Suggested title: Blood-epigenetic biomarker associations with tumor immunophenotype in patients with urothelial carcinoma from JAVELIN Bladder 100

Authors: Thomas Powles,[1]* Srikala S. Sridhar,[2] Joaquim Bellmunt,[3] Cora N. Sternberg,[4] Petros Grivas,[5] Ewan Hunter,[6] Matthew Salter,[6] Ryan Powell,[6] Ann Dring,[6] Jayne Green,[6] Alexandre Akoulitchev,[6] Roy Ronen,[7] Janusz Dutkowski,[7] Robert Amezquita (ORCiD: 0000-0001-6868-7193),[8] Chao-Hui Huang,[8] Diane Fernandez,[8] Robbin Nameki,[8] Keith A. Ching,[8] Jie Pu,[9] Shibing Deng,[8] Alessandra di Pietro,[10] Craig B. Davis[8]*

Affiliations: [1]Barts Cancer Institute, Experimental Cancer Medicine Centre, Queen Mary University of London, St. Bartholomew's Hospital, London, UK; [2]Princess Margaret Cancer Centre, University of Toronto, Toronto, ON, Canada; [3]Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, USA; [4]Englander Institute for Precision Medicine, Division of Hematology and Medical Oncology, Weill Cornell Medicine, Meyer Cancer Center, New York, NY, USA; [5]University of Washington, Fred Hutchinson Cancer Center, Seattle, WA, USA; [6]Oxford BioDynamics Plc, Oxford, UK; [7]Data4Cure, Inc., Waltham, MA, USA; [8]Pfizer, San Diego, CA, USA; [9]Pfizer, Tucson, AZ, USA; [10]Pfizer srl, Milano, Italy

27   ***Corresponding authors**

28   Thomas Powles, MD

29   Barts Cancer Institute, Experimental Cancer Medicine Centre, Queen Mary University of

30   London, St Bartholomew's Hospital, Charterhouse Square, London EC1M 6BQ, UK

31   Email: thomas.powles1@nhs.net

32   Phone: +44 793 204 8109

33

34   Craig B. Davis, PhD

35   Pfizer, San Diego, CA, 92121, USA

36   Email: craig.davis@pfizer.com

37   Phone: +1 860-389-3591

38

39   **Journal**: *Nature Communications*

	Journal requirements	Current draft
Title	15	15
Abstract	150 words	149 words
Main text (excluding abstract, Methods, references and figure legends)	5,000 words	2567
Online methods	3000 words	1649 words
Figures and tables	10	4 figures / 2 tables
References	70	50
Supplementary materials (Figures/tables)	Permitted	4 figures / 1 table

40

Abstract (149/150 words)

Response to immune checkpoint inhibitors (ICIs) is associated with several biological pathways, including tumor immunogenicity and antitumor immunity. Identifying host factors involved in these pathways may guide personalized ICI treatment. We describe the application of chromatin conformation assays to blood from patients with advanced urothelial carcinoma from the phase 3 JAVELIN Bladder 100 trial (NCT02603432). This trial demonstrated a significant survival benefit with avelumab maintenance plus best supportive care (BSC) vs BSC alone following non-progression with platinum-based chemotherapy as first-line therapy. Blood-based chromatin conformation markers (CCMs) were screened for associations with high/low immune effector gene expression in tumors and for interactions with outcomes and tumor mutation burden. Candidate CCMs included genes involved in several immune response pathways, such as *POU2F2*, which encodes a transcription factor that regulates B-cell maturation. Our findings suggest that polygenic host factors may affect response to ICIs and support further investigation of chromatin conformation assays.

Introduction

Exploratory analyses of biomarkers from clinical trials of immune checkpoint inhibitors (ICIs) have highlighted diverse biological pathways associated with response to ICI treatment; however, identifying biomarker assays that provide optimal predictive response value remains a major challenge [1, 2]. In the phase 3 JAVELIN Bladder 100 trial, avelumab first-line (1L) maintenance plus best supportive care (BSC) significantly prolonged overall survival (OS) and progression-free survival (PFS) vs BSC alone in patients with advanced urothelial carcinoma (UC) without disease progression after 1L platinum-based chemotherapy [3]. Based on these results, avelumab 1L maintenance was established as a standard of care in patients with advanced UC [4, 5].

Exploratory tumor biomarker analyses from JAVELIN Bladder 100 revealed complex relationships with treatment outcomes, with both innate and adaptive immune biomarkers showing enrichment for response [2]. Assays for individual biomarkers such as programmed death ligand 1 (PD-L1) and tumor mutation burden (TMB) did not optimally identify patients who were significantly more likely to benefit from avelumab 1L maintenance [2]. Exploratory (hypothesis-generating) analysis suggested that patients with low PD-L1 expression may not have significant OS benefit with avelumab 1L maintenance plus BSC (avelumab plus BSC) vs BSC alone if TMB scores were low (hazard ratio [HR],1.27 [95% CI, 0.799-2.006]), but had significant benefit if TMB scores were high (HR, 0.44 [95% CI, 0.251-0.768]). In contrast, patients with high PD-L1 expression had significant OS benefit even with low TMB (HR, 0.60 [95% CI, 0.382-0.955]), suggesting that elevated anti-tumor immunity (high PD-L1 expression) could, in part, compensate for low tumor immunogenicity (low TMB) [2]. However, no molecular biomarkers are currently used to select patients for avelumab 1L maintenance in clinical practice.

83   As highlighted in the recently updated cancer immunity cycle [6], non-transformed cell types in

84   the tumor microenvironment, including stromal and myeloid cells, have emerged as important

85   modulators of anti-tumor immunity released by ICIs. These findings, made possible through the

86   application of novel technologies, such as spatial transcriptomics, support hypotheses that

87   response to ICI may be contingent on host factors such as germline single nucleotide

88   polymorphisms [7], microbiomes [8], and epigenetic modification of the 3-dimensional (3D)

89   structure of chromatin [9-13]. EpiSwitch is an established 3D genomic biomarker platform,

90   based on the original chromosome conformation capture approach, that has demonstrated

91   promising utility in various disease settings [14-19], including response to avelumab observed in

92   a retrospective study [20].

93

94   We employed the EpiSwitch platform to evaluate the hypothesis that epigenetic host factors

95   measurable in peripheral blood may modulate immune activity in the tumor microenvironment of

96   patients enrolled in the JAVELIN Bladder 100 trial. A 3-stage process was used: (1)

97   identification of chromatin conformation markers (CCMs) associated with tumor immune gene

98   expression signature; (2) statistical modeling to rank CCMs' association with TMB and OS

99   following avelumab plus BSC vs BSC alone; and (3) characterization of the expression and

100   function of genes covered by the highest-ranked CCMs.

