Original Research

A comprehensive analysis of immune landscape of Indian triple negative breast cancer

Aruna Korlimarla[1#], Hari PS[1,2], Jyoti Prabhu[1], Chanthirika Ragulan[2], Yatish Patil[2], Snijesh VP[1], Krisha Desai[2], Aju Mathews[3], Maggie Cheang[2], Anguraj Sadanandam[2,4,#,*]

[1]St. John's Research Institute, Bangalore, India
[2]Division of Molecular Pathology, The Institute of Cancer Research, London, UK
[3] MOSC Medical College, Kolenchery, Kerala, India
[4]Centre for Global Oncology, Division of Molecular Pathology, The Institute of Cancer Research, London, UK

[#]These authors contributed equally as first authors
***Corresponding author:** Anguraj Sadanandam, Ph.D., Centre for Global Oncology, Division of Molecular Pathology, The Institute of Cancer Research (ICR), 15 Cotswold Road, Sutton, SM2 5NG, United Kingdom; Anguraj.sadanandam@icr.ac.uk; +0044-2034376440.

Running title: Immune landscape of Indian TNBC

Word count: 4994

Abstract

Background: Triple-negative breast cancer (TNBC) is a heterogeneous disease with a significant clinical challenge. TNBC alarmingly comprise 25-30% of breast cancers compared with only 10-15% among women in the West. Based on recent success in treating TNBCs with immunotherapies, one of the long-standing questions is whether Indian TNBC biologically and clinically resembles Western TNBC such that they respond similarly to immunotherapies. Here we sought to elucidate the immune landscape of Indian TNBCs for the first time, compare them to Western disease and associate them with clinical parameters and immunotherapy response.

Methods: We profiled 730 immune genes in 88 retrospective Indian TNBC samples using NanoString platform, clustered them into subtypes using a machine-learning approach and compared with Western TNBCs (n=422; public datasets). Subtype-specific gene signatures were identified, followed by clinicopathological, immune cell type, and pathway (multiomics) analyses. We also assessed responses to (cross-cancer) immunotherapy. Tumor-infiltrating lymphocytes

(TILs) and pan-macrophage marker were evaluated using hematoxilin-eosin staining and immunohistochemistry, respectively.

Results: We identified three robust and similarly distributed TNBC immune transcriptome subtypes (Subtypes-1-3) in Indian women, and they are represented correspondingly in Western TNBCs and associated with their well-known subtypes. Irrespective of the ethnicity, Subtype-1 harbored tumor microenvironmental and anti-tumor immune events associated with smaller tumors, younger age and a better prognosis. Subtype-1 mainly represented basal-like/claudin-low breast cancer and immunomodulatory TNBC subtypes. Subtype-1 showed an increase in cascade of events, including damage-associated molecular patterns, acute inflammation, Th1 responses, T-cell receptor-related and chemokine-specific signaling, antigen presentation, and viral-mimicry pathways. Subtype-1 was significantly ($p<0.05$) associated with pre-menopausal women, dense TILs and responses and/or improved prognosis to anti-PD-L1 and MAGEA3 immunotherapies. Subtype-2 was enriched for Th2/Th17 responses, $CD4^+$ regulatory cells, basal-like/mesenchymal subtypes and an intermediate prognosis. Subtype-3 patients expressed innate immune genes/proteins including those representing macrophages and neutrophils and had poor survival.

Conclusion: We identified three immune-specific TNBC subtypes in Indian (middle-income country) patients with differential clinical and immune behaviors, which largely overlapped with the Western TNBC cohorts. This study suggests cancer immunotherapy in TNBC may work similarly in both populations. Hence, this may expedite pharmaceutical adoption of immunotherapy to Indian TNBC patients.

Keywords: Immunity; immunotherapy; India; molecular stratification; triple-negative breast cancer; low-middle income countries; global oncology; DAMP; TCR signaling; immune landscape; Western countries; health disparity

INTRODUCTION

Breast cancer is the foremost cause of cancer-related deaths in women worldwide.[1] Ten to fifteen percent of breast cancers in Western women are triple-negative breast cancers (TNBCs), which do not express the estrogen receptor (ER), progesterone receptor (PR), or human epidermal growth factor receptor 2 (HER2) and have a poor overall prognosis and high recurrence risk.[2] TNBCs remain a clinical management challenge due to their aggressive characteristics and lack of targeted therapies.[3] Chemotherapy is used to treat primary and metastatic diseases, and while targeted therapies, including immunotherapy, are slowly emerging to treat TNBCs in the West.[3]

The incidence of TNBC is 20-31% in India, which is classified as a middle-income country by The World Bank based on its per capita income, than in Western populations (high-income-countries (HIC); 10-15%).[4 5] In India, TNBC affects younger, pre-menopausal women; and has a five-year survival rate of <30% when metastatic.[4-6] Approximately 100,000 women are diagnosed with TNBC each year in India, with a case fatality rate of 40%, contributing to India having one of the highest breast cancer-related death rates worldwide.[4-6] One of the proposed hypotheses for these increased death rates in India is that these TNBC patients have more aggressive disease with different immune biology than those from the West.[4 5] However, to our knowledge, no study has explored to address this hypothesis of immune infiltration differences between Indian and Western populations. If geographically/ethnically these diseases are different, it warrants distinct treatment opportunities for TNBC in India compared to the West. Therefore, understanding the molecular heterogeneity of TNBC and the biology of its response to existing therapies in this population are essential for improving outcomes both in low-middle-income countries (LMIC) and HIC.

To characterize TNBC biology, multiple groups have defined TNBC subtypes in the West using gene expression or integrated molecular profiles.[7-9] Immune cell infiltration into tumors is now recognized as intrinsic to tumor biology and immunotherapy responses, so a more granular analysis of immunity in TNBC is required for prediction and prognostication, especially given recent data showing that a proportion of patients respond well to immunotherapy - atezolizumab in combination with nab-paclitaxel for advanced PD-L1^{+} TNBCs (IMPassion130 study).[10] This combination therapy has been approved in India in 2020 based on IMPassion130 study,[11] which is

85 no longer an approved drug in the United States (US) after the outcomes from the follow-up trial
86 – IMPassion131 study.[12 13] Moreover, pembrolizumab (anti-PD1 immunotherapy), combined with
87 neoadjuvant chemotherapy, has been approved in the US for treating high-risk early-stage TNBCs
88 (based on KEYNOTE-522 study).[14] However, the unaddressed question is whether Indian and
89 Western TNBCs have similar biology and clinical outcomes; therefore, combination
90 immunotherapies from the West may benefit Indian TNBC women.

91 Hence, we sought to understand India-specific TNBC heterogeneity and functionally distinct
92 immune cell types infiltrating tumors in these patients using gene expression profiling and
93 compare this to Western TNBC patients and well-known TNBC subtypes[7-9]. The study goals were
94 to determine: (i) whether the immune composition and their clinical associations in Indian (LMIC)
95 TNBC women differed to that of TNBCs in Western populations (HIC); and (ii) whether differences
96 in the cellular composition of the immune infiltrate in TNBCs were associated with pathways,
97 immune signaling, patient survival and responses to therapy. We identified three robust immune
98 subtypes based on consensus clustering of the expression of immune-related genes in Indian TNBC
99 samples and found them to be represented similarly in two independent Western TNBC meta-
100 cohorts and associated with well-known TNBC subtypes from the West. Overall, our analysis
101 compared Indian TNBCs with those of the West at subtype levels (bi-directional) and suggests
102 avenues to rationally understand the immune and clinical landscapes to potentially translate and
103 personalize successful immunotherapy opportunities (from the West) to TNBC patients in India.

104

105 METHODS:

106 Patient samples, clinical characteristics and RNA isolation

107 TNBC patient samples were retrospectively collected as two independent cohorts from different
108 tertiary cancer care hospitals in Bangalore, India. Fifty out of 92 TNBC samples from one hospital,
109 St. John's Research Institute, were identified to be qualified for the study (**Figure 1A**). Similarly, 38
110 out of 48 TNBC samples were qualified for the study from Shri Shankara Cancer Hospital and
111 Research Centre. Both the studies were approved by the respective institutional ethics review
112 board and informed consent was obtained from the subjects.

113 Tumor tissue samples were collected at the time of surgery (before treatment), fixed in buffered
114 neutral formalin and processed as paraffin embedded blocks. Sections were cut from these blocks,
115 stained with hematoxylin and eosin (H&E) and examined by a pathologist to confirm the presence
116 of tumor. Immunohistochemistry for ER, PR and HER2 were done for determining receptor status
117 by standard procedures[15]. Only formalin-fixed paraffin embedded (FFPE) blocks containing pre-
118 treatment tissue with greater than 50% cancer epithelial cells were used for molecular analysis.
119 All clinical, histopathological and demographic details were collected from the hospital medical
120 records (**Figure 1B**). Radiologically recorded distant metastases or histologically confirmed local
121 recurrence and death related to disease had been documented.
122
123 RNA from all FFPE specimens was extracted by the Trizol method according to the manufacturer's
124 instructions and as described.[16] Briefly, 2x20 um sections were deparaffinized using xylene,
125 washed in absolute ethanol and overnight digested with Proteinase K. Lysate was taken for
126 extraction and RNA recovered from the aqueous phase.
127
128 nCounter profiling and subtype analysis
129 nCounter Immune profiling (NanoString Technologies)[17], subtype analysis[18] and the PPCCA
130 machine learning (ML) method[19] were performed as described. A detailed description has been
131 provided in Supplementary Methods.
132
133 TIL Scoring
134 Assessment of the TILs was done according to guidelines established by the International TIL
135 working group.[20] Large areas of central necrosis or fibrosis are not included for evaluation, tumor
136 was focused only on tumor-stroma at low magnification. Only mononuclear infiltrate of
137 lymphocytes and plasma cells were included and granulocytic infiltrate in areas of tumor necrosis
138 was not included. Based on the percentage of stromal TILs present, they were grouped into mild
139 (no or minimal immune cells), moderate (tumor with intermediate/heterogenous infiltrate) and
140 dense (tumor with high immune infiltrate) infiltrates.
141

142 Immunohistochemistry and scoring

143 IHC was done for CD68 according to standard procedures as described in Supplementary Methods.

144

145

146 RESULTS

147 Clinical characteristics of the Indian TNBC cohort

148 To our knowledge, this is the first comprehensive immune gene, cell type, signaling and
149 therapeutic analysis of retrospectively collected Indian TNBCs (n=88). A study flow chart and
150 collection source are shown in **Figure 1A**. The median age of the study population was 51 years,
151 with half ≤50 years of age and 44% pre-menopausal (**Figure 1B**). Most patients (77.3%) had grade-
152 3 tumors, the median tumor size was 2-5 cm (77.2%), 67% had stage II disease, and 60% were
153 lymph-node (LN) negative. Forty-two percent of samples had mild, 25% moderate, and 28% dense
154 TILs.

155 Indian TNBCs are heterogenous but can be divided into distinct subtypes with different prognoses

156 To characterize immune-based gene expression profiles of TNBCs in Indian women, 88 TNBC
157 samples were profiled for 730 immune genes using the PanCancer Immune gene panel
158 (NanoString Technologies). Batch effects in gene expression profiles between cohorts were
159 diagnosed and corrected using our exploBATCH[19] machine-learning (ML) tool (see Methods;
160 **Supplementary Figure 1A-D**). After selecting 392 variable genes (standard deviation ≥1), three
161 distinct TNBC immune gene expression subtypes were defined using unsupervised nonnegative
162 matrix factorization (NMF)-based consensus clustering (**Figure 2A).** The robustness of the clusters
163 and samples (or the likelihood of samples staying in the same cluster with re-iterative/consensus
164 clustering analysis) were defined by cophenetic coefficient and silhouette statistical metrics, as
165 previously described[18] (**Supplementary Methods; Supplementary Figure 1E-G and Supplementary**
166 **Table 1A**). The three subtypes were present in approximately equal proportions (27-38%; **Figure**
167 **2B**). 281 genes were significantly differentially expressed across the three subtypes according to
168 supervised significance analysis of microarrays (SAM) analysis (**Supplementary Table 1B**; see
169 Methods). Furthermore, prediction analysis for microarrays (PAM) statistical tool was used to

170 define summary statistics for each gene per subtype (centroids; **Supplementary Table 1C; see**
171 **Supplementary Methods**). Based on PAM centroids, we derived subtype-specific genes and
172 derived scores (mean expression of genes across samples; **Figures 2C-E**). Among the 281 SAM
173 genes, 204 were subtype-specific (non-zero) highest PAM centroid scores (**Supplementary Table**
174 **1D; Figure 2F**). A majority (79% and 19%) of 204 subtype-specific genes belonged to Subtype-1
175 and -2, respectively, whereas only 2% genes belonged to Subtype-3 (**Figure 2C**). Accordingly,
176 Subtype-1/-2 showed significantly ($p<0.001$) higher gene diversity as measured using the Shannon
177 Entropy method, as previously described[21]; **Figure 2G**), an index assessing immune gene
178 expression patterns. Hence, Subtype-1 shows immune-high gene expression patterns, whereas
179 Subtype-3 shows an immune-dormant or -exclusion pattern.

