2, *]

[1]State Key Laboratory of Esophageal Cancer Prevention & Treatment and Henan Key Laboratory for Esophageal Cancer Research of the First Affiliated Hospital, Zhengzhou University, Zhengzhou 450052, China.

[2]Department of Immunology, Genetics and Pathology, Uppsala University and Science for Life Laboratory, Uppsala, Sweden

[3]Department of Medical Oncology, the Affiliated Cancer Hospital of Zhengzhou University, Zhengzhou, Henan Province, 450008, China

[4]Henan key Laboratory of Tumor Molecular Biology, Medicine of First People's Hospital of Nanyang, Nanyang, 473004, China.

[5]Deaprtment of Thoracic Surgery, the Affiliated Cancer Hospital of Zhengzhou University, Zhengzhou, Henan Province, 450008, China

[6]Department of Pathology, Women and Infants Hospital of Zhengzhou, Zhengzhou, Henan Province, 450012, China

[7]Deaprtment of Pathology, the Affiliated Cancer Hospital of Zhengzhou University, Zhengzhou, Henan Province, 450008, China

[8]Department of Pathology, Linzhou Tumor Hospital, Linzhou, Henan Province, 456500, China

[9]Department of Thoracic Surgery, Linzhou People's Hospital, Linzhou, Henan Province, 456500, China

[10]Henan Key Laboratory of Microbiome and Esophageal Cancer Prevention and Treatment; Henan Key Laboratory of Cancer Epigenetics; Cancer Hospital, The First Affiliated Hospital, College of Clinical Medicine of Henan University of Science and Technology, Luoyang, 471003, China

*Correspondence: xingqi.chen@igp.uu.se, ldwang2007@126.com, gsg112258@163.com

Abstract:

Therapeutic targets for Gastric Cardia Adenocarcinoma (GCA), particularly for HER2-negative patients, are lacking. Here, we conducted multi-omics profiling on 128 GCA patients using mass spectrometry, whole-exome sequencing, RNA-Seq, and metabolomics. We found HER2 to be a favorable prognostic marker for GCA. Employing molecular counting, we categorized patients into HER2-high, -low, and -negative groups. We uncovered an enrichment of DNA repair features in the HER2-high group, while HER2-low and -negative groups exhibited strong inflammation. We found that tumor mutation burden may not be the distinguishing factor among these three groups. We revealed that the HER2-negative and -low groups have a tumor-suppressive immune microenvironment, and HER2 expression is associated with fatty acid metabolic profiles and inflammation in the blood of patients. Our study revealed anti-inflammatory and immune checkpoint inhibition, targeting PD-L2 and the CD47/SIRPA pair, as therapeutic strategies for HER2-negative GCA patients. These findings highlight promising avenues for personalized treatment for GCA.

Keywords: Gastric cardia adenocarcinoma, HER2, ERBB2, HER2-negative, Prognosis, Tumor microenvironment, Inflammation, Immune inhibitory checkpoints, SIRPA, CD47, DNA repair

63 **Introduction**

64 Gastric cardia adenocarcinoma(GCA) is a type II adenocarcinoma at the gastroesophageal
65 junction [1]. GCA exhibits distinct features compared to esophageal and gastric cancers [2-5], yet
66 therapeutic approaches often overlap. Consequently, specific therapeutic targets for GCA are
67 lacking. Human Epidermal Growth Factor Receptor 2 (HER2), the protein product of the Erb-
68 B2 Receptor Tyrosine Kinase 2 (*ERBB2*) gene, plays a crucial role in regulating cell growth
69 and division[6]. In certain cancers, such as breast cancer [7,8] and gastroesophageal cancer [9], HER2
70 can become overexpressed or amplified, leading to uncontrolled cell growth [10]. Targeted
71 therapies, such as trastuzumab (Herceptin) and others, have been developed to specifically
72 inhibit HER2 signaling and are effective in treating HER2-positive breast [11] and
73 gastroesophageal [12] cancers. Currently, HER2 targeted therapy has become a vital treatment
74 option for certain types of gastroesophageal cancers, including esophageal Adenocarcinoma
75 (EAC), Esophageal squamous-cell carcinoma (ESCC), GCA, and stomach adenocarcinoma
76 (STAD), particularly in advanced or metastatic cases [9,13-15]. Anti-HER2 targeted therapies have
77 improved survival outcomes in advanced gastroesophageal cancer patients with HER2
78 overexpression [13,14,16]. However, the relatively low prevalence (less than 30%) of HER2-
79 positive cases poses challenges [17-19], particularly for HER2-negative patients. The first-line
80 treatment for advanced HER2-negative gastric cancer and esophagogastric junction
81 adenocarcinoma mainly involves immunotherapy combined with chemotherapy, without
82 specific targets[9]. In solid malignancies, the overexpression or amplification of HER2 is
83 associated with an adverse prognosis [20,21], paradoxically, independent studies including our
84 recent study, have shown that *ERBB2* gene focal amplification or HER2 overexpression might
85 confer a favorable prognostic impact in patients with esophagogastric junction adenocarcinoma,
86 EAC [22] and GCA [23] . Hence, the role of HER2 as either a therapeutic target or a prognostic
87 marker in GCA remains uncertain. Consequently, conducting a comprehensive multi-omics
88 molecular profiling of GCA patients, stratified by the degree of HER2 expression, is imperative
89 for elucidating this paradox and gaining a more profound understanding that can inform
90 targeted therapy strategies for these malignancies, particularly the HER2-negative patients.
91

92 In this investigation, we assembled a cohort comprising 128 patients diagnosed with GCA. We
93 conducted an extensive array of analyses on our GCA cohort, including RNA-Seq for
94 transcriptome profiling, whole exome sequencing (WES) for examining genetic alterations,
95 mass spectrometry for the identification and quantification of proteins, single-cell RNA-Seq
96 (scRNA-Seq), and spatial transcriptomics for dissecting the tumor microenvironment. With a

97 highly sensitive molecular counting approach, we stratified GCA patients into HER2-high, -
98 low, and -negative groups. This method enabled us to undertake a comprehensive and detailed
99 characterization of molecular features for each patient group, including the exploration of gene
100 expression patterns, profiling of proteins, and examination of genetic alterations. We found
101 HER2 as a favorite prognostic marker in GCA and EAC, but not in gastric cancers.
102 Furthermore, our investigations unveiled a substantial enrichment of DNA repair pathways in
103 the HER2 high-expression group, while the HER2 low-expression and negative group
104 prominently exhibited features indicative of inflammation. We found that tumor mutation
105 burden (TMB) may not be the primary distinguishing factor among these three groups, but the
106 HER2 high group showed substantially higher genome instability compared with the HER2
107 negative group. We also found *ARID1A* mutations were significant for prognosis, especially in
108 the HER2-negative group. Notably, scRNA-Seq and spatial transcritons revealed the HER2-
109 high group displayed stronger growth factor signaling, whereas the HER2-low and HER2-
110 negative groups showed pronounced inflammation and higher M2 macrophage levels, and
111 variations in immune inhibitory checkpoints across the distinct HER2 expression groups. With
112 validation from multiple angles, our study suggests that anti-inflammatory therapy and immune
113 checkpoint inhibition targeting the CD47/SIRPA axis may be potential therapeutic strategies
114 for HER2-negative GCA patients.
115

116 **Results**

117 **HER2/*ERBB2* is a favorable prognostic marker in GCA and EAC**

118 Previously, we and others found that *ERBB2* gene focal amplification or HER2 overexpression
119 might confer a favorable prognostic impact on patients with esophageal adenocarcinoma (EAC)
120 and GCA [22,23]. This conclusion is derived from immunohistochemistry (IHC) staining of HER2
121 or gene amplification detection using whole-genome sequencing. These measurements involve
122 positive and negative HER2 staining or amplification assessments of patients. The HER2
123 scoring system typically employs immunohistochemistry (IHC) and/or fluorescence in situ
124 hybridization (FISH) techniques to evaluate the quantity and distribution of HER2 protein and
125 gene amplification in cancer tissue samples [24]. In the whole-genome sequencing method, HER2
126 amplification is examined within amplification regions[24]. It is well known there is a big
127 heterogeneous expression of HER2 protein in different regions of the tumor from the same
128 patient [17]. RNA-Seq is a molecular counting method used to calculate the number of transcripts.
129 To better understand the relationship between HER2 expression and patient prognosis in GCA,
130 we performed RNA-seq and mass spectrometry on both tumors and NATs from 54 GCA

patients (**Supplementary Figure 1a-c**). We observed a strong correlation between *ERBB2* RNA expression levels and HER2 protein levels in tumors (Pearson correlation = 0.7559), but amost no correlation in NATs (Pearson correlation = -0.03) (**Figure 1a**). Based on the median expression level of *ERBB2* RNA, we stratified GCA patients into high and low groups (**Figure 1b**, **Supplementary Figure 2)**. We also analyzed RNA-Seq data from TCGA for EAC, ESCC, and STAD (Stomach adenocarcinoma), stratifying patients into HER2-high and HER2-low groups based on the median *ERBB2* RNA expression levels. Significant differences in survival probabilities were observed among GCA (p = 0.011) and EAC (p = 0.0096) patients (**Figure 1c**). Specifically, the *ERBB2* high group demonstrated significantly better survival probabilities compared to the *ERBB2* low group. No significant survival difference was found in ESCC (p = 0.59) or STAD (p = 0.57). Further analysis within STAD subtypes revealed no difference in diffuse or intestinal carcinoma types (**Figure 1c**). Although not statistically significant, there was a trend for better survival in the *ERBB2* high group for STAD tubular adenocarcinoma (p = 0.19) and STAD NOS adenocarcinoma (p = 0.4) (**Figure 1c**). We also used quantile and natural separation as cutoffs for patient grouping (**Supplementary Figure 2**). In the quantile grouping, we still observed that high levels of HER2 RNA expression were associated with better prognosis in GCA (p = 0.049) and EAC (p = 0.05), but not in ESCC or STAD, further indicating that HER2 could be a potential prognostic marker for GCA and EAC. Considering genetic background variations, we also calculated the HER2 RNA level ratio between tumors and the corresponding NATs and divide patients into high and low groups (**Methods, Figure 1d**), we still observed significantly better survival in the HER2 high group (p = 0.021, **Figure 1e**). Next, we expanded our cohort to 128 GCA patients for protein quantification of tumors and NATs (**Supplementary Figure 1**). 31 patients showed HER2 negativity in both tumors and NATs, while 97 patients (75.78%) were HER2 positive. We calculated the HER2 level ratio between tumor and corresponding NATs to quantify HER2 protein levels across patients. Patients were then categorized into HER2-high (n = 48), HER2-low (n = 49), and HER2-negative (n = 31) groups based on this ratio (**Figure 1f**, **Methods**). The HER2-high group showed the best survival rates among the three groups (p = 0.014, **Figure 1g**).

The grouping assignment of patients based on HER2 protein levels showed a significant association with the HER2 RNA-based grouping assignment (**Supplementary Figure 3a**). The HER2 scoring system was applied to the tumor samples in our GCA cohort (**Supplementary Figure 3b**), where 6.3% of patients were HER2 3+. The HER2-high group showed the largest proportion of HER2 3+ cases (14%), compared to 2% in the HER2-low group and none in the

165 HER2-negative group. HER2 scoring significantly correlated with the MS HER2 grouping
166 method (p = 0.034), suggesting that MS groups may accurately reflect HER2 expression
167 (**Supplementary Figure 3c**). Post-surgery therapeutic strategies were unrelated to HER2
168 grouping (p = 0.320, **Supplementary Figure 3d**) or tumor stage (**Supplementary Figure 3e**)
169 ansd did not affect prognosis of patients (p = 0.54, **Supplementary Figure 3f**). Correlation
170 analysis showed that HER2 grouping was significantly associated with age (p = 0.005), family
171 history (p = 0.049), alcohol consumption (p = 0.031), tumor differentiation (p = 0.029), and
172 Lauren classification (p = 0.022) but was independent of other clinicopathological features
173 (**Supplementary Figure 4a** and **4b, Table S1**). We also analyzed the association between
174 prognosis and Lauren classification and found that Lauren classification alone was not a
175 prognostic marker (**Supplementary Figure 5**). However, the HER2-high group showed better
176 prognosis across Lauren classifications, though this was not statistically significant. These
177 findings suggest that while Lauren classification alone may not predict prognosis, it gains
178 prognostic value when combined with HER2 grouping. In summary, integrating TCGA RNA-
179 Seq data with our findings confirms that *ERBB2* RNA expression and HER2 protein levels are
180 favorate prognostic markers for both GCA. This novel observation in GCA and EAC suggests
181 that HER2-positive and high patient groups could be very different from those in other tumor
182 types, such as breast cancer, where HER2-positive patients usually have a poor prognosis [15].
183 Even though anti-HER2 therapies have started to be used for treatment in gastric and
184 esophageal cancers[13,14,16], following the strategy used in breast cancer, the specific features of
185 HER2-positive patients in GCA may differ significantly from those in breast cancer. Detailed
186 molecular characterization of GCA patients based on HER2 expression is crucial for better
187 understanding GCA and identifying therapeutic targets for HER2-negative patients. Motivated
188 by this, we conducted comprehensive analyses on our GCA cohort, including WES, RNA-Seq,
189 scRNA-Seq, and spatial transcriptomics (**Figure 2a**).
190

191 **Enrichment of DNA repair pathways in HER2-High tumors and inflammatory signatures**
192 **in HER2-Low/Negative tumors**

193 A similar number of proteins were detected in both tumors (median = 5,088) and NATs
194 (median = 5,089) (**Table S2**). To compare HER2 groups effectively, we used two strategies to
195 identify signature proteins: comparing tumor protein expression across HER2 groups and
196 evaluating protein levels by comparing tumors to NATs within each HER2 group. In the tumor-
197 only comparison, we identified 23 proteins unique to the High group, 6 to the Low group, and
198 92 to the Negative group. Additionally, 31 proteins were shared by the Low and Negative

199 groups, 11 by the High and Negative groups, and 132 by the High and Low groups (Fold
200 Change > 1.5, $p < 0.05$) (**Supplementary Figure 6a**, **Figure 2b**, **Table S3**). The HER2-
201 negative group had the most differentially expressed proteins. Gene set enrichment analysis
202 (GSEA) of hallmark gene sets for each HER2 group revealed that the HER2-high group was
203 enriched in proliferation-related sets, including Myc targets, mTORC1 signaling, and E2F
204 targets (**Figure 2c**, **Supplementary Figure 6b**, **Table S4**). Similarly, E2F targets indicated
205 proliferation features in the HER2-low group. Notably, the HER2-low group showed strong
206 inflammatory features, with enrichment in Interferon alpha response, Interferon gamma
207 response, and TNFA signaling via NFκB. Comparing tumors with their corresponding NATs
208 (**Figure 2d**, **Table S5**) revealed that Myc targets v1 were more enriched in the HER2-high and
209 low groups than in the HER2-negative group. Inflammatory hallmarks were present in all
210 groups, with the interferon-alpha response notably enriched in the HER2-low group.
211 Additionally, DNA repair and angiogenesis hallmarks were uniquely enriched in the HER2-
212 high group. We analyzed differential protein expression between tumors and NATs for each
213 HER2 group (Fold Change > 1.5, $p < 0.05$) (**Supplementary Figure 7a, 7b**, **Table S6**). In the
214 HER2-high group, we identified 167 up-regulated and 299 down-regulated proteins; the HER2-
215 low group had 99 up-regulated and 251 down-regulated proteins; and the HER2-negative group
216 had 154 up-regulated and 305 down-regulated proteins. Gene ontology (GO) enrichment
217 analysis of significant up-regulated proteins revealed distinct biological processes
218 (**Supplementary Figure 7c**, **Table S7**). The HER2-high group showed enrichment in DNA
219 replication terms, while the HER2-low group enriched immune response terms, and the HER2-
220 negative group enriched terms related to viral response.
221
222 We extracted common and unique proteins among the three HER2 groups (**Supplementary**
223 **Figure 8**). The HER2-high group had 60 specific proteins, including HER2, MCM2, MCM3,
224 and MCM7, highlighting HER2's role in cell cycle regulation and DNA replication. In contrast,
225 the HER2-low group contained 15 specific proteins enriched in inflammatory-related proteins
226 like TAP1, TAP2, IFT1, and IFIT3 [25], consistent with the inflammation features observed in
227 GSEA and GO term enrichment analyses. The HER2-negative group included 47 specific
228 proteins with HLA-DRA and HLA-A, crucial for immune response [26]. Overall, our analysis
229 indicated that HER2-high groups exhibited stronger proliferation and DNA repair features,
230 while HER2-negative and low groups displayed more pronounced inflammatory features.
231