101

102   **Results**

103   ***Study cohort and methodology***

104   The JAVELIN Bladder 100 study included 496 patients with tumor transcriptome data whose

105   blood could be analyzed by EpiSwitch (**Supplementary Figure 1a**). A 26-gene signature (JAV-

106   Immuno), identified in JAVELIN Renal 101 and validated in JAVELIN Bladder 100 [2, 21],

107   displayed modulation of the interaction between TMB, OS, and treatment (avelumab plus BSC

108   vs BSC alone) comparable to that of PD-L1 (**Figure 1a**). In the avelumab plus BSC arm, a

109   significant OS benefit was associated with JAV-Immuno levels (high vs low) in patients with low

110   TMB (HR, 0.52 [95% CI, 0.351-0.773]; p=0.0012) but not in patients with high TMB (HR, 1.22

111   [95% CI, 0.756-1.957]; p=0.42). No association was found between either biomarker and OS in

112   the BSC alone arm. JAV-Immuno was therefore chosen as the measure of tumor immune

113   response to probe for CCMs.

114

115   ***Identification of candidate CCMs***

116   The CCM discovery process was conducted as previously described (**Figure 1b**; see Methods

117   for details) [22]. Patient samples were classified as JAV-Immuno^{hi} if the corresponding tumor

118   expression was greater than the sample median and JAV-Immuno^{lo} if it was less or equal than

119   the median. Nuclei templates from 20 JAV-Immuno^{hi} and 20 JAV-Immuno^{lo} specimens were

120   hybridized to an array of >900K oligonucleotides spanning the human genome, thereby

121   identifying 150 loci that could be either positively or negatively associated with the JAV-Immuno

122   classification. Nested polymerase chain reaction (PCR) assays designed to give a positive

123   reading only if the PCR primer targets were juxtaposed by chromatin conformation were applied

124   to 40 JAV-Immuno^{hi} and 40 JAV-Immuno^{lo} samples (**Supplementary Figure 1a**). These nested

125   PCR data were used as a training set to build JAV-Immuno$^{hi/lo}$ classifier models. PCR data from

126   the remaining 416 samples were used to refine and test the classifier models. The models with

127   optimal test accuracy (0.68) contained 25 CCMs (**Figure 1c**). Univariate analysis of CCM

128   association with JAV-Immuno scores indicated that individual CCMs could be either positively or

129   negatively associated with JAV-Immuno, depending on the detection titer dilution, suggesting

130   complex relationships between the CCMs and expression of individual genes. The candidate

131   CCMs were further evaluated for association with TMB and OS using Cox proportional hazards

132   and elastic net models (see Methods for details). The 15 CCMs selected in the final model were

133   subjected to further biological annotation (**Supplementary Table 1**).

134

135 ***Biological annotation of genes covered by the CCMs***

136 The 15 CCMs in the final model were distributed across the human genome (**Supplementary**

137 **Table 1**). These CCMs were detected in <20% to >70% of the patients in the analysis

138 population (**Supplementary Figure 1b-i**), and no linkage disequilibrium was observed

139 (**Supplementary Figure 1b-ii**). Twelve of the 15 CCMs covered distinct regions, and 3

140 overlapped in a ≈290kb region on chromosome 22:20707691-20999032 (**Supplementary Table**

141 **1 and Supplementary Figure 1c**). Forty-four major genes in the Human Protein Atlas [23] were

142 mapped to these 15 CCMs, of which 29 genes (66%) were covered by >1 CCM in the final

143 model. In contrast, only 95 of the 323 genes (29%) mapping to the remaining CCMs in the

144 original set were covered by >1 CCM. Thus, the final model enriched for loci that were covered

145 by multiple CCMs, consistent with possible functional relevance. Only 1 of the genes in the

146 original 26-gene JAV-Immuno signature (*CD6*) [21] was present in the 367 genes mapped to the

147 original set, and it was not present in the 15 CCMs from the final model. Individual genes from

148 these 15 CCMs (**Table 1**) showed both positive and negative associations with JAV-Immuno

149 (**Figure 1d**), suggesting that host factors may either enhance or diminish anti-tumor immunity.

150

151 We evaluated potential expression of genes mapping to the 15 CCMs using the Tabula Sapiens

152 single-cell transcriptomic atlas [24], which contains single-cell sequencing data from almost

153 500,000 cells, representing over 400 cell types and 4 major cell lineages (immune, endothelial,

154 stromal, and epithelial) assessed in approximately 22 distinct anatomical locations collected

155 from multiple donors. **Supplementary Figure 2** shows expression patterns by cell type for 6

156 genes mapping to chromosome 22:20707691-20999032, which was covered by 3 CCMs in the

157 final model (**Supplementary Table 1**). Substantial variability in gene expression across cell

158 types was observed, suggesting that cell-specific gene-regulatory machinery operates in

159 conjunction with regulation of chromatin conformation.

160

161 We performed a comprehensive analysis of associations between genes and cell types reported

162 in the literature and public databases using the Data4Cure knowledge graph and platform [25].

163 Genes in the JAV-Immuno signature were predominantly associated with lymphocytes,

164 including T cells, B cells, and natural killer (NK) cells (**Figure 2a**), as previously described [21].

165 Genes covered by the selected CCMs were associated with endothelial and fibroblast cell types,

166 in addition to lymphocytes (**Figure 2b**). Myeloid lineages, including monocytes, neutrophils, and

167 macrophages, were associated with genes in both signatures (**Figure 2a and 2b**). JAV-Immuno

168 genes were predominantly associated with Gene Ontology Biological Processes related to

169 lymphocyte activation (**Figure 2c**) [26]; in contrast, genes mapping to the selected CCM set

170 were associated with a wide range of biological processes (**Figure 2d**). These findings suggest

171 that the chromatin conformations represented by the selected CCMs may exert both distal and

172 proximal effects on lymphocyte activation.

