Title

Distinct Sarcoma Microenvironments Predict Benefit from Addition of Pembrolizumab to Preoperative Radiotherapy and Surgery in SU2C-SARC032

Running Title

Sarcoma Microenvironment Predictors in SU2C-SARC032

Authors and Affiliations

Stefano Testa[1,#], Jonathon E. Himes[2,#], Ajay Subramanian[3], Serey C.L. Nouth[3], Karla V. Ballman[4,5], Rachel S. Heise[6], Matthew Pierpoint[7], Neda Nemat-Gorgani[3], Timothy J. Sears[3], Michael S. Binkley[3], Anusha Kalbasi[3], David L. Corcoran[8], Angela M. Hong[9,10], Brian E. Brigman[11,12], Richard F. Riedel[12,13], Matt van de Rijn[14], Yvonne M. Mowery[12,15], Kent J. Weinhold[12,16], David G. Kirsch[2,12,17,18,19,*], Everett J. Moding[3,20,*]

[1]Division of Hematology and Oncology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts, USA

[2]Department of Radiation Oncology, Duke University, Durham, North Carolina, USA

[3]Department of Radiation Oncology, Stanford University, Stanford, California, USA

[4]Department of Quantitative Health Sciences, Mayo Clinic, Rochester, Minnesota, USA

[5]Biostatistics Core, Sarcoma Alliance for Research through Collaboration, Ann Arbor, MI, USA

[6]Department of Population Health Sciences, Weill Cornell Medicine, New York, New York, USA

[7]Department of Pharmacology & Cancer Biology, Duke University, Durham, North Carolina, USA

[8]Center for Genomic and Computational Biology, Duke University, Durham, North Carolina, USA

[9]Department of Radiation Oncology, Chris O'Brien Lifehouse-Royal Prince Alfred Hospital, Sydney, Australia

[10]Faculty of Medicine and Health, The University of Sydney, Sydney, New South Wales, Australia

[11]Department of Orthopaedic Surgery, Duke University, Durham, North Carolina, USA

[12]Duke Cancer Institute, Duke University, Durham, North Carolina, USA

[13]Division of Medical Oncology, Department of Medicine, Duke University, Durham, North Carolina, USA

[14]Department of Pathology, Stanford University, Stanford, California, USA

[15]Department of Radiation Oncology, UPMC Hillman Cancer Center-University of Pittsburgh, Pittsburgh, Pennsylvania, USA

[16]Department of Surgery, Duke University, Durham, North Carolina, USA

[17]Radiation Medicine Program, Princess Margaret Cancer Centre, University Health Network, Toronto, Ontario, Canada

[18]Department of Radiation Oncology, University of Toronto, Toronto, Ontario, Canada

[19]Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada

[20]Stanford Cancer Institute, Stanford University, Stanford, California, USA

[#]These authors contributed equally

*Corresponding authors

Contact Information

Everett J. Moding, M.D., Ph.D., Assistant Professor, Department of Radiation Oncology, Stanford Cancer Institute, 875 Blake Wilbur Drive, Stanford, California, 94305-5847, United States of America. Email: emoding@stanford.edu. Phone: 650-498-1625.

and

David G. Kirsch, M.D., Ph.D., Professor, Department of Radiation Oncology, University of Toronto, 610 University Avenue, Toronto, Ontario, M5G 2M9, Canada. Email: david.kirsch@uhn.ca. Phone: 416-946-2094.

52

53 **Conflicts of Interest**

54 Y.M.M., K.V.B., K.J.W., D.G.K., and E.J.M. report grant funding from Stand Up to Cancer through
55 the Merck Catalyst Program to their institutions to support this study. A.M.H has served as a paid
56 consultant for Bayer and Telix. R.F.R. reports the following disclosures: Ownership: Limbguard,
57 LLC (Spouse/Owner); Institutional Clinical Research Support: AADi, Adaptimmune, AROG, Ayala,
58 BioAtla, Blueprint, Cogent, Daiichi-Sankyo, Deciphera, GlaxoSmithKline, InhibRx, Intensity
59 Therapeutics, Oncternal, PTC Therapeutics, SARC, Servier, SpringWorks, Tracon;
60 Consultant/Advisor: AADi, Adaptimmune, Bayer, Blueprint, Boehringer Ingelheim, Daiichi-
61 Sankyo, Deciphera, EMD Seroro, GlaxoSmithKline, Ipsen, NanoCarrier, Replimmune,
62 SpringWorks Therapeutics; Travel: Deciphera, SpringWorks Therapeutics. D.G.K. is a member of
63 the scientific advisory board and owns stock in Lumicell Inc, a company commercializing
64 intraoperative imaging technology. D.G.K. is a co-founder of and stockholder in XRAD
65 Therapeutics, which is developing radiosensitizers. None of these affiliations represents a conflict
66 of interest with respect to the work described here. D.G.K. is a coinventor on a patent for a
67 handheld imaging device and is a coinventor on a patent for radiosensitizers. XRAD Therapeutics,
68 Bristol Myers Squibb, and Varian Medical Systems provided research support in the past to
69 D.G.K., but this did not support the research described here. E.J.M. has served as a paid
70 consultant for Guidepoint and GLG. The other authors declare no competing interests.

Abstract

The addition of pembrolizumab to preoperative radiotherapy (RT) improved disease-free survival (DFS) for patients with stage III undifferentiated pleomorphic sarcoma (UPS) and dedifferentiated/pleomorphic liposarcoma (LPS) in the randomized SU2C-SARC032 trial. To precisely identify patients who benefit from pembrolizumab and RT, we performed comprehensive multi-omics profiling of pre- and post-treatment tumor and blood samples, including bulk RNA-seq, flow cytometry, and cytometry by time of flight. Additionally, we built a single-cell RNA-seq atlas spanning 65,786 cells from UPS and LPS to recover single-cell states in bulk tumor samples using digital cytometry. Two opposing tumor microenvironments (TMEs), immune-cold sarcoma ecotype 1 (SE1) and immune-hot sarcoma immune class E (SIC E), benefited from pembrolizumab. Pembrolizumab combined with RT depleted PD-1+ exhausted T cells in SIC E sarcomas and increased effector memory CD4+ T cells in SE1 sarcomas with an overall increase in CD8+ early activated T cells, CD4+ follicular helper T cells, and T cell receptor diversity. Matrix-remodeling stromal and epithelial-like sarcoma cell programs were associated with worse outcomes and diminished with pembrolizumab and RT. Our findings identify different mechanisms of response to pembrolizumab in localized, high-risk UPS/LPS and suggest that sarcoma TME signatures may identify patients most likely to benefit from adding pembrolizumab to preoperative RT.

Introduction

Soft tissue sarcomas (STS) are a rare and heterogeneous group of mesenchymal malignancies[1]. Management for patients with high-risk, localized disease usually involves a combination of preoperative radiotherapy (RT) followed by surgery to optimize local control[2]. However, approximately half of patients with high-risk features, including large tumor size (> 5 cm) and high grade, ultimately develop metastatic disease[3] with median overall survival of less than 2 years[4]. To reduce risk of metastasis, cytotoxic chemotherapy has been administered perioperatively to patients with high-risk localized STS with mixed results[5,6]. In the metastatic setting, immune checkpoint inhibition (ICI) targeting programmed cell death protein 1 (PD-1) and cytotoxic T-lymphocyte antigen-4 (CTLA-4) has shown promising activity in select STS histologies such as undifferentiated pleomorphic sarcoma (UPS) and pleomorphic or dedifferentiated liposarcoma (LPS)[7–9]. Furthermore, preclinical and clinical data have suggested that ICI can synergize with RT to improve local and distant anti-tumor responses[10,11].

Recently, the SU2C-SARC032 phase 2 randomized trial investigated the benefit of adding perioperative ICI to preoperative RT and surgery for patients with Stage III (>5 cm and grade 2 or 3) UPS or LPS of the extremity or limb-girdle[12]. In the trial, 143 patients were randomized to preoperative RT then surgery (control) or preoperative pembrolizumab with RT then surgery and postoperative pembrolizumab (experimental). Remarkably, pembrolizumab improved the 2-year disease-free survival (DFS) by 15% from 52% in the control arm to 67% in the experimental arm, mainly driven by a reduction in distant disease recurrence. However, pembrolizumab increased the risk of grade 3 adverse events from 31% to 56%, and most patients did not benefit from pembrolizumab. Improvement in DFS was predominantly driven by patients with grade 3 disease, but additional biomarkers will be critical to identify the patients with stage III STS most likely to benefit from pembrolizumab.

Due to their mesenchymal origin, STS cells form complex interactions with other stromal cells in the tumor microenvironment (TME), which might explain the poor performance of PD-L1

115 expression and tumor mutational burden for predicting ICI response in patients with STS[13].
116 Several prior studies have demonstrated that the sarcoma TME has prognostic value and can
117 potentially predict response to ICI[14,15]. For example, a previous study estimated the abundance
118 of nine stromal cell populations by deconvoluting bulk gene expression data to define five
119 sarcoma immune classes (SICs)[14,16]. Of these, SIC E, characterized by an inflamed TME with
120 abundant B-cells and tertiary lymphoid structures (TLS), was associated with improved
121 outcomes in patients with metastatic STS receiving pembrolizumab. We previously used a
122 machine learning method called EcoTyper[17] to identify three STS-specific cellular communities
123 of co-occurring transcriptional cell states (sarcoma ecotypes, SEs) that were associated with
124 prognosis and predicted response to immunotherapy across independent cohorts in metastatic
125 STS[15]. However, it is unclear if these predictors can identify patients with localized sarcomas
126 who could benefit from the addition of perioperative ICI to preoperative RT.
127 Here, we perform a comprehensive multi-omics analysis of samples from patients
128 enrolled in the SU2C-SARC032 trial to elucidate sarcoma TME dynamics in response to RT with
129 or without ICI and to identify predictors of benefit from pembrolizumab. We identify two distinct
130 TME signatures associated with better outcomes after pembrolizumab and provide evidence
131 that RT promotes intratumoral inflammation that may synergize with ICI to trigger anti-tumor
132 immune responses in immune cold STS. If validated in future studies, our findings could help to
133 identify patients with localized STS most likely to benefit from pembrolizumab to prevent the
134 development of metastatic disease.
135
136 **Results**
137 *Correlative multi-omics analysis of SU2C-SARC032 patient samples*
138 Patients in the SU2C-SARC32 trial were enrolled between November 28, 2017 and
139 November 14, 2023 and randomized to neoadjuvant RT (50 Gy in 25 fractions) followed by
140 surgical resection (control) or neoadjuvant RT with pre- and postoperative pembrolizumab

141 (experimental, **Fig. 1**). A total of 143 patients were enrolled in the trial, and 127 evaluable
142 patients were included in the modified intention-to-treat analysis (**Supplementary Fig. S1a**).
143 Three evaluable patients did not undergo surgery, two had disease progression before surgery,
144 and one patient on the experimental arm discontinued treatment due to adverse events. No
145 evaluable patients received adjuvant chemotherapy. Formalin-fixed paraffin-embedded (FFPE)
146 pre-treatment tumor biopsies and surgical resection specimens were analyzed with bulk RNA-
147 sequencing (RNA-seq). Overall, bulk RNA-seq data was available for 119 of 127 (94%)
148 evaluable patients with 70/119 (59%) patients having both pre-treatment biopsy and surgical
149 resection RNA-seq data. RNA-seq was available only for the pre-treatment time point in 41/119
150 (34%) and only on the surgical resection sample in 8/119 (7%). Fresh post-treatment surgical
151 resection samples were obtained for 56/127 (44%) patients and processed through flow
152 cytometry. Peripheral blood mononuclear cells (PBMC) samples were collected longitudinally for
153 Cytometry by Time of Flight (CyTOF) analysis in 55 patients before starting treatment ("pre-
154 treatment"), one week after starting treatment ("1 week on-treatment"), during the 3rd week of RT
155 ("Week 3 of RT"), after neoadjuvant treatment and before surgery ("Post-RT/Pre-Surgery"), and
156 at 3 and 12 months after surgery. Follow up data was updated on February 24, 2025 for the
157 correlative analyses.
158
159 *Baseline sarcoma TME signatures correlate with outcomes*
160 Based on the established association between SICs and SEs with patient outcomes in
161 localized STS and response to ICI in patients with metastatic STS[14,15], we investigated if these
162 STS TME signatures were associated with outcomes in patients enrolled in SU2C-SARC032. To
163 answer this question, we assigned each patient to an SIC and an SE based on bulk RNA-seq of
164 their pre-treatment sample (**Supplementary Table S1**). Of the 111 patients with pre-treatment
165 RNA-seq, 10/111 (9%) patients were classified as SIC A, 39/111 (35%) as SIC B, 11/111 (10%)
166 as SIC C, 25/111 (22%) as SIC D, and 26/111 (23%) as SIC E (**Fig. 2a**). Consistent with the

167 prior report[14], patients classified at baseline as SIC E, which is characterized by a rich pro-
168 inflammatory immune infiltrate, had better DFS than patients classified as any other SIC (hazard
169 ratio [HR]: 0.36; 95% confidence interval [CI] 0.19, 0.67; nominal *P* = 0.01, **Fig. 2b**,
170 **Supplementary Fig. S2a**). Additionally, when considering only SIC E patients, those receiving
171 pembrolizumab had a better DFS compared to those receiving RT alone (HR: 0.37; 95% CI
172 0.07, 1.85; nominal *P* = 0.23, **Fig. 2c**). In contrast, among patients belonging to any other SIC
173 than SIC E, there was little difference in DFS between those in the control versus experimental
174 arm (HR: 0.86; 95% CI 0.47, 1.55; nominal *P* = 0.59, **Fig. 2d**).
175 Next, we applied the EcoTyper framework to assign patients to one of the three
176 previously defined SEs based on the pre-treatment tumor RNA expression profile (**Fig. 2e**). Of
177 note, a subset of patients (*n* = 66) had a high abundance of multiple SEs and were not assigned
178 to a dominant SE. Consistent with our prior report[15], patients assigned to SE3 had worse DFS
179 than patients assigned to SE2 (HR: 5.4; 95% CI 1.3, 22.8; nominal *P* = 0.02, **Fig. 2f**). This is
180 consistent with SE2 usually being enriched for activated pro-inflammatory immune cell states,
181 whereas SE3 tends to be enriched for immunosuppressive cell states such as *SPP1*+ M2-like
182 tumor-associated macrophages. Interestingly, there was improved DFS in patients with SE1 that
183 received pembrolizumab compared to those receiving RT only (HR: 0.39; 95% CI 0.12, 1.3;
184 nominal *P* = 0.10, **Fig. 2g**), while there was minimal difference in DFS between non-SE1
185 patients treated with pembrolizumab and those treated with RT only (HR: 0.95; 95% CI 0.51,
186 1.79; nominal *P* = 0.88, **Fig. 2h**). Notably, SE1 sarcomas are immune cold tumors with low
187 proportions of activated immune cells, and we previously observed low response rates to ICI in
188 patients with metastatic SE1 tumors[15]. Because patients with grade 3 tumors appeared to derive
189 most of the benefit from pembrolizumab in SU2C-SARC032[12], we evaluated whether there was
190 a difference in the proportion of grade 2 and 3 tumors based on SIC or SE classification. Grade
191 2 tumors were more abundant in SIC A compared to SIC B patients (**Fig. S2b**, 60% versus 23%,
192 respectively, nominal *P* = 0.049), and SE2 compared to SE3 patients (**Fig. S2c**, 56% versus

193 0%, respectively, nominal $P = 0.02$). However, neither SIC E nor SE1 were enriched for grade 3
194 versus grade 2 tumors.
195 Next, we measured the expression of 18 genes comprising a T cell-inflamed gene
196 expression profile (GEP), which was previously associated with response to pembrolizumab in
197 patients across 9 different cancers[18] (**Supplementary Table S2**). The pre-treatment T cell-
198 inflamed GEP score was associated with longer DFS in patients that received pembrolizumab
199 (HR: 0.62; 95% CI 0.4, 0.6; nominal $P = 0.02$, **Fig. S2d**). Interestingly, the T cell-inflamed GEP
200 score was higher in patients belonging to SIC E compared to any other SIC except SIC D (mean
201 T cell-inflamed GEP: SIC E 4.7, SIC D 4.11, SIC C 3.3, SIC B 2.9, SIC A 1.9; nominal $P <$
202 0.0001, **Fig. S2e**), but the T cell-inflamed GEP score was lower for patients classified as SE1
203 compared to SE2 or SE3 (mean T cell-inflamed GEP 2.5, 4.7, 3.9, respectively; nominal $P <$
204 0.0001, **Fig. S2f**). Lastly, we determined if pre-treatment expression of either immune
205 checkpoint molecules or T cell co-stimulatory molecules was associated with DFS in patients in
206 the control and experimental arms (**Fig. S2g**). Interestingly, we found that high *CTLA4* (HR: 0.3;
207 95% CI 0.1, 0.8; nominal $P = 0.02$) and high *TIGIT* (HR: 0.3; 95% CI 0.1, 0.8; nominal $P = 0.01$)
208 expression correlated with longer DFS only in patients that received pembrolizumab, but not in
209 those that received RT only (**Fig. S2h** and **S2i**). Overall, these findings confirm the association
210 of SICs and SEs with prognosis in high-risk, localized STS and suggest that two TME
211 signatures (immune hot SIC E and immune cold SE1) identify different subsets of patients who
212 benefit from the addition of perioperative pembrolizumab to preoperative RT and surgery.
213
214 *Flow cytometry reveals post-treatment TME architecture*
215 For patients with viable fresh tumor specimens collected after neoadjuvant therapy at the
216 time of surgical resection, flow cytometry was used to measure treatment effect on tumor
217 immune cell infiltration (**Fig. 3**, **Supplementary Tables S3** and **S4**). Within the $CD4^+$ T cell
218 population, there was a lower proportion of $CD45RA^+$ $CCR7^+$ naïve T cells (median 0.008 vs