232 **Comparable mutation burden across groups with elevated genome gains in HER2-High**
233 **patients**
234 Next, we examined genome mutations and gains among the three HER2 groups (**Figure 3**,
235 **Supplementary Figure 9**, **Table S8**). Although there was a slight reduction in SNV and InDel
236 frequencies in the HER2-negative group, statistical analysis revealed no significant differences
237 (**Supplementary Figure 10a**, **10b**). All three groups showed a high degree of similarity in
238 mutation types (**Supplementary Figure 10c**). Additionally, there are no significant differences
239 in non-synonymous or SNV frequencies across the groups (**Supplementary Figure 10d**). We
240 compared the mutation frequencies of the top 20 mutated cancer genes across each group
241 (**Figure 3a**, **Supplementary Figure 10e**, **Table S9**). Notably, ARID1A was the only gene with
242 a significantly different mutation frequency among the three groups (p = 0.016), with higher
243 mutations in the HER2-negative and HER2-low groups. Patients with ARID1A mutations had
244 a worse prognosis (p = 0.00069) (**Figure 3b**). While prognosis varied among the three HER2
245 groups, those with ARID1A mutations generally fared worse, with significant differences in
246 the HER2-high ($p < 0.0001$) and HER2-low groups (p = 0.036), but not in the HER2-negative
247 group (p = 0.35) (**Figure 3b**). These findings suggest that HER2 expression and ARID1A
248 mutation signatures together impact the prognosis of GCA patients. This conclusion is
249 supported by examining patient prognosis differences across the three HER2 groups in both
250 mutated and unmutated contexts (**Figure 3c**). Other top mutated genes, such as FAT4, may
251 also affect patient prognosis in conjunction with HER2 grouping (**Supplementary Figure 11**).
252 We assessed chromosomal instability (CIN) in the three groups using the genome integrity
253 index (GII) [27] and weighted genome instability index (wGII) [28]. The mean GII values were
254 0.102 for the HER2-high group, 0.078 for the HER2-low group, and 0.071 for the HER2-
255 negative group. The HER2-high group's GII was 1.44 times higher than that of the HER2-
256 negative group (p = 0.015) (**Figure 3d**). Similarly, the mean wGII values were 0.060 for the
257 HER2-high group, 0.027 for the HER2-low group, and 0.021 for the HER2-negative group,
258 with the HER2-high group showing a 2.86-fold increase over the HER2-negative group (p =
259 0.0096) (**Figure 3d**).We also calculated chromosomal breakpoints, represented as the CIN
260 Index (CINindex) and weighted CIN Index (wCINindex). The HER2-high group exhibited
261 significantly higher CINindex values, 1.36 times greater than the HER2-negative group (p =
262 0.011), and wCINindex values, 1.37 times higher (p = 0.0064) (**Figure 3e**). While the HER2-
263 high group showed more copy number alterations (CNA), the difference in CNA across the
264 three HER2 groups was not statistically significant (p = 0.064) (**Figure 3f**).
265

266 We calculated genome gain frequencies using 1 Mbp bins among the HER2 groups (**Figure**
267 **3g**, **Supplementary Figure 12**, **Table S10**). Notably, 142 regions exhibited higher gains in the
268 HER2-high group, while 16 regions were more amplified in the HER2-negative group
269 compared to both the high and low groups. Subregions of Chromosomes 13, 15, and 17 showed
270 increased gains in the HER2-high group (**Figure 3g**, **3h**). Interestingly, Chromosome 19
271 subregions displayed high gains in both the HER2-high and negative groups but not in the low
272 group. We also examined oncogenes and tumor suppressor genes (TSGs) in the gained regions
273 (**Figure 3g**, **3h**). *CCNE1*, *MLLT1*, *CCND3*, and *FGFR3* were significantly gained in the HER2-
274 high group. Notably, a subregion of Chromosome 13 containing the TSG *BRCA2* showed
275 greater gain in the HER2-high group compared to the others. Additionally, ERBB2 gene gains
276 were markedly more frequent in the HER2-high group (31%) compared to the low and negative
277 groups (0%). Another TSG, *BRCA1*, located near ERBB2, also showed increased gain in the
278 HER2-high group, though not statistically significant. In summary, our findings indicate that
279 mutation is not a main distinguishing factor among these groups; however, the HER2-high
280 group shows a significantly higher frequency of genome gains.
281
282 **Stronger inflammation features are observed in HER2-negative groups**
283 Clear distinctions were observed between the RNA expression profiles of tumors and NATs
284 (**Supplementary Figure 13a**). GSEA of tumor RNA-Seq data revealed significant enrichment
285 in six cancer hallmark terms in the HER2-high group (**Figure 4a**, **Table S11**), particularly in
286 DNA repair features. The HER2-low group enriched 20 cancer hallmarks, with inflammation-
287 related terms like TNFA signaling via NFκB and Interferon gamma response. The HER2-
288 negative group mainly exhibited enrichment in inflammatory response and allograft rejection.
289 A differential gene expression analysis identified 572 significantly differentially expressed
290 genes (Fold Change > 1.5, FDR < 0.05) across the HER2 groups (**Figure 4b**, **Supplementary**
291 **Figure 13b**, **Table S12**). DNA repair-related genes, such as H2AFX [29], were enriched in both
292 HER2-high and low groups (**Figure 4b**), while infection and immune response genes like
293 *TNFSF4*, *TNFSF18*, and *TRIM50* were uniquely enriched in the HER2-negative group. We
294 proceeded with GSEA analysis within each HER2 group comparing tumor samples with their
295 corresponding NATs (**Figure 4c**, **Table S13**). Distinct enrichment patterns were evident, with
296 unique terms like TNFA signaling via NFκB, Apical junction, IL-6/JAK2/STAT3 signaling,
297 and Apoptosis enriched in HER2-low and negative groups. Notably, DNA repair hallmark was
298 more pronounced in HER2-high and low groups. Subsequently, a differential gene expression
299 analysis comparing tumors with NATs (Fold Change > 1.5, FDR < 0.05) revealed 5084, 5338,

300 and 4697 differentially expressed genes in the HER2-high, -low, and -negative groups,
301 respectively (**Supplementary Figure 14Aa-c**, **Table S14, Table S15**).
302
303 Notably, MHC class II genes, including HLA-DMA, HLA-DMB, HLA-DRA, HLA-DQB1,
304 and HLA-DQA1, were active only in the HER2-low and -negative groups (**Figure 4d**). GO
305 term analysis using the list of up-regulated genes in tumors compared to NATs showed that
306 cell cycle and DNA repair-related GO terms were strongly enriched in both the HER2-high
307 and HER2-low groups, while inflammatory terms were prominently enriched in the HER2-
308 negative group (**Supplementary Figure 15a-c, Table S16**). We also identified exclusively up-
309 regulated genes in each HER2 group for these targeted GO terms (**Table S17**). Notably, genes
310 like *H2AFX* and *BRCA2*, crucial for DNA repair, were significantly expressed in HER2-high
311 tumors compared to NATs. *BRCA1* expression was significantly higher in HER2-high and low
312 tumors compared to NATs. Additionally, MHC class II GO terms were exclusively enriched
313 in the negative group (**Supplementary Figure 15d**). In the HER2-high group, all the top 15
314 GO terms are associated with the mitotic cycle and chromosome segregation (**Supplementary**
315 **Figure 15e**). Similarly, in the HER2-low group, 9 of the top 15 GO terms focus on these
316 processes, with an additional layer related to inflammation. In the HER2-negative group, 10 of
317 the top 15 GO terms primarily revolve around inflammation. We analyzed the differential
318 expression of cytokine genes between tumor and adjacent tissue in each HER2 group (**Figure**
319 **4e**). The HER2-negative group showed the highest activity, with 33 cytokine genes, including
320 IL-6 and IFNG, exclusively active in HER2-negative tumors. IL-6 is linked to chronic
321 inflammation [30], while IFN-γ promotes inflammation [31], underscoring distinct inflammatory
322 responses in this group. We also evaluated 18 tumor inflammation genes [32] (**Supplementary**
323 **Figure 16**); 50% were significantly elevated in HER2-negative tumors, 44.4% in HER2-low
324 tumors, and only 16.6% in HER2-high tumors. These findings indicate distinct inflammatory
325 tumor microenvironments across HER2 groups.
326
327 **Single-cell RNA-Seq reveals that lower HER2 expression is linked to stronger**
328 **inflammatory features and higher M2 macrophage levels**
329 To better understand the tumor microenvironment in different HER2 groups, we first used
330 CIBERSORTx [33] to analyze immune cells in HER2 groups from bulk RNA-Seq data. We
331 identified 22 immune cell types, with no significant differences in NATs (**Supplementary**
332 **Figure 17a**, **17b**). In tumors, the HER2-negative group had significantly more M2
333 macrophages than the HER2-high group (p = 0.036), while the HER2-low group showed a

334 higher but not statistically significant proportion (p = 0.21) (**Figure 5a**, **5b**, **Supplementary**
335 **Figure 17c**). Staining for the M2 macrophage marker CD163 confirmed a higher presence in
336 the HER2-low (p = 0.016) and negative groups (p = 0.18) compared to the HER2-high group
337 (**Figure 5c**, **5d**, **Table S18**). Although the difference between the HER2-high and negative
338 groups was not significant, there was a trend toward more CD163-positive cells in the HER2-
339 negative. Our findings suggest distinct tumor microenvironments linked to HER2 expression
340 levels in GCA patients.
341
342 To further confirm our findings, we performed scRNA-Seq on tumors from seven GCA
343 patients, categorizing them into HER2 high (n = 4) and HER2 low groups (n = 3) based on
344 HER2 RNA expression (**Supplementary Figure 18a-d, Figure 5e**). The analysis identified
345 46,912 cells, revealing 19 distinct cell types heterogeneously distributed among the patients
346 (**Figure 5e-h**, **Supplementary Figure 18e**, **Figure S19**). Significant differences between the
347 HER2 high and low groups were found only in M2 macrophages (p = 0.0013) and dendritic
348 cells (p = 0.0015) (**Figure 5i**, **Supplementary Figure 20**), confirming higher M2 macrophage
349 levels in lower HER2 expression patients. Differential gene expression analysis of M2
350 macrophages identified 1,591 highly expressed genes for the HER2-high group and 86 genes
351 for the low group. Cytokine and chemokine genes like CXCL3 and S100A were specifically
352 enriched in the low group (**Supplementary Figure 21**). KEGG analysis showed stronger
353 inflammatory features, including TNF signaling, in the HER2-low group (**Figure 5j**, **Table**
354 **S20**), which is further supported by GO term enrichment (**Supplementary Figure 21b**, **21c**,
355 **Table S21**). Overall, lower HER2 expression is linked to stronger inflammatory features and
356 higher M2 macrophage levels.
357
358 **Single cell analysis uncovered that low HER2 expression correlates with stronger**
359 **inflammatory signaling pathways in epithelial tumor cells**
360 Next, we focused on epithelial tumor cells from scRNA-Seq to understand the communication
361 between tumor cells and their microenvironment. We identified seven subclusters of epithelial
362 tumor cells and found no significant differences in their distribution in the two groups (**Figure**
363 **6a**). We defined cells with high CNV and aneuploidy in epithelial cells (**Figure 6a**, **6b**;
364 **Supplementary Figure 22**). As expected, tumor cells exhibited significantly higher CNV
365 scores than normal epithelial cells (**Figure 6b**, p < 2.2e-16). We found no significant
366 differences in the distribution of aneuploid epithelial cells and high CNV cell numbers between
367 groups (**Figure 6a**, **6b**). We performed pseudotime trajectory analysis of epithelial tumor cells,

368 identifying three cell states (**Figure 6c**, **6d**, **Supplementary Figure 23**). Epithelial tumor cells
369 in the HER2 high group were more enriched in state 3, while those in the low group were more
370 prevalent in state 2 ($p < 2.2e-16$). Based on this analysis, we defined two cell fates (**Figure 6c**)
371 and extracted gene features for each fate (**Figure 6e**, **Table S24**). Cell fate 1, enriched in the
372 low group, exhibited high expression of cytokines and chemokines (e.g., *CCL11*, *CXCL5*)
373 (**Figure 6e**). Cell fate 2, enriched in the high group, showed elevated expression of growth
374 factors (e.g., *ERBB3*, *NRG1*). KEGG analysis revealed stronger inflammatory features in the
375 low group, with pathways like IL-17 signaling, while the high group showed enrichment in
376 growth factor pathways, including ERBB signaling (**Figure 6f**, **Table S23**).Our data suggest
377 that low HER2 expression correlates with stronger inflammatory features in the epithelial
378 tumor cells of GCA patients.
379
380 We observed stronger signaling communications in the HER2 high group compared to the low
381 group (**Figure 6g, Supplementary Figure 24a**, **24b**). In the high group, M1 macrophages
382 exhibited stronger communication with other cell types, while CD4+ and CD8+ T cells
383 interacted well with most. In the low group, M2 macrophages communicated with neutrophils,
384 endothelial cells, and other M2 macrophages, whereas in the high group, they interacted with
385 CD4+ T cells, M1 macrophages, and others (**Figure 6g**, **6h**, **Supplementary Figure 25**). M2
386 macrophages in the low group sent stronger cytokine and chemokine signals, while in the high
387 group, they sent stronger growth factor signals (**Figure 6i**, **Supplementary Figure 24c**).
388 Epithelial tumor cells received stronger growth factor signals in the high group and stronger
389 inflammatory signals in the low group (**Figure 6j**, **Supplementary Figure 25d**). Overall, the
390 high group exhibited stronger growth factor communication, whereas the low group displayed
391 stronger inflammatory signaling (**Figure 6k**, **Supplementary Figure 26**). NRG signaling was
392 exclusively contributed by epithelial tumor cells, endothelial cells, and fibroblasts in the high
393 group (**Figure 6l**). Although CCL and CXCL signaling was stronger in the low group,
394 communication was also present in the high group (**Supplementary Figure 27**). In the high
395 group, CXCL signaling was driven by interactions between regulatory T cells and M2
396 macrophages, while in the low group, it involved M2 macrophages, endothelial cells, and
397 fibroblasts (**Figure 6m**). Together, our data suggest distinct tumor microenvironments, with
398 lower HER2 expression linked to stronger inflammation, based on HER2 expression levels in
399 GCA patients. This conclusion is consistent with our findings from bulk RNA and protein
400 analyses.
401

402 **The CD47-SIRPA axis as a promising therapeutic option for GCA patients with low**
403 **HER2 expression**
404 Since there is a distinct tumor microenvironment in different HER2 groups, with stronger
405 inflammation in HER2-negative groups, we were motivated to identify potential therapeutic
406 targets for HER2-negative patients. Immunotherapy has become a promising therapeutic
407 treatment for esophageal and gastric cancers, often combined with anti-HER2 treatments. We
408 focused on identifying specific immune checkpoints for different HER2 groups. To identify
409 distinct immune inhibitory checkpoints among the three HER2 groups, we compared the RNA
410 expression levels of key immune checkpoint genes, ligands, and receptors between tumors and
411 adjacent normal tissue (NAT) in each group (**Figure 7a**). The HER2-negative group showed
412 significantly higher levels of FGL1 and TNFSF18 in tumors compared to NAT. Both HER2-
413 high and HER2-low groups had elevated HHLA2 expression, while the HER2-low group
414 exhibited increased PVR and CD276 levels. Additionally, CD70, PDCD1LG2 (PD-L2), and
415 SIRPA were higher in HER2-low and HER2-negative tumors. PD-L2 is a ligand for PD-1 and
416 a key immunotherapy target, while the CD47/SIRPA axis is also emerging as a potential
417 therapeutic target [34]. We further investigated the relationship between HER2 expression and
418 immune checkpoint levels, focusing on PD-L2 and CD47/SIRPA in different GCA patient
419 groups. We performed staining of PD-L2 and SIRPA in tumors from different HER2 groups
420 of GCA patients (**Figure 7b**). The proportion of PD-L2-positive cells was significantly higher
421 in the HER2-negative group than in the HER2-positive group ($p = 0.033$), and SIRPA-positive
422 cells were also more prevalent in HER2-negative patients ($p = 0.039$) (**Figure 7c**, **Table S24**).
423 These findings indicate elevated PD-L2 and CD47/SIRPA axis expression in the HER2-
424 negative group.
425
426 Due to the lack of GCA cell lines and similarities between esophageal adenocarcinoma (EAC)
427 and GCA, we employed the EAC cell line (OE19) to assess the relationship between ERBB2
428 expression and SIRPA and PD-L2 expression. Because CD47 and PD-L2 are not expressed in
429 this cell line, we compared SIRPA gene expression levels in OE19 with and without treatment
430 with the HER2 inhibitor lapatinib and after HER2 silencing via siRNA [35]. The results showed
431 a significant increase in SIRPA expression after HER2 inhibitor treatment and HER2 silencing
432 (**Supplementary Figure 26**), confirming that decreased ERBB2 expression correlates with
433 upregulated SIRPA.
434

435 We compared the strength of immune checkpoint interactions in different HER2 groups using
436 scRNA-Seq data (**Figure 7d**). Although we did not observe an enriched interaction between
437 PD-L2 and PD-1 due to low PD-L2 expression, we found a much stronger CD47-SIRPA
438 interaction in the low group compared to the high group, specifically between epithelial and
439 myeloid cells (**Figure 7d**, **Tables 25** and **26**). SIRPA expression was higher in the HER2-low
440 group than in the high group among epithelial cells and M2 macrophages, though this
441 difference was not statistically significant (**Figures 7e**, **7f**). Additionally, CD47 expression
442 levels were significantly elevated in the low group compared to the high group in both epithelial
443 cells (p = 6.0488e-09) and M2 macrophages (p = 8.1272e-07) (**Figures 7e**, **7f**). We also
444 observed stronger immune checkpoint interactions, such as BTLA and TNFRSF14 between
445 epithelial tumor cells and CD4+ T cells (**Figure 7d**). These findings indicate distinct immune
446 checkpoints across HER2 groups in GCA patients, particularly in pathways involving SIRPA
447 and CD47.
448
449 Next, we selected tumor-immune border regions from GCA patient samples to create tissue
450 microarrays (**Supplementary Figure 29a**) and conducted spatial transcriptomics on these
451 regions. High-quality data were obtained from 24 samples: 6 from the HER2-high group, 15
452 from the HER2-low group, and 3 from the HER2-negative group (**Supplementary Figure 29b-**
453 **f**). We identified 15 cell types from the spatial transcriptomics data (**Figure 7g**,
454 **Supplementary Figure 30**). Notably, the HER2-negative group had significantly higher levels
455 of dendritic cells (p = 0.024) and macrophages (p = 0.048) compared to the HER2-high group
456 (**Figure 7h**). Additionally, T cell levels were significantly higher in the HER2-low group
457 compared to the HER2-high group (p = 0.014). However, no significant differences were
458 observed in other cell types among the groups (**Supplementary Figure 31**). Since CD47 and
459 SIRPA primarily interact between myeloid and tumor cells, we calculated the colocalization
460 frequency of epithelial tumor cells and myeloid cells. This analysis revealed significantly
461 higher colocalization in the HER2-negative group compared to the HER2-high group (p =
462 0.043) (**Figure 7i**, **7j**, **Supplementary Figure 32a**). We further conducted a molecular
463 colocalization analysis of CD47-SIRPA at the interface of epithelial tumor and myeloid cells,
464 revealing significantly higher colocalization in the HER2-negative group compared to the
465 HER2-high group (p = 0.034) (**Figure 7i**, **7k**, **Supplementary Figure 32b**). Our findings
466 strongly suggest that blocking the CD47/SIRPA axis could be a promising therapeutic option
467 for GCA patients with low HER2 expression.
468