173

174 ***Spatial profiling***

175 Spatial transcriptomic data were used to explore the topographic expression of genes

176 represented by JAV-Immuno and the selected CCMs. A procured bladder cancer specimen was

177 selected for initial evaluation due to its robust representation of immune and non-immune cell

178 types [2]. **Figure 3a-i** shows a hematoxylin and eosin (H&E)–stained image of the specimen

179 identifying the position of lymphoid aggregates. Spatial indexing of transcriptomic data from this

180 specimen identified 8 clusters by BayesSpace (**Supplementary Figure 3a-i**). Similarities in the

181 histology and distribution of predefined gene signatures (**Supplementary Figure 3a-c**)

182 suggested that the 8 clusters represented 4 functional regions, labeled as epithelial, stromal,

183 transitional, and lymphoid aggregates (**Figure 3a-ii**). **Figure 3b-i** depicts the spatial distribution

184 of genes characteristic of tertiary lymphoid structures (TLS), originally identified in patients with

185 melanoma treated with ICIs [27]. The highest overall expression of the component genes was

186 observed in the functional region labeled as lymphoid aggregates (**Figure 3b-ii and -iii**),

187 suggesting that this region could be a locus of immune activation and maturation in the tumor

188 microenvironment. Enrichment of gene expression signatures characteristic of T follicular helper

189 cells, mature B-cells, cytotoxic T-cells, and dendritic cells recruitment further support the

190 lymphoid aggregate regions as a site for priming and maturation of anti-tumor lymphocytes

191 (**Supplementary Figure 3c-i-v**).

192

193 The spatial distribution of genes comprising the JAV-Immuno signature was strongly correlated

194 with those in the TLS signature, suggesting that JAV-Immuno is distinctly associated with TLS

195 function (**Figure 3c-i-iii**, **Supplementary Figure 3d**). To assess the potential generalizability of

196 these findings, expression of JAV-Immuno genes was evaluated in a collection of specimens

197 representing TLS from bladder, renal, and breast cancers. All 25 genes in the JAV-Immuno

198 signature were positively associated with the presence of TLS in these regions (**Figure 3c-iv**).

199 Overall, JAV-Immuno appears to be more closely associated with TLS than other signatures

200 associated with ICI response [2, 28-31].

201

202 The expression of 33 of the 35 major genes mapped to the 15 CCMs was assessed in the

203 bladder specimen and appeared to be distributed broadly throughout the specimen, with slightly

204 lower overall expression in the stromal region (**Figure 3d-i**). Hierarchical clustering of the genes

205 by distribution between the 4 annotated regions identified subsets of genes that were prominent

206 in the lymphoid aggregate and epithelial regions, as well as the stromal region (**Figure 3d-ii and**

207 **3d-iii**). Ten of the 15 selected CCMs in **Supplementary Table 1** contained ≥1 gene that was

208 more highly expressed in the lymphoid aggregate region relative to the epithelial region (**Figure**

209 **3d-ii**).

210

211 As for JAV-Immuno (**Figure 3b**), potential generalizability of the CCM findings was assessed by

212 evaluating expression of the genes in the TLS specimen collection (**Figure 3d-iv**). Several

213   genes positively associated with JAV-Immuno (**Figure 1d**) and expressed in the lymphoid

214   aggregate region (**Figure 3d-ii**) were associated with TLS regions in the collection (**Figure 3d-**

215   **iv**), including *POU2F2, MBNL1, ABI3BP*, *VPS13C,* and *CPEB1.* Other genes, including *CMPK2,*

216   *RSAD2*, *RNF144A*, and *SUCNR1*, were positively associated with JAV-Immuno but were

217   expressed more prominently outside the lymphoid aggregate regions (**Figure 3d-ii**) and were

218   not as strongly associated with TLS in the collection (**Figure 3d-iv**). Participation in antiviral and

219   antibacterial responses has been described for *RSAD2* [32, 33] and *CMPK2* [33, 34]. Reduced

220   levels of *RNF144A* have been associated with PD-L1 stabilization and tumorigenesis in mouse

221   bladder cancer models [35] and poor prognosis in patients with breast cancer [36]. *SUCNR1*

222   has been linked to altered immune infiltration of tumors in ovarian cancer [37] and renal cancer

223   [38]. The adaptive immune processes represented by JAV-Immuno may thus be tuned by

224   responses to stimuli in the tumor microenvironment that are outside of the lymphoid aggregates.

225

226   ***Associations between POU2F2, TMB, and OS***

227   Potential relationships between genes covered by the CCMs with TMB and OS with avelumab

228   plus BSC were assessed using a 3-way interaction test (**Figure 4a**). This analysis was done in

229   genes that were expressed in >50% of tumor samples and/or had a coefficient of variation of

230   >5% (**Supplementary Figure 4**). The strongest positive monogenic association was seen with

231   *POU2F2*, which also showed a positive association with JAV-Immuno (**Figure 1d**). *POU2F2* is a

232   transcription factor involved in B-cell differentiation and collaboration between B-cells and T-

233   cells and is expressed in multiple cell types involved in antigen presentation, including B cells

234   and monocytes (**Figure 2b**) [39, 40]. The presence of the corresponding CCM was associated

235   with low *POU2F2* gene expression and decreased expression of *POU2F2* target genes,

236   suggesting that the CCM negatively regulates *POU2F2* and downstream processes, including

237   generation of mature B cells and TLS formation (**Figure 4b**).

238

239 We then explored relationships between the *POU2F2* CCM, TMB, and OS in the avelumab plus

240 BSC and BSC arms. *POU2F2* CCM interactions with TMB and OS were more pronounced in

241 the avelumab plus BSC arm than in the BSC arm (**Table 2**). The absence vs presence of the

242 *POU2F2* CCM was associated with a positive effect on OS in patients treated with avelumab

243 plus BSC whose tumors displayed low TMB (less than the median value of 7.66

244 nonsynonymous single nucleotide variants per megabase); median OS was 37 vs 18 months

245 (HR, 0.46 [95% CI, 0.24-0.89]; p=0.02). These findings suggest that the effect of CCM

246 modulation on *POU2F2* expression may be strongest when the treatment effect is contingent on

247 immune responses to limited tumor immunogenicity.

248

249 **Discussion**

250 Data from the JAVELIN Bladder 100 trial are particularly valuable for investigating complex

251 biomarker interactions in advanced UC. This is because avelumab 1L maintenance was

252 associated with significant OS and PFS benefit and because avelumab was compared with BSC

253 rather than an active treatment control. Exploratory analyses of associations between immune

254 gene expression in tumor tissue, host chromatin structure, and outcomes from the JAVELIN

255 Bladder 100 trial demonstrate the potential for circulating epigenetic biomarkers to predict

256 response to ICI. In particular, systemic host chromatin structure may affect the composition of

257 the tumor microenvironment by influencing host expression of immune response genes,

258 including those affecting the formation of tertiary lymphoid structures. The interaction between

259 TMB and a CCM covering *POU2F2*, a master regulator transcription factor for maturation of

260 immune responses, in the avelumab plus BSC arm but not the BSC alone arm illustrates the

261 potential for host modulation of immune responses to tumor immunogenicity.