219 0.02, nominal P = 0.004, q = 0.02), and CD279$^+$ (PD-1$^+$) T cells (median 0.34 vs 0.48, nominal P
220 = 0.005, q = 0.02) in patients that received neoadjuvant pembrolizumab and RT compared to
221 those that received RT only (**Fig. 3a** and **3b**). Because pembrolizumab binds to CD279, it is
222 possible that the lower proportion of CD279$^+$ T cells was due to antigen masking by
223 pembrolizumab. Additionally, there was a higher proportion of CD152$^+$ (CTLA4$^+$) CD4$^+$ T cells
224 (median 0.33 vs 0.22, nominal P = 0.04, q = 0.12) and activated (CD38+ HLA-DR+) CD4$^+$ T
225 cells (median 0.016 vs 0.008, nominal P = 0.06, q = 0.15) in patients treated with
226 pembrolizumab plus RT compared to RT only. Considering only patients classified as SE1 at
227 baseline, at post-treatment there was lower abundance of CD45RA$^-$ CCR7$^+$ central memory
228 (Tcm) CD4$^+$ T cells (median 0.22 vs 0.67, nominal P = 0.004, q = 0.04) and higher abundance of
229 CD45RA$^-$ CCR7$^-$ effector memory (Tem) CD4$^+$ T cells (median 0.76 vs 0.24, nominal P = 0.004,
230 q = 0.04) in those receiving pembrolizumab compared to those treated with RT only (**Fig. 3c**
231 and **3d**). In contrast, in patients classified as SIC E at baseline, there was a post-treatment
232 reduction in PD-1$^+$ CD4$^+$ T-cells (median 0.28 vs 0.64, nominal P = 0.02, q = 0.14) and an
233 increase in CM CD4$^+$ T cells (median 0.47 vs 0.33, nominal P = 0.05, q = 0.22).
234 Similar to CD4$^+$ T cells, we observed a lower proportion of PD-1$^+$ T cells (median 0.37 vs
235 0.49, nominal P = 0.03, q = 0.23) and naïve T cells (median 0.01 vs 0.03, nominal P = 0.04, q =
236 0.23) among CD8$^+$ T cells in patients treated with pembrolizumab (**Fig. S3a** and **S3b**). Subgroup
237 analysis of patients categorized as SE1 at baseline again showed lower abundance in the Tcm
238 subset within the CD8$^+$ T cell population (median 0.11 vs 0.39, nominal P = 0.03, q = 0.62) and
239 an increase in Tem CD8$^+$ cells (median 0.78 vs 0.42, nominal P = 0.08, q = 0.65). In baseline
240 SIC E patients, there was no post-treatment difference in Tcm (median 0.19 vs 0.11, nominal P
241 = 0.61, q = 0.99) or Tem (median 0.63 vs 0.56, nominal P = 0.39, q = 0.92) CD8+ T cells by flow
242 cytometry (**Fig. 3e** and **3f**). Looking at subsets of T-regs in post-treatment specimens, there was
243 a lower proportion of PD-1$^+$ cells in patients in the experimental compared to the control arm,

244 which was true both when considering all patients (**Fig. S3c** and **S2d**) and when considering
245 only those classified as SIC E at baseline (**Fig. S3e** and **S3f**).
246 Next, we determined the association of post-treatment abundance of the different
247 immune cell subsets identified by flow cytometry with DFS (**Fig. 3g**). Interestingly, the
248 abundance of $CD8^+$ Tcm-cells was associated with worse DFS both when considering patients
249 in the experimental and control arms (HR: 12.7; 95% CI 2.2, 71.4; nominal $P = 0.004$) and when
250 considering only patients treated with pembrolizumab (HR: 10.2; 95% CI 1.1, 94.0; nominal $P =$
251 0.04). However, the fraction of CD8+ T cells among total CD45+ cells was associated with better
252 DFS in the control arm only (HR: 0.04; 95% CI 0.002, 0.92; nominal $P = 0.04$). Lastly, we found
253 that abundance of CD103+ tissue-resident memory T cells was associated with improved DFS
254 only in patients in the control group (HR: < 0.001; 95% CI 1e-67, 0.0003; nominal $P = 0.02$),
255 while PD-1+ T-regs were associated with worse DFS in patients treated with pembrolizumab
256 and RT (HR: 10.7; 95% CI 1.02, 112.7; nominal $P = 0.04$). Overall, these results indicate that
257 the addition of pembrolizumab to neoadjuvant RT might result in an overall shift of tumor
258 infiltrating lymphocytes towards an effector phenotype. However, there were different dynamics
259 in baseline immune hot tumors (SIC E), where we predominantly observed depletion of
260 exhausted and immunosuppressive PD-1+ T cell subsets, and immune cold tumors (SE1),
261 where immunotherapy was associated with an increase in the proportion of Tem cells.
262
263 *Longitudinal systemic immune monitoring through CyTOF*
264 Peripheral blood (PB) immune cell profiling has been shown to correlate with response
265 to immunotherapy across different cancers[19]. Based on this, we asked if neoadjuvant RT with or
266 without pembrolizumab could result in systemic immune changes detectable in the PB of STS
267 patients that is associated with outcomes. We obtained PB samples before and at different
268 intervals after treatment both for patients receiving RT alone and those receiving neoadjuvant
269 RT plus pembrolizumab. CyTOF was then used to measure changes in abundance of different

270 PBMC subsets over time (**Fig. 4**, **Supplementary Tables S5** and **S6**). Unsupervised clustering
271 revealed 9 major PBMC clusters including B cells, monocytes, CD4+ T cells, naïve CD8+ T
272 cells, effector CD8+ T cells, activated T cells, memory T cells, NK cells, and CD56+ NK cells,
273 and further sub-clustering of PB T cells revealed 11 sub-populations (**Fig. 4a** and **4b**).
274 Considering the PBMC clusters (**Fig. 4c**, **Supplementary Table S5**), there was a
275 reduction in the abundance of B cells between the pre-treatment time point and the 3rd week of
276 RT in both the control arm (mean centered log-ratio [CLR] 0.64 vs 0.24, nominal P = 0.03, q =
277 0.94) and the experimental arm (mean CLR 0.65 vs 0.20, nominal P = 0.01, q = 0.69). There
278 was a progressive reduction in circulating monocytes from the 3rd week of RT until one-year
279 post-surgery considering all patients (mean CLR 2.2 vs 1.6, nominal P = 0.0009, q = 0.01),
280 which was driven by those in the experimental arm (mean CLR 2.2 vs 1.6, nominal P = 0.009, q
281 = 0.69). In contrast, the abundance of Memory T cells increased at 1 week on-treatment in
282 patients in the RT plus pembrolizumab arm (mean CLR -2.9 vs -1.5, nominal P = 0.04, q = 0.92)
283 as opposed to patients that only received RT (mean CLR -2.7 vs -3.5, nominal P = 0.04, q =
284 0.94). Lastly, we observed a trend towards a higher abundance of circulating NK cells at 1 week
285 after initiating treatment in patients that received RT only compared to those that also received
286 pembrolizumab (mean CLR -0.6 vs -2.5, nominal P = 0.04, q = 0.36). Next, we determined if the
287 abundance of any of the identified PBMC clusters was associated with DFS at any of the time
288 points analyzed (**Fig. 4d**). Interestingly, abundance of circulating Effector CD8+ T cells at one-
289 year post-surgery was associated with higher progression risk when considering all patients
290 (HR: 10.4; 95% CI 1.4, 76.9; nominal P = 0.02).
291 Considering PB T cell subsets (**Fig. 4e**, **Supplementary Table S6**), activated CD4+ T
292 cells showed a decrease at 1 week on-treatment for patients that received pembrolizumab with
293 a subsequent return to pre-treatment levels at 3 months post-surgery (mean CLR -2.0 vs -1.5,
294 nominal P = 0.02, q = 0.98). In patients in the control arm, CD4+ Tcm cells decreased from the
295 3rd week of RT to the post-RT/pre-surgery time point (mean CLR -2.4 vs -3.7, nominal P = 0.03,

296 q = 0.98) with a subsequent increase at one-year post-surgery (mean CLR -3.7 vs -1.7, nominal
297 *P* = 0.003, q = 0.55). Also, the fraction of naïve CD4+ T cells was higher in the experimental
298 versus the control arm at 1-year post-surgery (mean CLR 2.5 vs 1.7, nominal *P* = 0.03, q =
299 0.40), while there was a trend for lower circulating T-regs in patients receiving pembrolizumab 1
300 week after starting treatment (mean CLR -0.3 vs 0.6, nominal *P* = 0.05, q = 0.59). Lastly, the
301 abundance of activated CD8+ T cells at 1 year post-surgery was associated with higher
302 progression risk when considering all patients (HR: 13.9; 95%CI 1.3, 147.0; nominal *P*=0.03,
303 **Fig. 4f**), while abundance of early activation CD8+ T cells at 3 months post-surgery (HR: 0.06;
304 95%CI 0.006, 0.6; nominal *P*=0.01) and abundance of a subset of naïve CD8+ T cells (c1) at 1
305 year post surgery (HR: 0.1; 95%CI 0.03, 0.7; nominal *P*=0.01) were associated with lower
306 progression risk, considering only patients in the experimental arm or all patients, respectively.
307 Overall, these results highlight both early and late post-treatment changes in relative abundance
308 of different PB immune cell subsets in patients with STS receiving immunotherapy and RT,
309 which might indicate activation of a systemic anti-tumor immune response.
310
311 *Digital cytometry dissects pre-treatment STS TME cell states*
312 To further dissect the cell states contributing to the STS TME before and after treatment,
313 we performed digital cytometry on the bulk gene expression obtained from the FFPE tumor
314 samples. We built a scRNA-seq atlas of UPS and LPS samples [20], identifying different malignant
315 and non-malignant TME cell states across 65,786 cells from 12 patients (**Fig. 5a**). Overall, we
316 identified 9 malignant cell states, 7 non-malignant mesenchymal cell states including cancer-
317 associated fibroblasts (CAFs) and endothelial cells, 17 lymphoid, and 12 myeloid immune cell
318 states (**Fig. S4a**, **Supplementary Tables S7** and **S8**). We derived a single-cell gene expression
319 signature matrix from the atlas to deconvolute the abundance of each cell state in bulk RNA-seq
320 samples with CIBERSORTx (**Supplementary Tables S9**, **S10**, and **S11**)[21]. To validate the
321 accuracy of our deconvolution strategy, we correlated the abundance of major immune cell

322 types in post-treatment tumor samples measured with CIBERSORTx versus flow cytometry and
323 observed good correlation despite the different tumor regions sampled and technical
324 approaches utilized (**Fig. S4b**).
325 Next, we determined the association of pre-treatment cell state abundance measured by
326 digital cytometry with DFS (**Fig. 5b**). Interestingly, a higher abundance of matrix CAFs, a subset
327 associated with extra-cellular matrix (ECM) remodeling and high expression of matrix proteins
328 including collagen genes (*COL1A1*, *COL1A2*, *COL3A1*) and matrix metalloproteinases
329 (*MMP2*)[22], was associated with worse DFS in the whole cohort (HR: 1.3; 95% CI 1.06, 1.7;
330 nominal P = 0.01) and in the control arm (HR: 1.3; 95% CI 1.01, 1.8; nominal P = 0.04).
331 Similarly, abundance of mesenchymal-like ECM remodeling sarcoma cells, a cell state with high
332 expression of ECM remodeling-related genes (*MMP2*, *COL6A3*, *COL1A2*), was associated with
333 worse DFS considering all patients (HR: 1.2; 95% CI 0.9, 1.4; nominal P = 0.05). Additionally,
334 we found that sarcoma cells with an epithelial-like expression profile were associated with worse
335 DFS both in the control (HR: 1.5; 95% CI 1.1, 2.0; nominal P = 0.01) and experimental (HR: 1.4;
336 95% CI 1.1, 1.9; nominal P = 0.02) arms and when considering all patients (HR: 1.4; 95% CI
337 1.2, 1.8; nominal P = 0.0007). Sarcoma cells with an undifferentiated mesenchymal-like
338 phenotype were associated with better DFS in all patients (HR: 0.7; 95% CI 0.6, 0.9; nominal P
339 = 0.001) and in the control arm (HR: 0.5; 95% CI 0.3, 0.8; nominal P = 0.006). Among immune
340 cell subsets, a higher abundance of B cells and plasma cells was associated with better DFS in
341 the whole cohort (HR: 0.5; 95% CI 0.4, 0.9; nominal P = 0.009), while macrophages with high
342 expression of the iron exporter *SLC40A1* and the selenium binding protein *SELENOP* were
343 associated with better DFS in all patients (HR: 0.6; 95% CI 0.4, 0.9; nominal P = 0.007) and
344 those in the control group (HR: 0.5; 95% CI 0.3, 0.9; nominal P = 0.01). Interestingly,
345 SELENOP$^+$ macrophages have been found to be enriched in patients with non-small cell lung
346 cancer responding to neoadjuvant chemoimmunotherapy, with spatial transcriptomics data
347 showing preferential localization of this macrophage subtype in TLSs[23].

348 Next, we asked if there was a difference in the abundance of the recovered cell states in
349 pre-treatment samples across SICs or SEs (**Fig. 5c** and **5d**, **Fig. S5a** and **S5b**). As expected,
350 pro-inflammatory immune cell states were enriched in SIC E compared to other SICs, including
351 B cells and plasma cells, activated CD8+ T cells (*HLA-DP/DQ/DR*, *CCL4/5*, and *NKG7*+),
352 *GZMK*+ CD8+ T memory cells, and CD56^{low} NK cells. We also observed enrichment of myeloid
353 cell subsets like amphiregulin-expressing dendritic cells (*AREG*+ DCs)[24] and *SELENOP*$^{+}$
354 *SLC40A1*$^{+}$ macrophages in SIC E and SIC D and higher levels of exhausted CD8+ T cells
355 (*PDCD1*, *LAG3*, *HAVCR2*, *TOX*, and *ENTPD1*+) in SIC E samples, possibly indicating the
356 establishment of a chronic anti-tumor immune response leading to T cell dysfunction[25,26]. Lastly,
357 matrix CAFs and both epithelial and ECM-remodeling sarcoma cell states tended to be enriched
358 in the immune-cold SIC A and SIC B. Looking at cell states abundance across SEs, we noticed
359 enrichment of immunosuppressive myeloid cells in SE3 compared to SE2 and SE1 samples,
360 including monocytes expressing high levels of epiregulin (*EREG*+), reported to promote
361 tumorigenesis and invasion in gliomas[27], and immunoregulatory polymorphonuclear cells
362 (PMNs) with high expression of *AQP9* and *LUCAT1*. As expected, SE2 showed a higher
363 abundance of activated CD8+ T cells compared to SE1. Lastly, SE1 showed higher proportion
364 of myofibroblastic sarcoma cells, characterized by high expression of genes associated with
365 motility and cytoskeleton activity (*ACTA2*, *MYL9*, *ACTG2*), and lower abundance of *SELENOP*$^{+}$
366 *SLC40A1*$^{+}$ macrophages compared to SE2 and SE3, possibly signaling the lack of a pre-
367 existing anti-tumor immune response.
368 Overall, these results further indicate how SE1 and SIC E tumors represent distinct
369 sarcoma TME states. Indeed, only one patient was classified as both SE1 and SIC E at baseline
370 (**Fig. 5e**). Additionally, we compared the abundance of the different deconvoluted cell states
371 between SIC E and SE1 and found that SIC E tumors were enriched for both activated and
372 stressed/exhausted CD4+ and CD8+ T cell states, B cells and plasma cells, and CD56^{high} and
373 CD56^{low} NK cells (**Fig. 5f**). In contrast, SE1 sarcomas were enriched for both epithelial-like and

374 mesenchymal-like ECM remodeling sarcoma cell states and matrix CAFs. These results further
375 show that SE1 tumors harbor an immune-desert phenotype with little to no immune cell
376 infiltration, opposite to the rich immune infiltration of SIC E tumors. These TME differences were
377 also reflected in outcomes, where baseline SE1 patients had worse DFS compared to SIC E
378 patients across both treatment arms (HR: 2.62; 95% CI 1.0, 6.9; nominal P = 0.048, **Fig. 5g**).
379
380 *Neoadjuvant RT plus ICI reshapes the sarcoma TME*
381 Next, we aimed to define the effects of neoadjuvant RT with or without ICI on the STS
382 TME. When comparing pathologic response in the resection specimen after neoadjuvant
383 therapy in the experimental vs. the control arm, the addition of pembrolizumab increased the
384 percent necrosis (median 90% vs 74%, nominal P = 0.01, **Fig. 6a**). However, pathological
385 response was not associated with DFS in either arm when measured as a continuous variable
386 or by defining a pathological complete response as ≥ 95% necrosis (**Fig. S6a**). When
387 determining the association of post-treatment abundance of the cell states recovered through
388 CIBERSORTx with DFS (**Fig. 6b**), we found that abundance of *CD36*+ CAFs was associated
389 with better DFS in both the experimental group (HR: 0.48; 95% CI 0.2, 0.9; nominal P = 0.02)
390 and the whole cohort (HR: 0.66; 95% CI 0.5, 0.9; nominal P = 0.02), while a higher abundance
391 of matrix CAFs was associated with worse DFS (HR: 1.18; 95% CI 0.9, 1.5; nominal P = 0.13).
392 Also, higher abundances of both activated CD8+ T cells (HR: 0.61; 95% CI 0.4, 0.9; nominal P =
393 0.02) and CD4+ T naïve (Tn) cells (*CCR7*, *FOXP1*, *BACH2*+) (HR: 0.72; 95% CI 0.5, 0.9;
394 nominal P = 0.04) were associated better DFS in the whole cohort, and a higher naïve CD4+ T
395 cell abundance was associated with better DFS in the control group (HR: 0.63; 95% CI 0.4, 0.9;
396 nominal P = 0.03).
397 We also investigated associations between post-treatment expression of immune
398 checkpoint genes and DFS to highlight possible mechanisms of resistance to ICI plus RT (**Fig.
399 S6b**). Interestingly, we found that higher post-treatment expression of *PDCD1* (PD-1) was

400 associated with better DFS in patients that received pembrolizumab (HR:0.44; 95% CI 0.2, 0.8;
401 nominal *P* = 0.007), but not in those treated with RT alone (HR: 0.69; 95% CI 0.4, 1.0; nominal
402 *P* = 0.09, **Fig. S6b, S6c**). Higher expression of *NCR3LG1* was associated with worse DFS only
403 in patients that received pembrolizumab and RT (HR: 1.8; 95% CI 1.1, 2.8; nominal *P* = 0.02,
404 **Fig. S6b, S6d**). *NCR3LG1* codes for the immune checkpoint molecule B7-H6, which acts as an
405 activating ligand for the NK cell receptor NCR3[28] and has been found to be over-expressed on
406 mesenchymal cell lines sensitive to NK cell-mediated cytotoxicity[29]. Importantly, B7-H6
407 expression has also been shown to promote cytolysis of activated T cells by NK cells and might
408 be implicated in resistance to ICI[30]. Indeed, we found that patients with high post-treatment
409 expression of *NCR3LG1* also tended to have lower expression of *B2M* (a proxy for HLA class I
410 expression)[31] and *HLA-E,* indicating a state of increased sensitivity to NK cell killing and
411 reduced susceptibility to ICI[32] (**Fig. S6e** and **S6f**). Overall, these findings suggest avenues for
412 novel therapeutic combinations to improve response to pembrolizumab in combination with RT.
413 To better understand the effect of RT and pembrolizumab on STS, we performed
414 differential gene expression analysis and gene set enrichment analysis on the pre-treatment
415 and post-treatment samples from both arms. When accounting for the effect of pembrolizumab,
416 post-RT samples showed enrichment for gene sets associated with an inflammatory response,
417 including IFN-γ, IL-2, and IL-6 signaling, as well as gene sets associated with an anti-tumor
418 responses like allograft rejection and apoptosis (**Fig. 6c**). In patients treated with
419 pembrolizumab, we observed a down-regulation of genes associated with epithelial to
420 mesenchymal transition and MYC-regulated genes and an enrichment for genes associated
421 with normal myogenesis, which might reflect reduced proliferation and decreased abundance of
422 cancer cells (**Fig. 6d**).
423 We then queried if the addition of pembrolizumab to RT changed the overall composition
424 of the TME. First, we measured the treatment-related change in abundance of the cell states
425 recovered through CIBERSORTx (**Fig. 6e**). Interestingly, there was a decrease in matrix CAFs

426 and ECM remodeling mesenchymal-like sarcoma cells only in patients that received
427 pembrolizumab but not in the control arm and a more marked decrease in epithelial-like
428 sarcoma cells in the experimental versus the control arm. Also, CD8+ early activated T cells
429 (*CRTAM*, *EOMES*, and *ZEB2*+) and CD4+ T follicular helper cells (*CXCL13*, *BTLA*, *TNFRSF4*,
430 and *TNFRSF18*+) increased after treatment in patients treated with pembrolizumab, but not in
431 those that received RT only. CD4+ memory T cells (*CCR7*, *SELL*, *TCF7*, and *LEF1*+) and
432 *SELENOP*+ *SLC40A1*+ macrophages increased in both treatment arms. To see if these
433 differential changes in cell states were matched by changes in abundance of TME cellular
434 ecosystems, we measured abundance of sarcoma ecotypes before and after treatment in the
435 control and treatment arm (**Fig. 6f**, **Supplementary Table S12**). We observed a marked
436 decrease in the abundance of the immune cold SE1 in patients that received pembrolizumab
437 but not in those treated with RT alone and an increase in the immune hot SE2 that was more
438 pronounced in the experimental compared to the control arm.
439 Next, we measured the pre- and post-treatment T cell-inflamed GEP score in patients in
440 the control and experimental arms (**Fig. S6g**, **S6h**) and found a post-treatment increase in the T
441 cell-inflamed GEP score only in baseline SE1 patients receiving pembrolizumab plus RT (**Fig.**
442 **6g**). Similar results were observed when patients were grouped by SIC, where there was a
443 more pronounced increase in the T cell-inflamed GEP score post-treatment for SIC A or SIC B
444 tumors and mainly among patients receiving pembrolizumab (**Fig. S6i**). To test if these
445 differences in immune activation profiles also corresponded to differences in the T cell
446 repertoire, we reconstructed the T cell receptor (TCR) sequences from pre- and post-treatment
447 bulk RNA-seq samples using TRUST4[33] and found that patients categorized as SE1 at baseline
448 had lower TCR diversity compared to either SE2 or SE3 (**Fig. 6h**) and a trend towards lower
449 TCR diversity for SIC A versus SIC E (**Fig. S6j**). We observed an increase in TCR diversity post-
450 treatment only in patients receiving pembrolizumab and RT (**Fig. 6i**).