180 Subtype-3 patients had significantly ($p<0.05$; log-rank test) poorer overall survival (OS) and
181 disease-free survival (DFS; **Figures 2H-I**) than Subtype-1 (favorable prognosis) and Subtype-2
182 (intermediate prognosis) patients. In a multivariate analysis of tumor-infiltrating-lymphocyte (TIL)
183 groups, tumor size, menopausal status, stage, and molecular subtypes, Subtype-3 and stage-3
184 were independently associated with DFS (**Figure 2J**).

185 Next, we investigated whether these subtypes were present in TNBC samples from two Western
186 cohorts - The Cancer Genome Atlas (TCGA, n=123)[22] and Molecular Taxonomy of Breast Cancer
187 International Consortium (METABRIC; n=299)[23], which mainly represent Caucasian ethnic group
188 from HIC. There was no significant difference in the distribution of subtypes in the Indian TNBC
189 cohort compared with the Western cohorts, suggesting that the Indian and Western TNBCs are
190 similar (**Figure 2K, Supplementary Table 1E-F**). There was also a similar trend in OS according to
191 subtype in the METABRIC data as in the Indian cohort (**Figure 2L**).

192 Since the incidence of TNBC is reported to be higher in pre-menopausal women world-wide,
193 including in India[2 6], we sought to further understand the association between TNBC subtypes and
194 menopausal status. Subtype-1 TNBCs were more common in pre-menopausal patients and
195 Subtype-3 TNBCs were more common in post-menopausal patients (**Figure 2M**; $p<0.16$). A similar,
196 but not-significant, trend of menopausal status was observed in the TCGA ($p<0.16$) and METABRIC
197 ($p<0.28$) data (**Supplementary Figures 2A-B**).

198 Related to these data and as expected, Subtype-1 patients were younger (median age 44) than
199 Subtype-2 (median age 50) and Subtype-3 (median age 54) patients (**Supplementary Figures 2C;**
200 $p<0.05$, Chi-square test), and this trend (not statistically significant) was similar in TCGA and
201 METABRIC data (**Supplementary Figures 2D-E)**. Interestingly, Subtype-1 tumors were smaller than
202 Subtype-2 and -3 tumors (**Figure 2N**), and Subtypes-1 and -3 were associated with a greater
203 number of involved LNs (from METABRIC data) than Subtype-2 tumors (**Figure 2O**). With respect
204 to intrinsic subtypes, Subtype-1 shows a similar distribution of claudin-low and basal-like subtypes,
205 whereas Subtypes-2 and -3 predominantly show basal-like. Among the subtypes, Subtype-3
206 reveals increased HER2-like subtype, as assessed using METABRIC data (**Figure 2P**). There was no
207 significant association between METABRIC[23] and our subtypes (**Supplementary Figure 2F**).
208 Interestingly, a majority of Subtype-1 is associated with Vanderbilt's immune TNBC subtype[7] and
209 Subtype-2 with mesenchymal. In contrast, Subtype-3 was a mixture of all Vanderbilt's subtypes
210 (**Figure 2Q**). Although TCGA immune subtypes[24] show a significant association with our immune
211 subtypes, there is no specific association of each subtype with TCGA subtypes. Intriguingly,
212 Subtype-2 has more than 50% samples representing wounding healing C1 TCGA immune subtypes
213 (**Supplementary Figure 2F**). Overall, we identified three immune-specific TNBC subtypes in the
214 Indian cohort with different immune gene expression, prognoses, and menopausal status, which
215 largely overlapped with Western TNBC and certain molecular classifications.

216 **Subtype-1 is enriched for TILs, Th1 cellular immunity and a cascade of immune changes**

217 We next investigated whether our subtypes from the Indian cohort, representing the overall TNBC
218 immune profiles, were associated with histological assessment of TILs. In our cohort, 44% had
219 mild, 30% had dense, and 26% had moderate TILs. Most tumors with dense TILs were Subtype-1
220 tumors (68%; n=25), and those with moderate (82%; n=22) and mild (81%; n=37) TIL infiltrates
221 were mainly Subtype-2 and -3 tumors (**Figure 3A**). With respect to immune gene and cellular
222 composition, Subtype-1 had higher expression of Th1-specific genes (*IFNG* and *IL12A*; **Figure 3B-**
223 **C**), reflecting increased cellular immunity. Interestingly, most chemokines represented in the 204
224 subtype-specific genes were only highly expressed in Subtype-1 (**Figures 3D**). These patterns
225 apparent in the Indian cohort were also seen in TCGA data, with increased Th1 cell, leukocyte,
226 stromal, and TIL regional fractions and lymphocyte infiltration signature scores in Subtype-1

227 (**Figure 3E**). Hence, Subtype-1 appears to have a gene expression profile balanced in favor of Th1
228 responses and chemokine expression (**Figures 3F**).

229 We next explored the proportion and types of different immune cells within the three subtypes
230 using single-sample geneset enrichment analysis (ssGSEA)[25] and the entire 730 genes. The geneset
231 distribution in the three subtypes and the inferred cell populations significantly associated (FDR
232 <0.2) with the three subtypes are shown in **Figure 3G-K**. Subtype-1 showed increased $CD8^+$
233 effector T-cells (and their genes *CD8A*/*CD8B*) and the cytolytic activity genes *GZMA and PRF1*, as
234 assessed using cytolytic scores and ssGSEA (see Methods; **Figures 3G-I**). This increase in $CD8^+$
235 effector T-cell activity was also associated with patterns representing increased plasmacytoid
236 dendritic cells, major histocompatibility class (MHC) class-I, and co-stimulation of APC and T-cells
237 and associated antigen processing and presentation (APP) genes, including genes representing
238 transporters (TAP), immunoproteases, MHC class-II, and human leukocyte antigen (HLA) (**Figure**
239 **3I-K**). Subtype-1 also showed increased expression of immune checkpoint and T-cell exhaustion
240 genes such as PD-1 (*PDCD1*), PD-L1 (*CD274*), *CTLA4*, and *LAG3* (**Figure 3L**), suggesting that
241 persistent stimulation of T-cells potentially promoted T-cell exhaustion through increased
242 expression of these genes in Subtype-1 tumors, as described[26]. Further interrogation of TCGA data
243 showed increased INF-γ response genes and macrophage regulation, specifically anti-tumor M1
244 macrophages, in Subtype-1 samples (**Figure 3M**). We also examined T-cell receptor (TCR) and B
245 cell receptor (BCR) repertoires in our subtypes using TCGA data. As expected, TCR and BCR
246 Shannon indices and richness scores were significantly higher in Subtype-1 tumors, which might
247 represent a highly reactive immune-associated stroma in the tumor (**Figure 3E and M**). Our
248 subtypes capture an immune landscape and heterogeneity in TNBC beyond those represented by
249 TIL grouping alone.

250 **Subtype-1 immune changes are associated with DAMPs, acute inflammation, and viral-mimicry**

251 The cascade of adaptive immune responses and activation of dendritic cells is known to be
252 associated with damage-associated molecular patterns (DAMPs), which are molecules released by
253 acute inflammatory, hypoxic, or stressed cells linked to tumor progression or treatment.[27] Hence,
254 we examined a score derived from the average of 15 genes representing various processes
255 involved in DAMPs.[28] Interestingly, the DAMP gene score was significantly ($p<0.05$) higher in

256 Subtype-1 tumors compared with other subtypes (**Figure 4A**). Unlike DAMPs linked to hypoxia and
257 adaptive immunity in our previous study of pancreatic neuroendocrine tumors,[21] DAMPs were not
258 linked to hypoxia in Subtype-1 TNBCs, which had significantly lower hypoxia scores than Subtypes-
259 2 and -3 as assessed using ssGSEA and the hypoxia signature from MSigDB (**Figure 4B**). Hence, we
260 hypothesized that these significant increases in DAMP scores and Th1 responses are associated
261 with acute inflammation[27] in Subtype-1 tumors. As predicted, the acute inflammation gene score
262 (mean expression of *IL15*, *IL21*, and *LTA* in each sample) was significantly ($p<0.05$) higher in
263 Subtype-1 tumors compared with the other subtypes (**Figure 4C**).

264 To further explore if enrichment of DAMP genes is associated with (statistically dependent on)
265 acute inflammation and Subtype-1, we applied our previously described ML approach,
266 probabilistic principal component and covariate analysis (PPCCA),[19] to the Indian TNBC cohort. This
267 method combines probabilistic principal component and multivariate regression analyses to infer
268 statistical associations or dependencies between three parameters (subtypes, acute
269 inflammation, and DAMP gene expression), which cannot be assessed by Pearson or other
270 correlation analyses. For this purpose, we categorized the severity of acute inflammation into high
271 and low based on the median acute inflammation score as a cut-off. Then, we removed the
272 expression differences in 15 DAMP genes between the high and low acute inflammation samples
273 by statistically adjusting or normalizing the expression data using PPCCA. This adjusted data led to
274 the loss of association or dependency of Subtype-1 on DAMP expression in TNBCs ($p<0.05$;
275 modified negative $\log_{10}$ p-value > 0 in **Figure 4D**), as assessed using the first probabilistic principal
276 component (pPC1; with higher proportion of variability) and linear regression analysis with the
277 PPCCA model. Before adjustment of acute inflammation by group, both acute inflammation
278 groups and subtypes were significantly (modified negative $\log_{10}$ p-value > 0 in **Figure 4D**)
279 associated with DAMP expression. This was also reflected as a partial change in DAMP gene
280 expression between Subtype-1 vs. Subtypes-2 and -3 after adjusting the acute inflammation
281 dependency in DAMP genes (**Figure 4F**). Conversely, there was no loss of a significant association
282 between acute inflammation and DAMP gene expression in pPC1 in the reverse analysis of
283 adjusting subtype dependency on DAMP gene expression, followed by regression analysis of acute
284 inflammation groups in pPC1 (**Figure 4E and G**). Moreover, we observed enrichment of host viral-

285 infection mimicry pathways and genes in Subtype-1 (**Figure 4H**), in a similar way to our rectal
286 cancer study[29]. These results demonstrate partial dependency of Subtype-1 on acute
287 inflammation for DAMP enrichment and provide clues that the progression of Subtype-1 is related
288 to viral-mimicry during infection.

289 **Subtype-1-specific pathways and associated immunotherapies**

290 Pathway analysis showed enrichment of Th1, IL2, IL12, T-cell receptor (TCR), B cell receptor (BCR),
291 cytolytic activity, PD-1, NFκB, chemokine/cytokine signaling and other related immune pathways
292 in Subtype-1 tumors using Indian samples (**Figure 5A-B; Supplementary Table 1G-H**). These
293 analyses allowed us to postulate a Subtype-1-specific pathway in $CD4^+$ Th1 lymphocytes (based on
294 **Figures 5A-B** and published studies[30 31]), where MEF2D pathway genes are activated through the
295 TCR and LCK/FYN signaling (**Figure 5C**). In this pathway, $CD4^+$ T-cell co-stimulation, potentially via
296 increased chemokines/cytokines and IL2, IL12/STAT4, and LCK/FYN signaling, further triggers
297 TNFR2-specific NFκB signaling to increase interferon-γ signaling, connecting multiple studies[32-36].
298 To further validate this $CD4^+$ Th1 lymphocyte signaling independently, we used TCGA multi-omics
299 profiles, including RNA-seq-based gene expression and mass spectrometry- and reverse phase
300 protein array (RPPA)-based protein expression, and associated them with our subtypes. These
301 independent analyses showed significant enrichment of all genes and majority of (total and/or
302 phosphorylated) proteins from these representative pathways, suggesting activation of this
303 pathway in Subtype-1 samples (**Figure 5D**).

304 Based on the above data and immune checkpoint gene expression (**Figure 3L**), we hypothesized
305 that Subtype-1 tumors might be amenable to targeting with immune checkpoint inhibitors. Hence,
306 we applied the expanded interferon-γ response gene signature from Ayers, et al. to our data,
307 which has been shown to be associated with responses to anti-PD-1 therapy in multiple cancer
308 types.[37] We detected a significant association between the Ayers, et al., signature and Subtype-1,
309 suggesting that Subtype-1 may respond to anti-PD-1 immunotherapy (**Figure 5E**). Moreover, we
310 applied Subtype-1 scores to Mariathasan *et al.*, gene expression profiles, where melanoma
311 patients were treated with anti-PD-L1 therapy.[38] Despite the difference in cancer types, we
312 observed that melanoma patients with high Subtype-1 scores had a better prognosis and enriched

313 significantly ($p<0.05$) for compete response and stable disease than patients with low scores
314 (**Figure 5F-G**).

315 Similarly, a signature derived from MAGE-A3 vaccination therapy[39] was significantly associated
316 with Subtype-1 tumors (**Figure 5H-I**) and, correspondingly, a majority of samples from melanoma
317 patients responding to MAGE-A3 vaccination were significantly enriched for the Subtype-1
318 signature as assessed using our PAM centroids. These results suggest mechanisms of immune
319 evasion in Subtype-1 breast tumors that may be targeted by anti-checkpoint therapy or MAGE-A3
320 vaccination, which warrants further assessment in TNBCs.