262 Such modulation may explain why tumor types with similar tumor mutation profiles might

263 respond differently to immunotherapy [41]. While the identification of epigenetic host factors

264 influencing response to ICIs adds a new level of complexity to biomarker analyses, it also

265 provides an opportunity for therapeutic targeting. Studies with combinations of epigenetic

266 agents and ICIs are ongoing, and measurement of chromatin biomarkers may potentially help

267 predict response [6]. Our work also underlines the great opportunity to address the challenges

268 of using single tumor-based biomarkers, such as PD-L1 or TMB, by incorporating assessment

269 of host factors.

270

271 Our findings are subject to known caveats related to exploratory post hoc analyses, including

272 the possibility of false discovery and the lack of statistical power to detect all meaningful

273 associations. As such, our results are hypothesis-generating; the clinical validity and utility of the

274 *POU2F2*-containing CCM and other CCMs reported in this study should be evaluated further in

275 prospective clinical studies. In addition, the generalizability and clinical utility of the spatial

276 transcriptomic observations need to be confirmed in larger datasets. However, our observations

277 support further evaluation of epigenetic and other blood-based biomarkers and substantiate the

278 hypothesis that polygenic host factors may contribute to ICI activity. This knowledge could

279 potentially help optimize the design of ICI-based treatment regimens for patients with advanced

280 UC.

281

282 **Online Methods**

283 *Study cohort*

284 JAVELIN Bladder 100 (NCT02603432) was a randomized phase 3 trial evaluating avelumab 1L

285 maintenance (administered every 2 weeks) plus BSC vs BSC alone in patients with advanced

286 UC whose cancer had not progressed after 4-6 cycles of 1L platinum-based chemotherapy [3].

287 Patients received treatment until disease progression, unacceptable toxicity, or withdrawal of

288 consent. OS (primary endpoint) and PFS (secondary endpoint) were significantly prolonged with

289 avelumab plus BSC vs BSC alone. Peripheral blood was collected from patients 4-10 weeks

290 after the date of administration of the last dose of chemotherapy but before treatment with

291   avelumab plus BSC or BSC alone. Blood samples from 496 patients satisfied quality and

292   consent criteria and were subjected to further evaluation, as shown in **Supplementary Figure**

293   **1a**. Characteristics and outcomes in the biomarker analysis set compared with the full analysis

294   set are shown in **Supplementary Table 2** and **Supplementary Figure 5,** respectively. The

295   biomarker analysis set did not differ significantly from the full analysis set.

296

297   ***Identification of candidate CCMs using the EpiSwitch platform***

298   *Preparation of 3D genomic templates.* EpiSwitch 3D libraries, with chromosome conformation

299   analytes converted to sequence-based tags, were prepared from frozen whole-blood samples

300   using EpiSwitch protocols following the manufacturer's instructions for EpiSwitch Explorer Array

301   kits (Oxford BioDynamics Plc). Samples were processed on the Freedom EVO 200 robotic

302   platform (Tecan Group Ltd). Briefly, the whole-blood sample was diluted and fixed with a

303   formaldehyde-containing EpiSwitch buffer. Density cushion centrifugation was used to purify

304   intact nuclei. Following a short detergent-based step to permeabilize the nuclei, restriction

305   enzyme digestion and proximity ligation were used to generate the 3D libraries. Samples were

306   centrifuged to pellet the intact nuclei before purification with an adapted protocol from the

307   QIAamp DNA FFPE Tissue kit (Qiagen) eluting in 1× TE buffer (pH 7.5). 3D libraries were

308   quantified using the Quant-iT Picogreen dsDNA Assay kit (Invitrogen) and normalized to

309   5 ng/mL prior to interrogation by PCR.

310

311   *Array design.* Custom microarrays were designed using the EpiSwitch pattern recognition

312   algorithm, which operates on Bayesian modeling and provides a probabilistic score for a

313   region's involvement in long-range chromatin interactions. The algorithm was used to annotate

314   the GRCh38 human genome assembly across approximately 1.1 million sites with the potential

315   to form long-range chromosome conformations. The most probable interactions were identified

316   and filtered on probabilistic score and proximity to protein, long non-coding RNA, or microRNA

317 coding sequences. Predicted interactions were limited to EpiSwitch sites that were ≥10 kb and

318 <300 kb apart. Repeat masking and sequence analysis was used to ensure unique CCM

319 sequences for each interaction. The EpiSwitch Explorer Array (Agilent Technologies, Product

320 Code X-HS-AC-02) containing 60-mer oligonucleotide probes was designed to interrogate

321 potential 3D genomic interactions. In total, 964,631 experimental probes and 2,500 control

322 probes were added to a 1 × 1 comparative genomic hybridization microarray slide design. The

323 experimental probes were placed on the design in singlicate, with the controls in groups of 250.

324 The control probes consisted of 6 different EpiSwitch interactions that are generated during the

325 extraction processes and used for monitoring library quality. An additional 4 external inline

326 control probe designs were added to detect a non-human (*Arabidopsis thaliana*) spike in DNA

327 added during the sample labeling protocol to provide a standard curve and control for labeling.

328 The external spike DNA consists of 400-bp single-stranded DNA fragments from genomic

329 regions of *Arabidopsis thaliana*. Array-based comparisons were performed as described

330 previously, with the modification of only 1 sample being hybridized to each array slide in the Cy3

331 channel.

332

333 *Translation of array-based 3D genomic markers to PCR readouts.* Stepwise diagnostic

334 biomarker discovery process using EpiSwitch technology and data analysis EpiSwitch

335 technology platform (Oxford Biodynamics, Oxford, UK) pairs high-resolution chromosome

336 conformation capture results with regression analysis and a machine learning algorithm to

337 develop disease classifications [42-44]. To select epigenetic biomarkers, patient samples for

338 each class were screened for statistically significant differences in conditional and stable profiles

339 of genome architecture. All samples used for the nested PCR biomarker discovery were

340 processed per the manufacturer's instructions using the proprietary EpiSwitch reagents and

341 standard protocols [45]. A PCR product was detected only if the corresponding CCM was

342   detected. To account for variable marker stability in assay conditions, 3 template concentrations

343   (1✕, 2✕, and 4✕) were evaluated.

344

345   *Genomic mapping.* Mapping was performed using the Bedtools closest function [42] for the 3

346   closest protein-coding loci: upstream, downstream, and within the long-range chromosome

347   interaction (Gencode v33). All markers were visualized using the EpiSwitch analytical portal.