469 **Discussion**

470 GCA exhibits distinct features compared to non-cardia gastric adenocarcinoma, yet specific
471 therapeutic targets are lacking. To revela the gene signature of GCA patients and their
472 therapeutic targets, we employed a highly sensitive molecular counting method using mass
473 spectrometry to assess HER2 expression in 128 GCA patients. This assessment categorized
474 patients into HER2-high, HER2-low, and HER2-negative groups, yielding novel insights into
475 GCA and identifying new therapeutic targets, particularly within the HER2-negative subgroup.
476 We discovered that the ratio of HER2 expression between tumors and their corresponding
477 normal adjacent tissues (NAT) at both the RNA and protein levels serves as a valuable
478 prognostic marker for GCA patients. We also found that the RNA expression of HER2 is a
479 favorable prognostic marker for EAC, but not for gastric cancer. Although no significant
480 differences in mutation burdens were observed among the three HER2 groups, we identified
481 heightened genome instability in the HER2-high group. Notably, DNA repair pathways were
482 significantly enriched in this group, likely due to the amplification and overexpression of tumor
483 suppressor genes such as BRCA1 and BRCA2, which may contribute to the improved
484 prognosis in this subgroup. Strong proliferation features in the HER2-high group could serve
485 as potential therapeutic targets. Conversely, the HER2-low and HER2-negative groups
486 exhibited significantly stronger inflammatory features. This suggests that anti-inflammatory
487 drugs, in combination with other treatments, could be considered as part of the therapeutic
488 regimen for these patients. Since anti-inflammatory cytokines can enhance anti-tumor
489 immunity by modulating the inflammatory response [36], careful therapeutic design is essential.
490 Additionally, we found that HER2 and ARID1A mutations together influence the prognosis of
491 GCA patients. As the second most frequently mutated tumor suppressor gene [37], ARID1A
492 presents therapeutic opportunities, particularly for HER2-low and HER2-negative patients.
493

494 Single-cell RNA-Seq and spatial transcriptomics revealed distinct tumor microenvironments
495 and immune signaling pathways: the HER2-high group had stronger growth factor signaling,
496 while the HER2-low and HER2-negative groups exhibited pronounced inflammation and
497 higher M2 macrophage levels. Furthermore, our results suggest that the HER2-negative and
498 HER2-low groups possess a tumor-suppressive immune microenvironment compared to the
499 HER2-high group. Our work explores the response variability of GCA patients to specific
500 immune inhibitory checkpoint treatments [13,14]. By identifying distinct tumor
501 microenvironments and variations in immune inhibitory checkpoints across HER2 expression
502 groups, we propose the CD47/SIRPA axis as potential targets for the HER2-negative and

503 HER2-low groups. These newly identified immune inhibitory checkpoints provide promising
504 therapeutic avenues for individualized treatment in GCA patients. Overall, our study offers
505 valuable insights and resources for GCA and paves the way for precision medicine in its
506 treatment.
507
508 **Acknowledgments**
509 This work is supported by the National Key R&D Program of China (2012AA02A503,
510 2016YFC0901403 to L.D.W.), the National Natural Science Foundation of China (81872032,
511 U1804262, 30025016 to L.D.W.; 81972571,81472234 to S.G.G.), the State Key Laboratory of
512 Esophageal Cancer Prevention & Treatment Key R&D Program of Henan province
513 (202305425 to L.D.W.), Henan Medical Science and Technology Research Project
514 (LHGJ20190001 to X.K.Z. and L.D.W.), Henan Province Key R&D and Promotion Special
515 (Science and Technology) Project (232102311032 to X.S. and L.D.W.), the Wallenberg
516 Academy Fellow in Medicine from Knut and Alice Wallenberg foundation (2023.0046 and
517 2024.0166 to X. C.), the Swedish Research Council (2022-00658 to X.C.), the Swedish Cancer
518 Foundation (21 1449Pj, 22 0491 JIA, 24 3484 Pj to X.C.). The funders played no role in the
519 study design, data acquisition and interpretation, or decision to publish.
520
521 **Author contributions**
522 L.D.W. and X.C conceived and designed the study. P.X analyzed all omics data. M.X.W.,
523 X.K.Z and X.S analyzed the correlation of omics and clinic-pathological data and follow-up
524 data. X.M.L., X.K.Z and L.L.L analyzed the immunohistochemistry data. All the other authors
525 collected clinical samples and related clinical information, performed experiments, and
526 interpreted data. P.X and M.X.W wrote the materials and methods. X.C wrote the manuscript
527 by collecting information from all authors. L.D.W. made critical comments and suggestions to
528 finalize the manuscript. All authors read the manuscript.
529
530 **Declaration of interests**
531 There is no conflict of interest from all authors.
532
533 **Declaration of generative AI and AI-assisted technologies**
534 Neither AI nor AI-assisted technologies were used in the writing of this manuscript.
535
536

537 **Supplemental information**

538 Supplementary Figures 1–32, and Tables S1–S26 are in separate files.

539

540 **Materials and Methods**

541 **Patient and Public Involvement**

542 Patients or the public were not involved in the design, or conduct, or reporting, or
543 dissemination plans of our research.

544

545 **Ethics Statement**

546 The ethics permit in this study about human patients was approved by the Ethics Committee of
547 Zhengzhou University (Approval No: 2021-0019).

548

549 **Study population, tissue and blood collection**

550 All the 128 patients clinically diagnosed with GCA from January 2019 to February 2020 were
551 enrolled in this study, including 102 males with an mean age ± SD of 64.6 ± 7.5, and 26 females
552 with an mean age of 64.7 ± 6.3. All the patients were from Henan, Hebei, and Shanxi provinces,
553 northern China, the highest incidence areas for both GCA and esophageal cancer. None of these
554 patients had undergone preoperative radiation therapy or chemotherapy, and they had no
555 history of other malignant tumors. All the patients were performed radical cardiectomy of GCA.
556 The tumor tissues and normal tissue adjacent to the tumor (NAT) to the tumor from GCA
557 patients were all collected in the operating room, and all the samples were processed within
558 half an hour after removal, and each sample of tumors and NATs (taken from tissue 3cm beyond
559 the edge of the cancer tissue) was certified through pathological examination by two well-
560 trained pathologists separately. Hematoxylin and eosin (H&E) stainings of NATs and their
561 corresponding tumors were further examined to valide by two pathologists. The clinical
562 characteristics and survival data were summarized in the analysis (**Table S1**). Staging of GCA
563 was determined by two pathologists based on the eighth edition of the AJCC (American Joint
564 Committee on Cancer's Staging System). The last follow-up was at the end of March 2024.
565 Fasting blood samples were meticulously obtained from GCA patients before surgical
566 interventions during their hospitalizations. Each patient contributed two whole blood samples,
567 with distinct processing purposes. The sample was drawn into an anticoagulant tube containing
568 ethylenediaminetetraacetic acid (EDTA) for Whole Exome Sequencing (WES). This study
569 implemented an integrative approach, encompassing whole exome sequencing, RNA

570 sequencing and proteomics. The investigation utilized formalin-fixed paraffin-embedded
571 (FFPE) tumor tissue blocks for immunohistochemistry and immunofluorescence staining.
572
573 **DNA extraction, library preparation, and whole exome sequencing**
574 Whole exome sequencing (WES) was performed on 113 pairs of fresh-frozen tissues and blood
575 (n = 112) or NATs (n = 1) from GCA. Genomic DNA from tumor tissue and blood was isolated
576 using the DNeasy Blood & Tissue Kit (69504, Qiagen), following the manufacturer's
577 instructions. Subsequently, precise quantification of DNA concentration was carried out using
578 the Qubit 2.0. DNA samples with a concentration of ≥20 ng/μL. A total amount of 0.6 μg
579 genomic DNA per sample was fragmented to an average size of 180–280 bp and subjected to
580 DNA library preparation using Illumina TruSeq DNA sample preparation kit. The Agilent
581 SureSelect Human All ExonV5 Kit (5190-6209, Agilent Technologies) was used for exome
582 capture according to the manufacturer's instruction. In brief, DNA libraries were hybridized
583 with liquid phase with biotin labeled probes from the Agilent SureSelect Human All ExonV5
584 Kit, then magnetic streptavidin beads were used to capture the exons of genes. Captured DNA
585 fragments were enriched in a PCR reaction with index barcodes for sequencing. Final libraries
586 were purified using AMPure XP beads (A63880, Beckman Coulter) and quantified using the
587 Agilent high sensitivity DNA kit (5067-4626, Agilent Technologies). WES libraries were
588 sequenced on Illumina Novaseq 6000 (Illumina) with 150 bp paired end mode according to the
589 manufacturer instruction.
590
591 **RNA isolation, library preparation, and sequencing**
592 RNA sequencing involved 54 pairs of fresh-frozen tissues and corresponding normal adjacent
593 tissues (NAT) from GCA patients. Total RNA extraction from each sample employed the RNA
594 Easy Fast Cell Total RNA Extraction Kit (DP451, Tiangen), following the manufacturer's
595 protocol. RNA integrity and total amount were assessed using the Agilent 2100 Bioanalyzer
596 (G2939A, Agilent). For mRNA library construction, the NEB Next® Ultra RNA Library Prep
597 Kit for Illumina (E7530L, NEB) was used. Polyadenylated mRNA was enriched from total
598 RNA using oligo(dT) magnetic beads, followed by random fragmentation. The first cDNA
599 strand was synthesized using random oligonucleotides as primers in the M-MuLV reverse
600 transcriptase system. Subsequently, RNaseH degraded the RNA strand, and the second cDNA
601 strand was synthesized using dNTPs in the DNA polymerase I system. The double-stranded
602 cDNA underwent end-repair, A-tailing, and ligation to sequencing adapters. AMPure XP beads
603 were employed to select cDNA fragments of approximately 370~420 bp, followed by PCR

604 amplification and purification to obtain the final library. Following library construction, initial
605 quantification used the Qubit 2.0 Fluorometer, with subsequent library dilution to 1.5 ng/μl.
606 The Agilent 2100 Bioanalyzer was then utilized to confirm the insert size of the library. Upon
607 meeting the specified criteria, qRT-PCR quantified the effective library concentration (ensuring
608 it exceeded 2nM) for quality verification. The final library underwent sequencing on the
609 Illumina NovaSeq6000, generating 150 bp paired-end reads.
610

611 **Protein extraction and proteomics**

612 Qualitative and quantitative assessment of proteins were conducted on 128 pairs of fresh-frozen
613 tissues and corresponding NATs from GCA. Initial processing involved the cryogenic grinding
614 of GCA tissues into cell powder using liquid nitrogen, followed by centrifugation.
615 Subsequently, the cell powder underwent sonication on ice, employing a high-intensity
616 ultrasonic processor (Scientz), and the resultant lysate was treated with lysis buffer (8 M urea,
617 1% Protease Inhibitor Cocktail). Centrifugation was employed to eliminate residual debris, and
618 the supernatant, containing proteins, was collected. Protein concentration was determined
619 using a BCA kit following the manufacturer's protocol. The subsequent steps encompassed
620 trypsin digestion, TMT/iTraq labeling, high-performance liquid chromatography (HPLC)
621 fractionation utilizing a Thermo Betasil C18 column (5 μm particles, 10 mm ID, 250 mm
622 length), and liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis. The
623 acquired data were processed using the MaxQuant search engine (v.1.5.2.8) against the
624 Homo_sapiens_9606 database.
625

626 **Immunohistochemistry (IHC) of proteins**

627 IHC of HER2, SIRPA, and PDCD1LG2 proteins was according to the same method as our
628 previous study[23]. In brief, 5 μm thick formalin fixed paraffin-embedded GCA tissue sections
629 were first deparaffined with xylene 15 min for three times, then were dehydrated through 100%
630 alcohol, 85% alcohol and 75% alcohol for 5 min each, followed by distilled water rinsing for
631 5 min. The epitope retrieval is performed in the microware by putting the tissue into citrate
632 buffer (pH 6.0). After the epitope retrieval, the tissue section is rinsed in Phosphate-Buffered
633 Saline buffer (PBS, pH 7.4). After blocked with 3% bovine serum albumin (BSA) 30 min at
634 room temperature, the tissues were incubated with primary antibodies, Rabbit anti-ERBB2
635 antibody (1:100 dilution, K00125, Roche), Rabbit anti-SIRPA (1:100, ab53721, Abcam), or
636 Rabbit anti-PDCD1LG2 (1:100, ab200377, Abcam), overnight at 4 °C. In the next day, the
637 washing is performed with PBS buffer for three times, 15 min each. The secondary antibody

638 (1:1000 dilution, Horseradish Peroxidase, HRP marked, PV-9000, ZSGB-BIO) was incubated
639 for 50 min at room temperature. After the secondary antibody incubation, the washing is
640 performed with PBS buffer three times on shaker, 15 min each. The tissue is stained with the
641 Harris Hematoxylin for 3 min. At last, the tissue section was mounted and imaged. The
642 identification of HER2, SIRPA and PDCD1LG2 positive cells was validated by two senior
643 pathologists.
644
645 **HER2 Immunohistochemical Scoring Criteria**
646 The HER2 scoring criteria were established and validated by two senior pathologists. Tumor
647 specimens were evaluated for membranous HER2 staining intensity using the following criteria:
648 HER2 0+: No reaction or less than 10% of tumor cells exhibited membranous HER2 staining.
649 HER2 1+: Faint or faintly visible membranous HER2 staining in at least 10% of tumor cells,
650 with only partial cell membrane HER2 staining observed.
651 HER2 2+: Weak to moderate basolateral, lateral, or complete membranous HER2 staining
652 observed in at least 10% of tumor cells. HER2 3+: Basolateral, lateral, or complete
653 membranous HER2 staining present in at least 10% of tumor cells.
654 The HER2 scoring criteria were meticulously applied to assess membranous HER2 staining
655 intensity in tumor cells, ensuring consistency through validation by two senior pathologists.
656
657 **Immunofluorescence staining of CD163**
658 The immunofluorescence assay was conducted on paraffin-embedded (FFPE) sections
659 obtained from GCA specimens. Initially, FFPE slices of GCA paraffin-embedded sections
660 underwent baking in an oven for paraffin fixation. Immediate dewaxing of the paraffin slices
661 followed, involving sequential treatments with an eco-friendly dewaxing solution, anhydrous
662 ethanol, 90% alcohol, 70% alcohol, and distilled water. Dewaxed slices were then subjected to
663 low-temperature fixation. The fixing solution was meticulously prepared in a dedicated fixing
664 box and heated in a microwave until boiling, ensuring the absence of air bubbles. Following
665 fixation, rapid cooling of the slices was achieved using ice cubes. Subsequent steps included
666 the removal of the fixing solution with TBST, application of an appropriate sealing solution,
667 removal of excess sealing solution, and sequential addition of primary and secondary
668 antibodies targeting CD163 (ab182422, Abcam, UK). Incubation at room temperature allowed
669 for antibody binding. Afterwards, residual primary and secondary antibody solutions were
670 washed off using TBST. The cleaned slides were treated with corresponding fluorescent dyes,
671 followed by incubation at room temperature. Subsequent TBST washes removed any excess

672 dye. DAPI working solution was applied for nuclear staining, with thorough cleaning to
673 eliminate residual DAPI, dewaxing solution, xylene, or debris. The slides were then sealed for
674 further analysis.
675
676 The image acquisition was performed using the FUISON model full-spectrum scanner from
677 the American company Akoya, capturing all image information of the entire tissue on the slides.
678 Image analysis was conducted using VISIOPHARM software, initially focusing on tissue
679 contour recognition. Experienced pathology experts further delineated tumor tissue regions by
680 overlaying the mIHC-stained images onto the corresponding HE images. During this process,
681 the software calculated the areas of the entire tissue and tumor regions. The VISIOPHARM
682 software employed its inherent deep learning algorithms for cell segmentation, facilitating the
683 determination of total cell counts within the tissue. The Phenplex Guided Workflow
684 functionality in VISIOPHARM software was utilized for cell phenotype identification,
685 enabling the quantification of individual positive cells for each channel and the determination
686 of co-positive cells between channels. Subsequently, statistical calculations leveraging the
687 obtained data on the entire tissue and tumor area, total cell count, positive cell counts for each
688 channel, and co-positive cell counts between channels were conducted. These calculations
689 included the determination of positivity rates, area density, as well as mean, standard deviation,
690 and inter-batch P-values among samples.
691
692 **TCGA data processing**
693 Transcriptome profiling data from The Cancer Genome Atlas Program (TCGA) were
694 downloaded from the GDC Data Portal with the project IDs ESCA (Esophageal Carcinoma)
695 and STAD (Stomach Adenocarcinoma). The expression data, represented by TPM (Transcripts
696 Per Million) normalized values, were extracted from each sample. Subsequently, all samples
697 were merged into a unified TPM matrix.
698
699 **Patient Stratification Based on ERBB2/HER2 Expression**
700 **ERBB2 RNA Expression Level**: For TCGA patients (ESCC, EAC, STAD) and GCA patients
701 (n = 54), individuals were stratified based on ERBB2 RNA levels (TPM) using three strategies:
702 upper vs. lower median, top vs. bottom quantile, and natural separation (very high vs. rest) to
703 distinguish between ERBB2 high- and low-expression groups.