321 **Subtype-2 is associated with Th2-based humoral and Th-17 immunity**

322 In contrast to Subtype-1, the characteristics of Subtype-2 included Th2 and Th17 gene expression,
323 representing a shift towards humoral and Th17 immunity (**Figures 6A-E**). As expected, this Th2
324 response was associated with an increased proportion of $CD4^+$ regulatory cells in Subtype-2
325 tumors (**Figure 6F**). When the tumor cell fraction was evaluated using TCGA data, Subtype-2
326 showed increased tumor purity (enriched for cancer cells rather than stroma) compared with the
327 other two subtypes, opposite to Subtype-1 (**Figure 6G** and **Figure 3E**).

328 Next, we evaluated pathway regulation in Subtype-2 in the Indian cohort. Subtype-2 was enriched
329 for cytokine signaling, specifically via IL17 (Th17) and lectin via the ficolin-based complement and
330 coagulation cascade pathways (**Figure 6H-J, Supplementary Table 1I**). Also, the data suggests a
331 balance tipping towards GATA3-induced Th2 cytokines in this subtype (**Figure 6J**). Overall, the Th2
332 response in Subtype-2 differs from Subtype-1.

333 **Subtype-3 is an immune desert subtype but enriched for innate immune cells**

334 Subtype-3 samples significantly expressed only five subtype-specific immune genes and therefore,
335 represents an immune desert subtype. However, there was high expression of five innate immune
336 genes and cell types, specifically macrophage and neutrophil genes (**Figure 7A-B**). Interestingly,
337 Subtype-3, along with Subtype-2, was relatively hypoxic compared with Subtype-1 (**Figure 4B**). We
338 confirmed the presence of macrophages using the pan-macrophage marker CD68 in subtype-
339 specific samples (n=39), which were indeed significantly more numerous in Subtype-3 samples
340 (**Figure 7C-D**). CD68 positivity was confined to the tumor stroma and not epithelial nests. There

341 was also a strong correlation between Subtype-3 and CD68-positive macrophages, with 31% of
342 the >50% positively stained tumors belonging to Subtype-3 ($p<0.05$). As opposed to Subtype-1,
343 there is an increased chronic inflammation signature in Subtype-3 patients (**Figure 7E**). The
344 Subtype-3 signature was associated with a worse prognosis (borderline significance) in melanoma
345 patients in Mariathasan et al. dataset[38] (**Figure 7F**). The overall summary of all the subtypes is
346 provided in **Figures 7G-H**.

347

348 DISCUSSION

349 Approximately 20-31% of all breast cancers in Indian women are TNBCs[4 5] compared with
350 approximately 15% in Western populations,[2] and the incidence of TNBC is higher in pre-
351 menopausal and young women in India.[5] Ethnic differences in cancer prevalence and type are
352 known to exist.[40] For instance, about 40% of breast cancers in African American women in the US
353 are TNBCs, a much higher prevalence than in the White ethnic groups.[41] A better understanding
354 of ethnicity-related disparities in cancer between Indian and Western populations would be useful
355 for developing diagnosis and management guidelines tailored to specific populations. Given the
356 increased prevalence of TNBC in India, there is a particular need to better understand the
357 molecular characteristics of TNBCs in the Indian context to improve clinical outcomes by
358 identifying potential population-specific biomarkers (if they exist) and personalizing affordable
359 treatment regimens. Suppose the differences between Indian and Western TNBCs are minor,
360 which is the case here, it will make it easy to adapt existing therapies, including potential
361 immunotherapy options tailored towards TNBC, in India.

362 In this study, we investigated the immune landscapes of 88 breast cancer samples from South
363 Indian women. To our knowledge, this is the first comprehensive analysis of immune
364 transcriptome analysis of Indian TNBCs. To avoid bias, we retrospectively analyzed two distinct
365 cohorts of retrospective samples from two different hospitals. Although these samples were
366 collected from one major metropolitan city in South India (Bangalore), the clinical parameters
367 were comparable to other Indian breast cancer studies, in that half of patients were aged 50 or
368 under and over two-fifths were premenopausal.[5 6] The age and menopausal symptom

369 characteristics more closely resembled an African and African-American populations but were
370 distinct from those usually seen in Western populations.[41 42]

371 TNBC is increasingly recognized as a heterogeneous disease with high recurrence and mortality
372 rates that more often displays aggressive clinical features such as high grade and lymph node
373 positivity.[2] Despite international efforts to identify effective therapeutic targets in TNBCs by
374 employing multi-omics strategies including proteomics, genomics, transcriptomics, and
375 methylomics, only a handful of targeted therapies exist or under consideration for this subgroup.[3]
376 Breast cancer was one of the first cancer types to undergo molecular classification.[43] Multiple
377 groups have defined subtypes in Western TNBCs,[7-9] and here, using Indian samples, we define
378 three immune TNBC subtypes with distinct gene and cell type profiles and prognoses (**Figures 7G-**
379 **H**). These immune subtypes defined in an Indian population were also present in Western TNBC
380 samples and in approximately the same proportions. There is an association between our's and
381 Vanderbilt's subtypes, however, our subtypes are specifically defined based on gene expression
382 representing immune cell infiltrations that display different characteristics from Vanderbilt's
383 subtypes. Nevertheless, the observed geographic differences are probably less to do with the
384 intrinsic immune biology of TNBC.

385 Although the incidence of TNBC is higher in young, premenopausal women,[5] most of these women
386 with TNBC belonged to the good prognosis Subtype-1 with active Th1 responses. In contrast, most
387 post-menopausal women with TNBC had poor prognosis, immune desert Subtype-3 tumors. This
388 increased risk of poor prognostic Subtype-3 in this age group may be attributable to compromised
389 immune function, specifically dysregulated adaptive Th1 responses, with aging; the capacity to
390 develop memory T-cells against cancer antigens is reduced in older people, compromising the
391 ability of the immune system to reject cancer cells[44]. Therefore, subsets of pre-menopausal and
392 post-menopausal women with TNBC may require distinct immunotherapies depending on their
393 subtype.

394

395 Here, we have also attempted to understand the mechanistic insights of these immune subtypes
396 associated with different immune infiltrations. The cytolytic activity of $CD8^+$ T-cells seen in

397 Subtype-1 depends on dedicated APCs, i.e., dendritic cells and, consistent with this, there was
398 increased MHC-I expression and antigen processing and HLA genes in these tumors, which leads
399 to co-stimulation of T-cell genes via TCR signaling in CD4 T-helper cells. Using pathway and
400 literature[30-36] analysis, we found that Subtype-1 was enriched for DAMP and CD4 T-helper TCR
401 signaling, and we further validated TCR and related signaling pathways using multi-omic gene and
402 protein data. This pathway may be associated with increased cytolytic activity in $CD8^+$ T-cells,
403 potentially reducing tumor progression and contributing to the better prognosis seen in patients
404 with Subtype-1 TNBCs. There was also increased expression of MHC-II-associated HLA genes
405 representing potentially activation of CD4, dendritic cells and M1 macrophages in Subtype-1
406 tumors, which are known to be associated with a good prognosis in several tumor types.[45] In
407 contrast, Subtype-3, with its low proportion of M0 and M1 macrophages and higher proportion of
408 M2 macrophages (validated using CD68 IHC) and neutrophils, was associated with poorer
409 outcomes. This is consistent with a study by Gentles et al.[46], who found that macrophages,
410 neutrophils, and plasma cells are poor prognostic markers in both breast and lung cancer.

411 Collectively, our findings show that there is considerable immune infiltrate variability in TNBCs,
412 which is partly determined by the molecular characteristics of the primary tumor and that
413 influences clinical outcomes. For example, we found that Subtype-1 may be primarily associated
414 with an acute immune response, as described previously in other cancers[47]. Remarkably, using our
415 PPCCA ML approach[19], we demonstrated a significant multivariate association between acute
416 immunity and subtypes with DAMP gene expression, suggesting increased acute immunity may be
417 associated with or driving Th1-enriched Subtype-1 TNBCs, a hypothesis that requires validation.
418 On the other hand, Subtype-3 may be associated with chronic inflammation and hypoxia, leading
419 to a poorer prognosis. While the roles of Th1 responses in acute inflammation and Th2 and M2
420 macrophages in chronic inflammation during cancer development are well described,[47] here we
421 relate this to differential stratification, prognosis, signaling, and therapeutic responses. Our
422 findings show that there is a complex relationship between acute/chronic immunity, intratumoral
423 immune cell heterogeneity, molecular subtype, and disease progression in TNBC. Treatments that
424 aim to enhance Th1-based cellular immunity against tumors are only effective in a subset of

425 patients[48], and our findings may help to identify this population using our Indian TNBC signature
426 for specific targeting and further therapeutic development.

427 The success of immunotherapy in a subset of TNBC patients has raised hope of efficacy in TNBCs,[10]
428 [14] and TNBCs with dense TILs are expected to have a better prognosis.[49] Nevertheless, the
429 objective response rate (ORR) for immune checkpoint inhibitors (ICI) monotherapy in metastatic
430 TNBC is only modest, and the IMpassion130 trial for locally advanced and metastatic TNBC
431 demonstrated that combined atezolizumab and nab-paclitaxel was more effective than
432 monotherapy with nab-paclitaxel alone.[10] In contrast, a follow-up trial, IMPassion131 showed no
433 improvement in patient survival with the same therapy.[13] Furthermore, several small studies have
434 reported only marginal ORRs with ICIs, with poor responses likely to be multifactorial and include
435 molecular heterogeneity, the use of ICIs as monotherapy, and their use in non-first-line settings.[3]
436 This is likely to be due, at least in part, to overall low PD-L1 expression (biomarker) in TNBCs,[50] so
437 it is essential to develop robust biomarkers to select patients who respond to immunotherapy.

438 Our data suggest Subtype-1 TNBC is potentially targetable with immunotherapy based on cross-
439 cancer analysis using melanoma data. Specifically, Subtype-1 was associated with immune
440 checkpoint inhibitor (ICI) responses and prognosis, as assessed using the anti-PD-1 treatment
441 prediction signature from Ayers et al.[37] and prognosis analysis using an anti-PD-L1 treated
442 melanoma dataset.[39] Our study also suggests an association between MAGEA3 immunotherapy
443 and Subtype-1 tumors. Recently, a phase-III trial assessing MAGEA3 immunotherapy in stage III
444 melanoma patients in the adjuvant setting failed.[51] Although MAGEA3-positive cutaneous
445 melanoma patients were selected for this trial, they were not further selected based on immune
446 landscape such as those defined here. Similarly, a recent study has implicated the role of MAGEA3
447 in hepatocellular carcinoma (HCC) and suggested as a novel therapeutic avenue of targeting
448 MAGEA3 for HCC.[52] In contrast, It is currently uncertain what therapy would best suit patients with
449 Subtype-2 tumors characterized by Th2/Th17 responses, infrequent LN involvement, an
450 intermediate prognosis, and increased tumor purity; further investigation is required to identify
451 potential immunotherapy targets associated with the immune complement cascade and GATA3.
452 The increased wound healing response in Subtype-2 is congruent with the association of Th2
453 response with wound healing[53], which might qualify them for antiangiogenic immunotherapies as

454 described[54]. However, this subtype may benefit from bioengineered immunotherapy using
455 collagen-binding domain (CBD)-IL12 to reactivate Th1 pathways,[55] followed by treatment with an
456 ICI. Subtype-3 was also associated with increased poor prognostic basal-like and HER2-like intrinsic
457 subtypes compared with Subtype-1. Based on our data, Subtype-3 may not respond to first-line
458 ICI-based immunotherapy.

459 In conclusion, here we characterized the immune heterogeneity in TNBCs in Indian women. In
460 doing so, we identified three subtypes with distinct immune cell infiltrates, immune cell signaling,
461 and gene signatures associated with prognosis and responses to immunotherapy. Overall, immune
462 gene expression in Indian and Western TNBCs appears to be largely similar. This immune-
463 transcriptome study suggests that therapies targeting immune microenvironment in Western
464 populations may be as effective in Indian populations, however, genetic changes may need to be
465 considered for geography-specific personalized therapy. This may accelerate pharmaceutical
466 adoption to Indian TNBC patients.

467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483

484 CONTRIBUTIONS

485 AS conceived and developed the idea; AS and AK further developed this idea; AK, JP and SS
486 provided patient samples and clinical follow up data; AS, AK, HPS, CR, YP, and KD performed the
487 experiments; AK and JP performed histopathology and related data analysis; AS, MC and SS
488 supervised the study.

489

490 ACKNOWLEDGEMENTS

491 We thank the Centre for Global Oncology at the ICR for the support.

492 COMPETING INTERESTS

493 None

494 FUNDING

495 We thank Newton Fund and Global Challenge Research Fund (GCRF)-based funding provided
496 through Research England to the ICR, which partly supported this work related to molecular
497 profiling and data analysis. We thank Nadathur Holdings Pvt Ltd and Bagaria Trust whose
498 generous funding supported patient recruitment, clinical follow up and analysis.