348

349   *Identification of top 25 CCMs associated with tumor JAV-Immuno scores.* Analyses for marker

350   prioritization were performed using libraries developed for the R Statistical Language (R version

351   4.1.2). JAV-Immuno includes the following genes: *CCL5, CD2, CD244, CD247, CD3E, CD3G,*

352   *CD6, CD8B, CD96, CST7, EOMES, GFI1, GPR18, GRAP2, IL7R, ITK, KCNA3, KLRD1,*

353   *NLRC3, PRF1, PSTPIP1, SH2D1A, SIT1, THEMIS, TRAT1,* and *XCL2.* Candidate markers

354   were identified using EpiSwitch whole-genome arrays comparing 20 patients whose tumors

355   were JAV-Immuno^{hi} (greater than the median) and 20 patients whose tumors were JAV-

356   Immuno^{lo} (less than the median). The data were normalized, quality controlled, and statistically

357   analyzed using the EpiSwitch array pipeline powered by the LIMMA package. The top 150

358   EpiSwitch markers were selected based on their corrected p value (false discovery rate) and

359   directionality/interaction frequency for JAV-Immuno. Feature engineering with the selected

360   EpiSwitch CCMs was performed using the same samples used in the array discovery, with an

361   additional 20 samples each of the JAV-Immuno^{hi} and JAV-Immuno^{lo} classes. A reduced set of

362   25 CCMs was generated using the univariate EXACT test (contingency testing), binary

363   weighting and multivariant analyses, permutated glmnet, VSURF, Boruta, and varSelRF. An

364   XGBoost algorithm model was used for further refinement of the CCM list [43]. Shapley additive

365   explanations values, used to characterize the contribution of each feature in a model to the final

366   prediction and final model selection, were used to confirm marker prioritization [44].

367

368 ***Prioritization of candidate CCMs based on estimated interactions with TMB and OS***

369 The starting feature set comprised 25 CCMs, each measured at 3 dilution levels for a total of 75

370 features, and was assessed in 496 patients for its association with OS.

371 *Model development*. In step 1, features were filtered out prior to modeling if <40 subjects were

372 included in either the present or absent categories and <15 subjects in any of the 4 subgroups

373 (avelumab present, avelumab absent, BSC present, BSC absent). This filter removed 15

374 features.

375

376 In step 2, due to high correlation among CCMs measured at multiple dilution levels, further

377 screening was performed to either select 1 dilution level per CCM or remove the CCM from

378 subsequent multivariate modeling. The screening assessed each feature in association with OS

379 using 3 nested Cox proportional hazards models. The first model assessed the feature as a

380 main effect; the second included the main effect and 2-way interaction with treatment; and the

381 third included the main effect, 2-way interactions with both treatment and TMB, and the 3-way

382 interaction among feature, treatment, and TMB. For each CCM, this resulted in 3 models for

383 each dilution level. The p values derived from the 3-way interaction, 2-way interaction, and main

384 effect terms were sequentially evaluated for each CCM. The dilution level with the lowest 3-way

385 interaction p value of <0.1 was selected. If none were <0.1, the process was repeated for the 2-

386 way interaction and main effect p values. If none of these p values were <0.1, then the CCM

387 was eliminated from downstream modeling.

388

389 Finally, in step 3, the remaining 16 features were fed into an elastic net Cox proportional

390 hazards model–building procedure to select a final feature set. The elastic net [46], a

391 regularized regression method using a weighted sum of the L1 and L2 penalty terms for the

392 lasso [47] and ridge [48] regression, was used for statistical model building (R package: glmnet).

393 The hyperparameters α and λ, which control the weight applied to each of the penalty terms and

394   the amount of shrinkage on the coefficients, were tuned using repeated 5-fold cross-validation

395   (CV). For each round of 5-fold CV, a pair of α and λ associated with the optimal model

396   performance (C-index) was determined (**Supplementary Figure 6**). The same procedure was

397   iterated 1,000 times using different random data partitions to account for potential CV error. The

398   most frequently selected hyperparameter pairs were chosen as optimal and applied to the full

399   dataset to obtain the final models. All but 1 of the 16 input features were retained by this

400   process, generating the list of 15 CCMs (**Supplementary Table 1**).

401

402   ***Assessment of potential CCM gene expression***

403   Public data on genes labeled using Human Genome Organization nomenclature have been

404   compiled into a knowledge graph [49] by Data4Cure (www.data4cure.com). Knowledge graph

405   associations between genes and specific biomedical entity domains (eg, cell types or pathways)

406   summarize evidence from thousands of publicly available datasets and millions of publications

407   linking biomedical entities.

408

409   ***Visium data processing***

410   SpaceRanger 2.0.1 was used to align raw reads to the GrCh38 genome and feature count

411   generation using the Human Transcriptome probe set v2.0. Visium data resolution was

412   enhanced 6-fold by inferring pseudospots using the BayesSpace method [50]. Briefly, using the

413   BayesSpace R package Visium Space Ranger, the outs folder was processed with seed 102,

414   spatialPreprocess(sceobj, platform="Visium", n.PCs=15, n.HVGs=2000, log.normalize=T),

415   followed by spatialCluster(sceobj, q=8, d=15, platform="Visium", init.method="mclust",

416   model="t", gamma=1, nrep=10000, burn.in=100, save.chain=F), and then

417   spatialEnhance(sceobj, q=8, platform="Visium", d=15, model="t", gamma=1, jitter_prior=0.3,

418   jitter_scale=3.5, nrep=100000, burn.in=100, save.chain=F). Enhanced features and spatial

419   coordinates for the full expression matrix were computed using enhanceFeatures and output for

420   subsequent visualization using the Data4Cure platform [25] (www.data4cure.com).

421   For the expression signatures of each pseudospot, the Data4Cure platform was used to

422   compute the gene signature score as the mean normalized average expression value across

423   the signature genes. The spot-level signature scores were plotted heat map style using the

424   pseudospot coordinates and overlayed on the approximate corresponding region of the H&E

425   image as a background image.

426   An unmarked, blinded H&E image was annotated by a pathologist for lymphoid aggregates

427   (numbered and circled in green).

428
429

Author information

Author contributions:

Conceptualization: TP, CD

Data curation: JP, KC, RN, EH, RR, JD

Formal Analysis: JP, EH

Investigation: TP, SS, JB, CS, PG, AD

Methodology: EH, MS, RP, AD, JG, AA, RN, KC, JP, SD, CD, RR, JD, DF, RA

Project administration: CD

Software: EH, RR, JD

Validation: EH, MS, RP, JG, AA, CH

Writing – original draft: TP, CD

Writing – review & editing: TP, SS, PG, KC, CD

Ethics declarations

Competing interests:

478 **Nameki, K. A. Ching, S. Deng, A. di Pietro,** and **C. B. Davis** are employees of and report stock

479 or other ownership interests in Pfizer.