704 **Ratio of ERBB2 RNA Expression Between Tumor and normal tissue adjacent to the**
705 **tumor (NAT)**: For GCA patients (n = 54), individuals were divided into two groups—high and
706 low—based on the median ratio of ERBB2 RNA expression levels between the tumor NAT.
707 **Ratio of HER2 Protein Expression Between Tumor and normal tissue adjacent to the**
708 **tumor (NAT)**: For GCA patients (n = 128), individuals with NA values for HER2 expression
709 were assigned to the HER2-negative group. The remaining patients were divided into two
710 groups—high and low— based on the median value of the ratio between tumor and NAT. All
711 omic data analyses were based on this stratification.
712
713 **Differential gene analysis in EAC cell line OE19**
714 RNA-Seq data for the EAC cell line OE19, both with and without treatment with the HER2
715 inhibitor lapatinib, as well as following HER2 siRNA silencing, were obtained from
716 ArrayExpress (E-MTAB-10304)[35]. The gene expression levels were further normalized using
717 transcripts per million (TPM). Differential gene analysis was performed with DESeq2[38]. The
718 significantly changes genes are with a cutoff of |log2(FoldChange)| >= 1 and FDR <= 0.
719
720 **Proteomics data processing and statistical analysis**
721 The raw LC-MS datasets were initially searched against a database and converted into matrices
722 containing the reporter intensity of peptides across samples. Subsequently, the relative
723 quantitative value for each modified peptide was determined through the following procedures:
724 First, the intensities of modified peptides (I) were centralized and transformed into relative
725 quantitative values (U) for each modified peptide in each sample: Uij = Iij / Mean(Ij), where i
726 represents the sample, and j denotes the modified peptide. Second, the relative quantitative
727 values (Q) of modified peptides were corrected by the median value: Qij = Uij / Median(Ui).
728 Proteins with 50% NA values across all samples were excluded. The protein expression ratio
729 (R) between tumor samples (T) and normal adjacent samples (NAT) was calculated with Rmn
730 = Tmn/NATmn, where m denotes each protein, and n denotes each sample. Differential analysis
731 among the three groups with the "One v.s. Rest" strategy was conducted on the Rmn matrix
732 using the wilcox.test function in R. Additionally, differential analysis between tumors and
733 NATs was performed based on tumor and NATs using the wilcox.test function in R. Proteins
734 were considered differentially expressed if their |FoldChange| was greater than or equal to 1.5,
735 and the p-value was less than or equal to 0.05.
736
737 **Whole exome sequencing data processing**

738 The pair-end sequencing data generated from Illumina NovaSeq 6000 were trimmed using fastp
739 v0.23. Subsequently, the trimmed data were mapped to the hg38 reference genome using bwa
740 mem v0.7.17[39]. The SAM files were converted into BAM format and sorted by samtools
741 v1.16.1. Duplicate reads were marked using "GATK4.3 MarkDuplicates." Base quality scores
742 were recalibrated using "GATK4.3 BaseRecalibrator" and "GATK4.3 ApplyBQSR,"
743 employing known variants from the dbSNP database (build 146) [8] and the 1000 Genomes
744 project[40].
745
746 **Somatic short variant calling**
747 The somatic short variants, including SNVs and InDels, were called using GATK Mutect2[41],
748 leveraging matched normal samples as reference background and utilizing the "af-only-
749 gnomad" file from the Gnomad datasets as germline resources. Subsequently, "GATK
750 GetPileupSummaries" and "GATK CalculateContamination" were employed to compute cross-
751 sample contamination. Finally, "GATK FilterMutectCalls" was used to filter short variants with
752 the contamination table, retaining mutations with a "PASS" filter flag for subsequent analysis.
753 Copy number alterations (CNAs) were identified using both CNVkit/0.9.8[42] and FREEC-11.6[43],
754 utilizing tumor-matched normal samples as the reference background. Copy numbers equal to
755 or less than 1 were categorized as CNA loss, while copy numbers equal to or greater than 3
756 were considered as CNA gain. CNAs identified by both software tools were retained for
757 subsequent analysis.
758
759 **SNVs and InDels analysis**
760 All SNVs and InDels identified by GATK were annotated using snpEff v5.1. Subsequently, the
761 annotated SNVs and InDels were transformed into Mutation Annotation Format (MAF) using
762 a custom script which is deposited in https://github.com/pengweixing/GCA_multi-omics,
763 following the guidelines provided in
764 https://docs.gdc.cancer.gov/Encyclopedia/pages/Mutation_Annotation_Format_TCGAv2/.
765 For each sample, variant types including "Missense_Mutation", "Nonsense_Mutation",
766 "Splice_Site", "Frame_Shift_Del", "Frame_Shift_Ins", "In_Frame_Del", and "In_Frame_Ins"
767 were selectively extracted. The identified SNVs and InDels from all samples were subsequently
768 aggregated. The distribution of mutated genes within three groups was visualized using the R
769 package "ComplexHeatmap" through a waterfall plot. Only mutated genes that align with
770 entries in the Cancer Gene Database (COSMIC) were retained for further analysis. The tumor
771 mutation burden (TMB) was quantified as the number of SNVs per million nucleotides in the

coding region. Synonymous alterations were also counted separately for the purpose of reducing sampling noise.

CNA analysis

Copy number alterations (CNAs) across the entire genome for each group were depicted using the R package Circlize. Additionally, CNAs in specific regions for each group were visualized utilizing the R package karyoploteR. The genome was divided into bins with a 1Mb window, and the CNAs were counted in each bin for three groups, with a subsequent application of the Chi-square test. CIN was estimated through genome integrity index (GII), the proportion of altered genome segments[44]. Weighted Genome Instability Index (wGII) was also caculated following a previous report[28].

The formula is as follows (where 'n' represents all autosomes):

$$wGII = \frac{\sum_{i=1}^{n} \frac{Total\ length\ of\ Gained\ and\ loss\ region}{Length\ of\ chromosome(i)}}{n}$$

The number of breakpoints representing the Chromosomal Instability Index (CINindex) was calcuated with following formula:

$$CINindex = Total\ number\ of\ breakpoints$$

wCINindex (Weighted Chromosomal Instability Index) was calcuated by considering the number of breakpoints per million base pairs in each chromosome.

The formula is as follows (where 'n' represents all autosomes):

$$wCINindex = \frac{\sum_{i=1}^{n} \frac{The\ number\ of\ breakspoints}{Length\ of\ chromosome(i)} * 1Million}{n}$$

RNA-seq data processing and differential analysis

The paired-end sequencing data were processed through the nf-core/rnaseq 3.10.1 pipeline. Briefly, fastq files were subjected to trimming using Trim Galore v0.6.10, followed by

803 alignment to the hg38 genome using STAR v2.7.0. Sorting and indexing of the resulting BAM
804 file were performed using samtools v1.16.1 as. Subsequently, Salmon was employed to
805 quantify the expression levels of individual genes. The gene expression levels were further
806 normalized using transcripts per million (TPM).
807 The categorization of groups in RNA-seq data aligned with the HER2-High, HER2-Low, and
808 HER2-Negative groups identified in proteomics. Differential analysis was conducted within
809 tumor samples with DESeq2[38], comparing HER2-High versus Rest, HER2-Low versus Rest,
810 and HER2-Negative versus Rest. A cutoff of |FoldChange| greater than or equal to 1.5 and a
811 false discovery rate (FDR) less than or equal to 0.05 was applied to identify differentially
812 expressed genes. Additionally, differential analysis was performed between HER2-High versus
813 NATs, HER2-Low versus NATs, and HER2-Negative versus NATs, with a cutoff of
814 |log2(FoldChange)| >= 1 and FDR <= 0.05 to filter differentially expressed genes.
815
816 **Cellular composition analysis**
817 CIBERSORTx[33] analysis was conducted in accordance with previously established procedures.
818 The docker-based tool of CIBERSORTx was utilized to estimate the relative fractions of 22
819 immune cell sub-populations, employing the LM22 reference gene expression signatures.
820 Comparisons were conducted among three groups for each cell type. Statistical differences
821 were calculated with "t-test" function in R.
822
823 **Gene Ontology (GO) enrichment and Gene set enrichment analysis (GSEA)**
824 Differentially expressed genes and proteins identified in transcriptomics and proteomics
825 datasets were chosen for Gene Ontology (GO) enrichment analysis. The enrichGO function
826 within the R package clusterProfiler v4.0 was employed for conducting gene ontology analysis,
827 with FDR of 0.05 applied to filter significant GO terms. For Gene Set Enrichment Analysis
828 (GSEA), 50 hallmark gene sets from the Human MSigDB collections were adopted. The genes
829 for input were ranked based on their weights which calculated with "$Gene_{weight}$ =
830 $sign(foldchange)$ * -log10(FDR)". "GSEA" function in clusterProfiler package was adopted
831 for GSEA analysis. A significance threshold of FDR = 0.05 was applied to filter GSEA
832 hallmarks.
833
834 ***ERBB2* grouping patients based on scRNA-seq data**
835 Raw reads (R2) from the scRNA-seq data were treated as bulk RNA-seq, as the 10x barcodes
836 and UMIs were located in R1. We used R2 as input for the bulk RNA-seq pipeline, employing

837 nf-core rnaseq (version 3.14.0)[45]. Gene expression was normalized using TPM, and samples
838 were divided into *ERBB2*-high and *ERBB2*-low groups based on expression levels.
839
840 **Tissue dissociation and preparation of single-cell RNA-Seq**
841 The tumor samples obtained were washed 2-3 times in 4°C precooled 1× PBS on ice,
842 transferred to a sterile RNase-free culture dish, and cut into small pieces (about 0.5 mm^2) using
843 10 cm surgical scissors. The tissues were then washed with 1× PBS to remove as much non-
844 target tissue as possible, such as blood stains, fatty layers, and connective tissue. The tissue
845 samples were incubated in a constant-temperature water bath or shaker at 37°C with either a
846 MACS Human Tumor Dissociation Kit (130-095-929, Miltenyi Biotec). Incubation was
847 generally terminated once the digestion solution became turbid and the tissue mass had
848 dissolved. Next, a 40 μm cell sieve was used to filter the suspension, or a 100 μm/70 μm cell
849 filter if the tissue mass was larger. The cell suspension was typically centrifuged at 300 rpm at
850 4°C for 5 minutes. The supernatant was discarded, and 1 mL of pre-cooled 1× PBS was added
851 to re-suspend the cells. Then, 3 mL of pre-cooled erythrocyte lysis buffer was added to the tube,
852 mixed evenly, and incubated at 4°C for 5-10 minutes. After incubation, the suspension was
853 centrifuged at 300 rpm at 4°C for 5 minutes, the supernatant was removed, and 1 mL of
854 precooled 1× PBS was added to fully re-suspend the cells. The cell concentration was measured
855 using a cell counter, and adjusted as needed to reach a target concentration of 700-1200 cells/μL,
856 with cell viability above 90% and cell clumping below 15%. Once the final cell concentration
857 and viability were achieved, the cells were placed on ice, and a 10X Genomics single-cell
858 transcriptome chip experiment was initiated within 30 minutes.
859
860 **Single-cell RNA-seq library preparation and sequencing**
861 We prepared single cell RNA-seq libraries with Chromium Next GEM Single Cell 3′ Reagent
862 Kits v3.1 on the Chromium Controller(10× Genomics).We prepared single cell suspensions
863 from cultured cell lines and single cells were suspended in PBS containing 0.04% BSA.The
864 cell suspension was loaded onto the Chromium Next GEM Chip G and ran the Chromium
865 Controller to generate single-cell gel beads in the emulsion(GEMs) according to the
866 manufacturer's recommendation.captured cells were lysed and the released RNA was barcoded
867 through reverse transcription in individual GEMs. Barcoded, full-length cDNA was generated
868 and libraries were constructed according to the performer's protocol .The quality of libraries
869 was assessed by Qubit 4.0 and the Agilent 2100 . Sequencing was performed on the Illumina

870 NovaSeq 6000 with a sequencing depth of at least 50,000 reads per cell and 150 bp (PE150)
871 paired-end reads.
872
873 **Quality control, data normalization and intergration of scRNA-seq data**
874 Raw reads from the 10x Genomics single-cell RNA-seq platform were demultiplexed and
875 aligned to the human reference genome (GRCh38) using CellRanger (version 7.2.0). The
876 filtered cell-to-gene matrix was processed in Seurat v5.0.3[46]. Cells were included if they
877 expressed between 500 to 7,000 genes, had UMIs between 100 and 50,000, and less than 20%
878 mitochondrial reads. Data normalization was performed with Seurat's 'NormalizeData'
879 function, followed by scaling using 'ScaleData' function. Dimensionality reduction was
880 conducted via 'RunPCA', and samples were integrated to remove biological variations using
881 the 'IntegrateLayers function.
882
883 **Cell type identification from single cell RNA-Seq**
884 We conducted cell clustering using Seurat's 'FindNeighbors' and 'FindClusters' functions. To
885 visualize the clusters, we applied the 'RunUMAP' function with "dims = 1:30". For cell type
886 annotation, we utilized esophageal adenocarcinoma (EAC) scRNA-seq data from the
887 LUD2015-005 project (https://zenodo.org/records/8083316) as a reference[47]. This reference
888 dataset was used to identify key cell compartments—Epithelium, Lymphocytes, Phagocytes,
889 and Stroma—by applying the 'FindTransferAnchors' and 'TransferData' functions with "dims
890 = 1:30" and normalization.method = 'SCT'. Additionally, plasma cells, B cells, T cells,
891 endothelial cells, fibroblasts, mast cells, pericytes, muscle cells, and myeloid cells were
892 annotated using the EAC reference. We extracted T cells and myeloid cells for further
893 dimensionality reduction and clustering. Cell-type specific marker genes were identified using
894 Seurat's 'FindAllMarkers' function. By combining marker genes and leveraging GPT-4's
895 advanced annotation capabilities[48], we refined the analysis of T cells and myeloid cells,
896 identifying subtypes such as macrophages, monocytes, neutrophils, dendritic cells, CD8+ T
897 cells, CD4+ T cells, regulatory T cells, and NK cells.
898
899 **Differential gene analysis in single cell RNA-seq**
900 For specific cell types, such as CD4+ T cells, CD8+ T cells, myeloid cells, M2 macrophages,
901 and epithelial tumor cells, we downsampled the cell numbers to balance the *ERBB2*-High and
902 *ERBB2*-Low groups. Differential gene expression analysis was performed using Seurat's[46]
903 'FindMarkers' function. Genes with an adjusted p-value < 0.05 and an absolute log2(fold

904 change) > 0.25 were considered significant. Gene ontology and KEGG pathway analyses were
905 conducted using the 'enrichGO' and 'enrichKEGG' functions from the clusterProfiler v4.7.1
906 package. GSEA analysis was performed using ranked genes based on fold changes with the
907 'GSEA' function of clusterProfiler v4.7.1.
908
909 **Pseudotime analysis**
910 Monocle2 was employed to infer epithelial cell developmental trajectories[49]. Seurat objects
911 were converted to CellDataSet for input into Monocle2, and differential genes between normal
912 and tumor epithelial cells were identified using 'FindAllMarkers' (p-adjusted < 0.05, log2(fold
913 change) > 1). Dimensionality reduction was performed with 'DDRTree' function, and cells
914 were ordered with the 'orderCells' function. Trajectories were visualized using
915 'plot_cell_trajectory', and cell fate was analyzed using the branched expression analysis
916 modeling (BEAM) function with branch_point=1. Heatmaps of cell fates were generated with
917 'plot_genes_branched_heatmap' with parameters of 'branch_point = 1, num_clusters = 4'. The
918 genes associated with each cell fate corresponded to the clusters shown in the heatmap. These
919 genes were extracted for Gene Ontology (GO) and KEGG pathway enrichment analysis using
920 the 'enrichGO' and 'enrichKEGG' functions from the clusterProfiler package v4.7.16.
921
922 **CNV estimation in epithelial cells**
923 We used the InferCNV package (version 1.14.2)(https://github.com/broadinstitute/inferCNV)
924 with default parameters to infer copy number variations (CNVs) in epithelial cells, using
925 lymphocytes and phagocytes as reference. CNVs were defined by copy number values greater
926 than 1.05 or less than 0.95. A CNV score for each cell was calculated based on the proportion
927 of genes affected by CNVs. To classify epithelial tumor cells, we applied the 'optbin' function
928 in Optbin v1.3 package to determine an optimal threshold that minimized the squared error,
929 dividing the cells into CNV-high and CNV-low groups. Cells with CNVs above the threshold
930 were classified as CNV-high, and those below as CNV-low. For each sample, we calculated the
931 proportions of CNV-high and CNV-low cells. A two-sided Wilcoxon test was performed to
932 compare the proportions of CNV-high cells between *ERBB2*-High and *ERBB2*-Low groups.
933
934 **Cell-cell communication analysis**
935 Cell-cell communication analysis was conducted using CellChat v2.1.2[50] with the 'Secreted
936 Signaling' database. For both the *ERBB2*-High and *ERBB2*-Low groups, we first identified
937 overexpressed genes using the 'identifyOverExpressedGenes' function, followed by