499

500

501 Figure Legends

502 **Figure 1. Consolidated Standards of Reporting Trials (CONSORT) diagram and clinical**
503 **characteristics of Indian TNBC samples. A.** Consort diagram showing workflow of samples selected
504 for the study. **B.** A table of clinical characteristics of samples.

505 **Figure 2. Identification and clinical characterization of immune subtypes using the Indian TNBC**
506 **samples and comparison with Western TNBC cohorts. A.** Heatmap of immune subtypes from
507 Indian TNBC (n=88) identified by NMF clustering method. Subtypes are shown on the top bar. The
508 scales are shown at the bottom. **B.** Proportion of immune subtypes from 88 Indian TNBCs. **C-E**.
509 Subtype-specific scores. **F.** Proportion of subtype-specific genes out of 204 genes. **G.** Shannon

510 diversity index of three immune subtypes. **H-I.** Kaplan-Meier curves representing H) OS (n=83) and
511 I) DFS (n=79) from the Indian TNBC cohort. **J.** Multivariate analysis of DFS and other clinical
512 parameters. **K.** Proportion of three immune subtypes in different cohorts of samples from India
513 (n=88) and Western populations – TCGA (n=123) and METABRIC (n=299). **L.** Kaplan-Meier curve
514 representing OS from the METABRIC TNBC cohort (n=299). **M.** Proportion of pre- and post-
515 menopausal samples represented in three immune subtypes from the Indian TNBC cohort (n=81).
516 **N.** Tumor size differences in immune subtypes from the Indian TNBC cohort (n=86). **O.** Number
517 of LNs involved in different immune subtypes from the METABRIC cohort (n=299). **P-Q.** Proportion
518 of P) intrinsic and Q) Vanderbilt subtypes represented in immune subtypes from the METABRIC
519 cohort (n=299). Kruskal-Wallis statistical test was performed for most of the analysis, except for
520 survival analysis and proportion analysis, where the log-rank test was used for survival anlaysis and
521 chi-squared test was used for the proportions. p-value of <0.05 was considered significant.

522 **Figure 3. Immune and clinical characteristics specific to Subtype-1. A.** Proportion of dense,
523 moderate, and mild TILS represented in three immune subtypes from the Indian TNBC cohort
524 (n=84). Chi-squared test was used for p-value calculation. **B-C.** Gene expression of Th1 response
525 genes – B) *IFNG* and C) *IL12A* in immune subtypes from the Indian TNBC cohort (n=88). **D.** Heatmap
526 showing chemokine expression in immune subtypes. **E.** Heatmap validation of specific immune
527 characteristics in immune subtypes from TCGA data. **F.** Schematic representing balance in Th1
528 response and chemokine gene expression in subtypes. **G-L.** Heatmaps and boxplot showing
529 changes in G) CD8 T-cell-specific genes, H) cytolytic activity, I) T and B cell types and activities
530 based on ssGSEA analysis, J) antigen presenting and processing genes, K) MHC-I &II HLA genes, L)
531 T-cell exhaustion genes. **M.** Heatmap showing macrophages, IFNG response, TCR diversity and
532 intratumoral heterogeneity from TCGA data. All the figures used the Indian TNBC cohort (n=88),
533 except figures E) and M), where TCGA data was used. Kruskal-Wallis statistical test was performed
534 for p-value significance for those in boxplots.

535 **Figure 4. Machine-learning based association analysis of acute inflammation, subtypes and DAMP**
536 **gene expression in the Indian cohort. A-C.** Boxplots showing differential changes in DAMP, hypoxia
537 and acute inflammation scores from the Indian TNBC cohort (n=88). Kruskal-Wallis statistical test
538 was performed for p-value significance. **D-E.** Barplots showing significant (p<0.01; linear

539 regression) association of 15 DAMP genes with D) acute inflammation (high vs. low) and E)
540 subtypes (Subtype-1 vs. others) in principal component (PC1) as assessed by PPCCA method using
541 the Indian TNBC cohort (n=88). **F.** Heatmap showing change in 15 DAMP pathway genes (median
542 expression across samples from each subtype) in immune subtypes before and after statistically
543 adjusting acute inflammation using the *PPCCA* method using the Indian TNBC cohort (n=88). **G.**
544 Heatmap showing change in 15 DAMP pathway genes (median expression across samples from
545 each subtype) in acute inflammation low and high groups before and after statistically adjusting
546 subtypes using the *PPCCA* method using the Indian TNBC cohort (n=88). **H.** Dotplot showing
547 enrichment of anti-viral-mimicry pathways in Subtype-1 from hallmarks gene sets from the
548 Molecular Signature DataBase (MSigDB) database using Subtype-1-specific genes. False discovery
549 rate (FDR) was calculated from p-values from hypergeometric test using hypeR R package (see
550 methods).
551
552 **Figure 5. Association of gene enrichment and therapy response to immune TNBC subtypes. A-B.**
553 Barplot showing gene enrichment of A) BIOCARTA and B) Reactome pathways using Enrichr tool
554 (see Methods) in Subtype-1. **C.** Schematic showing TCR and downstream signalling that effects
555 immune cell types based on data from the Indian TNBC gene expression, pathway analysis in A-B)
556 and literature. **D.** Multiomics enrichment analysis validation specific pathways from C) using TCGA
557 TNBC samples (n=18). Kruskal-Wallis statistical test was performed for p-value significance. **E.**
558 Boxplot showing differential T-cell-inflammed GEP in subtypes in the Indian cohort (n=88). **F.**
559 Kaplan-Meier curve and median survival data showing differential OS in melanoma samples (pre-
560 treatment; Mariathasan et al.[38]; n=348) with high and low enrichment of Subtype-1 genes and
561 their association with immunotherapy response. Log-rank test was performed for p-value
562 significance. **G.** Barplot showing the association of immune subtypes with clinical RECIST response
563 to immunotherapy in melanoma (Mariathasan et al.[38]; n=298). * - represents a significant ($p<0.05$;
564 Kruskal-Wallis test) enrichment in samples with high Subtype-1 score compared to those with low
565 score. **H.** Boxplot showing differential MAGEA3 therapy response signature in subtypes in the
566 Indian cohort (n=88). **I.** Proportion of melanoma samples (n=56) showing immune TNBC subtypes
567 with differential MAGEA3 therapy response, as a cross-cancer comparison analysis. Kruskal-Wallis

568 statistical test was performed for p-value significance for E) and H). Chi-squared test was
569 performed for p-value calculations for G) and I).
570
571 **Figure 6. Immune characteristics specific to Subtype-2. A-D.** Gene expression of Th2 and Th17
572 response genes – A) *IL4,* B) *IL5*, C) *IL17B* and D) *IL17A* in immune subtypes using the Indian TNBC
573 cohort (n=88). **E.** Schematic representing balance in Th2 and Th17 response and chemokine gene
574 expression in immune subtypes. **F.** Boxplot showing changes in CD4+ T regulatory cells in immune
575 subtypes as assessed by ssGSEA analysis using the Indian TNBC cohort (n=88). **G.** Boxplot showing
576 changes in tumor purity in immune subtypes from the TCGA data (n=101). **H-I.** Barplot showing
577 enrichment of A) KEGG and B) Reactome pathways using Enrichr tool (see Methods) in subtype-2.
578 **J.** A table showing BIOCARTA enrichment analysis in Subtype-2. Kruskal-Wallis statistical test was
579 performed for p-value significance for A-G).
580 **Figure 7. Immune characteristics and validation of macrophage markers in Subtype-3. A.** Heatmap
581 showing Subtype-3 specific genes associated with macrophages and neutrophils. **B.** Boxplot
582 showing changes in macrophages in immune subtypes as assessed by ssGSEA analysis using the
583 Indian TNBC cohort (n=88). **C-D.** IHC and quantitation (n=39) of pan-macrophage marker – CD68
584 in immune subtype samples using the Indian cohort. **E.** Boxplot showing differential changes in
585 chronic inflammation scores in immune subtypes using the Indian cohort (n=88). **F.** Kaplan-Meier
586 curve and median survival data showing differential OS in samples with high and low enrichment
587 of immune TNBC Subtype-3 genes in melanoma samples (pre-treatment; Mariathasan et al.[38];
588 n=348). **G.** Molecular characteristics of subtypes. **H.** Clinical characteristics of subtypes. Kruskal-
589 Wallis statistical test was performed for p-value significance for A), D) and E). Log-rank test was
590 performed for p-value significance for F).
591
592
593
594
595

596 REFERENCES

597 1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and
598 Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* 2021;71(3):209-49. doi:
599 10.3322/caac.21660 [published Online First: 2021/02/05]
600 2. Boyle P. Triple-negative breast cancer: epidemiological considerations and recommendations. *Ann*
601 *Oncol* 2012;23 Suppl 6:vi7-12. doi: 10.1093/annonc/mds187 [published Online First: 2012/10/04]
602 3. Bianchini G, De Angelis C, Licata L, et al. Treatment landscape of triple-negative breast cancer -
603 expanded options, evolving needs. *Nat Rev Clin Oncol* 2021 doi: 10.1038/s41571-021-00565-2
604 [published Online First: 2021/11/11]
605 4. Sandhu GS, Erqou S, Patterson H, et al. Prevalence of Triple-Negative Breast Cancer in India: Systematic
606 Review and Meta-Analysis. *J Glob Oncol* 2016;2(6):412-21. doi: 10.1200/JGO.2016.005397
607 [published Online First: 2017/07/19]
608 5. Kulkarni A, Kelkar DA, Parikh N, et al. Meta-Analysis of Prevalence of Triple-Negative Breast Cancer and
609 Its Clinical Features at Incidence in Indian Patients With Breast Cancer. *JCO Glob Oncol*
610 2020;6:1052-62. doi: 10.1200/GO.20.00054 [published Online First: 2020/07/09]
611 6. Akhtar M, Dasgupta S, Rangwala M. Triple negative breast cancer: an Indian perspective. *Breast Cancer*
612 *(Dove Med Press)* 2015;7:239-43. doi: 10.2147/BCTT.S85442 [published Online First: 2015/09/01]
613 7. Lehmann BD, Bauer JA, Chen X, et al. Identification of human triple-negative breast cancer subtypes
614 and preclinical models for selection of targeted therapies. *J Clin Invest* 2011;121(7):2750-67. doi:
615 10.1172/JCI45014 [published Online First: 2011/06/03]
616 8. Burstein MD, Tsimelzon A, Poage GM, et al. Comprehensive genomic analysis identifies novel subtypes
617 and targets of triple-negative breast cancer. *Clin Cancer Res* 2015;21(7):1688-98. doi:
618 10.1158/1078-0432.CCR-14-0432 [published Online First: 2014/09/12]
619 9. Jezequel P, Loussouarn D, Guerin-Charbonnel C, et al. Gene-expression molecular subtyping of triple-
620 negative breast cancer tumours: importance of immune response. *Breast Cancer Res* 2015;17:43.
621 doi: 10.1186/s13058-015-0550-y [published Online First: 2015/04/19]
622 10. Schmid P, Adams S, Rugo HS, et al. Atezolizumab and Nab-Paclitaxel in Advanced Triple-Negative
623 Breast Cancer. *N Engl J Med* 2018;379(22):2108-21. doi: 10.1056/NEJMoa1809615 [published
624 Online First: 2018/10/23]
625 11. Roche's Atezolizumab receives DCGI approval for treatment of metastatic triple-negative breast
626 cancer in India 2020 [Available from: https://www.rocheindia.com/en/media/Roche-metastatic-
627 triple-negative-breast-cancer-in-India.html.
628 12. Correction: Society for Immunotherapy of Cancer (SITC) clinical practice guideline on
629 immunotherapy for the treatment of breast cancer. *Journal for ImmunoTherapy of Cancer*
630 2021;9(12):e002597corr1. doi: 10.1136/jitc-2021-002597corr1
631 13. Miles D, Gligorov J, Andre F, et al. Primary results from IMpassion131, a double-blind, placebo-
632 controlled, randomised phase III trial of first-line paclitaxel with or without atezolizumab for
633 unresectable locally advanced/metastatic triple-negative breast cancer. *Ann Oncol*
634 2021;32(8):994-1004. doi: 10.1016/j.annonc.2021.05.801 [published Online First: 2021/07/06]
635 14. Schmid P, Cortes J, Pusztai L, et al. Pembrolizumab for Early Triple-Negative Breast Cancer. *N Engl J*
636 *Med* 2020;382(9):810-21. doi: 10.1056/NEJMoa1910549 [published Online First: 2020/02/27]
637 15. Prabhu JS, Korlimarla A, Desai K, et al. A Majority of Low (1-10%) ER Positive Breast Cancers Behave
638 Like Hormone Receptor Negative Tumors. *Journal of Cancer* 2014;5(2):156-65. doi:
639 10.7150/jca.7668 [published Online First: 2014/02/25]
640 16. Korlimarla A, Prabhu JS, Anupama CE, et al. Separate quality-control measures are necessary for
641 estimation of RNA and methylated DNA from formalin-fixed, paraffin-embedded specimens by