480

481 **Data Sharing Statement**

482 Any requests for data by qualified scientific and medical researchers for legitimate research

483 purposes will be subject to Merck's Data Sharing Policy. All requests should be submitted in

484 writing to Merck's data sharing portal (https://www.merckgroup.com/en/research/our-approach-

485 to-research-and-development/healthcare/clinical-trials/commitment-responsible-data-

486 sharing.html). When Merck has a co-research, co-development, or co-marketing or co-promotion

487 agreement, or when the product has been out-licensed, the responsibility for disclosure might be

488 dependent on the agreement between parties. Under these circumstances, Merck will endeavor

489 to gain agreement to share data in response to requests.

References

1. Banchereau, R., et al., *Molecular determinants of response to PD-L1 blockade across tumor types.* Nat Commun, 2021. **12**(1): p. 3969.
2. Powles, T., et al., *Avelumab maintenance in advanced urothelial carcinoma: biomarker analysis of the phase 3 JAVELIN Bladder 100 trial.* Nat Med, 2021. **27**(12): p. 2200-2211.
3. Powles, T., et al., *Avelumab Maintenance Therapy for Advanced or Metastatic Urothelial Carcinoma.* N Engl J Med, 2020. **383**(13): p. 1218-1230.
4. Powles, T., et al., *ESMO Clinical Practice Guideline interim update on first-line therapy in advanced urothelial carcinoma.* Ann Oncol, 2024. **35**(6): p. 485-490.
5. Grivas, P., et al., *Avelumab first-line maintenance treatment for advanced urothelial carcinoma: review of evidence to guide clinical practice.* ESMO Open, 2023. **8**(6): p. 102050.
6. Mellman, I., et al., *The cancer-immunity cycle: Indication, genotype, and immunotype.* Immunity, 2023. **56**(10): p. 2188-2205.
7. Khan, Z., et al., *Polygenic risk for skin autoimmunity impacts immune checkpoint blockade in bladder cancer.* Proc Natl Acad Sci U S A, 2020. **117**(22): p. 12288-12294.
8. Dizman, N., et al., *Nivolumab plus ipilimumab with or without live bacterial supplementation in metastatic renal cell carcinoma: a randomized phase 1 trial.* Nat Med, 2022. **28**(4): p. 704-712.
9. Cuartero, S., G. Stik, and R. Stadhouders, *Three-dimensional genome organization in immune cell fate and function.* Nat Rev Immunol, 2023. **23**(4): p. 206-221.
10. Johanson, T.M., et al., *Genome organization in immune cells: unique challenges.* Nat Rev Immunol, 2019. **19**(7): p. 448-456.
11. Tordini, F., et al., *The genome conformation as an integrator of multi-omic data: the example of damage spreading in cancer.* Front Genet, 2016. **7**: p. 194.
12. Bonev, B. and G. Cavalli, *Organization and function of the 3D genome.* Nat Rev Genet, 2016. **17**(11): p. 661-678.
13. Tien, F.-M., et al., *Epigenetic remodeling of the immune landscape in cancer: therapeutic hurdles and opportunities.* J Biomed Sci, 2023. **30**(1): p. 3.
14. Kempfer, R. and A. Pombo, *Methods for mapping 3D chromosome architecture.* Nat Rev Genet, 2020. **21**(4): p. 207-226.
15. Dekker, J., et al., *Capturing chromosome conformation.* Science, 2002. **295**(5558): p. 1306-11.
16. Jakub, J.W., et al., *A pilot study of chromosomal aberrations and epigenetic changes in peripheral blood samples to identify patients with melanoma.* Melanoma Res, 2015. **25**(5): p. 406-11.
17. Carini, C., et al., *Chromosome conformation signatures define predictive markers of inadequate response to methotrexate in early rheumatoid arthritis.* J Transl Med, 2018. **16**(1): p. 18.
18. Alshaker, H., et al., *Chromatin conformation changes in peripheral blood can detect prostate cancer and stratify disease risk groups.* J Transl Med, 2021. **19**(1): p. 46.
19. Pchejetski, D., et al., *Circulating chromosome conformation signatures significantly enhance PSA positive predicting value and overall accuracy for prostate cancer detection.* Cancers, 2023. **15**(3): p. 821.
20. Hunter, E., et al., *Development and validation of blood-based predictive biomarkers for response to PD-1/PD-L1 checkpoint inhibitors: evidence of a universal systemic core of 3D immunogenetic profiling across multiple oncological indications.* Cancers, 2023. **15**(10): p. 2696.
21. Motzer, R.J., et al., *Avelumab plus axitinib versus sunitinib in advanced renal cell carcinoma: biomarker analysis of the phase 3 JAVELIN Renal 101 trial.* Nat Med, 2020. **26**(11): p. 1733-1741.
22. Yan, H., et al., *Epigenetic chromatin conformation changes in peripheral blood can detect thyroid cancer.* Surgery, 2019. **165**(1): p. 44-49.
23. Karlsson, M., et al., *A single-cell type transcriptomics map of human tissues.* Sci Adv, 2021. **7**(31).