938 identification of interactions using the 'identifyOverExpressedInteractions' function. We then
939 computed communication probabilities and constructed the cellular communication network
940 with the 'computeCommunProb' function. Low-cell-number communications were filtered out
941 using the 'filterCommunication' function with a minimum cell threshold of 10. The
942 'computeCommunProbPathway' function was employed to calculate communication at the
943 signaling pathway level. To compare the two groups, we merged the *ERBB2*-High and *ERBB2*-
944 Low CellChat objects using the 'mergeCellChat' function. Differential interaction strengths
945 between the groups were visualized using circle plots with 'netVisual_diffInteraction' function
946 and heatmaps with 'netVisual_heatmap' function. We compared overall information flow
947 between cell types using a bar plot generated by the 'rankNet' function. Dysfunctional signaling
948 pathways were identified via differential expression analysis using the
949 'identifyOverExpressedGenes' function. Communication probabilities for ligand-receptor
950 pairs between cell types were visualized using bubble plots with 'netVisual_bubble' function,
951 and upregulated or downregulated ligand-receptor interactions were illustrated with a chord
952 diagram with 'netVisual_chord_gene'. Additionally, interactions of immune checkpoint
953 ligands and receptors across cell types were analyzed using CellphoneDB v5.0.1[51] with default
954 parameters. P-values and means for immune checkpoint gene pairs were extracted and
955 presented in a bubble plot.
956
957 **Spatial Transcriptomics**
958 Formalin-fixed, paraffin-embedded (FFPE) samples from GCA tumors were cut to perform
959 H&E staining (hematoxylin and eosin staining). The H&E staining of tumor sections was
960 reviewed by two senior pathologists, and tumor-immune border regions were punched to create
961 a tissue microarray for Spatial Transcriptomics. Nine tissue microarrays were included in each
962 Spatial Transcriptomics analysis. The tissue microarrays were prepared according to the
963 Visium CytAssist Spatial Gene Expression for FFPE-Tissue Preparation Guide (CG000518,
964 10× Genomics, Pleasanton, CA, USA). Tissue sections with a thickness of 5 μM were prepared
965 for subsequent experiments, including deparaffinization, H&E staining, and decrosslinking.
966 Library construction was performed using the Visium CytAssist Spatial Gene Expression
967 Reagent Kits (PN-1000520, 10x Genomics), following the manufacturer's protocol. Libraries
968 were sequenced using the Illumina NovaSeq 6000 platform with the PE150 sequencing mode.
969
970 **Quality control and normalization of 10x spatial transcriptomics**

971 Raw reads from the 10x Genomics Visium platform were demultiplexed and aligned to the
972 human reference genome (GRCh38) using SpaceRanger v3.0.0 software. The filtered cell-to-
973 gene matrix generated by SpaceRanger was then imported into Seurat for further analysis.
974 High-quality spots were defined as those with nCount_Spatial > 2000, nFeature_Spatial > 500,
975 and percent.mt < 20. Each 10x slide contained nine tissue sections arranged in a 3x3 grid. For
976 each section, we calculated the proportion of high-quality spots and retained sections with over
977 90% high-quality spots. One lymph node metastasis section was removed, along with another
978 section primarily composed of muscle cells and a section that was only partially covered with
979 tissue. The remaining tissue sections were normalized using Seurat's SCTransform function for
980 downstream analysis.
981
982 **Cell type decomposition of 10x spatial transcriptomics**
983 We performed cell type decomposition for each tissue section in the 10x spatial slide by
984 integrating fine-annotated scRNA-seq data with spatial transcriptomics data using the
985 cell2location v0.1.3 software[52]. Both datasets were normalized using Seurat's SCTransform
986 function, with the scRNA-seq data serving as the reference. Cell type signatures were estimated
987 from the reference data using a Negative Binomial (NB) regression model with a maximum of
988 250 epochs. Spatial mapping was then conducted to estimate cell types in each spot based on
989 the reference signatures, using a maximum of 10,000 epochs. The resulting cell abundance
990 estimates were loaded into Seurat for visualization. Cell type composition in each tissue section
991 was determined by normalizing the abundance of each cell type relative to the total cell
992 abundance. Comparisons of cell type abundances between the *HER2*-High, *HER2*-Low, and
993 *HER2*-Neg groups were conducted using a two-sided Wilcoxon test.
994
995 **Colocalization analysis**
996 To investigate ligand-receptor gene interactions in the 10x spatial transcriptomics data, we first
997 binarized the expression of each gene and the presence of each cell type. Macrophages,
998 dendritic cells, and monocytes were grouped together as myeloid cells. A cell type was
999 considered present in a spot if its cell2location-estimated abundance exceeded the median for
1000 that section. Similarly, a gene was considered expressed in a spot if its cell2location-estimated
1001 abundance was greater than one for that section. We counted the spots where both epithelial
1002 tumor cells and myeloid cells were either co-localized or not. Subsequently, in spots where
1003 these two cell types were co-localized, we identified spots where both ligand and receptor genes

1004 were co-expressed or not. Comparisons of co-localized spots among the HER2-High, HER2-
1005 Low, and HER2-Neg groups were made using a two-sided Wilcoxon test.
1006
1007 **Prognosis analysis**
1008 Survival analysis was conducted using the R package survival v.3.5-7 and survminer v.0.4.9.
1009 The Kaplan-Meier method was employed to estimate overall survival time, and statistical
1010 differences were assessed with log-rank test.
1011
1012 **Data processing and visualization**
1013 In addition to the previously mentioned R packages, the data processing and visualization also
1014 incorporated other tools, including dplyr v1.1.4, stringr v1.5.1, pheatmap v1.0.12, ggplot2
1015 v3.4.4, ggpubr v0.6.0, and ggvenn v0.1.10.
1016
1017 **Code availability**
1018 The code to perform all analyses and regenerate all the figures in this study is provided at
1019 https://github.com/pengweixing/GCA_multi-omics.
1020
1021 **Data availability**
1022 All raw data from this study is deposited in the China National Center for Bioinformation with
1023 access number of HRA006640.
1024 TCGA data that used in this study were downloaded from
1025 https://portal.gdc.cancer.gov/projects/TCGA-ESCA and
1026 https://portal.gdc.cancer.gov/projects/TCGA-STAD.
1027 RNA-Seq data for the EAC cell line OE19, both with and without treatment with the HER2
1028 inhibitor lapatinib, as well as following HER2 siRNA silencing, were obtained from
1029 ArrayExpress (E-MTAB-10304).
1030
1031 **References**

1032 1 Wang, L. D. *et al.* Genome-wide association study of esophageal squamous cell carcinoma in Chinese
1033 subjects identifies susceptibility loci at PLCE1 and C20orf54. *Nat Genet* **42**, 759-U746,
1034 doi:10.1038/ng.648 (2010).
1035 2 Wang, Z. *et al.* NNMT enriches for AQP5(+) cancer stem cells to drive malignant progression in early
1036 gastric cardia adenocarcinoma. *Gut* **73**, 63-77, doi:10.1136/gutjnl-2022-328408 (2023).
1037 3 Li, S. *et al.* Integrative proteomic characterization of adenocarcinoma of esophagogastric junction. *Nat*
1038 *Commun* **14**, 778, doi:10.1038/s41467-023-36462-8 (2023).
1039 4 Nakauchi, M. *et al.* Distinct Differences in Gastroesophageal Junction and Gastric Adenocarcinoma in
1040 2194 Patients: In Memory of Rebecca A. Carr, February 24, 1988-January 19, 2021. *Ann Surg* **277**,
1041 629-636, doi:10.1097/SLA.0000000000005320 (2023).

5 Rusch, V. W. Are cancers of the esophagus, gastroesophageal junction, and cardia one disease, two, or several? *Semin Oncol* **31**, 444-449, doi:10.1053/j.seminoncol.2004.04.023 (2004).

6 Moasser, M. M. The oncogene HER2: its signaling and transforming functions and its role in human cancer pathogenesis. *Oncogene* **26**, 6469-6487, doi:10.1038/sj.onc.1210477 (2007).

7 Goutsouliak, K. *et al.* Towards personalized treatment for early stage HER2-positive breast cancer. *Nat Rev Clin Oncol* **17**, 233-250, doi:10.1038/s41571-019-0299-9 (2020).

8 Tarantino, P. *et al.* Comprehensive genomic characterization of HER2-low and HER2-0 breast cancer. *Nat Commun* **14**, 7496, doi:10.1038/s41467-023-43324-w (2023).

9 Rha, S. Y. *et al.* The phase III, randomized, double-blind, placebo-controlled KEYNOTE-811 study of pembrolizumab plus trastuzumab and chemotherapy for HER2+metastatic gastric or gastroesophageal junction (mG/GEJ) adenocarcinoma. *Ann Oncol* **34**, S1521-S1522, doi:10.1016/j.annonc.2023.10.266 (2023).

10 Gupta, A. *et al.* EGFR-directed antibodies promote HER2 ADC internalization and efficacy. *Cell Rep Med* **5**, 101792, doi:10.1016/j.xcrm.2024.101792 (2024).

11 Slamon, D. J. *et al.* Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *New Engl J Med* **344**, 783-792, doi:Doi 10.1056/Nejm200103153441101 (2001).

12 Bang, Y. J. *et al.* Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. *Lancet* **376**, 687-697, doi:10.1016/S0140-6736(10)61121-X (2010).

13 Janjigian, Y. Y. *et al.* The KEYNOTE-811 trial of dual PD-1 and HER2 blockade in HER2-positive gastric cancer. *Nature* **600**, 727-730, doi:10.1038/s41586-021-04161-3 (2021).

14 Janjigian, Y. Y. *et al.* Pembrolizumab plus trastuzumab and chemotherapy for HER2-positive gastric or gastro-oesophageal junction adenocarcinoma: interim analyses from the phase 3 KEYNOTE-811 randomised placebo-controlled trial. *Lancet* **402**, 2197-2208, doi:10.1016/S0140-6736(23)02033-0 (2023).

15 Liu, D. *et al.* Efficacy and safety of KN026, a bispecific anti-HER2 antibody, in combination with KN046, an anti-CTLA4/PD-L1 antibody, in patients with advanced HER2-positive nonbreast cancer: a combined analysis of a phase Ib and a phase II study. *Signal Transduct Target Ther* **10**, 104, doi:10.1038/s41392-025-02195-x (2025).

16 Lv, H. F. *et al.* Expression of Human Epidermal Growth Factor Receptor-2 Status and Programmed Cell Death Protein-1 Ligand Is Associated With Prognosis in Gastric Cancer (vol 10, 580045, 2021). *Front Oncol* **11**, doi:ARTN 672599 10.3389/fonc.2021.672599 (2021).

17 Zhang, H. *et al.* Intratumoral and intertumoral heterogeneity of HER2 immunohistochemical expression in gastric cancer. *Pathol Res Pract* **216**, doi:ARTN 153229 10.1016/j.prp.2020.153229 (2020).

18 Van Cutsem, E. *et al.* HER2 screening data from ToGA: targeting HER2 in gastric and gastroesophageal junction cancer. *Gastric Cancer* **18**, 476-484, doi:10.1007/s10120-014-0402-y (2015).

19 Kim, W. H. *et al.* HER2 Status in Gastric and Gastroesophageal Junction Cancer: Results of the Large, Multinational HER-EAGLE Study. *Appl Immunohistochem Mol Morphol* **26**, 239-245, doi:10.1097/PAI.0000000000000423 (2018).

20 Slamon, D. J. *et al.* Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science* **235**, 177-182, doi:10.1126/science.3798106 (1987).

21 Lob, S. *et al.* Prognostic effect of HER2 evolution from primary breast cancer to breast cancer metastases. *J Cancer Res Clin Oncol* **149**, 5417-5428, doi:10.1007/s00432-022-04486-0 (2023).

22 Plum, P. S. *et al.* HER2/neu (ERBB2) expression and gene amplification correlates with better survival in esophageal adenocarcinoma. *BMC Cancer* **19**, 38, doi:10.1186/s12885-018-5242-4 (2019).

23 Zhao, X. K. *et al.* Focal amplifications are associated with chromothripsis events and diverse prognoses in gastric cardia adenocarcinoma. *Nat Commun* **12**, 6489, doi:10.1038/s41467-021-26745-3 (2021).

24 Boku, N. HER2-positive gastric cancer. *Gastric Cancer* **17**, 1-12, doi:10.1007/s10120-013-0252-z (2014).

25 Diamond, M. S. & Farzan, M. The broad-spectrum antiviral functions of IFIT and IFITM proteins. *Nat Rev Immunol* **13**, 46-57, doi:10.1038/nri3344 (2013).

26 Hiam-Galvez, K. J., Allen, B. M. & Spitzer, M. H. Systemic immunity in cancer. *Nat Rev Cancer* **21**, 345-359, doi:10.1038/s41568-021-00347-z (2021).

27 Birkbak, N. J. *et al.* Paradoxical relationship between chromosomal instability and survival outcome in cancer. *Cancer Res* **71**, 3447-3452, doi:10.1158/0008-5472.CAN-10-3667 (2011).

28 Endesfelder, D. *et al.* Chromosomal instability selects gene copy-number variants encoding core regulators of proliferation in ER+ breast cancer. *Cancer Res* **74**, 4853-4863, doi:10.1158/0008-5472.CAN-13-2664 (2014).

29 Mah, L. J., El-Osta, A. & Karagiannis, T. C. gammaH2AX: a sensitive molecular marker of DNA damage and repair. *Leukemia* **24**, 679-686, doi:10.1038/leu.2010.6 (2010).

30 Rose-John, S., Jenkins, B. J., Garbers, C., Moll, J. M. & Scheller, J. Targeting IL-6 trans-signalling: past, present and future prospects. *Nat Rev Immunol* **23**, 666-681, doi:10.1038/s41577-023-00856-y (2023).

31 Katze, M. G., He, Y. & Gale, M., Jr. Viruses and interferon: a fight for supremacy. *Nat Rev Immunol* **2**, 675-687, doi:10.1038/nri888 (2002).

32 Braun, D. A. *et al.* Interplay of somatic alterations and immune infiltration modulates response to PD-1 blockade in advanced clear cell renal cell carcinoma. *Nat Med* **26**, 909-918, doi:10.1038/s41591-020-0839-y (2020).

33 Newman, A. M. *et al.* Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat Biotechnol* **37**, 773-782, doi:10.1038/s41587-019-0114-2 (2019).

34 Hao, Y., Zhou, X. X., Li, Y. L., Li, B. L. & Cheng, L. The CD47-SIRPα axis is a promising target for cancer immunotherapies. *Int Immunopharmacol* **120**, doi:ARTN 110255 10.1016/j.intimp.2023.110255 (2023).

35 Ogden, S., Carys, K., Ahmed, I., Bruce, J. & Sharrocks, A. D. Regulatory chromatin rewiring promotes metabolic switching during adaptation to oncogenic receptor tyrosine kinase inhibition. *Oncogene* **41**, 4808-4822, doi:10.1038/s41388-022-02465-w (2022).

36 Hagemann, T., Balkwill, F. & Lawrence, T. Inflammation and cancer: a double-edged sword. *Cancer Cell* **12**, 300-301, doi:10.1016/j.ccr.2007.10.005 (2007).

37 Mullen, J., Kato, S., Sicklick, J. K. & Kurzrock, R. Targeting ARID1A mutations in cancer. *Cancer Treat Rev* **100**, 102287, doi:10.1016/j.ctrv.2021.102287 (2021).

38 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).

39 Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)* **25**, 1754-1760, doi:10.1093/bioinformatics/btp324 (2009).

40 Consortium, G. P. *et al.* A global reference for human genetic variation. *Nature* **526**, 68-74, doi:10.1038/nature15393 (2015).

41 Benjamin, D. *et al.* Calling Somatic SNVs and Indels with Mutect. (2019).

42 Talevich, E., Shain, A. H., Botton, T. & Bastian, B. C. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. *PLOS Computational Biology* **12**, e1004873, doi:10.1371/journal.pcbi.1004873 (2016).

43 Boeva, V. *et al.* Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. *Bioinformatics* **28**, 423-425, doi:10.1093/bioinformatics/btr670 (2012).

44 Chin, S. F. *et al.* High-resolution aCGH and expression profiling identifies a novel genomic subtype of ER negative breast cancer. *Genome Biology* **8**, R215, doi:10.1186/gb-2007-8-10-r215 (2007).

45 Ewels, P. A. *et al.* The nf-core framework for community-curated bioinformatics pipelines. *Nature Biotechnology* **38**, 276-278, doi:10.1038/s41587-020-0439-x (2020).

46 Hao, Y. *et al.* Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol* **42**, 293-304, doi:10.1038/s41587-023-01767-y (2024).

47 Carroll, T. M. *et al.* Tumor monocyte content predicts immunochemotherapy outcomes in esophageal adenocarcinoma. *Cancer Cell* **41**, 1222-+, doi:10.1016/j.ccell.2023.06.006 (2023).

48 Hou, W. & Ji, Z. Assessing GPT-4 for cell type annotation in single-cell RNA-seq analysis. *Nat Methods* **21**, 1462-1465, doi:10.1038/s41592-024-02235-4 (2024).

49 Qiu, X. J. *et al.* Reversed graph embedding resolves complex single-cell trajectories. *Nature Methods* **14**, 979-+, doi:10.1038/Nmeth.4402 (2017).

50 Jin, S. *et al.* Inference and analysis of cell-cell communication using CellChat. *Nat Commun* **12**, 1088, doi:10.1038/s41467-021-21246-9 (2021).

51 Efremova, M., Vento-Tormo, M., Teichmann, S. A. & Vento-Tormo, R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nature Protocols* **15**, 1484-1506, doi:10.1038/s41596-020-0292-x (2020).

52 Kleshchevnikov, V. *et al.* Cell2location maps fine-grained cell types in spatial transcriptomics. *Nat Biotechnol* **40**, 661-671, doi:10.1038/s41587-021-01139-4 (2022).

1160 **Figure legends**

1161 **Figure 1**: **The expression of *ERBB2*/HER2 is a favorable prognosis marker for esophageal**
1162 **adenocarcinoma (EAC) and gastric cardia adenocarcinoma (GCA)**.

1163 **a.** The Pearson correlation between *ERBB2* RNA expression and HER2 protein expression in
1164 normal tissue adjacent to the tumor (NAT) (left panel) and tumor (right panel) from the
1165 GCA cohort. N = HER2 protein expression levels detected by mass spectrometry.

1166 **b.** The distribution of normalized *ERBB2* RNA expression in the GCA cohort, as well as in
1167 EAC and STAD from TCGA. ESCC = Esophageal squamous-cell carcinoma; STAD =
1168 Stomach adenocarcinoma; TPM = Transcripts Per Million.

1169 **c.** Kaplan-Meier survival curves of two patient groups based on the median *ERBB2* RNA
1170 expression from TCGA. High = upper half of patients; Low = lower half of patients.

1171 **d.** The strategy for dividing the GCA cohort into two groups based on the median ratio of
1172 *ERBB2* RNA expression levels between tumors and NATs (Left) and prognosis analysis for
1173 each group (Right).