642 quantitative PCR. *J Mol Diagn* 2014;16(2):253-60. doi: 10.1016/j.jmoldx.2013.11.003 [published
643 Online First: 2014/01/15]
644 17. Ragulan C, Eason K, Fontana E, et al. Analytical Validation of Multiplex Biomarker Assay to Stratify
645 Colorectal Cancer into Molecular Subtypes. *Sci Rep* 2019;9(1):7665. doi: 10.1038/s41598-019-
646 43492-0 [published Online First: 2019/05/23]
647 18. Sadanandam A, Lyssiotis CA, Homicsko K, et al. A colorectal cancer classification system that associates
648 cellular phenotype and responses to therapy. *Nature Medicine* 2013;19:619-25. doi:
649 10.1038/nm.3175
650 19. Nyamundanda G, Poudel P, Patil Y, et al. A Novel Statistical Method to Diagnose, Quantify and Correct
651 Batch Effects in Genomic Studies. *Sci Rep* 2017;7(1):10849. doi: 10.1038/s41598-017-11110-6
652 [published Online First: 2017/09/09]
653 20. Salgado R, Denkert C, Demaria S, et al. The evaluation of tumor-infiltrating lymphocytes (TILs) in breast
654 cancer: recommendations by an International TILs Working Group 2014. *Ann Oncol*
655 2015;26(2):259-71. doi: 10.1093/annonc/mdu450 [published Online First: 2014/09/13]
656 21. Young K, Lawlor RT, Ragulan C, et al. Immune landscape, evolution, hypoxia-mediated viral mimicry
657 pathways and therapeutic potential in molecular subtypes of pancreatic neuroendocrine
658 tumours. *Gut* 2020 doi: 10.1136/gutjnl-2020-321016 [published Online First: 2020/09/05]
659 22. Ciriello G, Gatza ML, Beck AH, et al. Comprehensive Molecular Portraits of Invasive Lobular Breast
660 Cancer. *Cell* 2015;163(2):506-19. doi: 10.1016/j.cell.2015.09.033 [published Online First:
661 2015/10/10]
662 23. Curtis C, Shah SP, Chin SF, et al. The genomic and transcriptomic architecture of 2,000 breast tumours
663 reveals novel subgroups. *Nature* 2012;486(7403):346-52. doi: 10.1038/nature10983 [published
664 Online First: 2012/04/24]
665 24. Thorsson V, Gibbs DL, Brown SD, et al. The Immune Landscape of Cancer. *Immunity* 2018;48(4):812-
666 30 e14. doi: 10.1016/j.immuni.2018.03.023 [published Online First: 2018/04/10]
667 25. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based
668 approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy*
669 *of Sciences* 2005;102:15545-50.
670 26. Jiang W, He Y, He W, et al. Exhausted CD8+T Cells in the Tumor Immune Microenvironment: New
671 Pathways to Therapy. *Front Immunol* 2020;11:622509. doi: 10.3389/fimmu.2020.622509
672 [published Online First: 2021/02/27]
673 27. Hernandez C, Huebener P, Schwabe RF. Damage-associated molecular patterns in cancer: a double-
674 edged sword. *Oncogene* 2016;35(46):5931-41. doi: 10.1038/onc.2016.104 [published Online First:
675 2016/04/19]
676 28. Tang D, Kang R, Coyne CB, et al. PAMPs and DAMPs: signal 0s that spur autophagy and immunity.
677 *Immunol Rev* 2012;249(1):158-75. doi: 10.1111/j.1600-065X.2012.01146.x [published Online
678 First: 2012/08/15]
679 29. Wilkins A, Fontana E, Nyamundanda G, et al. Differential and longitudinal immune gene patterns
680 associated with reprogrammed microenvironment and viral mimicry in response to neoadjuvant
681 radiotherapy in rectal cancer. *J Immunother Cancer* 2021;9(3) doi: 10.1136/jitc-2020-001717
682 [published Online First: 2021/03/09]
683 30. Potthoff MJ, Olson EN. MEF2: a central regulator of diverse developmental programs. *Development*
684 2007;134(23):4131-40. doi: 10.1242/dev.008367 [published Online First: 2007/10/26]
685 31. Palacios EH, Weiss A. Function of the Src-family kinases, Lck and Fyn, in T-cell development and
686 activation. *Oncogene* 2004;23(48):7990-8000. doi: 10.1038/sj.onc.1208074 [published Online
687 First: 2004/10/19]

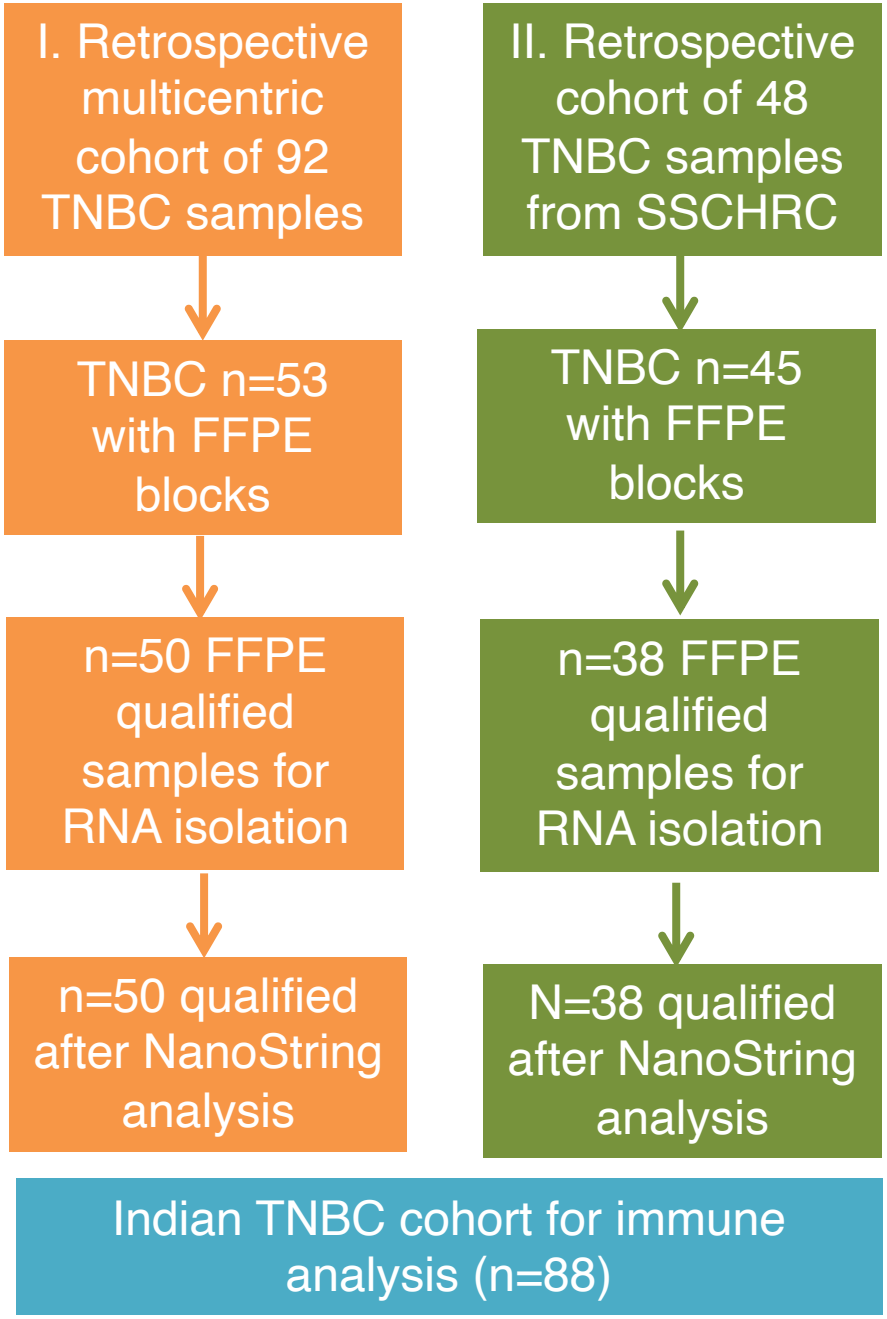
688 32. Goplen NP, Saxena V, Knudson KM, et al. IL-12 Signals through the TCR To Support CD8 Innate Immune
689 Responses. *J Immunol* 2016;197(6):2434-43. doi: 10.4049/jimmunol.1600037 [published Online
690 First: 2016/08/16]
691 33. Miller PG, Bonn MB, McKarns SC. Transmembrane TNF-TNFR2 Impairs Th17 Differentiation by
692 Promoting Il2 Expression. *J Immunol* 2015;195(6):2633-47. doi: 10.4049/jimmunol.1500286
693 [published Online First: 2015/08/14]
694 34. Ridge JP, Di Rosa F, Matzinger P. A conditioned dendritic cell can be a temporal bridge between a CD4+
695 T-helper and a T-killer cell. *Nature* 1998;393(6684):474-8. doi: 10.1038/30989 [published Online
696 First: 1998/06/12]
697 35. Eisel D, Das K, Dickes E, et al. Cognate Interaction With CD4(+) T Cells Instructs Tumor-Associated
698 Macrophages to Acquire M1-Like Phenotype. *Front Immunol* 2019;10:219. doi:
699 10.3389/fimmu.2019.00219 [published Online First: 2019/03/12]
700 36. Hagn M, Sontheimer K, Dahlke K, et al. Human B cells differentiate into granzyme B-secreting cytotoxic
701 B lymphocytes upon incomplete T-cell help. *Immunol Cell Biol* 2012;90(4):457-67. doi:
702 10.1038/icb.2011.64 [published Online First: 2011/08/03]
703 37. Ayers M, Lunceford J, Nebozhyn M, et al. IFN-gamma-related mRNA profile predicts clinical response
704 to PD-1 blockade. *J Clin Invest* 2017;127(8):2930-40. doi: 10.1172/JCI91190 [published Online
705 First: 2017/06/27]
706 38. Mariathasan S, Turley SJ, Nickles D, et al. TGFbeta attenuates tumour response to PD-L1 blockade by
707 contributing to exclusion of T cells. *Nature* 2018;554(7693):544-48. doi: 10.1038/nature25501
708 [published Online First: 2018/02/15]
709 39. Ulloa-Montoya F, Louahed J, Dizier B, et al. Predictive gene signature in MAGE-A3 antigen-specific
710 cancer immunotherapy. *J Clin Oncol* 2013;31(19):2388-95. doi: 10.1200/JCO.2012.44.3762
711 [published Online First: 2013/05/30]
712 40. Cho B, Han Y, Lian M, et al. Evaluation of Racial/Ethnic Differences in Treatment and Mortality Among
713 Women With Triple-Negative Breast Cancer. *JAMA Oncol* 2021;7(7):1016-23. doi:
714 10.1001/jamaoncol.2021.1254 [published Online First: 2021/05/14]
715 41. Dietze EC, Sistrunk C, Miranda-Carboni G, et al. Triple-negative breast cancer in African-American
716 women: disparities versus biology. *Nat Rev Cancer* 2015;15(4):248-54. doi: 10.1038/nrc3896
717 [published Online First: 2015/02/13]
718 42. Prakash O, Hossain F, Danos D, et al. Racial Disparities in Triple Negative Breast Cancer: A Review of
719 the Role of Biologic and Non-biologic Factors. *Front Public Health* 2020;8:576964. doi:
720 10.3389/fpubh.2020.576964 [published Online First: 2021/01/09]
721 43. Perou CM, Sørlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature*
722 2000;406:747-52. doi: 10.1038/35021093
723 44. Derhovanessian E, Solana R, Larbi A, et al. Immunity, ageing and cancer. *Immun Ageing* 2008;5:11. doi:
724 10.1186/1742-4933-5-11 [published Online First: 2008/09/26]
725 45. Axelrod ML, Cook RS, Johnson DB, et al. Biological Consequences of MHC-II Expression by Tumor Cells
726 in Cancer. *Clin Cancer Res* 2019;25(8):2392-402. doi: 10.1158/1078-0432.CCR-18-3200 [published
727 Online First: 2018/11/23]
728 46. Gentles AJ, Newman AM, Liu CL, et al. The prognostic landscape of genes and infiltrating immune cells
729 across human cancers. *Nature Medicine* 2015;21:938-45. doi: 10.1038/nm.3909
730 47. Grivennikov SI, Greten FR, Karin M. Immunity, inflammation, and cancer. *Cell* 2010;140(6):883-99. doi:
731 10.1016/j.cell.2010.01.025 [published Online First: 2010/03/23]
732 48. Xu HM. Th1 cytokine-based immunotherapy for cancer. *Hepatobiliary Pancreat Dis Int* 2014;13(5):482-
733 94. doi: 10.1016/s1499-3872(14)60305-2 [published Online First: 2014/10/14]
734 49. Dieci MV, Criscitiello C, Goubar A, et al. Prognostic value of tumor-infiltrating lymphocytes on residual
735 disease after primary chemotherapy for triple-negative breast cancer: a retrospective multicenter