451 Given the observed benefit of pembrolizumab in SU2C-SARC032 for patients with grade
452 3 disease[12], we asked if combining tumor grade with baseline SIC and SE classification could
453 better identify patients that would benefit the most from pembrolizumab. Interestingly, we found
454 that patients classified as either SIC E or SE1 that also had grade 3 tumors showed a DFS
455 benefit from the addition of pembrolizumab (**Fig. 6j**, HR: 0.26; 95% CI 0.09, 0.8; nominal P =
456 0.03), while those with grade 2 disease or classified as non-SIC E and non-SE1 did not appear
457 to benefit from immunotherapy (**Fig. 6k**, HR: 1.18; 95% CI 0.6, 2.3; nominal P = 0.62,
458 interaction P = 0.03). Taken together, these results support a model where the addition of
459 pembrolizumab to neoadjuvant RT induces immune activation in immune cold SE1 sarcoma
460 while also reigniting an established yet dysfunctional pre-existing anti-tumor response in
461 immune hot SIC E sarcomas.
462
463 **Discussion**
464 In this study, we performed a comprehensive multi-omics assessment of patients with
465 stage III LPS or UPS of the extremity receiving neoadjuvant RT with or without perioperative
466 pembrolizumab as part of SU2C-SARC032[12]. Our analyses revealed two distinct baseline STS
467 TME states associated with benefit from the addition of pembrolizumab to neoadjuvant RT. We
468 found that both an immune desert state, corresponding to SE1, and an immune hot TME,
469 represented by SIC E, correlated with clinical benefit from the addition of pembrolizumab to RT.
470 This observation likely indicates complementary mechanisms of benefit from ICI and RT[34].
471 Similar to other solid tumors that respond to ICI[35], the presence of an established anti-
472 tumor immune response in SIC E STSs was associated with benefit from the addition of
473 immunotherapy to RT. SIC E tumors are characterized by high expression of gene signatures
474 related to CD8+ and CD4+ T cells, NK cells, and B cells and are enriched for TLS on IHC,
475 indicating a robust chronic anti-tumor response[14]. Our deconvolution strategy provides a high-
476 resolution view of the cellular states that contribute to defining SIC E and is consistent with prior

477 observations that SIC E STSs might be primed for benefiting from anti-PD-1/PD-L1 agents[14,15].
478 Flow cytometry analyses on post-treatment samples in this study revealed that, in SIC E tumors,
479 the addition of ICI to RT mainly resulted in depletion of likely exhausted PD-1+ CD8+ and PD-1+
480 CD4+ T cells, which could lead to a reinvigorated T cell effector function[36,37]. Additionally, while
481 the association between an SIC E TME state and benefit from ICI was previously demonstrated
482 in the metastatic setting[14], our results also indicate a clinical benefit from immunotherapy in
483 localized STSs with an immune hot TME. This is also consistent with data from a recent phase 2
484 randomized non-comparative trial (NCT03307616), which studied neoadjuvant nivolumab or
485 ipilimumab/nivolumab in patients with resectable retroperitoneal dedifferentiated LPS (DDLPS)
486 and neoadjuvant nivolumab combined with RT for patients with UPS of the extremity[38]. This
487 study showed that patients with DDLPS that received neoadjuvant ICI and showed presence of
488 TLS at surgery had improved overall survival and that baseline immune cell infiltration, rather
489 than PD-L1 expression, correlated with outcomes in both cohorts, which is consistent with our
490 observations.

491 At the same time, our study suggested a benefit from the addition of immunotherapy for
492 patients with baseline SE1 tumors, which are typically immune cold with very little immune
493 infiltration, likely signaling lack of a pre-existing immune response[15]. Indeed, digital cytometry
494 revealed low baseline abundance of activated CD8+ T cells and *SLC40A1*+ *SELENOP*+
495 macrophages in SE1 tumors with a higher proportion of STS cell states, which is consistent with
496 the lack of anti-tumor immunity. After pembrolizumab and RT, SE1 tumors showed an increase
497 in CD4+ and CD8+ Tem cells by flow cytometry, indicating the induction of a *de novo* immune
498 response. Accordingly, we observed a post-treatment increase in the T cell-inflamed GEP score
499 only in baseline SE1 patients treated with the RT/ICI combination. This was accompanied by an
500 overall increase in TCR repertoire diversity and an increase in early activated CD8+ T cells and
501 CD4+ Tfh cells only in the experimental arm. Notably, SE3, which was associated with response
502 to ICI in patients with metastatic sarcomas[15], was not associated with benefit from

503 pembrolizumab in SU2C-SARC032. This suggests a different mechanism of response when ICI
504 is combined with radiotherapy or in patients with gross versus micrometastatic disease. Overall,
505 our findings support a model of synergism between RT and ICI, where RT induces immunogenic
506 cell death with release of damage associated molecular patterns (DAMPs) and cytokines, TME
507 inflammation, and improved antigen presentation, leading to accumulation of naïve T cells that
508 are then activated by ICI[39–41].

509 The clinical benefit of pembrolizumab in SU2C-SARC032 appeared to be driven by
510 patients with grade 3 tumors who have previously been shown to have a higher risk of
511 developing metastatic disease[42–44]. Our findings suggest that transcriptional biomarkers can be
512 integrated with tumor grade to more precisely identify patients benefiting from the addition of
513 pembrolizumab to preoperative RT in localized STS. Although this observation needs
514 prospective validation prior to incorporation into routine clinical management, an integrated
515 biomarker could help to balance the benefit-risk profile of ICI in the neoadjuvant setting due to
516 the increased incidence of grade 3 adverse events observed in the experimental arm of SU2C-
517 SARC032. Additionally, while this study only included LPS and UPS patients, our findings
518 demonstrate benefit from RT plus ICI in immune cold SE1 tumors, which have not been shown
519 to benefit from immunotherapy alone in the metastatic setting[15]. This suggests that it might
520 worthwhile to explore the synergism between RT and ICI in future studies in other high grade
521 STSs, which are generally considered to be refractory to immunotherapy in the setting of ICI
522 alone, such as leiomyosarcomas and synovial sarcomas[9]. Lastly, we acknowledge the
523 limitations of our study and correlative analyses, which are retrospective and exploratory in
524 nature. Our observations aim to be hypothesis-generating and will require validation in
525 independent cohorts and larger prospective clinical studies. A phase 3 randomized trial
526 (HARMONY) to validate the benefit of adding anti-PD-1 ICI to preoperative radiotherapy in
527 grade 3 UPS/LPS has been approved by the Canadian Clinical Trials Group (CCTG) and has

528 received funding from the Canadian Institute of Health Research (CIHR). This multi-institutional
529 clinical trial will provide an opportunity to validate the biomarkers identified in SU2C-SARC032.
530 In summary, we show that the sarcoma TME carries prognostic and predictive value in
531 patients with localized disease treated with preoperative RT and PD-1 blockade. If validated in
532 independent cohorts and prospective trials, sarcoma TME signatures could provide a
533 biologically relevant and sarcoma-specific biomarker to identify patients most likely to benefit
534 from the addition of ICI to neoadjuvant RT.
535
536 **Materials and Methods**
537 *Clinical trial design*
538 SU2C-SARC032 was an open-label, randomized, superiority phase II clinical trial.
539 Patients were enrolled at 20 academic institutions in Australia, Canada, Italy, and the United
540 States of America. The full original and final amended trial protocol and statistical analysis plan
541 with summaries of each amendment are available online in the appendix of the initial publication
542 of the trial results[12]. The trial was conducted by Sarcoma Alliance for Research Through
543 Collaboration (SARC) and funded by Stand Up to Cancer and Merck Sharp & Dohme.
544 Pembrolizumab was provided by Merck Sharp & Dohme. The study design was conceived by
545 the co-investigators and approved by Merck Sharp & Dohme. There was no patient or public
546 involvement in the design, conduct, or reporting of the trial. The funders of the study had no role
547 in data collection, data analysis, data interpretation, or writing of this report. SU2C-SARC032
548 was registered with ClinicalTrials.gov under NCT03092323 on March 27, 2017
549 (https://www.clinicaltrials.gov/study/NCT03092323).
550 Patients were eligible for enrollment if they were ≥12 years old with a resectable, non-
551 metastatic, grade 2 or 3 UPS (including myxofibrosarcoma) or dedifferentiated/pleomorphic LPS
552 of the extremity and an Eastern Cooperative Oncology Group performance status score of 0 or
553 1. Key exclusion criteria were prior therapy with an anti-PD-1, anti-PD-L1, or anti-PD-L2 agent,

554 prior RT to the site of the sarcoma, concurrent active malignancy, active autoimmune disease or
555 infection, diagnosis of immunodeficiency or systemic immunosuppressive therapy, and planned
556 receipt of neoadjuvant or adjuvant chemotherapy. Full inclusion and exclusion criteria are
557 provided in the clinical trial protocol.
558 Patients were randomly assigned (1:1, stratified by tumor grade) to receive either
559 neoadjuvant RT (50 Gy in 25 fractions) followed by surgery (control group) or neoadjuvant
560 intravenous pembrolizumab (200 mg) for three doses before, during, and after neoadjuvant RT
561 followed by surgery and up to 14 doses of adjuvant pembrolizumab. The first dose of
562 pembrolizumab was within 7 days of enrollment, and RT started within 14 days of enrollment for
563 the control group or 1-14 days after the first pembrolizumab infusion for the experimental group.
564 Randomization was coordinated by the central clinical trial office at SARC using permuted block
565 randomization with random block sizes. The enrolling sites did not have access to the random
566 allocation sequence. Patients, physicians, and data analysts were not blinded to group
567 allocation.
568
569 *Patient sample collection*
570 Written informed consent was obtained for all patient data and samples obtained and
571 analyzed in this study. All human research in this study was conducted in accordance with the
572 Declaration of Helsinki. The SU2C-SARC032 study protocol was reviewed and approved by
573 SARC, Merck Sharp & Dohme, and the institutional review board at each institution participating
574 in the trial. The correlative analyses were approved by the institutional review boards at
575 Stanford University and Duke University.
576
577 *Bulk RNA sequencing*
578 RNA was extracted from FFPE pre-treatment core needle biopsies and surgically
579 resected tumors using RNAstorm FFPE RNA Extraction Kits (Cell Data Sciences). Tissue from

580 pre-treatment core needle biopsies was provided as pre-cut 5 μm slides from the treating site,
581 and RNA was extracted from 2-3 slides per patient. Surgically resected tumors were provided as
582 tissue blocks, and RNA was extracted from 2-3 one mm cores obtained from the region with the
583 highest sarcoma content based on pathologist review of hematoxylin and eosin-stained slides
584 from the block. Libraries were prepared for sequencing using SMARTer Stranded Total RNA-
585 Seq Kits v2-Pico Input Mammalian (Takara Bio) and sequenced with 150-bp paired-end reads
586 on a NovaSeq 6000 (Illumina). FASTQ files were quasi-aligned using Salmon[45].
587
588 *Definition of sarcoma ecotypes and sarcoma immune classes*
589 SIC centroids were defined based on UPS and LPS samples from TCGA using the SIC
590 labels provided by the authors[14]. Pre-processed bulk RNA-seq profiles from TCGA[46] were
591 downloaded and scaled to transcripts per million (TPM). TPM counts were log2-tranformed,
592 then analyzed using MCP-Counter[16] to measure the abundance of 10 tumor microenvironment
593 cell types. MCP-Counter scores were Z-score transformed across all samples for each cell type,
594 excluding fibroblasts. Samples were grouped by their assigned SIC, and the mean Z-score
595 abundance for each cell population was computed to create centroids for each SIC. Pre-
596 treatment bulk RNA-seq samples from SU2C-SARC032 were processed in the same manner
597 and assigned to the closest centroid based on Euclidean distance. Sarcoma ecotypes were
598 recovered and quantified using the EcoTyper reference-guided annotation framework [17] by
599 applying the previously learnt non-negative matrix factorization model[15] with TPM normalized
600 expression values as input.
601
602 *Flow cytometry*
603 Freshly isolated post-treatment tumors were disaggregated using the Human Tumor
604 Dissociation Kit Protocol (Miltenyi Biotec). Tumor cell suspensions were immediately viably
605 cryopreserved and stored in vapor phase liquid nitrogen until thawing prior to flow cytometry

606 analysis. All flow cytometry assays were performed in the Duke Immune Profiling Core (DIPC)
607 according to established Standard Operating Procedures (SOP) and in a strictly 'blinded'
608 manner. Viably cryopreserved tumor cell suspensions were thawed, washed, and enumerated
609 using a Muse Cell Analyzer. High parameter immune profiling was then performed using a 25-
610 channel custom flow cytometry panel to profile components of the tumor microenvironment
611 (**Table S4**). Cells were stained with a vital dye (Zombie Yellow Fixable Viability Kit, BioLegend)
612 along with Fc block (BD Biosciences) for 15 minutes at room temperature in the dark. Cells
613 were washed with FACSWash Buffer (D-PBS, ThermoFisher) containing 0.5% FBS. Cells were
614 then stained using a Cytokine/Chemokine Reagent (CCR) staining cocktail containing saturating
615 amounts of antibodies, washed, and subsequently stained with a surface staining cocktail
616 containing saturating amounts of antibody. Following surface staining, cells were washed,
617 permeabilized, and incubated with Fc block (BD Biosciences). After permeabilization, cells were
618 washed with eBioscience 1X Permeabilization Solution (ThermoFisher), incubated with Fc block
619 (BD Biosciences), washed, and stained with an intracellular staining cocktail. Following
620 intracellular staining, cells were washed, fixed with D-PBS containing 1% formalin, and acquired
621 using a calibrated and optimized BD FACSymphony A5 Cell Analyzer (BD Biosciences) within 6
622 hours of fixation. Compensation Beads (BD Biosciences) were stained for each monoclonal
623 antibody and used to correct for spectral overlap prior to gating analysis. FlowJo version 10.10.0
624 (BD Biosciences) was used to analyze the data using a gating template. An internal PBMC
625 batch control sample was included with each batch to detect potential batch-to-batch variation.
626 Gating was performed according to best practices with all gates set based on marker
627 expression of the batch control. Analysis gates were drawn for each panel as follows: 1) Time
628 gate was used to exclude air bubbles and clogs that might have occurred during acquisition for
629 all files, 2) Zombie Yellow Fixable Viability dye was used to exclude non-viable cells, and 3)
630 Singlet gating of both Forward and Side Scatter was used to exclude aggregates.

631

632 *CyTOF*

633 SU2C-SARC032 patient PBMCs collected prior to treatment, one week after starting
634 treatment, during the 3rd week of RT, after neoadjuvant treatment and before surgery, and at 3
635 and 12 months after surgery were viably frozen and stored in liquid nitrogen in the Duke
636 Substrate Services Core and Research Support facility. PBMCs from 2 healthy donors were
637 similarly utilized for Phytohaemagglutinin-stimulated (PHA-P; Invivogen) and non-stimulated
638 staining controls and viably frozen in aliquots for inclusion with each batch of patient samples to
639 identify any batch effects. PBMC phenotyping was conducted using a 35-marker mass
640 cytometric approach (**Table S13**). PBMCs were gently thawed in RPMI with 10% FBS
641 (ThermoFisher) and 50 mg/mL benzonase nuclease (Millipore Sigma). After thawing, PBMCs
642 were treated with Cell-ID Cisplatin in Maxpar PBS (Standard BioTools) for live-dead
643 discrimination. PBMCs were washed and then stained with Fc block (BD Biosciences) in
644 Maxpar Cell Staining Buffer (Standard BioTools) followed by anti-human CD45 with different
645 isotope conjugates utilized as barcodes for the SU2C-SARC032 samples and the spike-in
646 controls. CD45-stained control PBMCs were washed, spiked-in to SU2C-SARC032 samples,
647 and stained with the extracellular antibodies in the panel diluted in Maxpar Cell Staining Buffer
648 (Standard BioTools). PBMC samples were then barcoded with palladium isotopes (Standard
649 BioTools) and pooled prior to fixation and permeabilization with the respective eBioscience 1X
650 solutions (ThermoFisher). Batching was avoided for samples with low post-thaw viability
651 (viability < 70%). Samples were then stained with the intracellular antibodies in the panel diluted
652 in eBioscience 1X permeabilization solution (ThermoFisher). After staining and washing, cells
653 were again fixed with 1.6% formaldehyde and stained overnight with Cell-ID Intercalator in
654 Maxpar Fix and Perm (Standard BioTools) for singlet discrimination. Data acquisition was
655 performed on a Helios mass cytometer (Standard BioTools) according to manufacturer
656 instructions.