1174 **e.** Kaplan-Meier survival curves of two patient groups based on ratio of ERBB2 RNA
1175 expression.

1176 **f.** The strategy for dividing the GCA cohort into three groups based on the ratio of HER2
1177 protein expression levels between tumors and NATs. "Negative" refers to patients with
1178 HER2-negative expression; "High" refers to the upper half of HER2-positive patients;
1179 "Low" refers to the lower half of HER2-positive patients.

1180 **g.** Kaplan-Meier survival curves of three patient groups based on ratio of HER2 protein
1181 expression.

1182

1183 **Figure 2**: **Protein expression landscape of different HER2 groups from GCA patients.**

1184 **a.** The illustration of multiple omics data from GCA patients.

1185 **b.** Unique protein profiles of different HER2 GCA patient groups from tumor only.

1186 **c.** Gene set enrichment analysis (GSEA) with protein of different HER2 GCA patient groups
1187 from tumor only comparison.

1188 **d.** GSEA of different HER2 GCA patient groups by comparing tumors with corresponding
1189 normal tissue adjacent to the tumor (NAT).

1190 **Figure 3**: **Somatic mutation and gain landscape of different HER2 groups from GCA**
1191 **patients.**
1192 **a.** The top 20 mutated oncogenes and tumor suppressor genes captured in the GCA cohort,
1193 along with their mutation frequency in different HER2 groups.
1194 **b.** Comparison of prognosis in GCA patients with and without *ARID1A* gene mutations.
1195 **c.** Comparison of prognosis in GCA patients among the three HER2 groups and *ARID1A* gene
1196 mutations
1197 **d.** Comparison of chromosome instability (CIN) among the three different HER2 groups
1198 using the genome integrity index.
1199 **e.** Comparison of the number of breakpoints among the three different HER2 groups.
1200 **f.** Comparison of copy number alterations (CNAs) among the three different HER2 groups.
1201 **g.** The significantly different CNA regions in the three HER2 groups with a bin size of 1 Mbp;
1202 oncogenes and tumor suppressor genes are labeled on the side.
1203 **h.** Examples of significant CNA regions in the three HER2 groups on chromosomes 13, 15,
1204 and 17.
1205
1206

1207 **Figure 4**: **Gene expression signature of different HER2 groups from GCA patients**

1208 **a.** Gene set enrichment analysis (GSEA) using RNA expression from tumors of different
1209 HER2 GCA patient groups.

1210 **b.** Gene expression signatures from tumors of different HER2 GCA patient groups.

1211 **c.** GSEA with RNA expression of different HER2 GCA patient groups by comparing tumors
1212 with corresponding normal tissue adjacent to the tumor (NAT).

1213 **d.** Differential RNA expression from genes of major histocompatibility complex (MHC) class
1214 I and II in different HER2 groups by comparing tumors with corresponding NATs.

1215 **e.** Differential RNA expression from genes of cytokines from different HER2 groups by
1216 comparing tumors with corresponding NATs.

1217 **f.** FDR = false discovery rate.

1218

1219

1220

1221

1222 **Figure 5**: **Deciphering tumor microenvironment of different HER2 groups from GCA**
1223 **patients**

1224 **a.** Comparison of cell components from tumor bulk RNA-seq predictions across different
1225 HER2 groups.
1226 **b.** Quantitative comparison of M2 macrophages in tumors only from different HER2 groups.
1227 **c.** Quantification of immunohistochemistry (IHC) staining of CD163 in tumors across
1228 different HER2 groups.
1229 **d.** Representative images of immunohistochemistry (IHC) staining of CD163 in tumors from
1230 different HER2 groups.
1231 **e.** Grouping of GCA patients into *ERBB2* high and low samples based on *ERBB2* gene
1232 expression levels.
1233 **f.** Proportional quantification of different cell types in seven GCA patients from scRNA-Seq.
1234 **g.** Cell type identification in seven GCA patients from scRNA-Seq.
1235 **h.** Distribution of different cell types in seven GCA patients.
1236 **i.** Quantitative comparison of dendritic cells (left panel) and M2 macrophages (right panel)
1237 in different HER2 groups from scRNA-Seq.
1238 **j.** The top 10 pathways enriched from KEGG (Kyoto Encyclopedia of Genes and Genomes)
1239 analysis in different HER2 groups from M2 macrophages.
1240

1241 **Figure 6**: **Unique signal pathways in epithelial cells of different HER2 groups from GCA**
1242 **patients**

1243 **a.** Cell type identification in epithelial tumor cells (first left), the distribution of cells in high
1244 and low groups (second left), the distribution of high and low CNV (copy number variants)
1245 in each cell cluster (second right), and the distribution of aneuploid and diploid cells in each
1246 cell cluster (first right).

1247 **b.** Identification of high CNV (copy number variants) cells (first left), comparison of CNV
1248 values between epithelial tumor cells and epithelial normal cells (second left), comparison
1249 of the proportion of high CNV cells in high and low groups (second right), and comparison
1250 of the proportion of aneuploid cells in high and low groups (first right).

1251 **c.** Pseudotime projection of epithelial tumor cells and the distribution of cells from high and
1252 low groups.

1253 **d.** Distribution of epithelial normal cells and tumor cells in different cell states (left) and
1254 different cell states in high and low groups (right).

1255 **e.** Extracted feature genes for different cell fates from epithelial tumor cells.

1256 **f.** The top 10 pathways enriched from KEGG (Kyoto Encyclopedia of Genes and Genomes)
1257 analysis for different cell fates from epithelial tumor cells.

1258 **g.** Comparison of signal communication among different cell types in high and low groups.

1259 **h.** Circle plots presenting signal communication sent by M2 macrophages (left) and received
1260 by epithelial cells (right).

1261 **i.** Representative cell signals sent by M2 macrophages from scRNA-Seq.

1262 **j.** Representative cell signals received by epithelial tumor cells from scRNA-Seq.

1263 **k.** Signal strength comparison from different signaling pathways in high and low GCA groups
1264 from scRNA-Seq.

1265 **l.** Circle plots presenting signal pathway communication of NRG in high group.

1266 **m.** Circle plots presenting CXCL pathway contributed by M2 macrophages in high and low
1267 groups.

1268

1269 **Figure 7**: **Spatial transcriptons uncovers distinct immune inhibitory checkpoints among**
1270 **three HER2 groups**

1271 **a.** Comparison of differential gene expression between tumors and corresponding normal
1272 tissue adjacent to the tumor (NAT) for paired immune checkpoint receptors and ligands in
1273 different HER2 groups; FC = Fold Change; white block = False Discovery Rate (FDR)
1274 greater than 0.05; an asterisk indicates $|\log_2(\text{Fold Change})| \geq 1$ and FDR less than 0.05.

1275 **b.** Representative immunohistochemistry (IHC) staining of positive and negative cells for PD-
1276 L2 (PDCD1LG2) and SIRPA.

1277 **c.** Quantitative comparison of PD-L2 and SIRPA positive cells detected in the three HER2
1278 groups.

1279 **d.** Interaction strength of paired immune checkpoint receptors and ligands in high and low
1280 groups detected in scRNA-Seq.

1281 **e.** Comparisons of *SIRPA* and *CD47* gene expression from M2 macrophages in high and low
1282 groups.

1283 **f.** Comparisons of *SIRPA* and *CD47* gene expression from epithelial tumor cells in high and
1284 low groups.

1285 **g.** Distribution of different cell types identified from spatial transcriptomics in the three GCA
1286 groups.

1287 **h.** Quantitative comparison of dendritic cells and macrophages in different HER2 groups from
1288 spatial transcriptomics.

1289 **i.** Representative images of H&E (Hematoxylin and Eosin) staining (top row), identified
1290 epithelial tumor cells (second row), myeloid cells (third row), co-localization of epithelial
1291 cells and myeloid cells (bottom second row), and co-localization of CD47-SIRPA in
1292 epithelial and myeloid cells (bottom row) from spatial transcriptomics in the three different
1293 HER2 groups.

1294 **j.** Quantitative comparison of the co-localization of epithelial cells and myeloid cells from
1295 spatial transcriptomics in the three different HER2 groups.

1296 **k.** Quantitative comparison of the co-localization of CD47-SIRPA in epithelial and myeloid
1297 cells from spatial transcriptomics in the three different HER2 groups.

1298

Figure 1

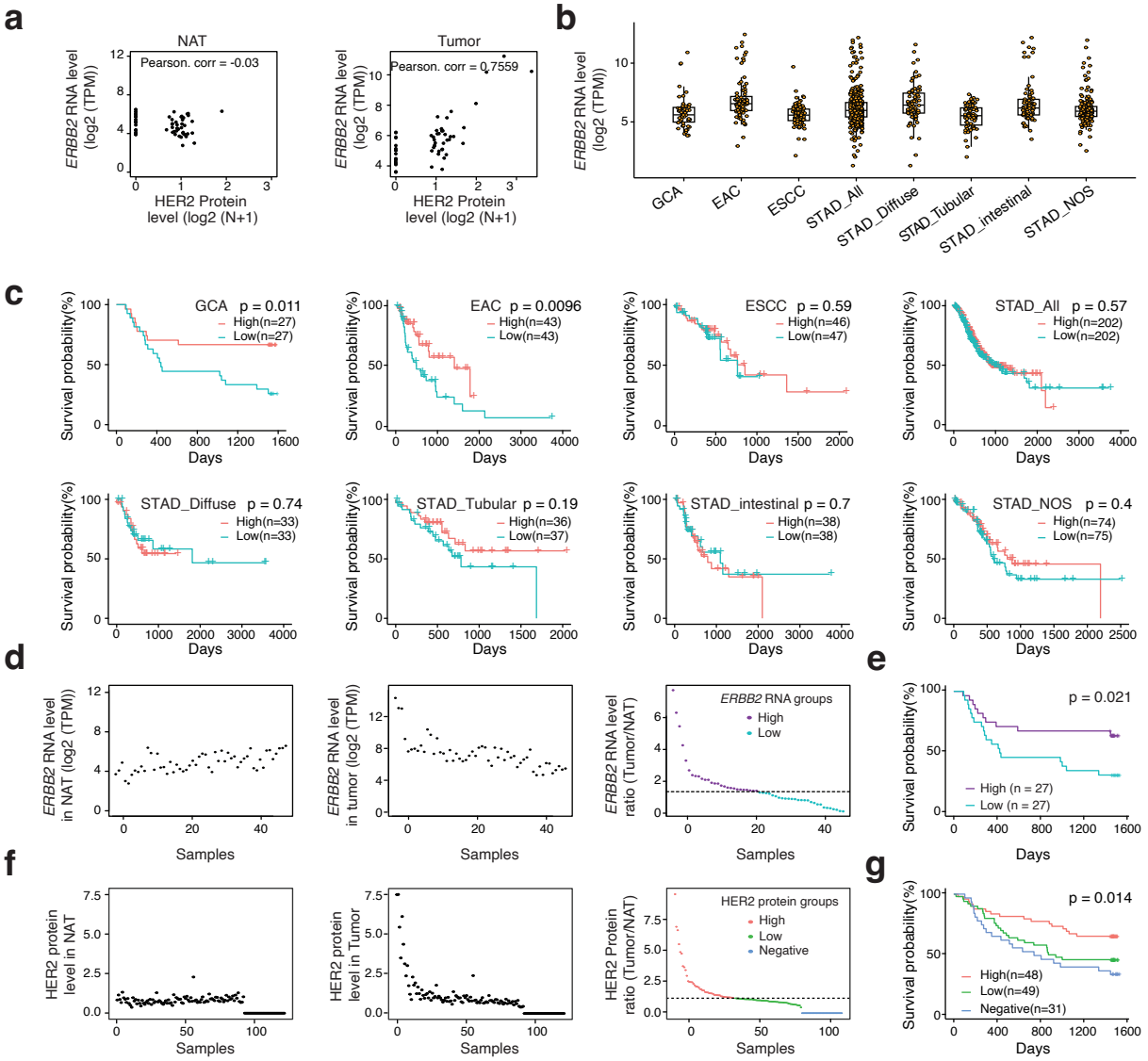

a
NAT
Pearson. corr = -0.03
ERBB2 RNA level (log2 (TPM))
HER2 Protein level (log2 (N+1))
Tumor
Pearson. corr = 0.7559
b
GCA
EAC
ESCC
STAD_All
STAD_Diffuse
STAD_Tubular
STAD_intestinal
STAD_NOS
c
GCA p = 0.011
High(n=27)
Low(n=27)
EAC p = 0.0096
High(n=43)
Low(n=43)
ESCC p = 0.59
High(n=46)
Low(n=47)
STAD_All p = 0.57
High(n=202)
Low(n=202)
STAD_Diffuse p = 0.74
High(n=33)
Low(n=33)
STAD_Tubular p = 0.19
High(n=36)
Low(n=37)
STAD_intestinal p = 0.7
High(n=38)
Low(n=38)
STAD_NOS p = 0.4
High(n=74)
Low(n=75)
Survival probability(%)
Days
d
ERBB2 RNA level in NAT (log2 (TPM))
ERBB2 RNA level in tumor (log2 (TPM))
ERBB2 RNA level ratio (Tumor/NAT)
ERBB2 RNA groups
High
Low
Samples
e
p = 0.021
High (n = 27)
Low (n = 27)
f
HER2 protein level in NAT
HER2 protein level in Tumor
HER2 Protein ratio (Tumor/NAT)
HER2 protein groups
High
Low
Negative
g
p = 0.014
High(n=48)
Low(n=49)
Negative(n=31)

Figure 2

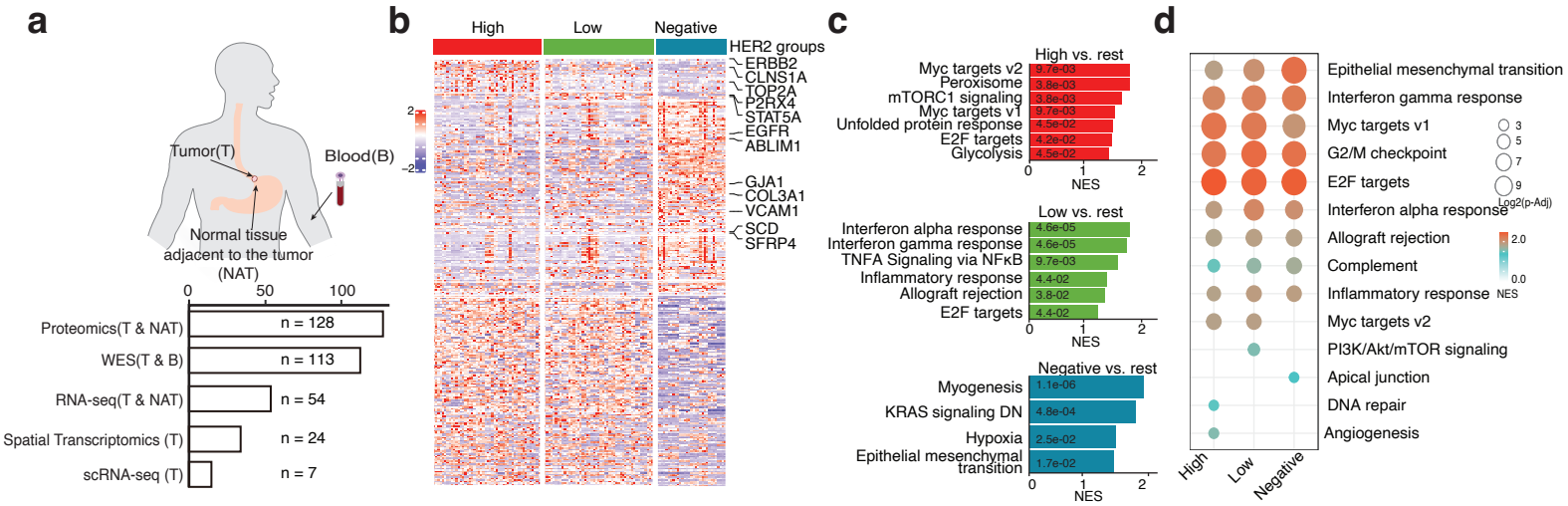

a
Tumor(T)
Blood(B)
Normal tissue adjacent to the tumor (NAT)
0
50
100
Proteomics(T & NAT)
n = 128
WES(T & B)
n = 113
RNA-seq(T & NAT)
n = 54
Spatial Transcriptomics (T)
n = 24
scRNA-seq (T)
n = 7
b
High
Low
Negative
HER2 groups
ERBB2
CLNS1A
TOP2A
P2RX4
STAT5A
EGFR
ABLIM1
GJA1
COL3A1
VCAM1
SCD
SFRP4
2
-2
c
High vs. rest
Myc targets v2
Peroxisome
mTORC1 signaling
Myc targets v1
Unfolded protein response
E2F targets
Glycolysis
9.7e-03
3.8e-03
3.8e-03
9.7e-03
4.5e-02
4.2e-02
4.5e-02
NES
Low vs. rest
Interferon alpha response
Interferon gamma response
TNFA Signaling via NFκB
Inflammatory response
Allograft rejection
E2F targets
4.6e-05
4.6e-05
9.7e-03
4.4-02
3.8-02
4.4-02
NES
Negative vs. rest
Myogenesis
KRAS signaling DN
Hypoxia
Epithelial mesenchymal transition
1.1e-06
4.8e-04
2.5e-02
1.7e-02
NES
d
Epithelial mesenchymal transition
Interferon gamma response
Myc targets v1
G2/M checkpoint
E2F targets
Interferon alpha response
Allograft rejection
Complement
Inflammatory response
Myc targets v2
PI3K/Akt/mTOR signaling
Apical junction
DNA repair
Angiogenesis
High
Low
Negative
Log2(p-Adj)
NES