736 study. *Ann Oncol* 2014;25(3):611-18. doi: 10.1093/annonc/mdt556 [published Online First:
737 2014/01/10]
738 50. Bai X, Ni J, Beretov J, et al. Immunotherapy for triple-negative breast cancer: A molecular insight into
739 the microenvironment, treatment, and resistance. *Journal of the National Cancer Center*
740 2021;1(3):75-87. doi: https://doi.org/10.1016/j.jncc.2021.06.001
741 51. Dreno B, Thompson JF, Smithers BM, et al. MAGE-A3 immunotherapeutic as adjuvant therapy for
742 patients with resected, MAGE-A3-positive, stage III melanoma (DERMA): a double-blind,
743 randomised, placebo-controlled, phase 3 trial. *Lancet Oncol* 2018;19(7):916-29. doi:
744 10.1016/S1470-2045(18)30254-7 [published Online First: 2018/06/18]
745 52. Craig AJ, Garcia-Lezana T, Ruiz de Galarreta M, et al. Transcriptomic characterization of cancer-testis
746 antigens identifies MAGEA3 as a driver of tumor progression in hepatocellular carcinoma. *PLoS*
747 *Genet* 2021;17(6):e1009589. doi: 10.1371/journal.pgen.1009589 [published Online First:
748 2021/06/25]
749 53. Gause WC, Wynn TA, Allen JE. Type 2 immunity and wound healing: evolutionary refinement of
750 adaptive immunity by helminths. *Nat Rev Immunol* 2013;13(8):607-14. doi: 10.1038/nri3476
751 [published Online First: 2013/07/06]
752 54. Hua Y, Bergers G. Tumors vs. Chronic Wounds: An Immune Cell's Perspective. *Front Immunol*
753 2019;10:2178. doi: 10.3389/fimmu.2019.02178 [published Online First: 2019/10/02]
754 55. Ishihara J, Ishihara A, Sasaki K, et al. Targeted antibody and cytokine cancer immunotherapies through
755 collagen affinity. *Sci Transl Med* 2019;11(487) doi: 10.1126/scitranslmed.aau3259 [published
756 Online First: 2019/04/12]

757

Figure 1

A



B

Clinical Table(N=88)		n(%)
Age(yrs)	Mean	50
	Median	51
	≤50	44(50)
	>50	44(50)
Menopausal status	Pre	39(44)
	Post	49(56)
Grade	2	17(19.3)
	3	68(77.3)
	NA	3(3.4)
Tumor Size	<2 cm	7(8)
	2-5 cm	68(77.2)
	>5 cm	11(12.5)
	NA	2(2.3)
Stage	I	9(10)
	II	59(67)
	III	20(23)
Lymph node status	Positive	33(38)
	Negative	53(60)
	NA	2(2)
Lymphocyte infiltrate groups	Dense	25(28)
	Mild	37(42)
	Moderate	22(25)
	NA	4(5)

Figure 2

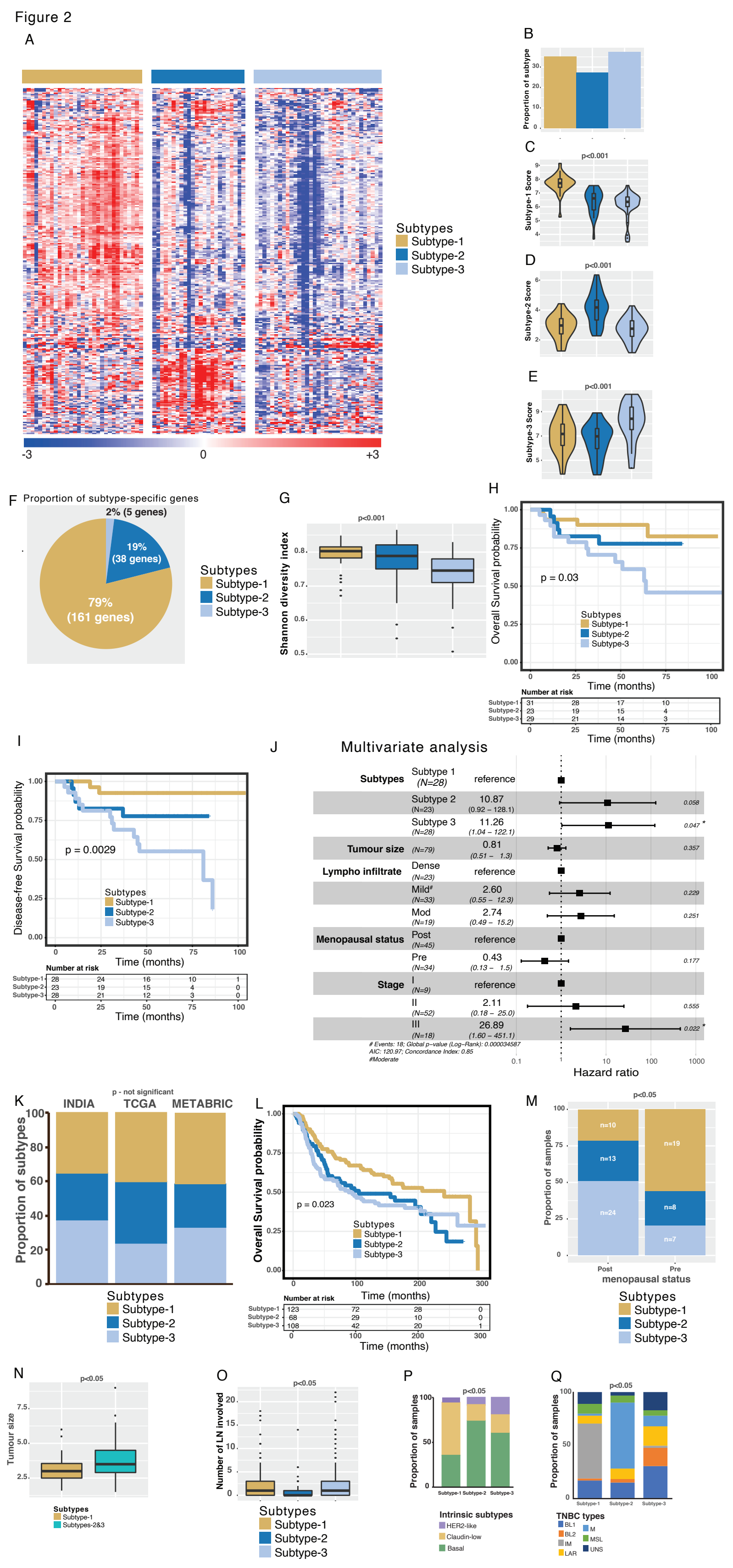

A
-3
0
+3
Subtypes
Subtype-1
Subtype-2
Subtype-3
B
Proportion of subtype
C
p<0.001
Subtype-1 Score
D
p<0.001
Subtype-2 Score
E
p<0.001
Subtype-3 Score
F
Proportion of subtype-specific genes
2% (5 genes)
19% (38 genes)
79% (161 genes)
Subtypes
Subtype-1
Subtype-2
Subtype-3
G
p<0.001
Shannon diversity index
H
Overall Survival probability
p = 0.03
Subtypes
Subtype-1
Subtype-2
Subtype-3
Time (months)
Number at risk
Subtype-1 31 28 17 10
Subtype-2 23 19 15 4
Subtype-3 29 21 14 3
I
Disease-free Survival probability
p = 0.0029
Subtypes
Subtype-1
Subtype-2
Subtype-3
Time (months)
Number at risk
Subtype-1 28 24 16 10 1
Subtype-2 23 19 15 4 0
Subtype-3 28 21 12 3 0
J
Multivariate analysis
Subtypes
Subtype 1 (N=28) reference
Subtype 2 (N=23) 10.87 (0.92 – 128.1) 0.058
Subtype 3 (N=28) 11.26 (1.04 – 122.1) 0.047 *
Tumour size (N=79) 0.81 (0.51 – 1.3) 0.357
Lympho infiltrate
Dense (N=23) reference
Mild# (N=33) 2.60 (0.55 – 12.3) 0.229
Mod (N=19) 2.74 (0.49 – 15.2) 0.251
Menopausal status
Post (N=45) reference
Pre (N=34) 0.43 (0.13 – 1.5) 0.177
Stage
I (N=9) reference
II (N=52) 2.11 (0.18 – 25.0) 0.555
III (N=18) 26.89 (1.60 – 451.1) 0.022 *
Events: 18; Global p–value (Log–Rank): 0.000034587
AIC: 120.97; Concordance Index: 0.85
#Moderate
Hazard ratio
K
p - not significant
INDIA
TCGA
METABRIC
Proportion of subtypes
Subtypes
Subtype-1
Subtype-2
Subtype-3
L
Overall Survival probability
p = 0.023
Subtypes
Subtype-1
Subtype-2
Subtype-3
Time (months)
Number at risk
Subtype-1 123 72 28 0
Subtype-2 68 29 10 0
Subtype-3 108 42 20 1
M
p<0.05
Proportion of samples
n=10
n=13
n=24
n=19
n=8
n=7
Post
Pre
menopausal status
Subtypes
Subtype-1
Subtype-2
Subtype-3
N
p<0.05
Tumour size
Subtypes
Subtype-1
Subtypes-2&3
O
p<0.05
Number of LN involved
Subtypes
Subtype-1
Subtype-2
Subtype-3
P
p<0.05
Proportion of samples
Subtype-1
Subtype-2
Subtype-3
Intrinsic subtypes
HER2-like
Claudin-low
Basal
Q
p<0.05
Proportion of samples
Subtype-1
Subtype-2
Subtype-3
TNBC types
BL1
BL2
IM
LAR
M
MSL
UNS

Figure 3

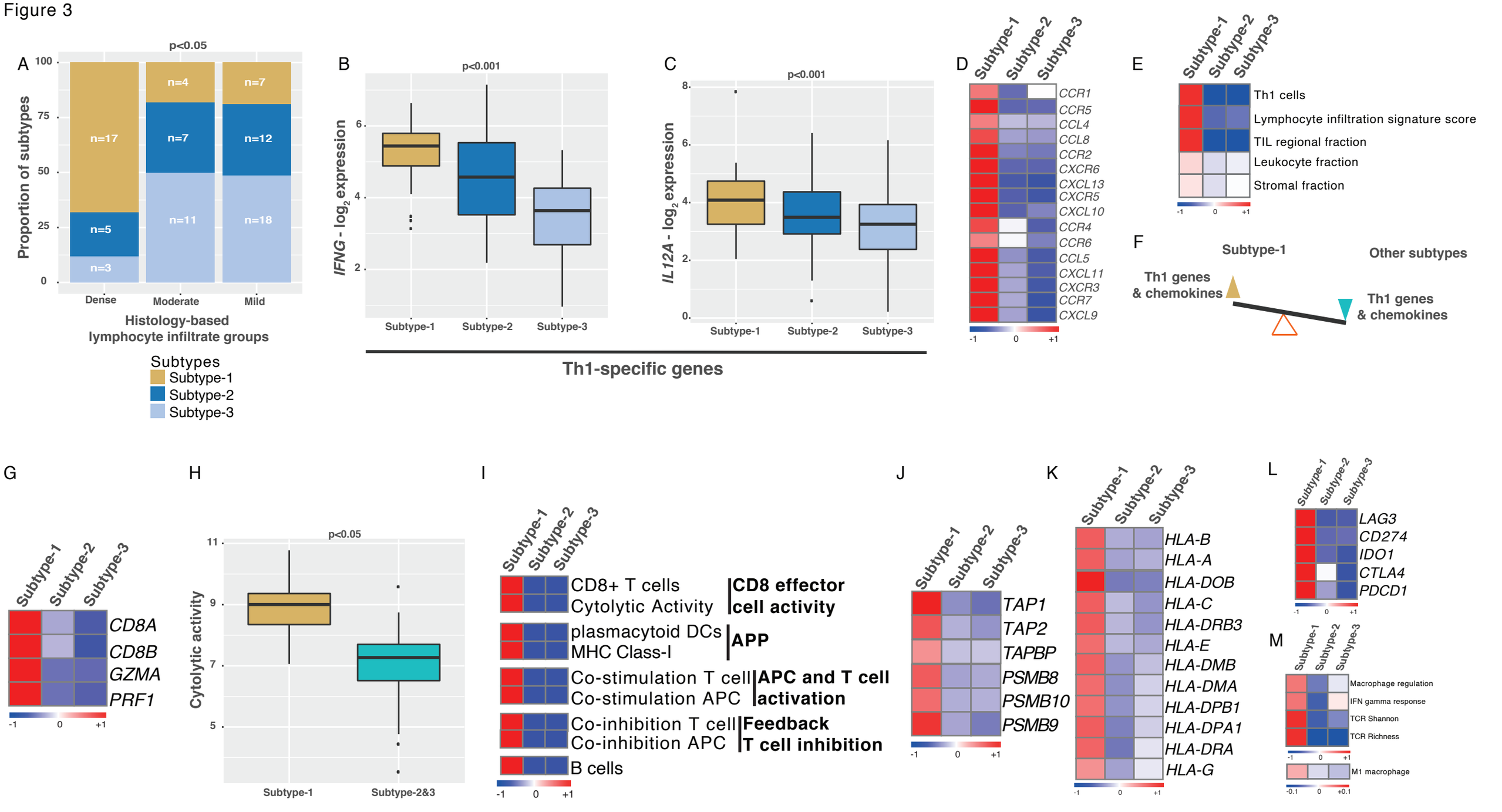
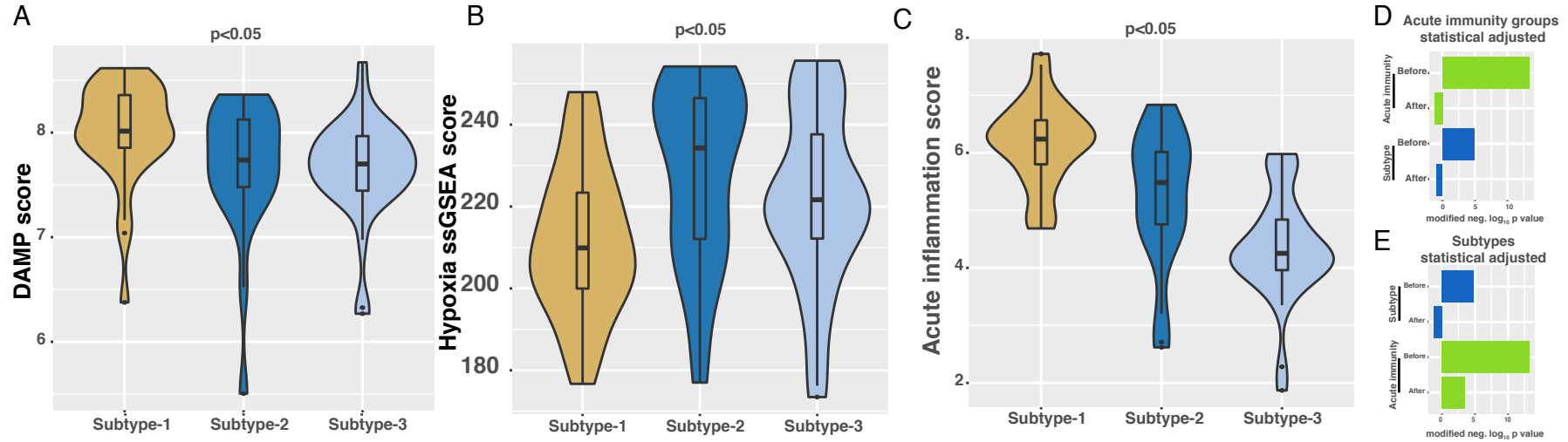