538 24. Tabula Sapiens Consortium, et al., *The Tabula Sapiens: a multiple-organ, single-cell*
539   *transcriptomic atlas of humans.* Science, 2022. **376**(6594): p. eabl4896.
540 25. Hu-Lieskovan, S., et al., *Association of tumor mutational burden and immune gene expression*
541   *with response to PD-1 blockade by sasanlimab across tumor types and routes of administration.*
542   Targeted Onco, 2021. **16**(6): p. 773-787.
543 26. Ashburner, M., et al., *Gene ontology: tool for the unification of biology. The Gene Ontology*
544   *Consortium.* Nat Genet, 2000. **25**(1): p. 25-9.
545 27. Helmink, B.A., et al., *B cells and tertiary lymphoid structures promote immunotherapy response.*
546   Nature, 2020. **577**(7791): p. 549-555.
547 28. Higgs, B.W., et al., *Interferon gamma messenger RNA signature in tumor biopsies predicts*
548   *outcomes in patients with non-small cell lung carcinoma or urothelial cancer treated with*
549   *durvalumab.* Clin Cancer Res, 2018. **24**(16): p. 3857-3866.
550 29. Fehrenbacher, L., et al., *Atezolizumab versus docetaxel for patients with previously treated non-*
551   *small-cell lung cancer (POPLAR): a multicentre, open-label, phase 2 randomised controlled trial.*
552   Lancet, 2016. **387**(10030): p. 1837-46.
553 30. Eckstein, M., et al., *Cytotoxic T-cell-related gene expression signature predicts improved survival*
554   *in muscle-invasive urothelial bladder cancer patients after radical cystectomy and adjuvant*
555   *chemotherapy.* J Immunother Cancer, 2020. **8**(1): p. e000162.
556 31. Ayers, M., et al., *IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade.* J Clin
557   Invest, 2017. **127**(8): p. 2930-2940.
558 32. Wiedemann, G.M., et al., *Cutting edge: STAT1-mediated epigenetic control of Rsad2 promotes*
559   *clonal expansion of antiviral NK cells.* J Immunol, 2020. **205**(1): p. 21-25.
560 33. Arnold, N., et al., *Genomic and functional analysis of the host response to acute simian varicella*
561   *infection in the lung.* Sci Rep, 2016. **6**: p. 34164.
562 34. Larsson, I., et al., *Genome-scale metabolic modeling of glioblastoma reveals promising targets*
563   *for drug development.* Front Genet, 2020. **11**: p. 381.
564 35. Ho, S.R., et al., *RNF144A deficiency promotes PD-L1 protein stabilization and carcinogen-induced*
565   *bladder tumorigenesis.* Cancer Lett, 2021. **520**: p. 344-360.
566 36. Yang, Y.L., et al., *RNF144A functions as a tumor suppressor in breast cancer through ubiquitin*
567   *ligase activity-dependent regulation of stability and oncogenic functions of HSPA2.* Cell Death
568   Differ, 2020. **27**(3): p. 1105-1118.
569 37. Zhang, J., et al., *Association between succinate receptor SUCNR1 expression and immune*
570   *infiltrates in ovarian cancer.* Front Mol Biosci, 2020. **7**: p. 150.
571 38. Najm, R., M.Y. Hachim, and R.K. Kandasamy, *Divulging a pleiotropic role of succinate receptor*
572   *SUCNR1 in renal cell carcinoma microenvironment.* Cancers (Basel), 2022. **14**(24): p. 6064.
573 39. Hodson, D.J., et al., *Regulation of normal B-cell differentiation and malignant B-cell survival by*
574   *OCT2.* Proc Natl Acad Sci U S A, 2016. **113**(14): p. E2039-46.
575 40. Corcoran, L., et al., *Oct2 and Obf1 as facilitators of B:T cell collaboration during a humoral*
576   *immune response.* Front Immunol, 2014. **5**: p. 108.
577 41. Powles, T., et al., *Clinical efficacy and biomarker analysis of neoadjuvant atezolizumab in*
578   *operable urothelial carcinoma in the ABACUS trial.* Nat Med, 2019. **25**(11): p. 1706-1714.
579 42. Quinlan, A.R., *BEDTools: the Swiss-army tool for genome feature analysis.* Curr Protoc
580   Bioinformatics, 2014. **47**: p. 11.12.1-34.
581 43. Chen, T. and C. Guestrin, *XGBoost: a scalable tree boosting system*, in *Proceedings of the 22nd*
582   *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2016,
583   Association for Computing Machinery: San Francisco, California, USA. p. 785–794.
584 44. Biecek, P. and T. Burzykowski, *Shapley Additive Explanations (SHAP) for Average Attributions*.
585   2021. p. 95-106.

586 45. Salter, M., et al., *Initial identification of a blood-based chromosome conformation signature for aiding in the diagnosis of amyotrophic lateral sclerosis.* EBioMedicine, 2018. **33**: p. 169-184.
588 46. Zou, H. and T. Hastie, *Regularization and variable selection via the elastic net.* J R Stat Soc Series B Stat Methodol, 2005. **67**(2): p. 301-320.
590 47. Tibshirani, R., *Regression shrinkage and selection via the Lasso.* J R Stat Soc Series B Stat Methodol, 1996. **58**(1): p. 267-288.
592 48. Hoerl, A.E. and R.W. Kennard, *Ridge regression: applications to nonorthogonal problems.* Technometrics, 1970. **12**(1): p. 69-82.
594 49. Tiddi, I. and S. Schlobach, *Knowledge graphs as tools for explainable machine learning: A survey.* Artificial Intelligence, 2022. **302**: p. 103627.
596 50. Zhao, E., et al., *Spatial transcriptomics at subspot resolution with BayesSpace.* Nat Biotechnol, 2021. **39**(11): p. 1375-1384.
598

Table 1. Compiled relationship between genes in the selected CCM set and JAV-Immuno, or entities in the CURIE knowledge graph. Values reported in the cell type, immune compartment, and immune process columns represent aggregated association scores between the individual genes and selected entities in the Data4Cure CURIE Knowledge Graph. The association score summarizes data and literature evidence from publicly available datasets and publications linking biomedical entities. **CCM**, chromatin conformation marker; **MHC**, major histocompatibility complex.

| Gene | Gene description | Association between tumor gene expression and JAV-Immuno score | | | Cell type | | | | | | | Immune compartment | | Immune process | | |
		Statistic value	p value	q value	Endothelial cell	Fibroblast	B cell	Monocyte	Natural killer cell	Dendritic cell	Cytotoxic T cell	Lymph node	Germinal center	MHC class I antigen presentation	MHC class II antigen presentation	Response to stress
POU2F2	POU class 2 homeobox 2	0.8141921	2.75E−135	4.96E−134	0.09011	0.2449	0.2558	0.2563	0.228	0.2125	0.02845	0.09261	0.09931	0	0.04072	0.004806
MBNL1	Muscle blindlike splicing regulator 1	0.6073972	2.31E−58	2.77E−57	0.2185	0.222	0.2186	0.2325	0.07344	0.04999	0.02845	0		0.008677	0.008505	0.1479
ABI3BP	ABI family member 3 binding protein	0.5517624	2.09E−46	1.88E−45	0.212	0.2435	0.2236	0.1964	0.1889	0.1974	0.02832	0.07382		0	0	0
VPS13C	Vacuolar protein sorting 13 homolog C	0.4595461	6.43E−31	4.63E−30	0.07422	0.1806	0.05065	0.06594	0.06424	0.05733	0.02213	0		0	0	0
SUCNR1	Succinate receptor 1	0.4390713	4.48E−28	2.69E−27	0.2147	0.2118	0.1985	0.2108	0.2092	0.208	0.15	0.09239		0.003451	0.00712	0.0858
RSAD2	Radical S-adenosyl methionine domain containing 2	0.3267927	1.49E−15	6.72E−15	0.2266	0.2148	0.234	0.2137	0.2234	0.229	0.1652	0.09274	0.07864	0	0.06117	0.1613
CMPK2	Cytidine/uridine monophosphate kinase 2	0.2787244	1.47E−11	5.89E−11	0.2066	0.2194	0.2099	0.2243	0.1865	0.2063	0.1437	0.07898	0.06812	0	0.0499	0.08855
PI4KA	Phosphatidylinositol 4-kinase alpha	0.152742	0.000265	0.000954	0.2101	0.2087	0.2186	0.1967	0.197	0.04704	0.02591	0		0.01603	0	0
RNF144A	Ring finger protein 144A	0.1471635	0.000444	0.00145	0.09449	0.2008	0.1847	0.1624	0.06401	0.1495	0.08942	0.07085	0.01967	0	0	0
CPEB1	Cytoplasmic polyadenylation element binding protein 1	0.1458114	0.000502	0.0015	0.1894	0.2084	0.1893	0.1863	0.07279	0.05963	0	0		0	0	0.01358
NPY4R	Neuropeptide Y receptor Y4	0.1201768	0.00419	0.00944	0.04041	0.1401	0.1273	0.1208	0.1318	0.03498	0	0		0.003384	0	0
ZNF573	Zinc finger protein 573	0.1156841	0.00586	0.0113	0.06559	0.1173	0.1437	0.1506	0.1212	0.09401	0.04908	0.02246		0.005744	0	0
SNAP29	Synaptosome associated protein 29	0.1154207	0.00598	0.0113	0.2049	0.1919	0.06076	0.06829	0.06407	0.04174	0.02845	0		0.006851	0.005367	0.1535
ZNF781	Zinc finger protein 781	0.1124934	0.00739	0.0121	0.1024	0.0865	0.1235	0.08621	0.09041	0.04321	0.005367	0.04858		0.003986	0.002624	0.000162
SLC38A7	Solute carrier family 38 member 7	0.0978701	0.0199	0.0286	0.1976	0.05722	0.192	0.06142	0.05257	0.03522	0.02277	0		0	0	0.08463