Figure 3

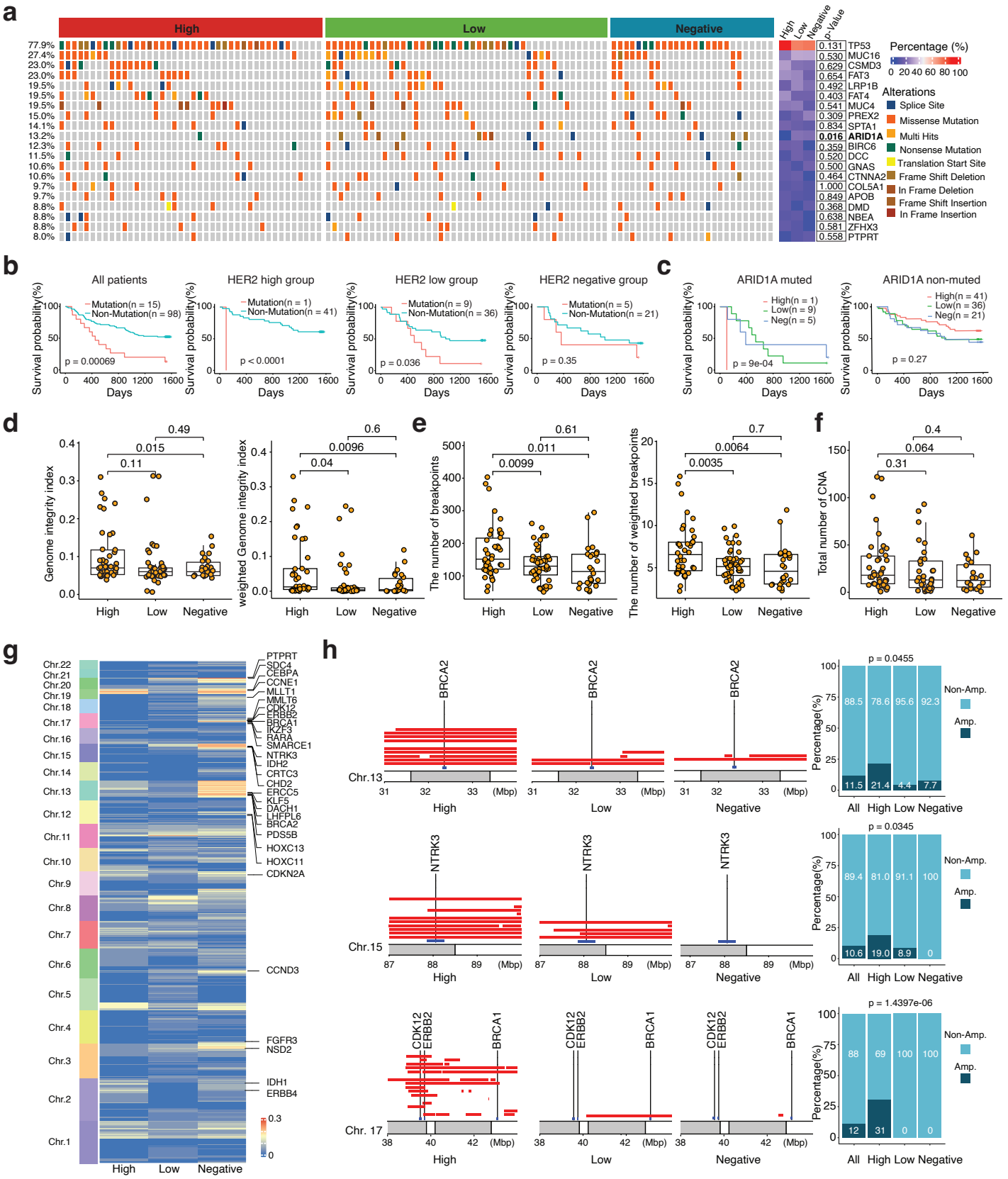

a
High
Low
Negative
High
Low
Negative
p-Value
77.9% 0.131 TP53
27.4% 0.530 MUC16
23.0% 0.629 CSMD3
23.0% 0.654 FAT3
19.5% 0.492 LRP1B
19.5% 0.403 FAT4
19.5% 0.541 MUC4
15.0% 0.309 PREX2
14.1% 0.834 SPTA1
13.2% 0.016 ARID1A
12.3% 0.359 BIRC6
11.5% 0.520 DCC
10.6% 0.500 GNAS
10.6% 0.464 CTNNA2
9.7% 1.000 COL5A1
9.7% 0.849 APOB
8.8% 0.368 DMD
8.8% 0.638 NBEA
8.8% 0.581 ZFHX3
8.0% 0.558 PTPRT
Percentage (%)
0 20 40 60 80 100
Alterations
Splice Site
Missense Mutation
Multi Hits
Nonsense Mutation
Translation Start Site
Frame Shift Deletion
In Frame Deletion
Frame Shift Insertion
In Frame Insertion
b
All patients
Mutation(n = 15)
Non-Mutation(n = 98)
p = 0.00069
HER2 high group
Mutation(n = 1)
Non-Mutation(n = 41)
p < 0.0001
HER2 low group
Mutation(n = 9)
Non-Mutation(n = 36)
p = 0.036
HER2 negative group
Mutation(n = 5)
Non-Mutation(n = 21)
p = 0.35
Survival probability(%)
Days
c
ARID1A muted
High(n = 1)
Low(n = 9)
Neg(n = 5)
p = 9e-04
ARID1A non-muted
High(n = 41)
Low(n = 36)
Neg(n = 21)
p = 0.27
d
Genome integrity index
0.49
0.015
0.11
weighted Genome integrity index
0.6
0.0096
0.04
e
The number of breakpoints
0.61
0.011
0.0099
The number of weighted breakpoints
0.7
0.0064
0.0035
f
Total number of CNA
0.4
0.064
0.31
High
Low
Negative
g
Chr.22
Chr.21
Chr.20
Chr.19
Chr.18
Chr.17
Chr.16
Chr.15
Chr.14
Chr.13
Chr.12
Chr.11
Chr.10
Chr.9
Chr.8
Chr.7
Chr.6
Chr.5
Chr.4
Chr.3
Chr.2
Chr.1
PTPRT
SDC4
CEBPA
CCNE1
MLLT1
MMLT6
CDK12
ERBB2
BRCA1
IKZF3
RARA
SMARCE1
NTRK3
IDH2
CRTC3
CHD2
ERCC5
KLF5
DACH1
LHFPL6
BRCA2
PDS5B
HOXC13
HOXC11
CDKN2A
CCND3
FGFR3
NSD2
IDH1
ERBB4
0.3
0
h
BRCA2
Chr.13
31 32 33 (Mbp)
p = 0.0455
Percentage(%)
88.5 78.6 95.6 92.3
11.5 21.4 4.4 7.7
Non-Amp.
Amp.
All High Low Negative
NTRK3
Chr.15
87 88 89 (Mbp)
p = 0.0345
89.4 81.0 91.1 100
10.6 19.0 8.9 0
CDK12
ERBB2
BRCA1
Chr. 17
38 40 42 (Mbp)
p = 1.4397e-06
88 69 100 100
12 31 0 0

Figure 4

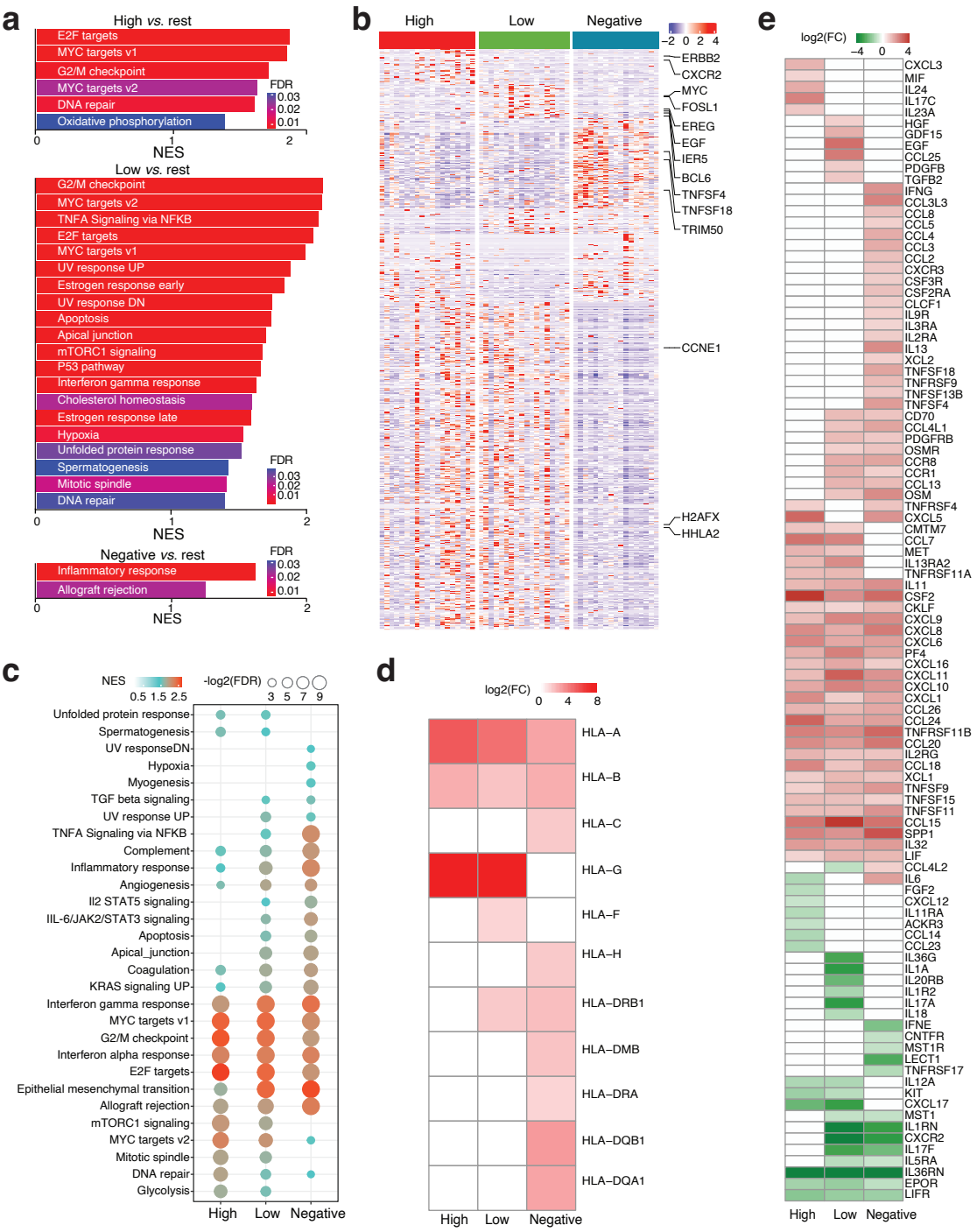

a
High vs. rest
E2F targets
MYC targets v1
G2/M checkpoint
MYC targets v2
DNA repair
Oxidative phosphorylation
FDR
0.03
0.02
0.01
0
1
2
NES
Low vs. rest
G2/M checkpoint
MYC targets v2
TNFA Signaling via NFKB
E2F targets
MYC targets v1
UV response UP
Estrogen response early
UV response DN
Apoptosis
Apical junction
mTORC1 signaling
P53 pathway
Interferon gamma response
Cholesterol homeostasis
Estrogen response late
Hypoxia
Unfolded protein response
Spermatogenesis
Mitotic spindle
DNA repair
Negative vs. rest
Inflammatory response
Allograft rejection
b
High
Low
Negative
−2 0 2 4
ERBB2
CXCR2
MYC
FOSL1
EREG
EGF
IER5
BCL6
TNFSF4
TNFSF18
TRIM50
CCNE1
H2AFX
HHLA2
c
NES
0.5 1.5 2.5
-log2(FDR)
3 5 7 9
Unfolded protein response
Spermatogenesis
UV responseDN
Hypoxia
Myogenesis
TGF beta signaling
UV response UP
TNFA Signaling via NFKB
Complement
Inflammatory response
Angiogenesis
Il2 STAT5 signaling
IIL-6/JAK2/STAT3 signaling
Apoptosis
Apical_junction
Coagulation
KRAS signaling UP
Interferon gamma response
MYC targets v1
G2/M checkpoint
Interferon alpha response
E2F targets
Epithelial mesenchymal transition
Allograft rejection
mTORC1 signaling
MYC targets v2
Mitotic spindle
DNA repair
Glycolysis
High
Low
Negative
d
log2(FC)
0 4 8
HLA−A
HLA−B
HLA−C
HLA−G
HLA−F
HLA−H
HLA−DRB1
HLA−DMB
HLA−DRA
HLA−DQB1
HLA−DQA1
High
Low
Negative
e
log2(FC)
−4 0 4
CXCL3
MIF
IL24
IL17C
IL23A
HGF
GDF15
EGF
CCL25
PDGFB
TGFB2
IFNG
CCL3L3
CCL8
CCL5
CCL4
CCL3
CCL2
CXCR3
CSF3R
CSF2RA
CLCF1
IL9R
IL3RA
IL2RA
IL13
XCL2
TNFSF18
TNFRSF9
TNFSF13B
TNFSF4
CD70
CCL4L1
PDGFRB
OSMR
CCR8
CCR1
CCL13
OSM
TNFRSF4
CXCL5
CMTM7
CCL7
MET
IL13RA2
TNFRSF11A
IL11
CSF2
CKLF
CXCL9
CXCL8
CXCL6
PF4
CXCL16
CXCL11
CXCL10
CXCL1
CCL26
CCL24
TNFRSF11B
CCL20
IL2RG
CCL18
XCL1
TNFSF9
TNFSF15
TNFSF11
CCL15
SPP1
IL32
LIF
CCL4L2
IL6
FGF2
CXCL12
IL11RA
ACKR3
CCL14
CCL23
IL36G
IL1A
IL20RB
IL1R2
IL17A
IL18
IFNE
CNTFR
MST1R
LECT1
TNFRSF17
IL12A
KIT
CXCL17
MST1
IL1RN
CXCR2
IL17F
IL5RA
IL36RN
EPOR
LIFR
High
Low
Negative