A
p<0.05
Proportion of subtypes
100
75
50
25
0
n=17
n=5
n=3
n=4
n=7
n=11
n=7
n=12
n=18
Dense
Moderate
Mild
Histology-based
lymphocyte infiltrate groups
Subtypes
Subtype-1
Subtype-2
Subtype-3
B
p<0.001
IFNG - log2 expression
Subtype-1
Subtype-2
Subtype-3
C
p<0.001
IL12A - log2 expression
Subtype-1
Subtype-2
Subtype-3
Th1-specific genes
D
Subtype-1
Subtype-2
Subtype-3
CCR1
CCR5
CCL4
CCL8
CCR2
CXCR6
CXCL13
CXCR5
CXCL10
CCR4
CCR6
CCL5
CXCL11
CXCR3
CCR7
CXCL9
-1
0
+1
E
Subtype-1
Subtype-2
Subtype-3
Th1 cells
Lymphocyte infiltration signature score
TIL regional fraction
Leukocyte fraction
Stromal fraction
-1
0
+1
F
Subtype-1
Other subtypes
Th1 genes
& chemokines
Th1 genes
& chemokines
G
Subtype-1
Subtype-2
Subtype-3
CD8A
CD8B
GZMA
PRF1
-1
0
+1
H
p<0.05
Cytolytic activity
11
9
7
5
Subtype-1
Subtype-2&3
I
Subtype-1
Subtype-2
Subtype-3
CD8+ T cells
Cytolytic Activity
CD8 effector
cell activity
plasmacytoid DCs
MHC Class-I
APP
Co-stimulation T cell
Co-stimulation APC
APC and T cell
activation
Co-inhibition T cell
Co-inhibition APC
Feedback
T cell inhibition
B cells
-1
0
+1
J
Subtype-1
Subtype-2
Subtype-3
TAP1
TAP2
TAPBP
PSMB8
PSMB10
PSMB9
-1
0
+1
K
Subtype-1
Subtype-2
Subtype-3
HLA-B
HLA-A
HLA-DOB
HLA-C
HLA-DRB3
HLA-E
HLA-DMB
HLA-DMA
HLA-DPB1
HLA-DPA1
HLA-DRA
HLA-G
-1
0
+1
L
Subtype-1
Subtype-2
Subtype-3
LAG3
CD274
IDO1
CTLA4
PDCD1
-1
0
+1
M
Subtype-1
Subtype-2
Subtype-3
Macrophage regulation
IFN gamma response
TCR Shannon
TCR Richness
-1
0
+1
M1 macrophage
-0.1
0
+0.1