657 Data was debarcoded with the Standard BioTools provided software package and data
658 was analyzed on Cytobank (Beckman). Geometric gating, iridium isotope gating, and cisplatin
659 gating were used to select for viable singlets. Raw CyTOF data was processed using the
660 "flowCore" Bioconductor R package[47]. Marker abundance was subset to the 35 measured
661 markers. Unstimulated and stimulated PBMC spike-in controls were segregated by different
662 CD45 isotope and excluded from clustering. The marker expression was arcsinh transformed
663 via the "CATALYST" package. Clustering was carried out using the Phenograph[48] based on the
664 top 30 principal components of marker expression. Hierarchical clustering of the median marker
665 expression of each initial cluster resulted in 9 distinct cell populations. This process was
666 repeated exclusively for the identified T cell clusters, which resulted in 11 distinct T cell
667 subtypes. UMAP embeddings were also derived from the principal components of the marker
668 data for visualization.
669 Given the compositional nature of PBMCs and T cells subsets fractions obtained through
670 CyTOF, where for each patient at each time point the fractions sum up to 1, centered log-ratio
671 transformation was applied on the raw fractions before downstream analysis using the
672 "compositions" R package[49].
673
674 *STS single-cell RNA sequencing atlas*
675 Previously published UPS, myxofibrosarcoma (MFS), and LPS (including both DDLPS
676 and well differentiated LPS) scRNA-seq samples were downloaded from GEO under accession
677 codes GSE279852 and GSE212527, and all analyses were performed in R using the "Seurat"
678 package[50]. Cells expressing hemoglobin genes, which usually represent artifacts or
679 contaminated cells, were removed from further analysis. Tumor and normal cells were
680 annotated for each sample as previously described[20], then malignant cells, CAFs, myeloid cells,
681 lymphoid cells, and endothelial cells (macro-groups) were identified separately for each

682 histology. Within each histology-specific macro-group, subclusters with an expression profile not
683 consistent with the macro-group of cells analyzed were removed from further analysis.
684 For each macro-group, the cells from each sample were normalized with SCTransform
685 using the glmGamPoi method[51] setting vst.flavor to "v2" followed by integration across samples
686 of all histologies using Harmony[52]. Dissociation induced genes (DIGs) obtained from Van Den
687 Brink et al.[53] and measured using Seurat's AddModuleScore, confounding from cycling cells
688 using Seurat's CellCycleScoring, and the percent of mitochondrial genes for each cell were
689 regressed out during SCTransform normalization. Objects were integrated using the top 3000
690 most variable features excluding DIGs, ribosomal protein genes, T cell receptor genes (for
691 lymphoid cell clustering), *MALAT1*, and cell cycle-specific genes to avoid confounding in
692 dimensionality reduction and clustering. Harmony was run using 15 dimensions and setting the
693 sample of origin as the blocking variable for batch effect correction. UMAP embeddings were
694 obtained using the first 15 Harmony dimensions, and subclusters were obtained by calculating
695 the k-nearest neighbors (k-NN) from a shared nearest-neighbor graph using the Louvain
696 algorithm implemented in the FindClusters function of the Seurat package at different
697 resolutions continuously from 0.4 to 2.0. The "clustree" R package[54] was used to determine the
698 optimal resolution for each macro-group. To label the final cell states, both FindAllMarkers and
699 limma-voom were used to perform differential gene expression testing and identify marker
700 genes for each subcluster. Seurat's FindAllMarkers was run with default parameters except for
701 min.pct = 0.25 and only.pos = TRUE. Limma-voom was run after converting the scRNA-seq data
702 to pseudobulks for each sample as previously described[55,56]. Cell states were defined by
703 comparing the identified marker genes within each cluster with the expression profile of known
704 cell states reported in the literature.
705
706 *CIBERSORTx deconvolution of bulk RNA-seq data*

707 Separate scRNA-seq signatures for immune and non-immune cells were built as above
708 and used as input to run CIBERSORTx in raw UMI counts format[21] with pre- and post-treatment
709 bulk RNA-seq tumor samples normalized to counts per million (CPM). CIBERSORTx was then
710 run in single cell mode using S-mode batch correction and setting the number of permutations
711 for p-value calculation to 500. The bulk RNA-seq data was then separately deconvoluted using
712 the TR4 signature matrix, which consists of immune, epithelial, endothelial, and fibroblast cells
713 derived from bulk RNA-seq of flow-sorted cellular populations[21]. The abundance of different cell
714 types recovered using the TR4 signature was used to normalize the abundance of immune and
715 non-immune single-cell states recovered using CIBERSORTx in single-cell mode, so that for
716 each sample, the total cell fractions summed to 1.0. Fractions were then centered log-ratio
717 transformed using the "compositions" R package[49] before downstream analysis.
718
719 *Differential gene expression and gene set enrichment analysis*
720 Differential gene expression analysis was performed using DESeq2[57] with treatment
721 arm, time point, and the interaction between the treatment arm and time point included in the
722 design. The effect of RT on gene expression (**Fig. 6B**) was analyzed by contrasting gene
723 expression in post-RT vs. pre-treatment samples, and the effect of pembrolizumab on gene
724 expression after treatment (**Fig. 6C**) was analyzed using the interaction term. Gene set
725 enrichment analysis for the Hallmark gene sets was performed for each comparison with the
726 "fgsea" R package with genes ranked by the $-\log_{10}$(p-value) multiplied by the sign of the $\log_2$(fold
727 change).
728
729 *T cell receptor repertoire deconvolution*
730 The T cell receptor repertoire for each sample was deconvoluted from raw bulk RNA-seq
731 FASTQ files using TRUST4[33]. A database of sequences for V, J, and C human genes was built
732 from the International ImMunoGeneTics information system and used for analysis[58].

733

734 *Statistical analysis*

735 The primary endpoint of SU2C-SARC032 was disease-free survival. SU2C-SARC032
736 was designed to detect a 22% increase in 2-year disease free survival from 50% in the control
737 group to 72% in the experimental group with 80% power at a one-sided 0.05 level of
738 significance. The protocol-defined primary statistical analysis plan was to compare disease-free
739 survival between the two treatment groups using a stratified log-rank test in a modified intention-
740 to-treat population that consisted of all evaluable patients. The prespecified significance level for
741 the primary analysis was a one-sided *P* value of 0.05. Secondary outcomes included distant
742 disease-free survival, local recurrence-free survival, overall survival, wound complications, and
743 treatment-related adverse events as assessed by CTCAE v4.0. Exploratory outcomes included
744 characterizing the peripheral and intra-tumoral immune response to combined pembrolizumab
745 and RT and identifying predictors of clinical response to pembrolizumab and RT. On April 15,
746 2020, the protocol was amended to include a preplanned subgroup analysis to compare primary
747 and secondary endpoints for patients in the control versus experimental arm for each sarcoma
748 immune class. The study was not powered for secondary or exploratory analyses. A run-in
749 safety analysis was performed after the first three patients were accrued to the experimental
750 arm and observed for 30 days after cycle 3 of pembrolizumab to assess for grade 4 or greater
751 toxicities. The primary endpoint analysis was triggered once 45 disease-free survival events
752 were observed, and outcomes were masked to study investigators until this time. The primary
753 and secondary outcomes, including all adverse events were reported previously[12]. Only
754 evaluable patients were included in exploratory analyses, and patients with missing data were
755 excluded.

756 The DFS interval was defined as the time from randomization to disease recurrence or
757 death. Patients who were alive without documented disease recurrence were censored at the
758 time of their last disease evaluation. For associations between cell subsets and DFS in samples

759 collected after randomization, we performed a landmark analysis to protect against guarantee-
760 time bias. Specifically, the DFS interval was calculated from the time of sample collection, and
761 patients that had a DFS event before the time of sample collection were excluded. In all cases,
762 DFS was calculated using the Kaplan-Meier method, and groups were compared using a two-
763 sided log-rank test. Hazard ratios were obtained by fitting univariable Cox proportional hazards
764 models using the coxph function from the "survival" R package, and the association of each
765 individual variable with DFS was calculated using two-sided Wald tests. Distributions of paired
766 samples were compared using paired Wilcoxon signed-rank tests, and distributions of unpaired
767 samples were compared using two-sided Wilcoxon rank-sum tests. When comparing
768 distributions of more than two groups, Kruskal-Wallis tests were used to test for overall
769 differences among groups, followed by post-hoc Dunn's tests for pairwise comparisons.
770 Proportions were compared using two-sided Fisher's exact tests. Correlations between
771 variables were measured using Spearman correlation. Multiple hypothesis testing correction
772 was performed using the Benjamini-Hochberg method. No threshold for significance was
773 prespecified due to the exploratory nature of the correlative analyses, and nominal *P* values
774 were reported. All analyses were performed either in R version 4.3.1 or using Prism 8
775 (GraphPad Software).
776
777 *Data and code availability*
778 The RNA sequencing data generated in this study have been deposited in the Database
779 of Genotypes and Phenotypes at phs003921.v1.p1 and will be made publicly available upon
780 publication. The scRNA-seq UPS/LPS atlas was built from previously published data available
781 from GEO through the accession code GSE279852 and GSE212527. The code used to perform
782 the analyses reported in the manuscript has been deposited at
783 https://github.com/StefTesta/SARC032_correlative. The EcoTyper code is available for nonprofit

784 academic use at https://github.com/ejmoding/sarcomaecotyper. All other datasets and
785 information necessary to replicate this work are provided in the supplementary tables. Any
786 additional information is available from the corresponding authors upon reasonable request.
787

788 **Acknowledgements**

789 We thank the patients and caregivers who participated in this study. We thank Denise
790 Reinke, Steven Young, Scott Okuno, and the research coordinators at SARC for supporting this
791 trial by collecting patient data and samples. In particular, we thank Erin Kozlowski and Meghan
792 Blalock who provided expert project management at SARC. We thank the UNC core for mass
793 cytometry that provided necessary equipment and materials for the CyTOF analysis. We also
794 thank the participating clinicians and clinical research teams at the participating sites for
795 enrolling patients to make this international trial and correlative studies possible. This work was
796 supported by a Stand Up To Cancer Catalyst Research Team grant with support from Merck
797 (DGK, KJW, EJM, and YMM), a research grant from the Investigator-Initiated Studies Program
798 of Merck Sharp & Dohme Corp (DGK), the National Cancer Institute (EJM: K08CA25542501;
799 DGK: R35CA197616), the Department of Defense (EJM: W81XWH-22-1-0161), the My Blue
800 Dots organization (EJM), the Tad and Diane Taube Family Foundation (MvR and EJM), and the
801 GPA Andrew Ursini Charitable Fund (AMH). The opinions expressed in this paper are those of
802 the authors and do not necessarily represent those of Merck Sharp & Dohme Corp. The
803 schematic in Fig.1 was created with BioRender.com.
804

805 **References**

806 1. Dufresne, A., Brahmi, M., Karanian, M. & Blay, J.-Y. Using biology to guide the treatment of
807 sarcomas and aggressive connective-tissue tumours. *Nat Rev Clin Oncol* **15**, 443–458
808 (2018).

809 2. Von Mehren, M. *et al.* Soft Tissue Sarcoma, Version 2.2022, NCCN Clinical Practice
810 Guidelines in Oncology. *Journal of the National Comprehensive Cancer Network* **20**, 815–
811 833 (2022).
812 3. Wang, D. *et al.* Significant Reduction of Late Toxicities in Patients With Extremity Sarcoma
813 Treated With Image-Guided Radiation Therapy to a Reduced Target Volume: Results of
814 Radiation Therapy Oncology Group RTOG-0630 Trial. *J Clin Oncol* **33**, 2231–2238 (2015).
815 4. Tap, W. D. *et al.* Doxorubicin plus evofosfamide versus doxorubicin alone in locally
816 advanced, unresectable or metastatic soft-tissue sarcoma (TH CR-406/SARC021): an
817 international, multicentre, open-label, randomised phase 3 trial. *Lancet Oncol* **18**, 1089–
818 1103 (2017).
819 5. Gronchi, A. *et al.* Histotype-tailored neoadjuvant chemotherapy versus standard
820 chemotherapy in patients with high-risk soft-tissue sarcomas (ISG-STS 1001): an
821 international, open-label, randomised, controlled, phase 3, multicentre trial. *Lancet Oncol*
822 **18**, 812–822 (2017).
823 6. Blay, J.-Y. & Le Cesne, A. Adjuvant chemotherapy in localized soft tissue sarcomas: still not
824 proven. *Oncologist* **14**, 1013–1020 (2009).
825 7. Tawbi, H. A. *et al.* Pembrolizumab in advanced soft-tissue sarcoma and bone sarcoma
826 (SARC028): a multicentre, two-cohort, single-arm, open-label, phase 2 trial. *Lancet Oncol*
827 **18**, 1493–1501 (2017).
828 8. Burgess, M. A. *et al.* Clinical activity of pembrolizumab (P) in undifferentiated pleomorphic
829 sarcoma (UPS) and dedifferentiated/pleomorphic liposarcoma (LPS): Final results of
830 SARC028 expansion cohorts. *JCO* **37**, 11015–11015 (2019).
831 9. Lee, A. Q., Hao, C., Pan, M., Ganjoo, K. N. & Bui, N. Q. Histologic and Immunologic Factors
832 Associated with Response to Immune Checkpoint Inhibitors in Advanced Sarcoma. *Clinical*
833 *Cancer Research* **31**, 678–684 (2025).

834 10. Sharabi, A. B., Lim, M., DeWeese, T. L. & Drake, C. G. Radiation and checkpoint blockade
835 immunotherapy: radiosensitisation and potential mechanisms of synergy. *Lancet Oncol* **16**,
836 e498-509 (2015).
837 11. Wisdom, A. J. *et al.* Single cell analysis reveals distinct immune landscapes in transplant
838 and primary sarcomas that determine response or resistance to immunotherapy. *Nat*
839 *Commun* **11**, 6410 (2020).
840 12. Mowery, Y. M. *et al.* Safety and efficacy of pembrolizumab, radiation therapy, and surgery
841 versus radiation therapy and surgery for stage III soft tissue sarcoma of the extremity
842 (SU2C-SARC032): an open-label, randomised clinical trial. *Lancet* **404**, 2053–2064 (2024).
843 13. Keung, E. Z. *et al.* Correlative Analyses of the SARC028 Trial Reveal an Association
844 Between Sarcoma-Associated Immune Infiltrate and Response to Pembrolizumab. *Clin*
845 *Cancer Res* **26**, 1258–1266 (2020).
846 14. Petitprez, F. *et al.* B cells are associated with survival and immunotherapy response in
847 sarcoma. *Nature* **577**, 556–560 (2020).
848 15. Subramanian, A. *et al.* Sarcoma microenvironment cell states and ecosystems are
849 associated with prognosis and predict response to immunotherapy. *Nat Cancer* **5**, 642–658
850 (2024).
851 16. Becht, E. *et al.* Estimating the population abundance of tissue-infiltrating immune and
852 stromal cell populations using gene expression. *Genome Biol* **17**, 218 (2016).
853 17. Luca, B. A. *et al.* Atlas of clinically distinct cell states and ecosystems across human solid
854 tumors. *Cell* **184**, 5482-5496.e28 (2021).
855 18. Ayers, M. *et al.* IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade. *J*
856 *Clin Invest* **127**, 2930–2940 (2017).
857 19. Dyikanov, D. *et al.* Comprehensive peripheral blood immunoprofiling reveals five
858 immunotypes with immunotherapy response characteristics in patients with cancer. *Cancer*
859 *Cell* **42**, 759-779.e12 (2024).

860 20. Testa, S. *et al.* SCAN-ACT: adoptive T cell therapy target discovery through single-cell
861 transcriptomics. *Genome Med* **17**, 89 (2025).
862 21. Newman, A. M. *et al.* Determining cell type abundance and expression from bulk tissues
863 with digital cytometry. *Nat Biotechnol* **37**, 773–782 (2019).
864 22. Cords, L. *et al.* Cancer-associated fibroblast classification in single-cell and spatial
865 proteomics data. *Nat Commun* **14**, 4294 (2023).
866 23. Cui, X., Liu, S., Song, H., Xu, J. & Sun, Y. Single-cell and spatial transcriptomic analyses
867 revealing tumor microenvironment remodeling after neoadjuvant chemoimmunotherapy in
868 non-small cell lung cancer. *Mol Cancer* **24**, 111 (2025).
869 24. Bles, N., Di Pietrantonio, L., Boeynaems, J.-M. & Communi, D. ATP confers tumorigenic
870 properties to dendritic cells by inducing amphiregulin secretion. *Blood* **116**, 3219–3226
871 (2010).
872 25. Khan, O. *et al.* TOX transcriptionally and epigenetically programs CD8+ T cell exhaustion.
873 *Nature* **571**, 211–218 (2019).
874 26. Blackburn, S. D. *et al.* Coregulation of CD8+ T cell exhaustion by multiple inhibitory
875 receptors during chronic viral infection. *Nat Immunol* **10**, 29–37 (2009).
876 27. Chen, Y. *et al.* Tumor-associated monocytes promote mesenchymal transformation through
877 EGFR signaling in glioma. *Cell Reports Medicine* **4**, 101177 (2023).
878 28. Brandt, C. S. *et al.* The B7 family member B7-H6 is a tumor cell ligand for the activating
879 natural killer cell receptor NKp30 in humans. *J Exp Med* **206**, 1495–1503 (2009).
880 29. Sheffer, M. *et al.* Genome-scale screens identify factors regulating tumor cell responses to
881 natural killer cells. *Nat Genet* **53**, 1196–1206 (2021).
882 30. Kilian, M. *et al.* The immunoglobulin superfamily ligand B7H6 subjects T cell responses to
883 NK cell surveillance. *Sci. Immunol.* **9**, eadj7970 (2024).
884 31. Torrejon, D. Y. *et al.* Antitumor Immune Responses in B2M-Deficient Cancers. *Cancer*
885 *Immunol Res* **11**, 1642–1655 (2023).

886 32. Li, Y. *et al.* Unlocking the therapeutic potential of the NKG2A-HLA-E immune checkpoint
887 pathway in T cells and NK cells for cancer immunotherapy. *J Immunother Cancer* **12**,
888 e009934 (2024).
889 33. Song, L. *et al.* TRUST4: immune repertoire reconstruction from bulk and single-cell RNA-
890 seq data. *Nat Methods* **18**, 627–630 (2021).
891 34. Gong, J., Le, T. Q., Massarelli, E., Hendifar, A. E. & Tuli, R. Radiation therapy and PD-1/PD-
892 L1 blockade: the clinical development of an evolving anticancer combination. *j.*
893 *immunotherapy cancer* **6**, 46 (2018).
894 35. Ribas, A. & Wolchok, J. D. Cancer immunotherapy using checkpoint blockade. *Science* **359**,
895 1350–1355 (2018).
896 36. He, X. & Xu, C. Immune checkpoint signaling and cancer immunotherapy. *Cell Res* **30**, 660–
897 669 (2020).
898 37. Kumagai, S. *et al.* The PD-1 expression balance between effector and regulatory T cells
899 predicts the clinical efficacy of PD-1 blockade therapies. *Nat Immunol* **21**, 1346–1358
900 (2020).
901 38. Roland, C. L. *et al.* A randomized, non-comparative phase 2 study of neoadjuvant immune-
902 checkpoint blockade in retroperitoneal dedifferentiated liposarcoma and extremity/truncal
903 undifferentiated pleomorphic sarcoma. *Nat Cancer* **5**, 625–641 (2024).
904 39. Galluzzi, L., Buqué, A., Kepp, O., Zitvogel, L. & Kroemer, G. Immunogenic cell death in
905 cancer and infectious disease. *Nat Rev Immunol* **17**, 97–111 (2017).
906 40. Golden, E. B., Pellicciotta, I., Demaria, S., Barcellos-Hoff, M. H. & Formenti, S. C. The
907 convergence of radiation and immunogenic cell death signaling pathways. *Front. Oncol.* **2**,
908 (2012).
909 41. Buchwald, Z. S. *et al.* Radiation, Immune Checkpoint Blockade and the Abscopal Effect: A
910 Critical Review on Timing, Dose and Fractionation. *Front. Oncol.* **8**, 612 (2018).