Figure 5

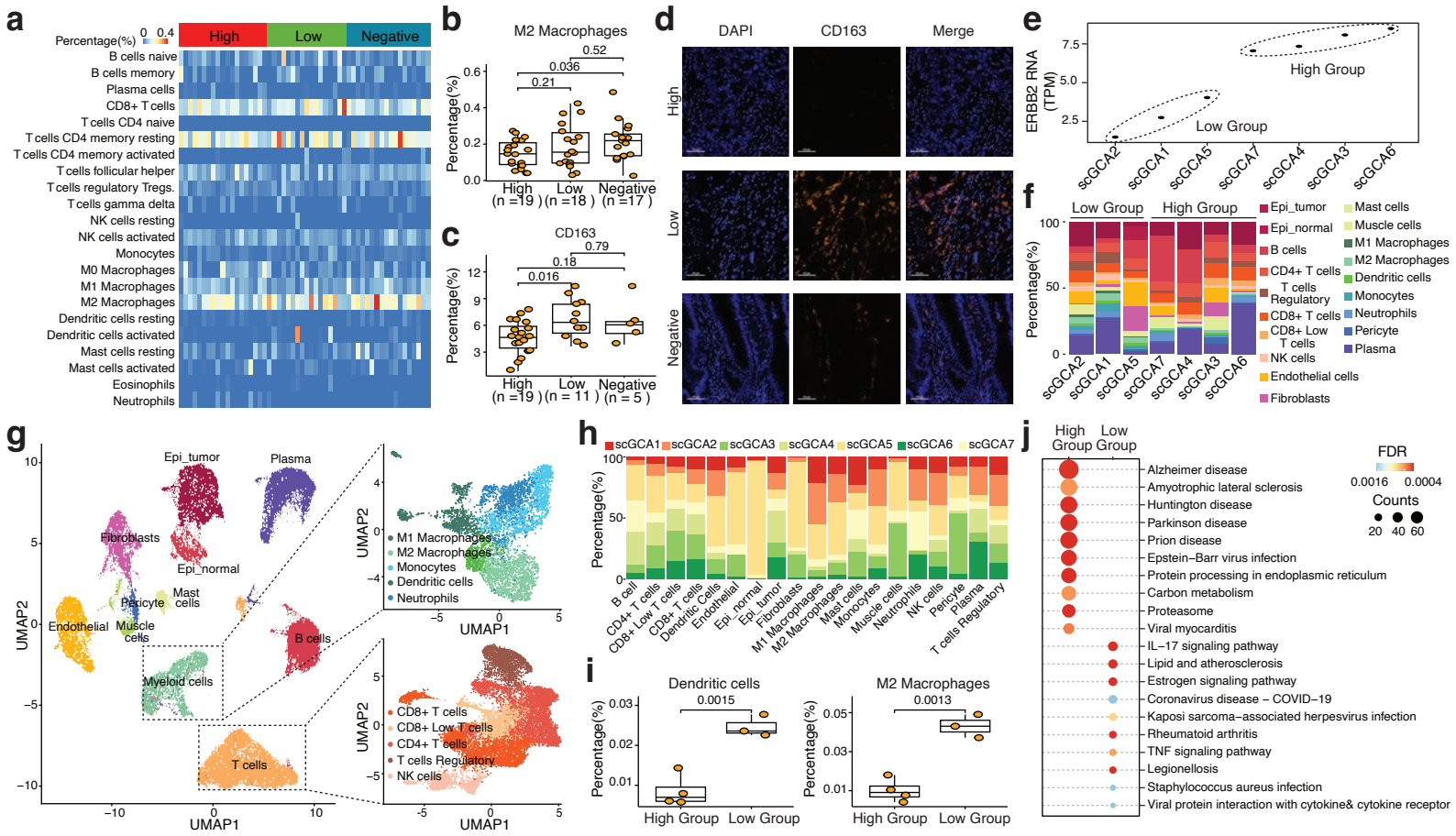

a
Percentage(%)
0 0.4
High
Low
Negative
B cells naive
B cells memory
Plasma cells
CD8+ T cells
T cells CD4 naive
T cells CD4 memory resting
T cells CD4 memory activated
T cells follicular helper
T cells regulatory Tregs.
T cells gamma delta
NK cells resting
NK cells activated
Monocytes
M0 Macrophages
M1 Macrophages
M2 Macrophages
Dendritic cells resting
Dendritic cells activated
Mast cells resting
Mast cells activated
Eosinophils
Neutrophils
b
M2 Macrophages
Percentage(%)
0.52
0.036
0.21
High (n =19)
Low (n =18)
Negative (n =17)
c
CD163
Percentage(%)
0.79
0.18
0.016
High (n =19)
Low (n = 11)
Negative (n = 5)
d
DAPI
CD163
Merge
High
Low
Negative
e
ERBB2 RNA (TPM)
High Group
Low Group
scGCA2
scGCA1
scGCA5
scGCA7
scGCA4
scGCA3
scGCA6
f
Low Group
High Group
Percentage(%)
scGCA2
scGCA1
scGCA5
scGCA7
scGCA4
scGCA3
scGCA6
Epi_tumor
Epi_normal
B cells
CD4+ T cells
T cells Regulatory
CD8+ T cells
CD8+ Low T cells
NK cells
Endothelial cells
Fibroblasts
Mast cells
Muscle cells
M1 Macrophages
M2 Macrophages
Dendritic cells
Monocytes
Neutrophils
Pericyte
Plasma
g
Epi_tumor
Plasma
Fibroblasts
Epi_normal
Pericyte
Mast cells
Endothelial
Muscle cells
B cells
Myeloid cells
T cells
UMAP1
UMAP2
M1 Macrophages
M2 Macrophages
Monocytes
Dendritic cells
Neutrophils
CD8+ T cells
CD8+ Low T cells
CD4+ T cells
T cells Regulatory
NK cells
h
scGCA1
scGCA2
scGCA3
scGCA4
scGCA5
scGCA6
scGCA7
Percentage(%)
B cell
CD4+ T cells
CD8+ Low T cells
CD8+ T cells
Dendritic Cells
Endothelial
Epi_normal
Epi_tumor
Fibroblasts
M1 Macrophages
M2 Macrophages
Mast cells
Monocytes
Muscle cells
Neutrophils
NK cells
Pericyte
Plasma
T cells Regulatory
i
Dendritic cells
0.0015
M2 Macrophages
0.0013
Percentage(%)
High Group
Low Group
j
High Group
Low Group
FDR
0.0016
0.0004
Counts
20
40
60
Alzheimer disease
Amyotrophic lateral sclerosis
Huntington disease
Parkinson disease
Prion disease
Epstein–Barr virus infection
Protein processing in endoplasmic reticulum
Carbon metabolism
Proteasome
Viral myocarditis
IL–17 signaling pathway
Lipid and atherosclerosis
Estrogen signaling pathway
Coronavirus disease – COVID–19
Kaposi sarcoma–associated herpesvirus infection
Rheumatoid arthritis
TNF signaling pathway
Legionellosis
Staphylococcus aureus infection
Viral protein interaction with cytokine& cytokine receptor

Figure 6

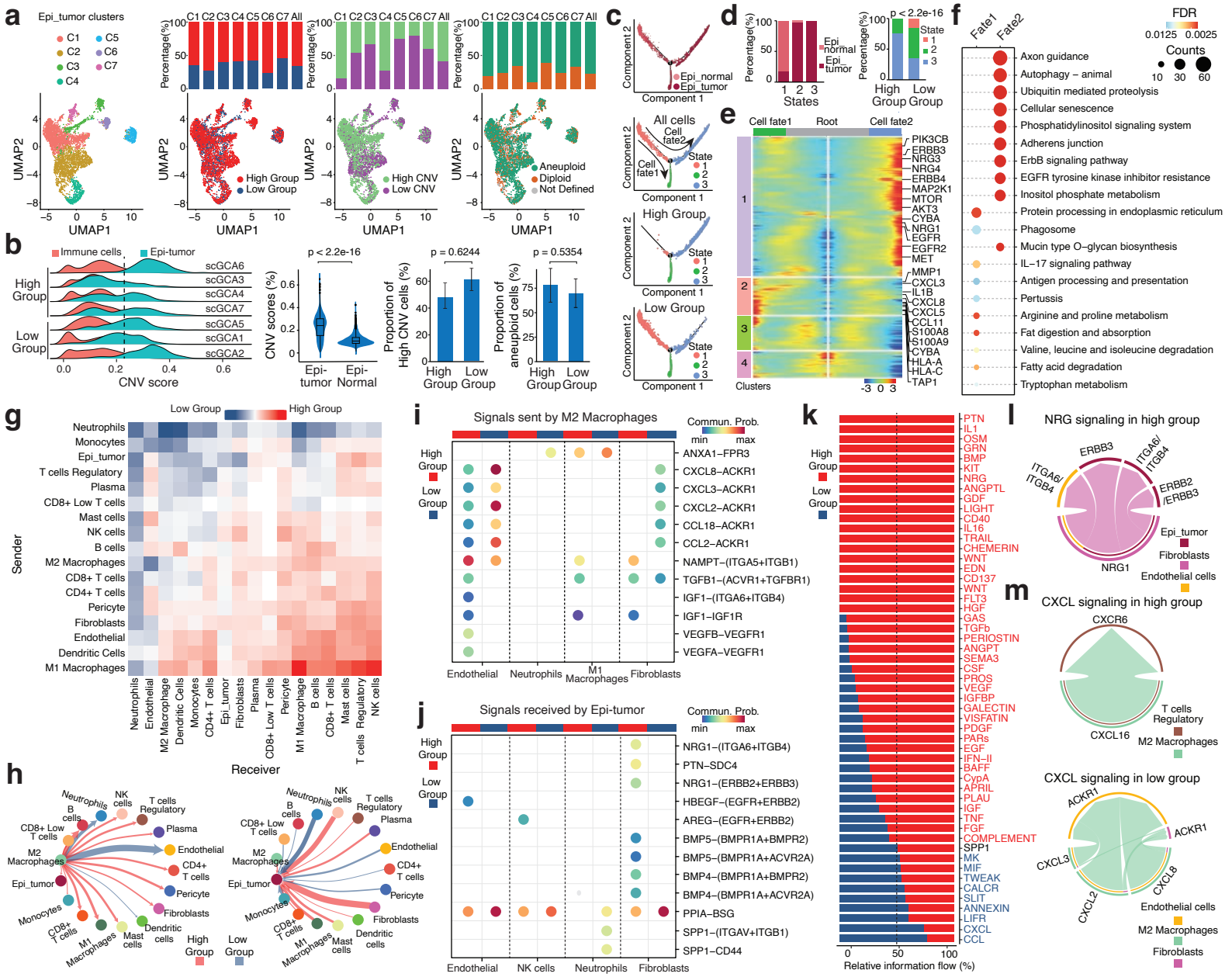
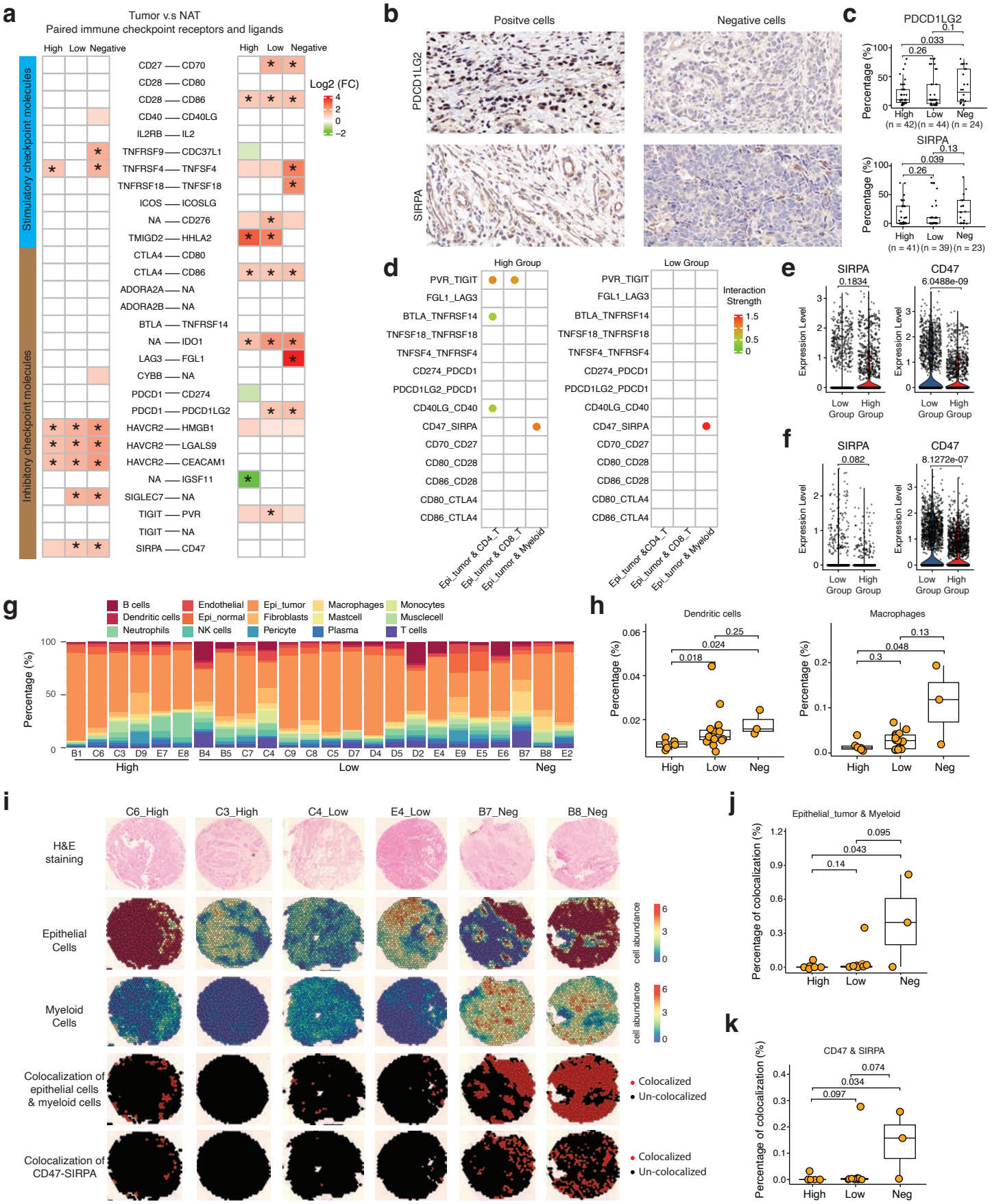

a
Epi_tumor clusters
C1
C2
C3
C4
C5
C6
C7
UMAP1
UMAP2
Percentage(%)
High Group
Low Group
High CNV
Low CNV
Aneuploid
Diploid
Not Defined
b
Immune cells
Epi-tumor
scGCA6
scGCA3
scGCA4
scGCA7
scGCA5
scGCA1
scGCA2
CNV score
CNV scores (%)
p < 2.2e-16
Epi-tumor
Epi-Normal
Proportion of High CNV cells (%)
p = 0.6244
Proportion of aneuploid cells (%)
p = 0.5354
c
Component 1
Component 2
Epi_normal
Epi_tumor
All cells
Cell fate1
Cell fate2
State
High Group
Low Group
d
Epi normal
Epi tumor
States
e
Cell fate1
Root
Cell fate2
Clusters
PIK3CB
ERBB3
NRG3
NRG4
ERBB4
MAP2K1
MTOR
AKT3
CYBA
NRG1
EGFR
EGFR2
MET
MMP1
CXCL3
IL1B
CXCL8
CXCL5
CCL11
S100A8
S100A9
CYBA
HLA-A
HLA-C
TAP1
f
Fate1
Fate2
FDR
Counts
Axon guidance
Autophagy – animal
Ubiquitin mediated proteolysis
Cellular senescence
Phosphatidylinositol signaling system
Adherens junction
ErbB signaling pathway
EGFR tyrosine kinase inhibitor resistance
Inositol phosphate metabolism
Protein processing in endoplasmic reticulum
Phagosome
Mucin type O-glycan biosynthesis
IL-17 signaling pathway
Antigen processing and presentation
Pertussis
Arginine and proline metabolism
Fat digestion and absorption
Valine, leucine and isoleucine degradation
Fatty acid degradation
Tryptophan metabolism
g
Low Group
High Group
Sender
Receiver
Neutrophils
Monocytes
Epi_tumor
T cells Regulatory
Plasma
CD8+ Low T cells
Mast cells
NK cells
B cells
M2 Macrophages
CD8+ T cells
CD4+ T cells
Pericyte
Fibroblasts
Endothelial
Dendritic Cells
M1 Macrophages
M2 Macrophage
Dendritic Cells
CD8+ Low T cells
M1 Macrophage
CD8+ T cells
T cells Regulatory
NK cells
h
Neutrophils
NK cells
T cells Regulatory
B cells
CD8+ Low T cells
Plasma
M2 Macrophages
Endothelial
CD4+ T cells
Epi_tumor
Pericyte
Monocytes
Fibroblasts
CD8+ T cells
M1 Macrophages
Mast cells
Dendritic cells
High Group
Low Group
i
Signals sent by M2 Macrophages
Commun. Prob.
min
max
ANXA1–FPR3
CXCL8–ACKR1
CXCL3–ACKR1
CXCL2–ACKR1
CCL18–ACKR1
CCL2–ACKR1
NAMPT–(ITGA5+ITGB1)
TGFB1–(ACVR1+TGFBR1)
IGF1–(ITGA6+ITGB4)
IGF1–IGF1R
VEGFB–VEGFR1
VEGFA–VEGFR1
Endothelial
Neutrophils
M1 Macrophages
Fibroblasts
j
Signals received by Epi-tumor
NRG1–(ITGA6+ITGB4)
PTN–SDC4
NRG1–(ERBB2+ERBB3)
HBEGF–(EGFR+ERBB2)
AREG–(EGFR+ERBB2)
BMP5–(BMPR1A+BMPR2)
BMP5–(BMPR1A+ACVR2A)
BMP4–(BMPR1A+BMPR2)
BMP4–(BMPR1A+ACVR2A)
PPIA–BSG
SPP1–(ITGAV+ITGB1)
SPP1–CD44
Endothelial
NK cells
Neutrophils
Fibroblasts
k
PTN
IL1
OSM
GRN
BMP
KIT
NRG
ANGPTL
GDF
LIGHT
CD40
IL16
TRAIL
CHEMERIN
WNT
EDN
CD137
WNT
FLT3
HGF
GAS
TGFb
PERIOSTIN
ANGPT
SEMA3
CSF
PROS
VEGF
IGFBP
GALECTIN
VISFATIN
PDGF
PARs
EGF
IFN-II
BAFF
CypA
APRIL
PLAU
IGF
TNF
FGF
COMPLEMENT
SPP1
MK
MIF
TWEAK
CALCR
SLIT
ANNEXIN
LIFR
CXCL
CCL
Relative information flow (%)
l
NRG signaling in high group
ERBB3
ITGA6/ ITGB4
ITGA6/ ITGB4
ERBB2 /ERBB3
Epi_tumor
Fibroblasts
Endothelial cells
NRG1
m
CXCL signaling in high group
CXCR6
T cells Regulatory
CXCL16
M2 Macrophages
CXCL signaling in low group
ACKR1
ACKR1
CXCL3
CXCL8
CXCL2
Endothelial cells
M2 Macrophages
Fibroblasts

Figure 7


a
Tumor v.s NAT
Paired immune checkpoint receptors and ligands
High Low Negative
Stimulatory checkpoint molecules
Inhibitory checkpoint molecules
CD27 — CD70
CD28 — CD80
CD28 — CD86
CD40 — CD40LG
IL2RB — IL2
TNFRSF9 — CDC37L1
TNFRSF4 — TNFSF4
TNFRSF18 — TNFSF18
ICOS — ICOSLG
NA — CD276
TMIGD2 — HHLA2
CTLA4 — CD80
CTLA4 — CD86
ADORA2A — NA
ADORA2B — NA
BTLA — TNFRSF14
NA — IDO1
LAG3 — FGL1
CYBB — NA
PDCD1 — CD274
PDCD1 — PDCD1LG2
HAVCR2 — HMGB1
HAVCR2 — LGALS9
HAVCR2 — CEACAM1
NA — IGSF11
SIGLEC7 — NA
TIGIT — PVR
TIGIT — NA
SIRPA — CD47
Log2 (FC)
b
Positve cells
Negative cells
PDCD1LG2
SIRPA
c
PDCD1LG2
Percentage (%)
High Low Neg
(n = 42) (n = 44) (n = 24)
SIRPA
(n = 41) (n = 39) (n = 23)
d
High Group
Low Group
Interaction Strength
PVR_TIGIT
FGL1_LAG3
BTLA_TNFRSF14
TNFSF18_TNFRSF18
TNFSF4_TNFRSF4
CD274_PDCD1
PDCD1LG2_PDCD1
CD40LG_CD40
CD47_SIRPA
CD70_CD27
CD80_CD28
CD86_CD28
CD80_CTLA4
CD86_CTLA4
Epi_tumor & CD4_T
Epi_tumor & CD8_T
Epi_tumor & Myeloid
e
SIRPA
CD47
Expression Level
Low Group
High Group
f
g
B cells
Dendritic cells
Neutrophils
Endothelial
Epi_normal
NK cells
Epi_tumor
Fibroblasts
Pericyte
Macrophages
Mastcell
Plasma
Monocytes
Musclecell
T cells
Percentage (%)
High
Low
Neg
h
Dendritic cells
Macrophages
i
C6_High
C3_High
C4_Low
E4_Low
B7_Neg
B8_Neg
H&E staining
Epithelial Cells
Myeloid Cells
Colocalization of epithelial cells & myeloid cells
Colocalization of CD47-SIRPA
cell abundance
Colocalized
Un-colocalized
j
Epithelial_tumor & Myeloid
Percentage of colocalization (%)
k
CD47 & SIRPA