Figure 4



Machine learning-based association analysis

Figure 3

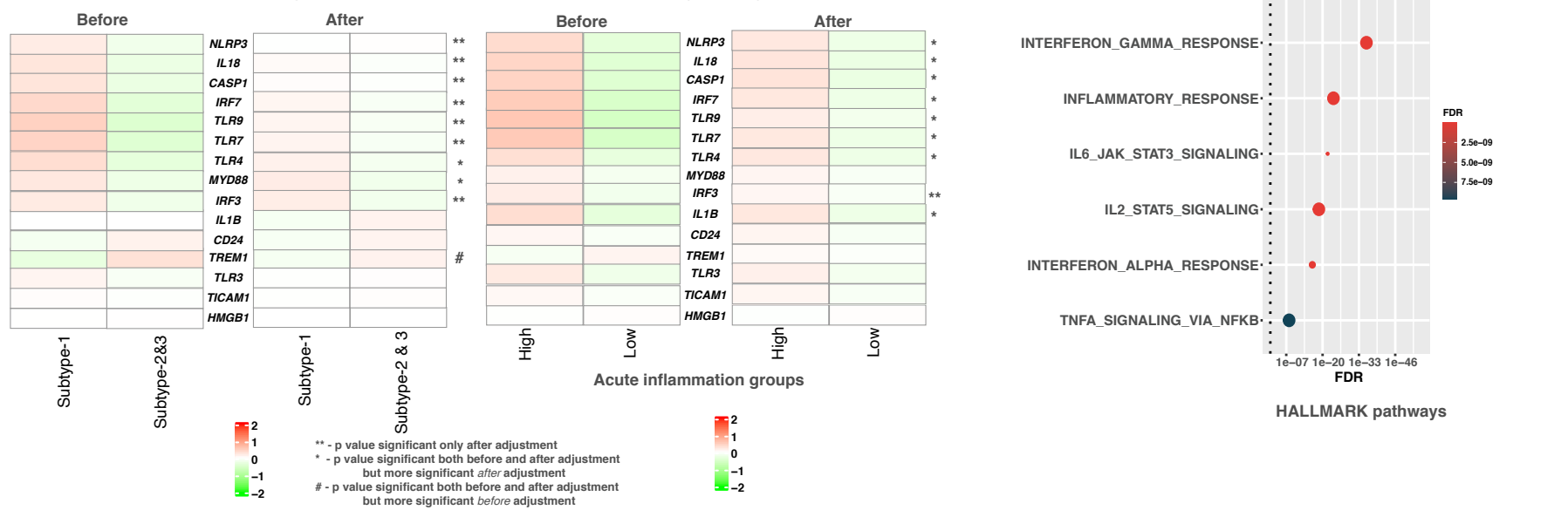


Figure 5

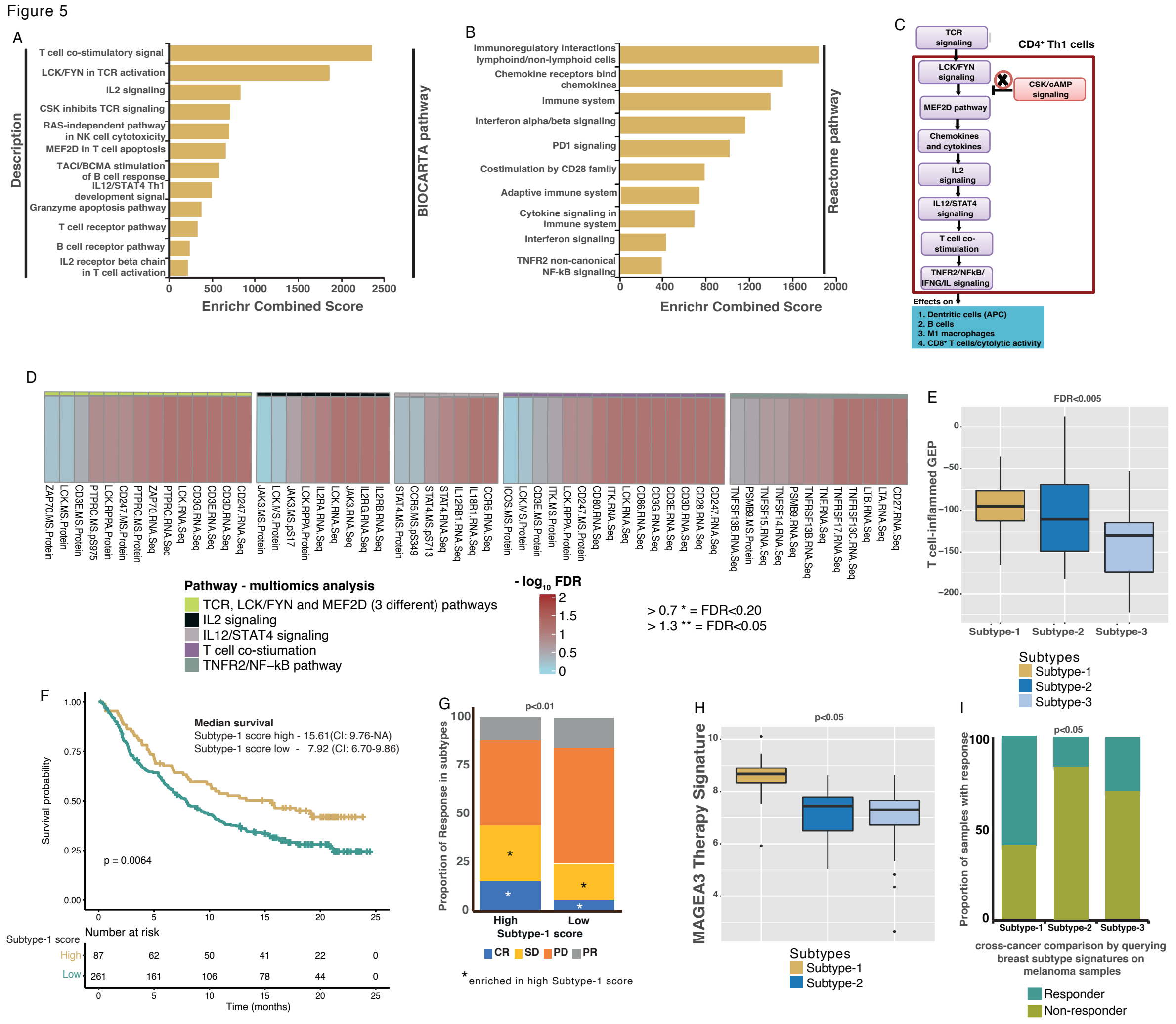
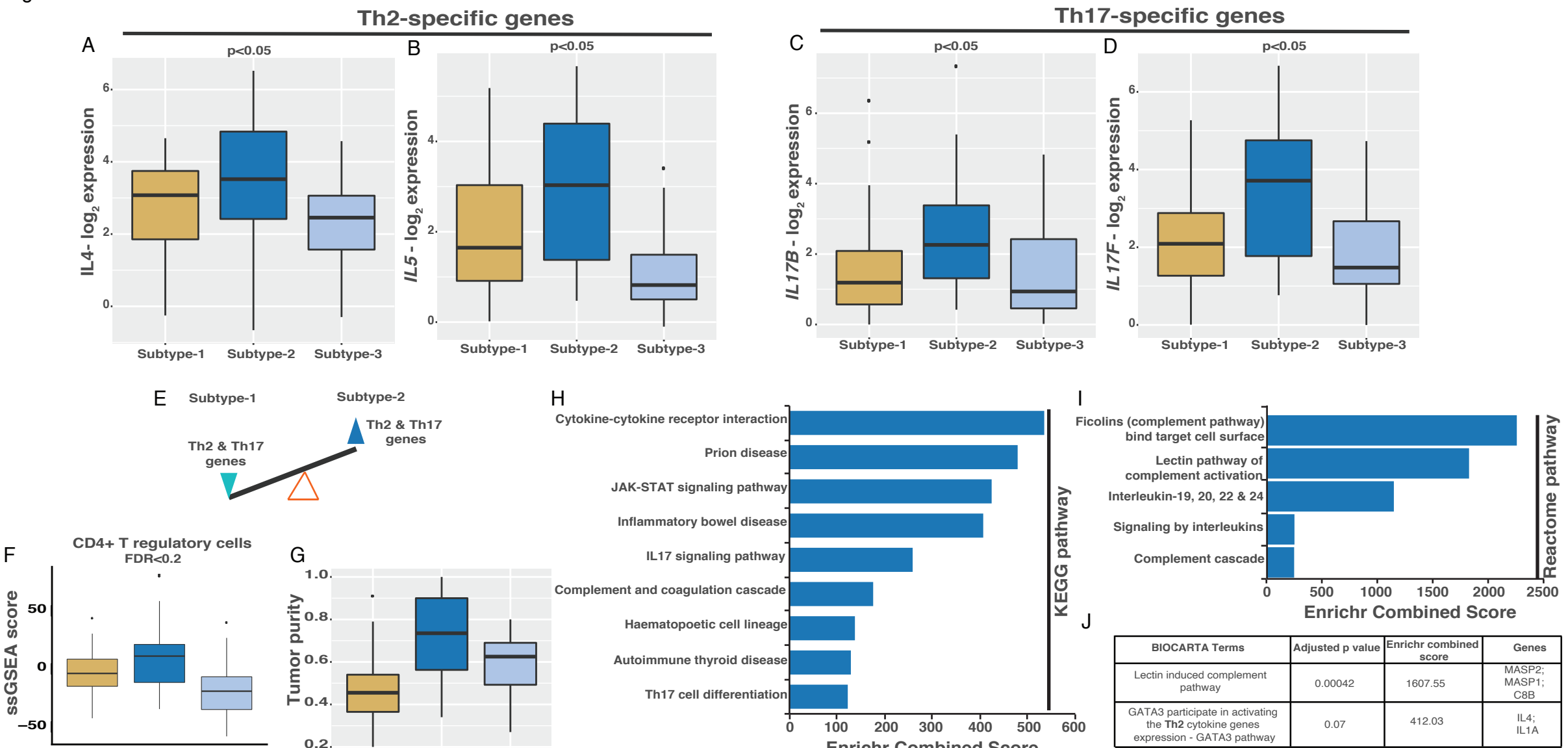
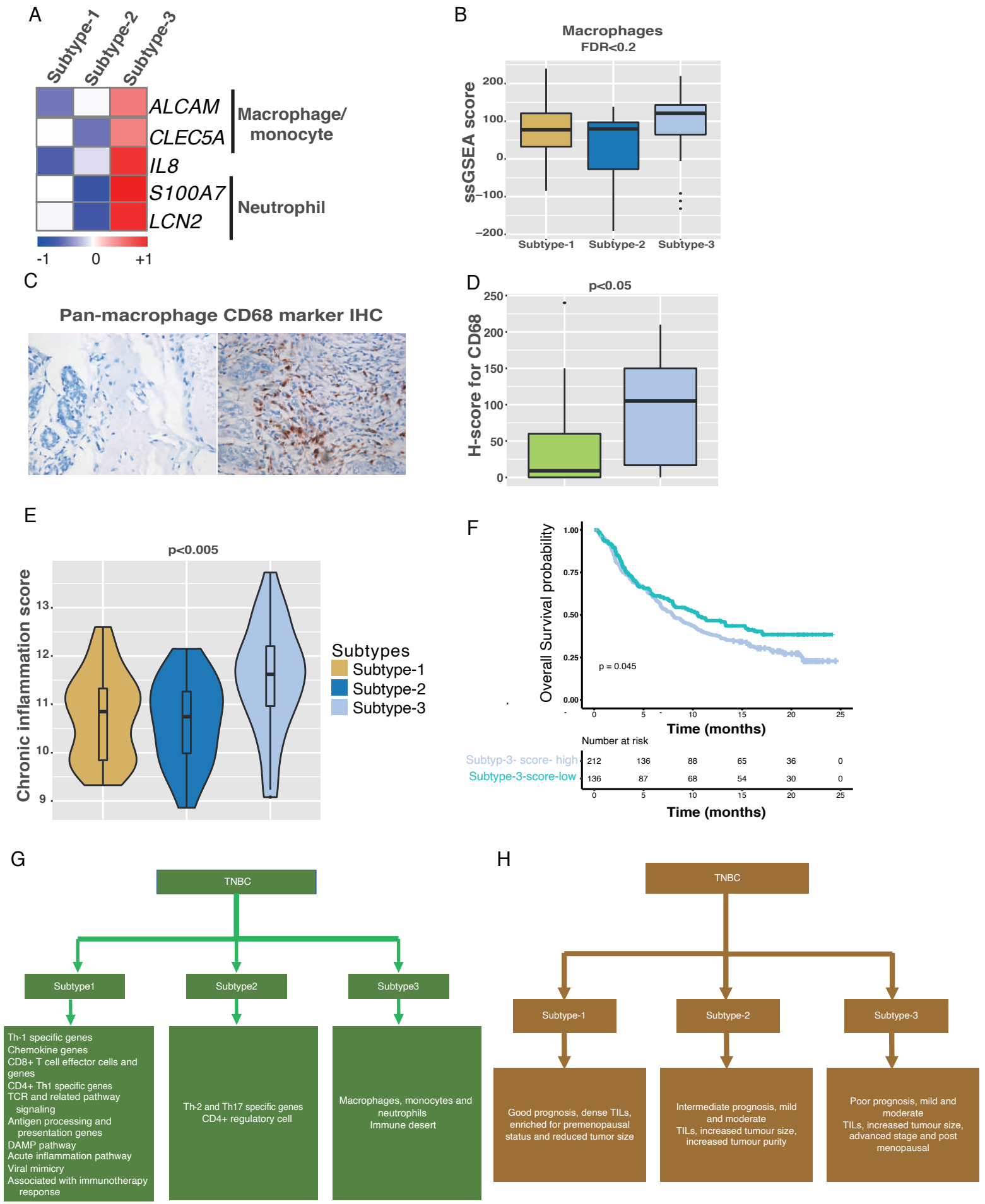

A
Description
T cell co-stimulatory signal
LCK/FYN in TCR activation
IL2 signaling
CSK inhibits TCR signaling
RAS-independent pathway in NK cell cytotoxicity
MEF2D in T cell apoptosis
TACI/BCMA stimulation of B cell response
IL12/STAT4 Th1 development signal
Granzyme apoptosis pathway
T cell receptor pathway
B cell receptor pathway
IL2 receptor beta chain in T cell activation
0 500 1000 1500 2000 2500
Enrichr Combined Score
BIOCARTA pathway
B
Immunoregulatory interactions lymphoind/non-lymphoid cells
Chemokine receptors bind chemokines
Immune system
Interferon alpha/beta signaling
PD1 signaling
Costimulation by CD28 family
Adaptive immune system
Cytokine signaling in immune system
Interferon signaling
TNFR2 non-canonical NF-kB signaling
0 400 800 1200 1600 2000
Enrichr Combined Score
Reactome pathway
C
CD4+ Th1 cells
TCR signaling
LCK/FYN signaling
CSK/cAMP signaling
MEF2D pathway
Chemokines and cytokines
IL2 signaling
IL12/STAT4 signaling
T cell co-stimulation
TNFR2/NFkB/IFNG/IL signaling
Effects on
1. Dentritic cells (APC)
2. B cells
3. M1 macrophages
4. CD8+ T cells/cytolytic activity
D
ZAP70.MS.Protein
LCK.MS.Protein
CD3E.MS.Protein
PTPRC.MS.pS975
LCK.RPPA.Protein
CD247.MS.Protein
PTPRC.MS.Protein
ZAP70.RNA.Seq
PTPRC.RNA.Seq
LCK.RNA.Seq
CD3G.RNA.Seq
CD3E.RNA.Seq
CD3D.RNA.Seq
CD247.RNA.Seq
JAK3.MS.Protein
LCK.MS.Protein
JAK3.MS.pS17
JAK3.RNA.Seq
LCK.RPPA.Protein
LCK.RNA.Seq
IL2RA.RNA.Seq
JAK3.RNA.Seq
IL2RG.RNA.Seq
IL2RB.RNA.Seq
STAT4.MS.Protein
CCR5.MS.pS349
STAT4.MS.pS713
STAT4.RNA.Seq
IL12RB1.RNA.Seq
IL18R1.RNA.Seq
CCR5.RNA.Seq
ICOS.MS.Protein
LCK.MS.Protein
CD3E.MS.Protein
ITK.MS.Protein
LCK.RPPA.Protein
CD247.MS.Protein
CD80.RNA.Seq
ITK.RNA.Seq
LCK.RNA.Seq
CD86.RNA.Seq
CD3G.RNA.Seq
CD3E.RNA.Seq
CD3D.RNA.Seq
CD28.RNA.Seq
CD247.RNA.Seq
TNFSF13B.RNA.Seq
PSMB9.MS.Protein
TNFSF15.RNA.Seq
TNFSF14.RNA.Seq
PSMB9.RNA.Seq
TNFRSF13B.RNA.Seq
TNF.RNA.Seq
TNFRSF17.RNA.Seq
TNFRSF13C.RNA.Seq
LTB.RNA.Seq
LTA.RNA.Seq
CD27.RNA.Seq
Pathway - multiomics analysis
TCR, LCK/FYN and MEF2D (3 different) pathways
IL2 signaling
IL12/STAT4 signaling
T cell co-stiumation
TNFR2/NF–kB pathway
- log10 FDR
2
1.5
1
0.5
0
> 0.7 * = FDR<0.20
> 1.3 ** = FDR<0.05
E
FDR<0.005
T cell-inflammed GEP
0
−50
−100
−150
−200
Subtype-1 Subtype-2 Subtype-3
Subtypes
Subtype-1
Subtype-2
Subtype-3
F
Median survival
Subtype-1 score high - 15.61(CI: 9.76-NA)
Subtype-1 score low - 7.92 (CI: 6.70-9.86)
Survival probability
1.00
0.75
0.50
0.25
0.00
p = 0.0064
0 5 10 15 20 25
Number at risk
Subtype-1 score
High 87 62 50 41 22 0
Low 261 161 106 78 44 0
Time (months)
G
p<0.01
Proportion of Response in subtypes
100
75
50
25
0
High
Low
Subtype-1 score
CR SD PD PR
* enriched in high Subtype-1 score
H
p<0.05
MAGEA3 Therapy Signature
10
8
6
4
Subtypes
Subtype-1
Subtype-2
I
p<0.05
Proportion of samples with response
100
50
0
Subtype-1 Subtype-2 Subtype-3
cross-cancer comparison by querying breast subtype signatures on melanoma samples
Responder
Non-responder

Figure 6



BIOCARTA Terms	Adjusted p value	Enrichr combined score	Genes
Lectin induced complement pathway	0.00042	1607.55	MASP2; MASP1; C8B
GATA3 participate in activating the **Th2** cytokine genes expression - GATA3 pathway	0.07	412.03	IL4; IL1A

Figure 7


A
Subtype-1
Subtype-2
Subtype-3
ALCAM
CLEC5A
IL8
S100A7
LCN2
Macrophage/ monocyte
Neutrophil
-1
0
+1
B
Macrophages
FDR<0.2
ssGSEA score
200
100
0
−100
−200
Subtype-1
Subtype-2
Subtype-3
C
Pan-macrophage CD68 marker IHC
D
p<0.05
H-score for CD68
250
200
150
100
50
0
E
p<0.005
Chronic inflammation score
13
12
11
10
9
Subtypes
Subtype-1
Subtype-2
Subtype-3
F
Overall Survival probability
1.00
0.75
0.50
0.25
0.00
p = 0.045
0
5
10
15
20
25
Time (months)
Number at risk
Subtyp-3- score- high 212 136 88 65 36 0
Subtype-3-score-low 136 87 68 54 30 0
Time (months)
G
TNBC
Subtype1
Subtype2
Subtype3
Th-1 specific genes
Chemokine genes
CD8+ T cell effector cells and genes
CD4+ Th1 specific genes
TCR and related pathway signaling
Antigen processing and presentation genes
DAMP pathway
Acute inflammation pathway
Viral mimicry
Associated with immunotherapy response
Th-2 and Th17 specific genes
CD4+ regulatory cell
Macrophages, monocytes and neutrophils
Immune desert
H
TNBC
Subtype-1
Subtype-2
Subtype-3
Good prognosis, dense TILs, enriched for premenopausal status and reduced tumor size
Intermediate prognosis, mild and moderate
TILs, increased tumour size, increased tumour purity
Poor prognosis, mild and moderate
TILs, increased tumour size, advanced stage and post menopausal