911 42. Haniball, J. *et al.* Prognostic factors and metastatic patterns in primary myxoid/round-cell
912 liposarcoma. *Sarcoma* **2011**, 538085 (2011).
913 43. Callegaro, D. *et al.* Development and external validation of two nomograms to predict
914 overall survival and occurrence of distant metastases in adults after surgical resection of
915 localised soft-tissue sarcomas of the extremities: a retrospective analysis. *Lancet Oncol* **17**,
916 671–680 (2016).
917 44. Peiper, M., Zurakowski, D., Knoefel, W. T. & Izbicki, J. R. Malignant fibrous histiocytoma of
918 the extremities and trunk: an institutional review. *Surgery* **135**, 59–66 (2004).
919 45. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and
920 bias-aware quantification of transcript expression. *Nat Methods* **14**, 417–419 (2017).
921 46. Tatlow, P. J. & Piccolo, S. R. A cloud-based workflow to quantify transcript-expression levels
922 in public cancer compendia. *Sci Rep* **6**, 39259 (2016).
923 47. Ellis, B. *et al. flowCore: flowCore: Basic Structures for Flow Cytometry Data*. (2025).
924 48. Levine, J. H. *et al.* Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells
925 that Correlate with Prognosis. *Cell* **162**, 184–197 (2015).
926 49. Boogaart, K. G. van den, Tolosana-Delgado, R. & Bren, M. *Compositions: Compositional*
927 *Data Analysis*. (2024).
928 50. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of
929 single-cell gene expression data. *Nature Biotechnology* **33**, 495–502 (2015).
930 51. Ahlmann-Eltze, C. & Huber, W. glmGamPoi: Fitting Gamma-Poisson Generalized Linear
931 Models on Single Cell Count Data. *Bioinformatics*
932 https://doi.org/10.1093/bioinformatics/btaa1009 (2020) doi:10.1093/bioinformatics/btaa1009.
933 52. Korsunsky, I. *et al. Harmony: Fast, Sensitive, and Accurate Integration of Single Cell Data*.
934 (2023).
935 53. Van Den Brink, S. C. *et al.* Single-cell sequencing reveals dissociation-induced gene
936 expression in tissue subpopulations. *Nat Methods* **14**, 935–936 (2017).

937 54. Zappia, L. & Oshlack, A. Clustering trees: a visualization for evaluating clusterings at
938 multiple resolutions. *GigaScience* **7**, (2018).
939 55. Testa, S. *et al.* SCAN-ACT: Adoptive T Cell Therapy Target Discovery Through Single-Cell
940 Transcriptomics. Preprint at https://doi.org/10.1101/2025.01.18.633736 (2025).
941 56. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and
942 microarray studies. *Nucleic Acids Res* **43**, e47 (2015).
943 57. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for
944 RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
945 58. Lefranc, M.-P. *et al.* IMGT, the international ImMunoGeneTics information system. *Nucleic*
946 *Acids Res* **37**, D1006-1012 (2009).
947

Figure 1

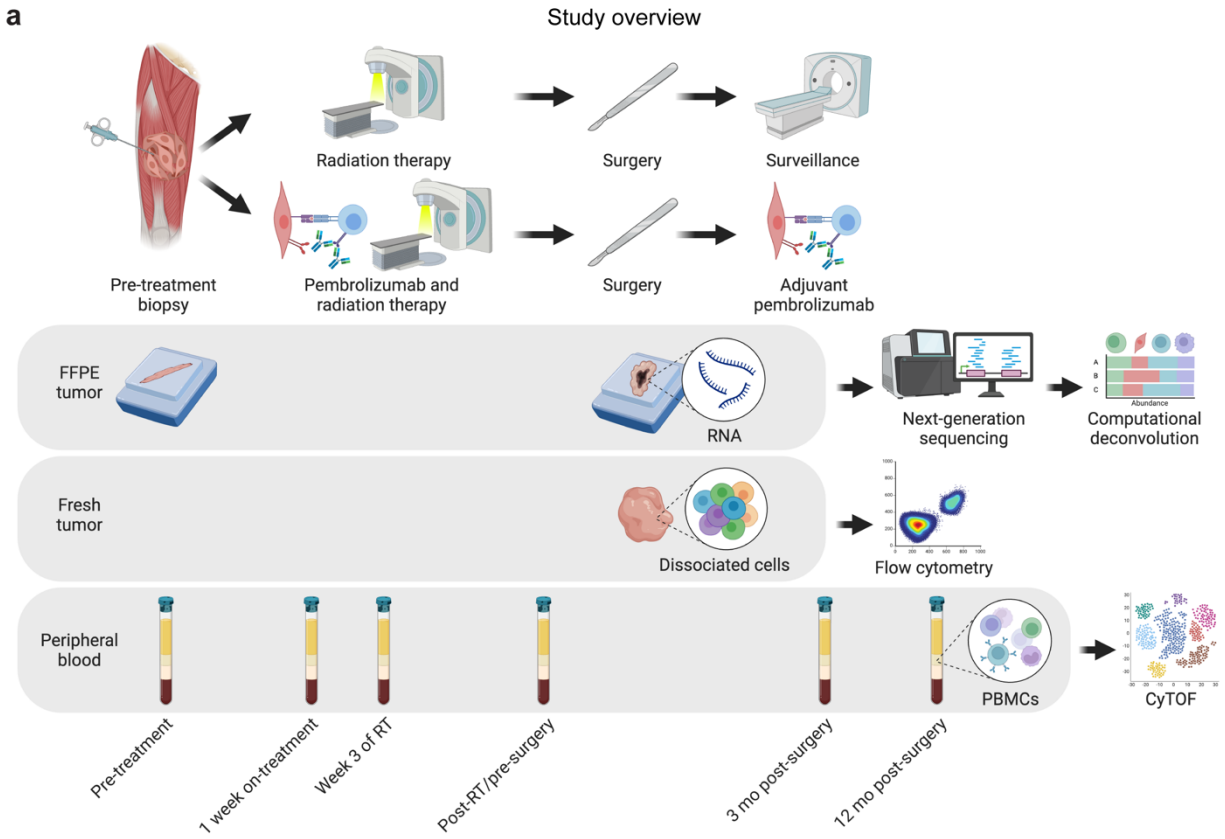


Fig. 1: Overview of multi-omics tumor and peripheral blood analyses of patients treated in SU2C-SARC032. **a**, Patients were randomized to neoadjuvant radiotherapy followed by surgery (control) or 3 cycles of neoadjuvant pembrolizumab and radiotherapy followed by surgery and up to 14 cycles of adjuvant pembrolizumab (experimental). Bulk RNA was isolated from formalin-fixed paraffin embedded (FFPE) pre-treatment tumor biopsies and surgical resection tumor specimens prior to next-generation sequencing and digital cytometry to measure cell state abundances. Fresh surgical resection specimens after preoperative radiotherapy with or without pembrolizumab were dissociated prior to analysis with flow cytometry. Peripheral blood mononuclear cells (PBMCs) were isolated from longitudinal peripheral blood samples for analysis using cytometry by time-of-flight (CyTOF).

Figure 2

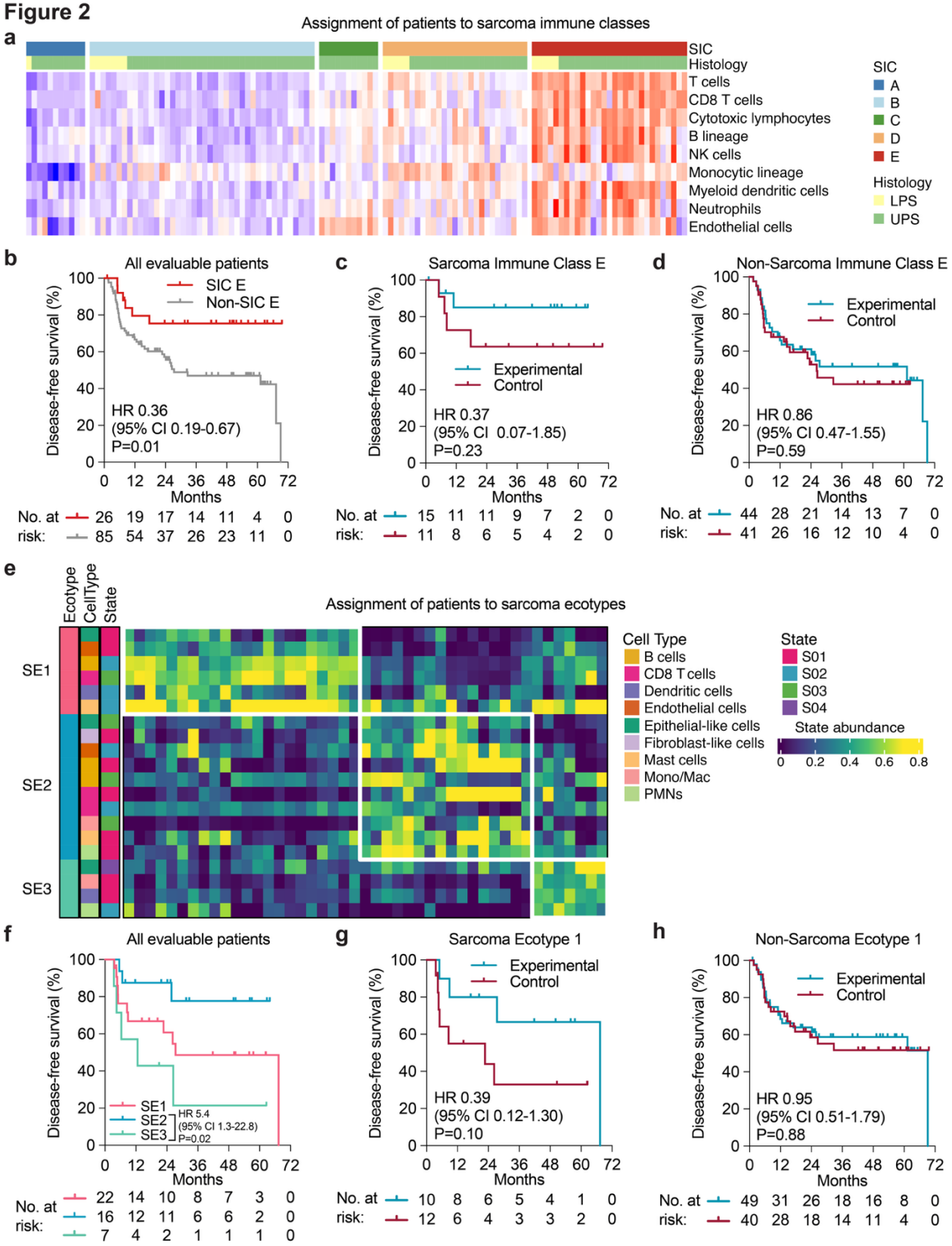

a
Assignment of patients to sarcoma immune classes
SIC
Histology
T cells
CD8 T cells
Cytotoxic lymphocytes
B lineage
NK cells
Monocytic lineage
Myeloid dendritic cells
Neutrophils
Endothelial cells
SIC
A
B
C
D
E
Histology
LPS
UPS
b
All evaluable patients
SIC E
Non-SIC E
Disease-free survival (%)
HR 0.36
(95% CI 0.19-0.67)
P=0.01
Months
No. at risk:
26 19 17 14 11 4 0
85 54 37 26 23 11 0
c
Sarcoma Immune Class E
Experimental
Control
HR 0.37
(95% CI 0.07-1.85)
P=0.23
No. at risk:
15 11 11 9 7 2 0
11 8 6 5 4 2 0
d
Non-Sarcoma Immune Class E
Experimental
Control
HR 0.86
(95% CI 0.47-1.55)
P=0.59
No. at risk:
44 28 21 14 13 7 0
41 26 16 12 10 4 0
e
Ecotype
CellType
State
Assignment of patients to sarcoma ecotypes
SE1
SE2
SE3
Cell Type
B cells
CD8 T cells
Dendritic cells
Endothelial cells
Epithelial-like cells
Fibroblast-like cells
Mast cells
Mono/Mac
PMNs
State
S01
S02
S03
S04
State abundance
0 0.2 0.4 0.6 0.8
f
All evaluable patients
SE1
SE2
SE3
HR 5.4
(95% CI 1.3-22.8)
P=0.02
No. at risk:
22 14 10 8 7 3 0
16 12 11 6 6 2 0
7 4 2 1 1 1 0
g
Sarcoma Ecotype 1
Experimental
Control
HR 0.39
(95% CI 0.12-1.30)
P=0.10
No. at risk:
10 8 6 5 4 1 0
12 6 4 3 3 2 0
h
Non-Sarcoma Ecotype 1
Experimental
Control
HR 0.95
(95% CI 0.51-1.79)
P=0.88
No. at risk:
49 31 26 18 16 8 0
40 28 18 14 11 4 0

Fig. 2: Association of pre-treatment sarcoma microenvironment signatures with outcomes. a, Heat map of cell type abundances in the tumor microenvironment of pre-treatment sarcomas measured using MCP-counter. Samples are grouped by sarcoma immune class (SIC) assignment. (**b**, **c**, **d**), Kaplan-Meier curves showing disease-free survival (DFS) for (**b**) evaluable patients in SIC E versus all other SICs, (**c**) SIC E patients in the experimental versus control arm, and (**d**) non-SIC E patients in the experimental versus control arm. **e**, Heat map of cell state abundances in the tumor microenvironment of pre-treatment sarcomas measured using EcoTyper. Samples are grouped by sarcoma ecotype (SE) assignment. (**f**, **g**, **h**), Kaplan-Meier curves showing DFS for evaluable patients based on (**f**) SE assignment, (**g**) SE1 patients in the experimental versus control arm, and (**h**) non-SE1 patients in the experimental versus control arm. Only patients assigned to an SE are displayed in **f**, and patients not assigned to an SE are included in **g** and **h**. P-values were calculated using two-sided log-rank tests.

Figure 3

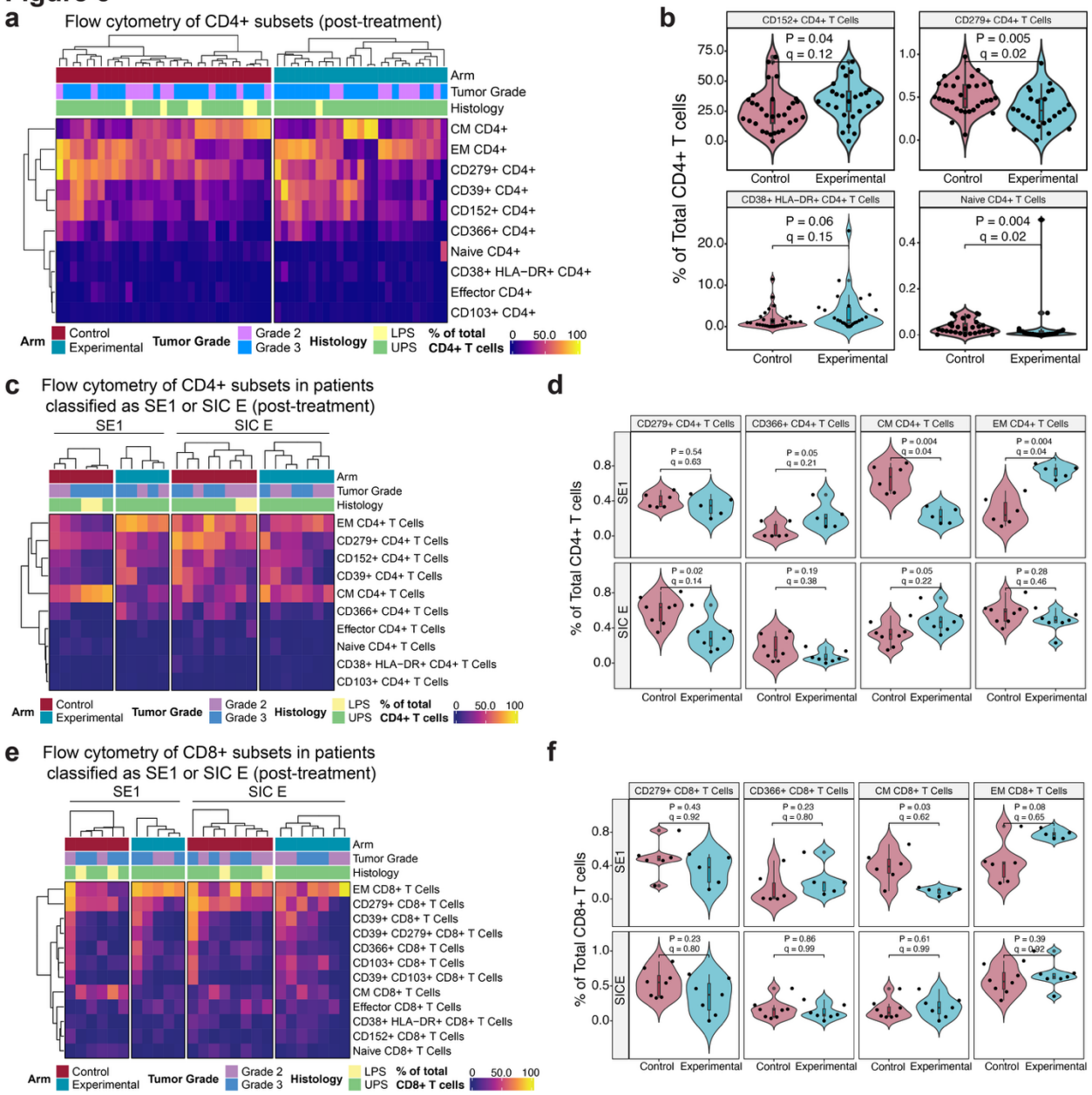

a
Flow cytometry of CD4+ subsets (post-treatment)
Arm
Tumor Grade
Histology
CM CD4+
EM CD4+
CD279+ CD4+
CD39+ CD4+
CD152+ CD4+
CD366+ CD4+
Naive CD4+
CD38+ HLA−DR+ CD4+
Effector CD4+
CD103+ CD4+
Arm Control Experimental
Tumor Grade Grade 2 Grade 3
Histology LPS UPS
% of total CD4+ T cells 0 50.0 100
b
CD152+ CD4+ T Cells
P = 0.04
q = 0.12
CD279+ CD4+ T Cells
P = 0.005
q = 0.02
CD38+ HLA−DR+ CD4+ T Cells
P = 0.06
q = 0.15
Naive CD4+ T Cells
P = 0.004
q = 0.02
% of Total CD4+ T cells
Control
Experimental
c
Flow cytometry of CD4+ subsets in patients classified as SE1 or SIC E (post-treatment)
SE1
SIC E
Arm
Tumor Grade
Histology
EM CD4+ T Cells
CD279+ CD4+ T Cells
CD152+ CD4+ T Cells
CD39+ CD4+ T Cells
CM CD4+ T Cells
CD366+ CD4+ T Cells
Effector CD4+ T Cells
Naive CD4+ T Cells
CD38+ HLA−DR+ CD4+ T Cells
CD103+ CD4+ T Cells
Arm Control Experimental
Tumor Grade Grade 2 Grade 3
Histology LPS UPS
% of total CD4+ T cells 0 50.0 100
d
CD279+ CD4+ T Cells
P = 0.54
q = 0.63
CD366+ CD4+ T Cells
P = 0.05
q = 0.21
CM CD4+ T Cells
P = 0.004
q = 0.04
EM CD4+ T Cells
P = 0.004
q = 0.04
SE1
P = 0.02
q = 0.14
P = 0.19
q = 0.38
P = 0.05
q = 0.22
P = 0.28
q = 0.46
SIC E
% of Total CD4+ T cells
Control Experimental
e
Flow cytometry of CD8+ subsets in patients classified as SE1 or SIC E (post-treatment)
SE1
SIC E
Arm
Tumor Grade
Histology
EM CD8+ T Cells
CD279+ CD8+ T Cells
CD39+ CD8+ T Cells
CD39+ CD279+ CD8+ T Cells
CD366+ CD8+ T Cells
CD103+ CD8+ T Cells
CD39+ CD103+ CD8+ T Cells
CM CD8+ T Cells
Effector CD8+ T Cells
CD38+ HLA−DR+ CD8+ T Cells
CD152+ CD8+ T Cells
Naive CD8+ T Cells
Arm Control Experimental
Tumor Grade Grade 2 Grade 3
Histology LPS UPS
% of total CD8+ T cells 0 50.0 100
f
CD279+ CD8+ T Cells
P = 0.43
q = 0.92
CD366+ CD8+ T Cells
P = 0.23
q = 0.80
CM CD8+ T Cells
P = 0.03
q = 0.62
EM CD8+ T Cells
P = 0.08
q = 0.65
SE1
P = 0.23
q = 0.80
P = 0.86
q = 0.99
P = 0.61
q = 0.99
P = 0.39
q = 0.92
SICE
% of Total CD8+ T cells
Control Experimental
g
Association of T cell subsets measured by flow cytometry in resected sarcomas with disease-free survival
Both
Control
Experimental
CD4+ T cells
CD8+ T cells
Total CD4+ (of CD45+)
CD279+ T−regs
CD103+ CD4+
CM CD8+
Total CD8+ (of CD45+)
−Log10P-value
Shorter DFS (P < 0.05)
Longer DFS (P < 0.05)
P = 0.05

Fig. 3: Flow cytometry analysis of post-treatment samples. **a-f**, Heat maps showing the abundance of (**a**, **c**) CD4+ T cell, and (**e**) CD8+ T cell subsets in resected sarcomas after radiotherapy with or without pembrolizumab measured by flow cytometry. Patient samples are grouped according to study arm. Panels **c** and **e** include only patients classified as SIC E or SE1 at baseline. Violin plots comparing the abundance of select (**b**, **d**) CD4+ T-cell, and (**f**) CD8+ T-cell subsets in resected sarcomas after radiotherapy with or without pembrolizumab measured by flow cytometry. Panels **d** and **f** include only patients classified as SIC E or SE1 at baseline. P-values were calculated using a Wilcoxon rank-sum test. Multiple hypothesis testing correction was applied using the Benjamini-Hochberg method (q-value). **g**, Bar plot showing the association of cell subset abundance with disease-free survival in all evaluable patients (both) and patients treated on the control and experimental arms calculated using univariable Cox proportional hazards models. P-values were calculated using two-sided Wald tests.

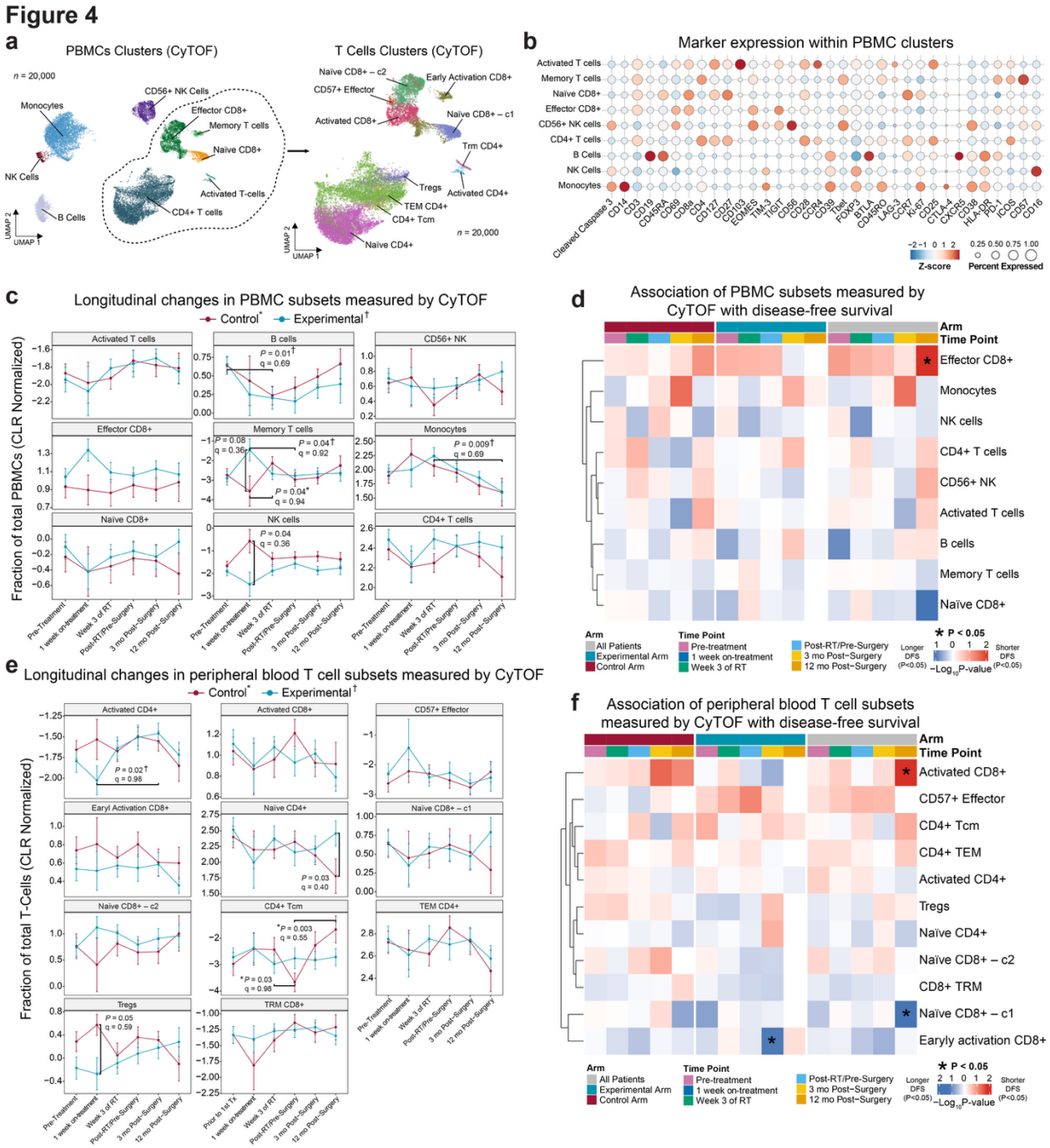


Fig. 4: Longitudinal CyTOF analysis of peripheral blood mononuclear cells (PMBCs). **a**, Unsupervised clustering of PBMCs (left) and T cell subsets (right) based on surface marker expression measured by CyTOF. 20,000 cells were randomly selected for plotting. **b**, Dot plot showing expression of surface markers in each cluster. **c-f**, Longitudinal changes in (**c**) PMBC subsets and (**e**) peripheral blood T cell subsets over time in the experimental and control arms. Plotted values represent centered log-ratio (CLR) normalized fractions of total PBMCs or T cells. P-values were calculated using two-sided Wilcoxon rank-sum tests. Multiple hypothesis testing

correction was applied using the Benjamini-Hochberg method (q-value). Heat maps showing the association of (**d**) PBMC subsets and (**f**) peripheral blood T cell subsets measured using CyTOF with disease-free survival in all evaluable patients and patients treated on the control and experimental arms calculated using univariate Cox proportional hazards models. P-values were computed using two-sided Wald tests without correction for multiple hypothesis testing. P-values < 0.05 are highlighted with an asterisk or a cross.

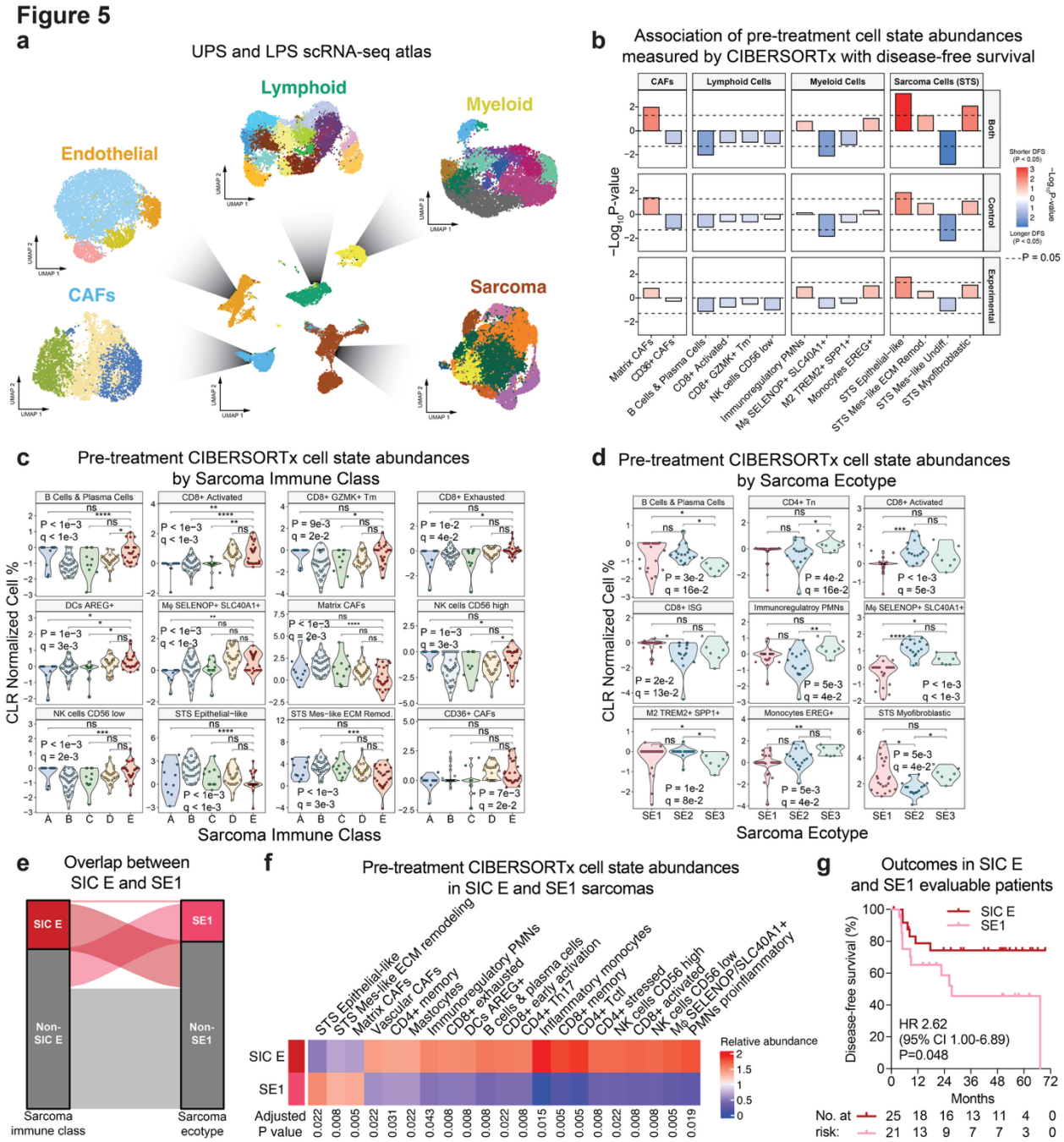


Fig. 5: Digital cytometry of the pre-treatment sarcoma microenvironment. **a**, UMAPs showing major cell type clusters and cell state subclusters defined from single cell RNA-sequencing (scRNA-seq) of undifferentiated pleomorphic sarcomas (UPS) and liposarcomas (LPS). This scRNA-seq atlas was used to create a signature matrix for digital cytometry of bulk RNA-seq data with CIBERSORTx. **b**, Bar plot showing the association of CIBERSORTx cell state abundances measured in pre-treatment biopsies with disease-free survival (DFS) in all evaluable patients (both) and patients treated on the control or experimental arms calculated using univariable Cox proportional hazards models. P-values were calculated using two-sided

Wald tests without correction for multiple hypothesis testing. **c** and **d**, Differences in CIBERSORTx cell states abundance in pre-treatment samples across (**c**) sarcoma immune classes and (**d**) sarcoma ecotypes. Plotted values represent centered log ratio (CLR) normalized fractions. P-values shown were computed using a Kruskal-Wallis test and pairwise comparisons between SICs or SEs were computed using a Dunn's test (ns: P > 0.05; *: P <= 0.05; **: P <= 1e-2; ***: P <= 1e-3; ****: P <= 1e-4). P-value adjustment for multiple hypothesis testing was done using the Benjamini-Hochberg method (q-value). **e**, River plot showing overlap between sarcoma immune class E (SIC E) and sarcoma ecotype 1 (SE1). The thickness of the bars represents the proportion of patients in each group. **f**, Heat map showing CIBERSORTx cell states with different pre-treatment abundances between SIC E and SE1 sarcomas. P-values were calculated using two-sided Wilcoxon rank-sum tests and corrected for multiple hypothesis testing using the Benjamini-Hochberg method. **g**, Kaplan-Meier curves showing the DFS for evaluable patients assigned to SIC E versus SE1 based on pre-treatment RNA-sequencing. P-value was calculated using a two-sided log-rank test. MΦ: macrophages.

Figure 6

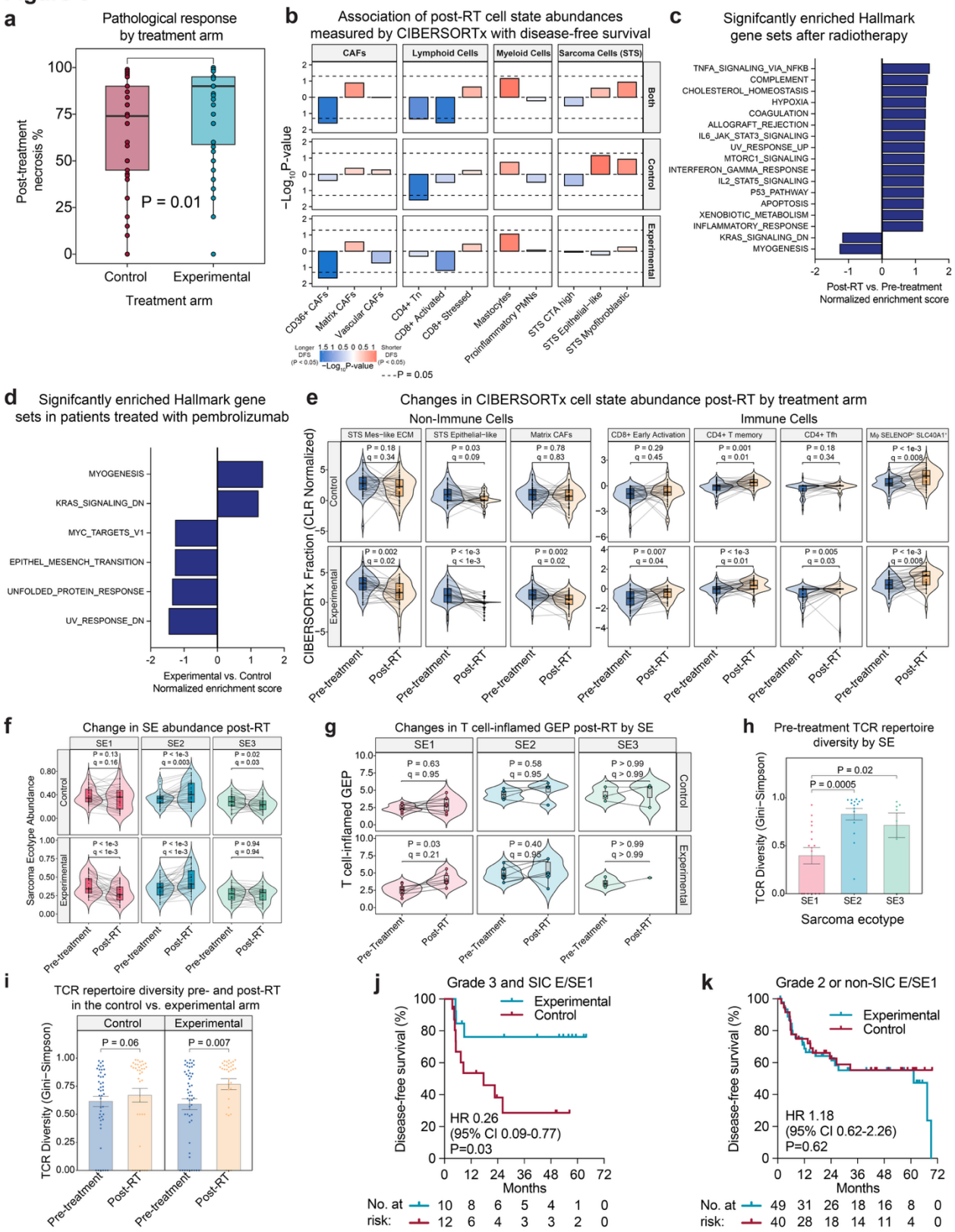

a
Pathological response by treatment arm
Post-treatment necrosis %
P = 0.01
Control
Experimental
Treatment arm
b
Association of post-RT cell state abundances measured by CIBERSORTx with disease-free survival
CAFs
Lymphoid Cells
Myeloid Cells
Sarcoma Cells (STS)
Both
Control
Experimental
−Log10 P-value
CD36+ CAFs
Matrix CAFs
Vascular CAFs
CD4+ Tn
CD8+ Activated
CD8+ Stressed
Mastocytes
Proinflammatory PMNs
STS CTA high
STS Epithelial−like
STS Myofibroblastic
Longer DFS (P < 0.05)
Shorter DFS (P < 0.05)
- - - P = 0.05
c
Signifcantly enriched Hallmark gene sets after radiotherapy
TNFA_SIGNALING_VIA_NFKB
COMPLEMENT
CHOLESTEROL_HOMEOSTASIS
HYPOXIA
COAGULATION
ALLOGRAFT_REJECTION
IL6_JAK_STAT3_SIGNALING
UV_RESPONSE_UP
MTORC1_SIGNALING
INTERFERON_GAMMA_RESPONSE
IL2_STAT5_SIGNALING
P53_PATHWAY
APOPTOSIS
XENOBIOTIC_METABOLISM
INFLAMMATORY_RESPONSE
KRAS_SIGNALING_DN
MYOGENESIS
Post-RT vs. Pre-treatment Normalized enrichment score
d
Signifcantly enriched Hallmark gene sets in patients treated with pembrolizumab
MYOGENESIS
KRAS_SIGNALING_DN
MYC_TARGETS_V1
EPITHEL_MESENCH_TRANSITION
UNFOLDED_PROTEIN_RESPONSE
UV_RESPONSE_DN
Experimental vs. Control Normalized enrichment score
e
Changes in CIBERSORTx cell state abundance post-RT by treatment arm
Non-Immune Cells
Immune Cells
STS Mes−like ECM
STS Epithelial−like
Matrix CAFs
CD8+ Early Activation
CD4+ T memory
CD4+ Tfh
Mϕ SELENOP+ SLC40A1+
CIBERSORTx Fraction (CLR Normalized)
Control
Experimental
Pre-treatment
Post-RT
f
Change in SE abundance post-RT
SE1
SE2
SE3
Sarcoma Ecotype Abundance
Control
Experimental
Pre-treatment
Post-RT
g
Changes in T cell-inflamed GEP post-RT by SE
SE1
SE2
SE3
T cell-inflamed GEP
Control
Experimental
Pre-Treatment
Post-RT
h
Pre-treatment TCR repertoire diversity by SE
P = 0.02
P = 0.0005
TCR Diversity (Gini−Simpson)
SE1
SE2
SE3
Sarcoma ecotype
i
TCR repertoire diversity pre- and post-RT in the control vs. experimental arm
Control
Experimental
P = 0.06
P = 0.007
TCR Diversity (Gini−Simpson)
Pre-treatment
Post-RT
j
Grade 3 and SIC E/SE1
Experimental
Control
Disease-free survival (%)
HR 0.26
(95% CI 0.09-0.77)
P=0.03
Months
No. at risk:
10 8 6 5 4 1 0
12 6 4 3 3 2 0
k
Grade 2 or non-SIC E/SE1
Experimental
Control
Disease-free survival (%)
HR 1.18
(95% CI 0.62-2.26)
P=0.62
Months
No. at risk:
49 31 26 18 16 8 0
40 28 18 14 11 4 0

Fig. 6: Effect of radiotherapy on the sarcoma microenvironment and an integrated predictor of benefit from pembrolizumab. a, Box plot comparing the post-treatment pathological response, measured as percent of tissue necrosis, in patients treated with RT alone versus those that received RT plus pembrolizumab. P value was obtained using a two-sided Wilcoxon rank-sum test. **b**, Bar plot showing the association of CIBERSORTx cell state abundances measured in post-radiotherapy (post-RT) surgically resected tumors with disease-free survival (DFS) in all evaluable patients (both) and patients treated on the control and experimental arms calculated using univariable Cox proportional hazards models. P-values were calculated using two-sided Wald tests without correction for multiple hypothesis testing. **c** and **d**, Bar plots showing gene set enrichment analysis (GSEA) results comparing (**c**) post-RT versus pre-treatment samples accounting for the effect of pembrolizumab and (**d**) experimental versus control samples measured as the interaction of pembrolizumab with radiotherapy. **e**, Box-violin plots showing the post-RT change in non-immune (*left*) and immune (*right*) cell state abundances measured using CIBERSORTx in patients treated in the control and experimental arm. Plotted values represent centered log ratio (CLR) normalized fractions. P-values were computed using Wilcoxon signed-rank tests and were corrected for multiple comparisons using the Benjamini-Hochberg method. **f**, Box-violin plots showing the abundance of sarcoma ecotypes before and after treatment in patients in the experimental and control arms. P-values were computed using a Wilcoxon signed-rank test and corrected for multiple comparisons using the Benjamini-Hochberg method. **g**, Box-violin plots displaying the change in T cell-inflamed GEP scores described by Ayers et al. (*Journal of Clinical Investigation*, 2017) from pre-treatment to post-RT grouped by sarcoma ecotype and treatment arm. P-values were calculated using a Wilcoxon signed-rank test and adjusted for multiple hypothesis testing using the Benjamini-Hochberg method. **h** and **i**, Bar plots showing the T cell receptor repertoire diversity as computed using TRUST4 in (**h**) pre-treatment samples grouped by SE assignment or (**i**) comparing pre-treatment to post-treatment samples in the control versus experimental arm. P-values were obtained using a two-sided Wilcoxon rank-sum test. **j** and **k**, Kaplan-Meier curves showing DFS by treatment arm for (**j**) patients with grade 3 tumors assigned to SIC E or SE1 and (**k**) patients with grade 2 tumors or tumors not assigned to SIC E or SE1. P-values were calculated using a two-sided log-rank tests. MΦ: macrophages.

Figure S1

a

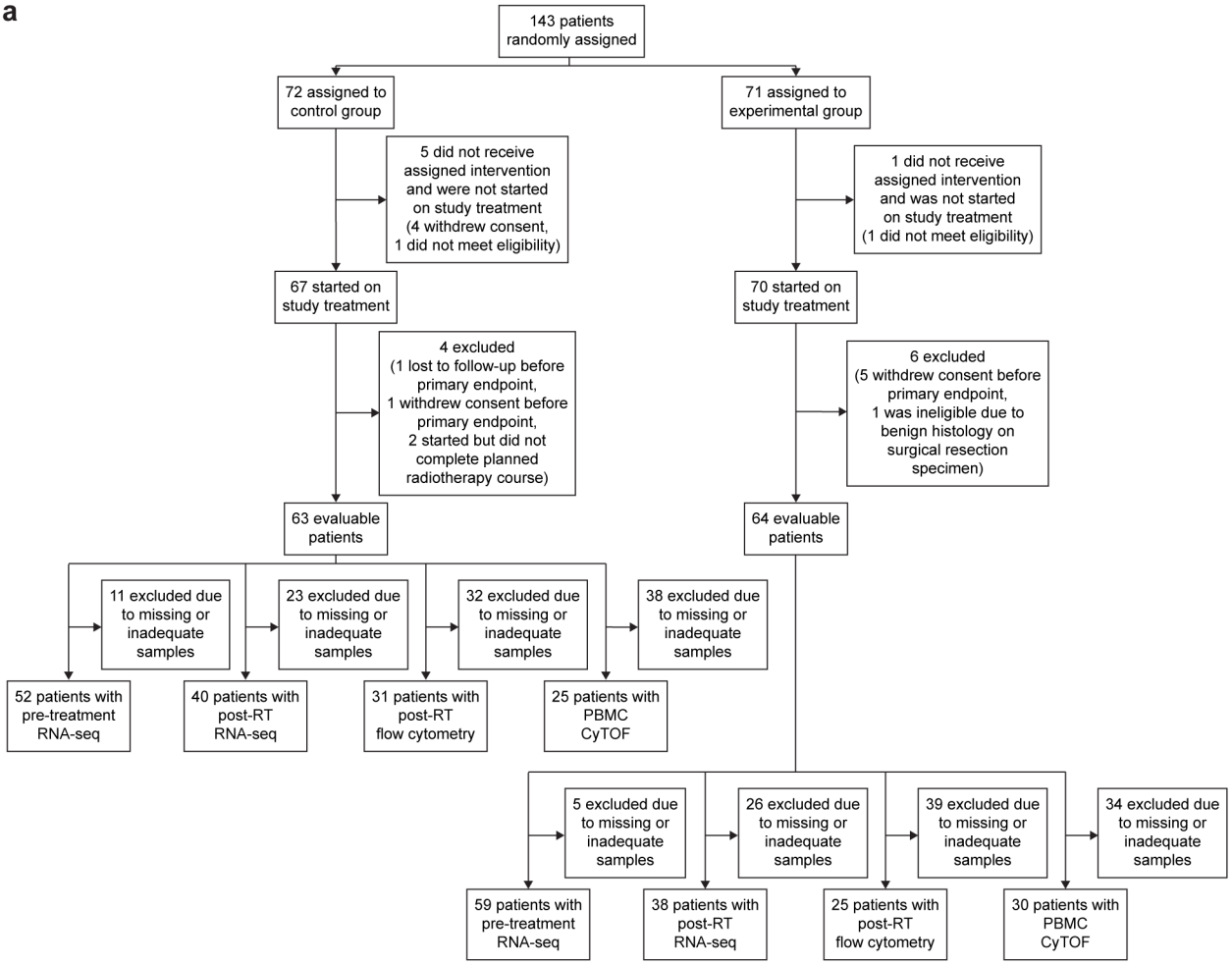
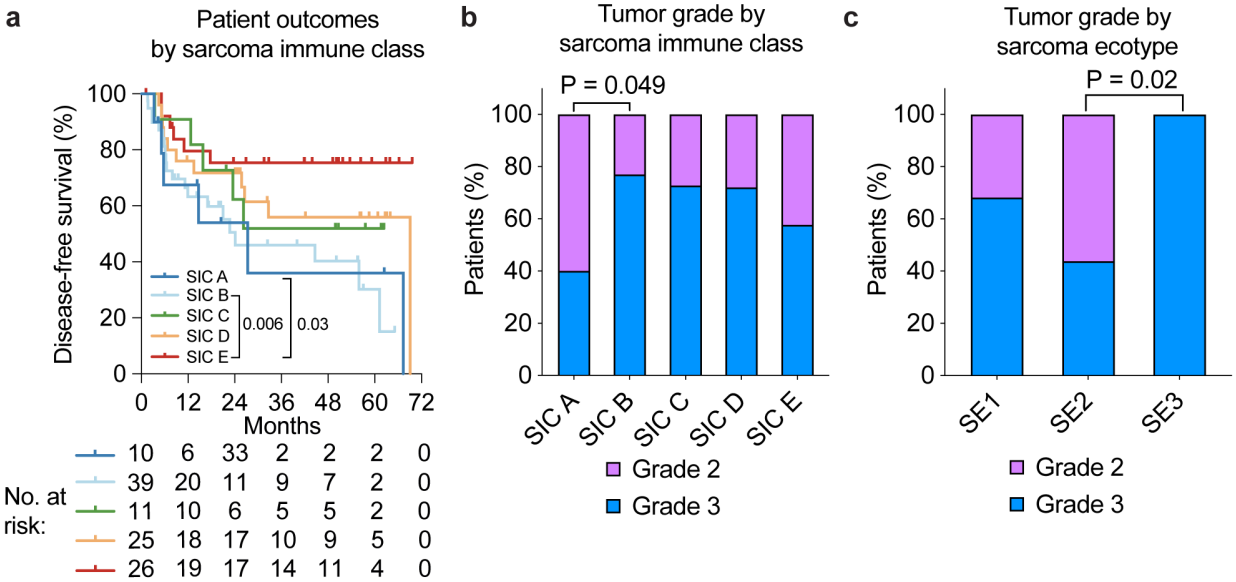
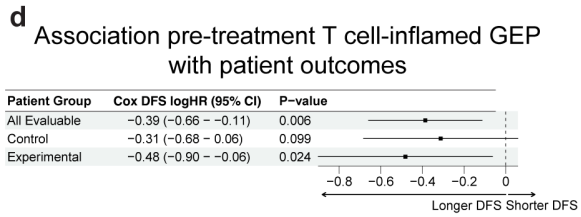
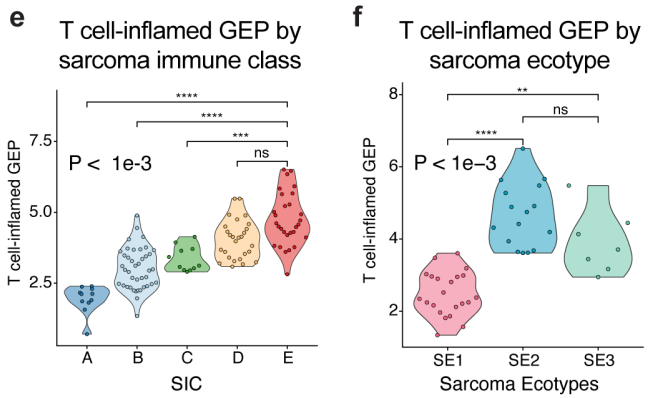


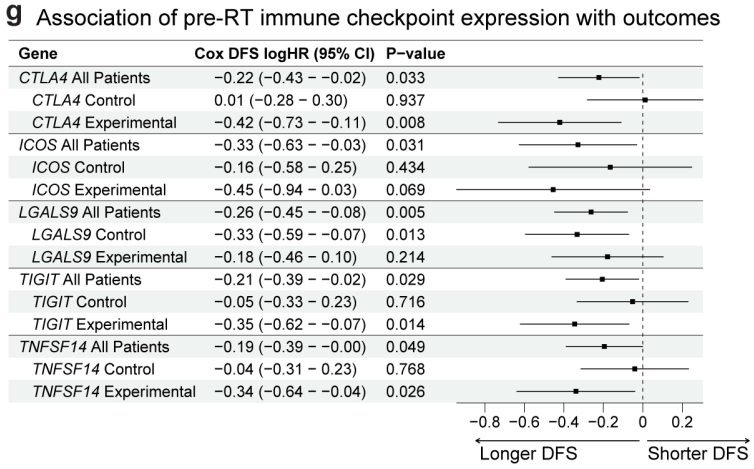
Fig. S1: CONSORT diagram. a, Flow diagram for the SU2C-SARC032 clinical trial. RT: radiotherapy, RNA-seq: RNA-sequencing, PBMC: peripheral blood mononuclear cells, CyTOF: Cytometry by Time of Flight.

Figure S2


a
Patient outcomes by sarcoma immune class
Disease-free survival (%)
Months
SIC A
SIC B
SIC C
SIC D
SIC E
0.006
0.03
No. at risk:
10 6 33 2 2 2 0
39 20 11 9 7 2 0
11 10 6 5 5 2 0
25 18 17 10 9 5 0
26 19 17 14 11 4 0
b
Tumor grade by sarcoma immune class
P = 0.049
Patients (%)
Grade 2
Grade 3
c
Tumor grade by sarcoma ecotype
P = 0.02
SE1
SE2
SE3


d
Association pre-treatment T cell-inflamed GEP with patient outcomes
Patient Group | Cox DFS logHR (95% CI) | P−value
All Evaluable | −0.39 (−0.66 − −0.11) | 0.006
Control | −0.31 (−0.68 − 0.06) | 0.099
Experimental | −0.48 (−0.90 − −0.06) | 0.024
Longer DFS Shorter DFS


e
T cell-inflamed GEP by sarcoma immune class
P < 1e-3
T cell-inflamed GEP
SIC
f
T cell-inflamed GEP by sarcoma ecotype
P < 1e−3
Sarcoma Ecotypes


g Association of pre-RT immune checkpoint expression with outcomes
Gene | Cox DFS logHR (95% CI) | P−value
CTLA4 All Patients | −0.22 (−0.43 − −0.02) | 0.033
CTLA4 Control | 0.01 (−0.28 − 0.30) | 0.937
CTLA4 Experimental | −0.42 (−0.73 − −0.11) | 0.008
ICOS All Patients | −0.33 (−0.63 − −0.03) | 0.031
ICOS Control | −0.16 (−0.58 − 0.25) | 0.434
ICOS Experimental | −0.45 (−0.94 − 0.03) | 0.069
LGALS9 All Patients | −0.26 (−0.45 − −0.08) | 0.005
LGALS9 Control | −0.33 (−0.59 − −0.07) | 0.013
LGALS9 Experimental | −0.18 (−0.46 − 0.10) | 0.214
TIGIT All Patients | −0.21 (−0.39 − −0.02) | 0.029
TIGIT Control | −0.05 (−0.33 − 0.23) | 0.716
TIGIT Experimental | −0.35 (−0.62 − −0.07) | 0.014
TNFSF14 All Patients | −0.19 (−0.39 − −0.00) | 0.049
TNFSF14 Control | −0.04 (−0.31 − 0.23) | 0.768
TNFSF14 Experimental | −0.34 (−0.64 − −0.04) | 0.026
Longer DFS
Shorter DFS

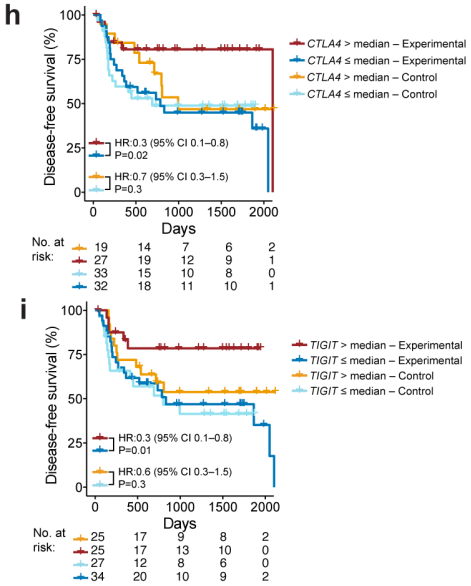

h
CTLA4 > median – Experimental
CTLA4 ≤ median – Experimental
CTLA4 > median – Control
CTLA4 ≤ median – Control
HR:0.3 (95% CI 0.1–0.8) P=0.02
HR:0.7 (95% CI 0.3–1.5) P=0.3
Disease-free survival (%)
Days
No. at risk:
19 14 7 6 2
27 19 12 9 1
33 15 10 8 0
32 18 11 10 1
i
TIGIT > median – Experimental
TIGIT ≤ median – Experimental
TIGIT > median – Control
TIGIT ≤ median – Control
HR:0.3 (95% CI 0.1–0.8) P=0.01
HR:0.6 (95% CI 0.3–1.5) P=0.3
No. at risk:
25 17 9 8 2
25 17 13 10 0
27 12 8 6 0
34 20 10 9 2

Fig. S2: Pre-treatment sarcoma microenvironment gene expression signatures. **a**, Kaplan-Meier curves showing disease-free survival (DFS) by SIC. P-values were calculated using two-sided log-rank tests. **b** and **c**, Proportion of patients with grade 2 or grade 3 tumors grouped by (**b**) sarcoma immune class (SIC) and (**c**) sarcoma ecotype (SE). P-values were calculated using two-sided Fisher's exact tests. **d**, Forest plot showing the association of the T cell-inflamed gene expression profile (GEP) described by Ayers et al. (*Journal of Clinical Investigation*, 2017) with DFS by treatment arm calculated using univariable Cox proportional hazards models. P-values were calculated using two-sided Wald tests. **e** and **f**, Violin plots displaying the distribution of T cell-inflamed GEP scores grouped by (**e**) SIC and (**f**) sarcoma ecotype. P-values were calculated using a Kruskal-Wallis test and pairwise comparisons between SICs or SEs were computed using a Dunn's test (ns: $p > 0.05$; *: $p <= 0.05$; **: $p <= 1e-2$; ***: $p <= 1e-3$; ****: $p <= 1e-4$). **g**, Forest plot showing the association of immune checkpoint gene expression with DFS by treatment arm calculated using univariable Cox proportional hazards models. P-values were calculated using two-sided Wald tests. **h** and **i**, Kaplan-Meier curves showing outcomes stratified based on pre-treatment (**h**) *CTLA4* and (**i**) *TIGIT* expression respectively. Gene expression is classified as high for a given sample if it is above the median for all pre-treatment samples. P-values were calculated using a two-sided Wald test.

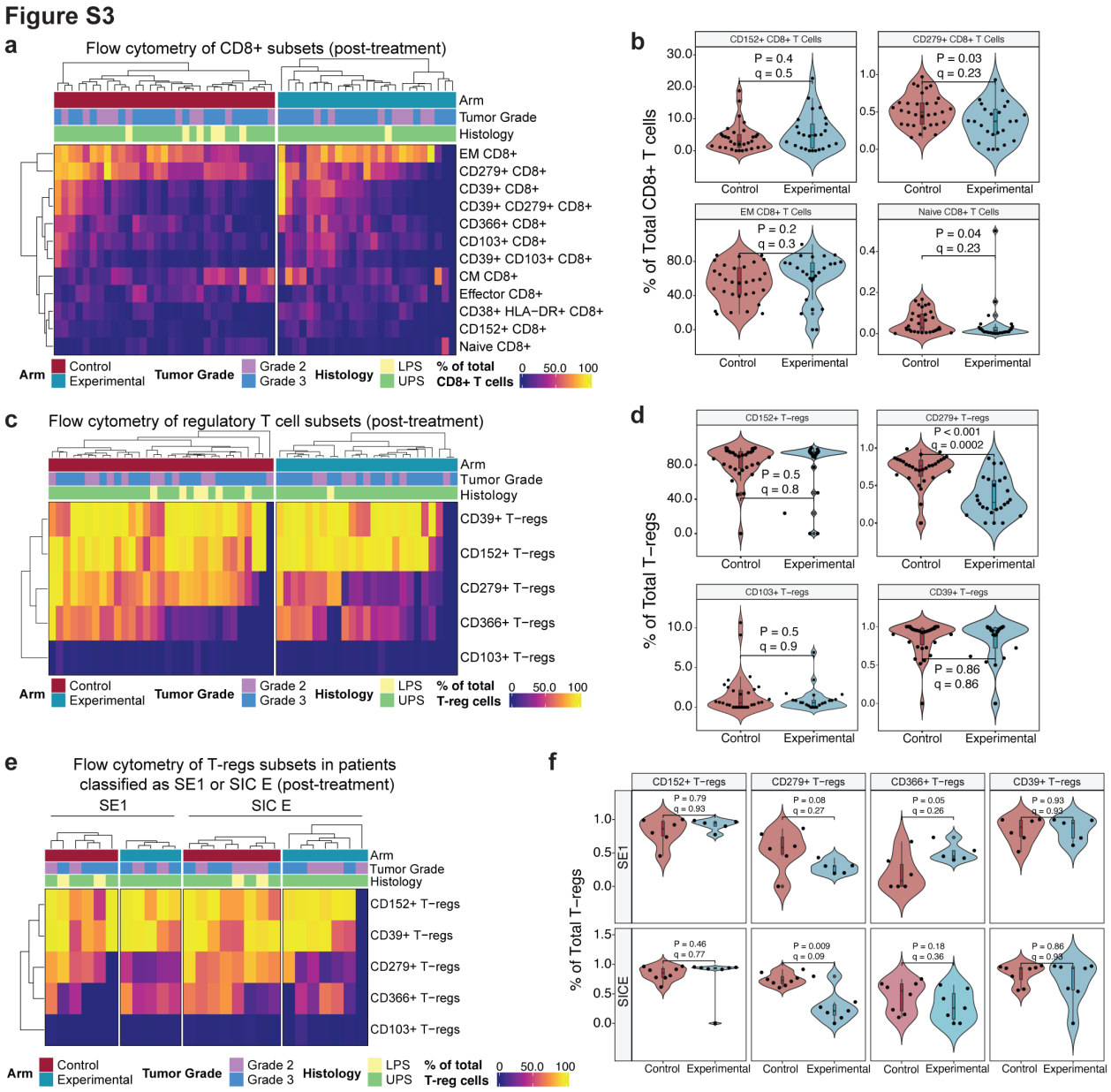


Fig. S3: **Flow cytometry of post-treatment samples**. **a-f**, Heat maps showing the abundance of (**a**) CD8+ T cell and (**c**, **e**) regulatory T cell subsets in resected sarcomas after radiotherapy with or without pembrolizumab measured by flow cytometry. Patient samples are grouped according to study arm. Panel **e** includes only patients classified as SIC E or SE1 at baseline. Violin plots comparing the abundance of select (**b**) CD8+ T cell, and (**d**, **f**) regulatory T cell subsets in resected sarcomas after radiotherapy with or without pembrolizumab measured by flow cytometry. Panel **f** shows only patients classified as SIC E or SE1 at baseline. P-values were calculated using two-sided Wilcoxon rank-sum tests. Multiple hypothesis testing correction was applied using the Benjamini-Hochberg method (q-value).

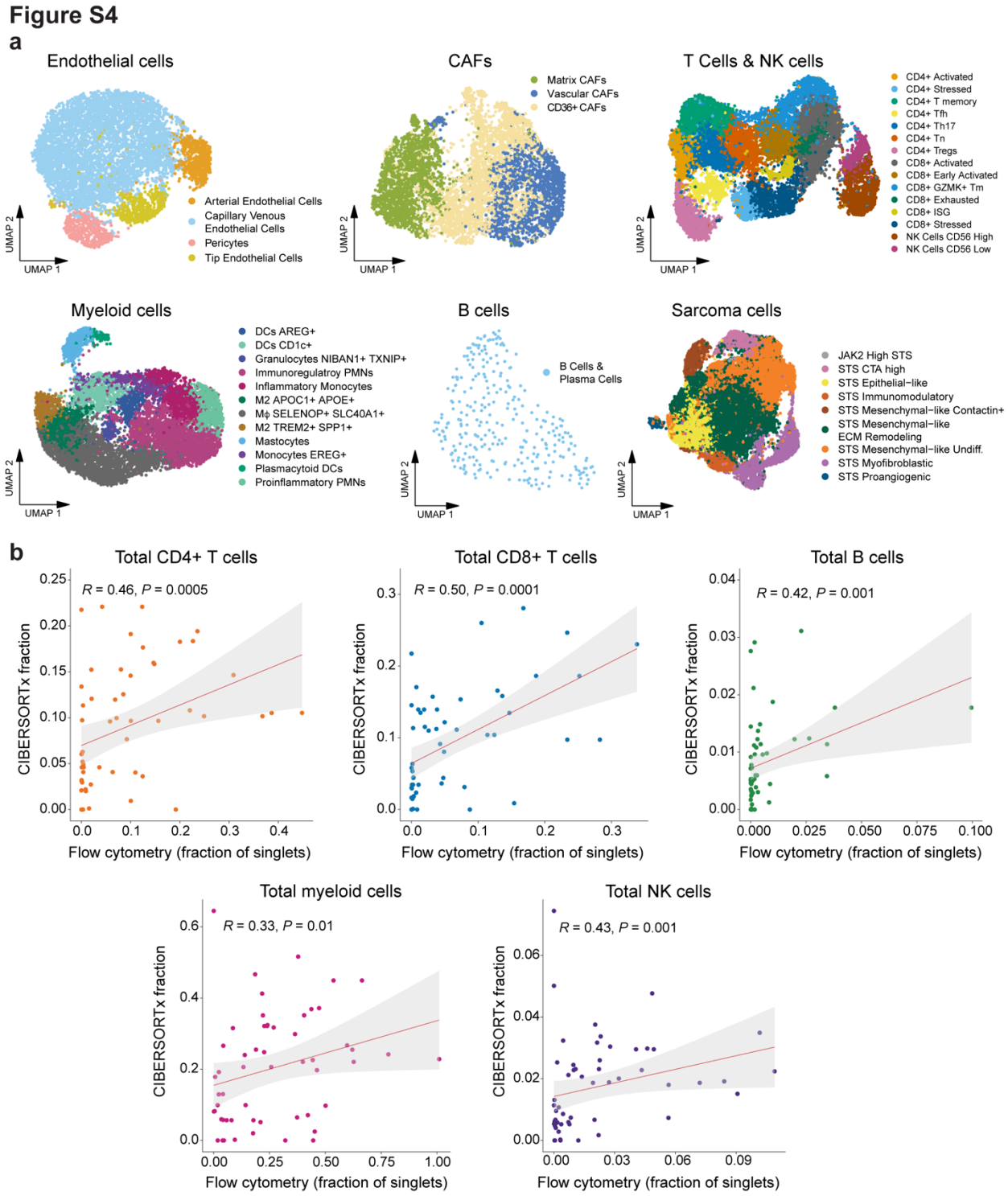


Fig. S4: A single cell RNA-seq (scRNA-seq) reference atlas for bulk RNA-seq deconvolution of undifferentiated pleomorphic sarcomas (UPS) and liposarcomas (LPS). **a**, UMAP plots showing the different non-malignant and malignant cell states identified for each major cell type in the STS scRNA-seq reference atlas, including endothelial cells, cancer associated fibroblasts (CAFs), B cells, lymphoid, and myeloid cells. **b**, Scatter plots showing the

correlation between the abundance of cell types recovered in bulk RNA-seq of post-treatment formalin-fixed paraffin embedded (FFPE) samples using CIBERSORTx and measured through flow cytometry in the matched dissociated fresh tumor samples. Spearman's correlation coefficients and two-sided P-values are displayed on the graph. The line of best fit by linear regression and 95% confidence intervals are shown on the graph.

Figure S5

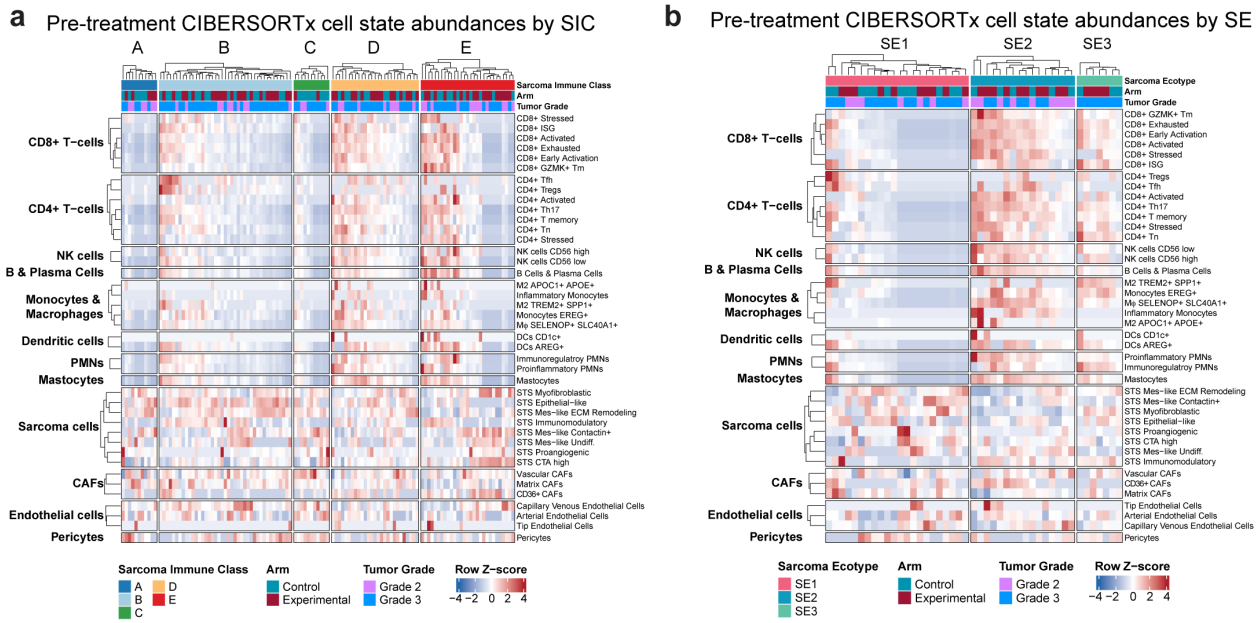
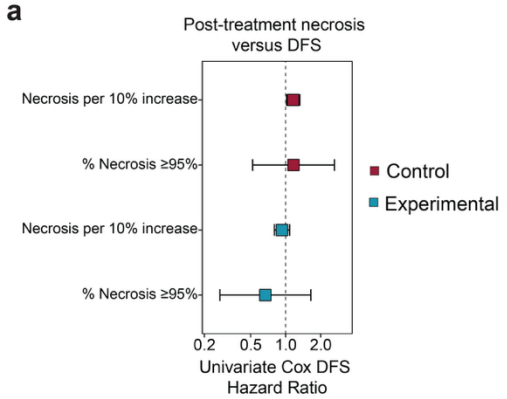


Fig. S5: Tumor microenvironment cell states across sarcoma immune classes and sarcoma ecotypes. **a** and **b**, Heat maps showing the abundance of cell states recovered with CIBERSORTx in pre-treatment samples grouped by (**a**) SIC and (**b**) SE. MΦ: macrophages.

Figure S6

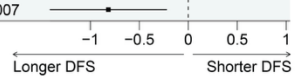
a



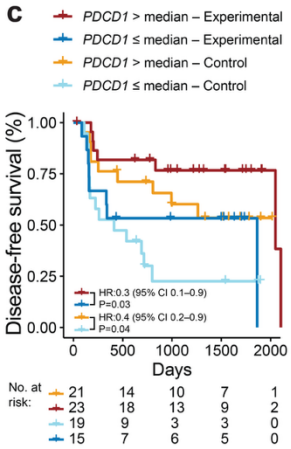
b

Association of post-RT immune checkpoint expression with outcomes

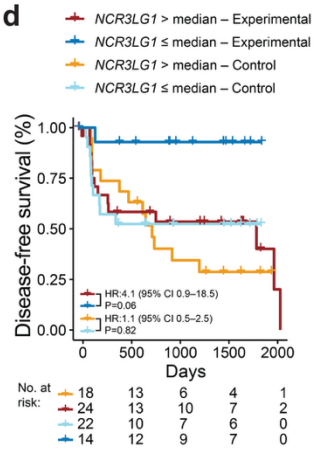
Gene	Cox DFS logHR (95% CI)	P−value
BTLA All Patients	−0.43 (−0.69 − −0.16)	0.002
BTLA Control	−0.39 (−0.76 − −0.03)	0.036
BTLA Experimental	−0.43 (−0.83 − −0.02)	0.041
CD276 All Patients	0.40 (0.02 − 0.78)	0.037
CD276 Control	0.32 (−0.14 − 0.77)	0.171
CD276 Experimental	0.35 (−0.29 − 0.99)	0.282
CTLA4 All Patients	−0.29 (−0.50 − −0.08)	0.007
CTLA4 Control	−0.32 (−0.60 − −0.04)	0.023
CTLA4 Experimental	−0.19 (−0.54 − 0.16)	0.288
NCR3LG1 All Patients	0.21 (−0.10 − 0.52)	0.183
NCR3LG1 Control	−0.04 (−0.44 − 0.35)	0.825
NCR3LG1 Experimental	0.56 (0.08 − 1.04)	0.023
PDCD1 All Patients	−0.56 (−0.88 − −0.23)	0.001
PDCD1 Control	−0.36 (−0.78 − 0.06)	0.090
PDCD1 Experimental	−0.82 (−1.41 − −0.22)	0.007



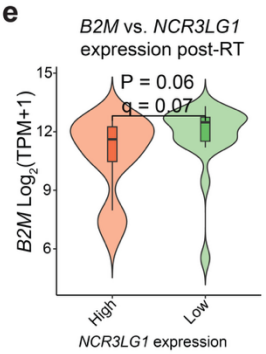
c



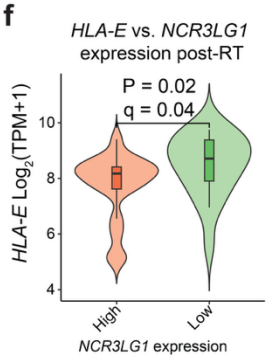
d



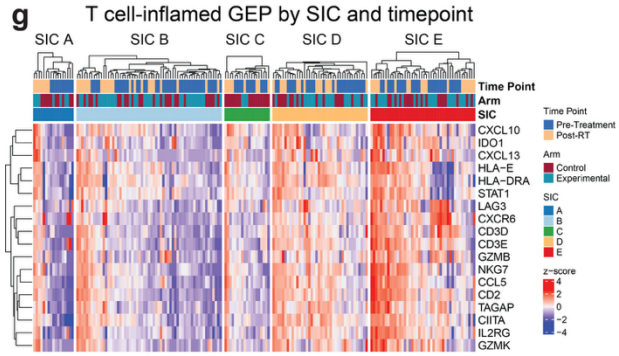
e



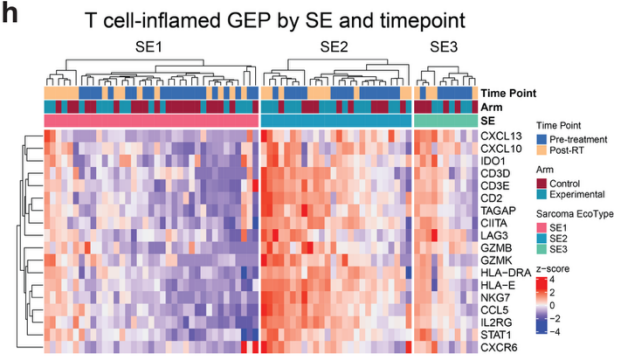
f



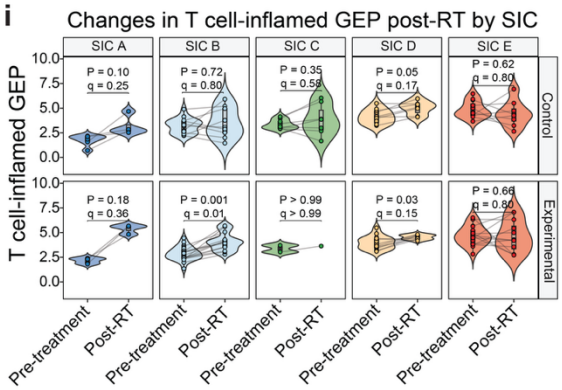
g



h



i



j

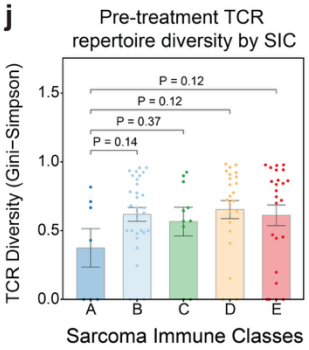


Fig. S6: Association between immune checkpoint expression and outcomes and changes in the T cell-inflamed gene expression profile (GEP) post-radiotherapy (post-RT). **a**, Univariable Cox proportional hazards analysis measuring association of pathological response with DFS in patients in the control versus the experimental arm. Pathological response is measured as post-treatment percent necrosis, and it was used as either a continuous variable or dichotomized as a categorical variable using a cutoff of 95%. P-values were calculated using two-sided Wald tests. **b**, Forest plot showing the association of post-RT immune checkpoint gene expression with DFS by treatment arm calculated using univariable Cox proportional hazards models. P-values were calculated using two-sided Wald tests. **c** and **d**, Kaplan-Meier curves showing outcome interaction between post-RT (**c**) *PDCD1* and (**d**) *NCR3LG1* expression and treatment arm. Gene expression is classified as high for a given sample if it is above the median for all pre-treatment samples. P-values were calculated using a two-sided Wald test. **e** and **f**, Expression of (**e**) *B2M* and (**f**) *HLA-E* in post-treatment samples from the experimental arm grouped based on *NCR3LG1* expression, classified as high or low based on median expression. P-values were calculated using a two-sided Wilcoxon rank-sum test. Multiple hypothesis testing correction was applied using the Benjamini-Hochberg method (q-value). **g** and **h**, Heat maps showing the expression of genes in the T cell-inflamed GEP in pre-treatment and post-RT samples grouped by pre-treatment (**g**) sarcoma immune class (SIC) and (**h**) sarcoma ecotype (SE). **i**, Box-violin plots displaying the change in T cell-inflamed GEP score described by Ayers et al. (*Journal of Clinical Investigation*, 2017) from pre-treatment to post-RT grouped by sarcoma immune class and treatment arm. P-values were calculated using a Wilcoxon signed-rank test and adjusted for multiple hypothesis testing using the Benjamini-Hochberg method. **j**, Bar plots showing the T cell receptor repertoire diversity as computed using TRUST4 in pre-treatment samples grouped by sarcoma immune class assignment. P-values were obtained using Wilcoxon rank-sum tests.