Symbol	Description																
DEDD2	Death effector domain containing 2	0.0966533	0.0215	0.0297	0.08156	0.1777	0.1756	0.1643	0.1515	0.146	0.1052	0.06802	0	0		0.02047	0.06925
C2CD4B	C2 calcium dependent domain containing 4B	0.0826085	0.0495	0.066	0.2162	0.1775	0.154	0.1431	0.1367	0.05737	0.01532	0.0178	0	0	0	0	0
TMEM14E	transmembrane protein 14E, pseudogene	0.0678954	0.107	0.137	0.03418	0.0232	0.0338	0.03111	0.04544	0.02271	0.004312	0	0	0	0	0	0
MMP16	Matrix metallopeptidase 16	0.0639516	0.129	0.16	0.2134	0.2523	0.2056	0.2038	0.2085	0.06526	0.007903	0.09234	0	0	0	0	0
LZTR1	Leucine zipper like transcription regulator 1	0.0550683	0.191	0.222	0.2021	0.1965	0.06215	0.1714	0.2034	0.03179	0.01318	0.08179	0		0.007257	0.008901	0.006873
RPS17	Ribosomal protein S17	0.0165515	0.694	0.757	0.2079	0.183	0.2076	0.2269	0.2022	0.179	0.1641	0			0.03428	0.07347	0.1058
CNBD1	Cyclic nucleotide binding domain containing 1	0.0040582	0.923	0.923	0.1355	0.1114	0.1151	0.1164	0.1073	0.09581	0.01741	0.04782	0		0.003208	0.003264	0.00432
C2CD4A	C2 calcium dependent domain containing 4A	−0.004789	0.909	0.923	0.171	0.1587	0.1331	0.1213	0.1156	0.05684	0.004867	0.04788	0			0.003982	0.08443
DCAF4L2	DDB1 and CUL4 associated factor 4 like 2	−0.008794	0.835	0.884	0.06324	0.03736	0.06659	0.09815	0.1282	0.05735	0	0.06016	0		0.00528	0.01013	0.00442
ZNF526	Zinc finger protein 526	−0.048678	0.248	0.279	0.1291	0.1437	0.07571	0.06573	0.04032	0.03047	0.004803	0	0	0	0	0	0
CNOT1	CCR4-NOT transcription complex subunit 1	−0.055119	0.19	0.222	0.0678	0.21	0.2023	0.2026	0.03454	0.1721	0.01716	0			0.01848	0.003042	0.09112
ZFP30	ZFP30 zinc finger protein	−0.098922	0.0186	0.0279	0.1219	0.1227	0.1167	0.1308	0.1257	0.07336	0.02253	0	0	0	0	0	0
ZNF607	Zinc finger protein 607	−0.110554	0.00848	0.0133	0.04603	0.0471	0.06989	0.08244	0.1002	0.03778	0.0391	0.02276	0		0.004409	0.003307	0.01889
GPRIN2	G protein regulated inducer of neurite outgrowth 2	−0.112585	0.00734	0.0121	0.06067	0.1465	0.1516	0.03462	0.1517	0.04801	0.00213	0.05896	0.02013		0.01973	0.00166	
TFG	Trafficking from ER to golgi regulator	−0.114673	0.00631	0.0114	0.2094	0.1879	0.2205	0.2067	0.06115	0.02826	0.02845	0.08952	0		0.003158	0.00423	0.005551
SERPIND1	Serpin family D member 1	−0.116083	0.00569	0.0113	0.2145	0.1994	0.189	0.1905	0.05945	0.1672	0	0.09033	0			0.004157	0.1498
CRKL	CRK like proto-oncogene, adaptor protein	−0.130329	0.00189	0.00454	0.2009	0.2036	0.2249	0.1969	0.2029	0.1667	0.0242	0.09401	0		0.02339	0.02136	0.02151
AIFM3	Apoptosis inducing factor mitochondria associated 3	−0.138936	0.000919	0.00236	0.19	0.1729	0.05	0.06931	0.0611	0.05562	0.006126	0.08887	0		0.0107	0.004403	0.1262
ANXA8L1	Annexin A8 like 1	−0.14415	0.000582	0.00161	0.1466	0.1922	0.1392	0.1212	0.1354	0.1276	0.02692	0.0607	0		0	0	0
GOT2	Glutamic-oxaloacetic transaminase 2	−0.381291	5.01E−21	2.58E−20	0.2346	0.2296	0.2188	0.2172	0.07141	0.1993	0.1728	0.09115	0.07472		0.006362	0.09641	

604
605

606 **Table 2. Associations between POU2F2 marker and OS.**

607 **BSC**, best supportive care; **HR**, hazard ratio; **NE**, not estimable; **OS**, overall survival; **TMB**, tumor mutation burden.

Treatment	TMB	*POU2F2* marker	No. of patients (N=457)	No. of events	OS, median (95% CI), months	HR (*POU2F2* marker absent vs present) (95% CI)	p value
Avelumab plus BSC	≤Median	Absent	22	10	36.99 (18.17-NE)	0.46 (0.240-0.894)	0.0218
Avelumab plus BSC	≤Median	Present	112	82	17.77 (13.34-22.34)		
Avelumab plus BSC	>Median	Absent	11	7	19.25 (17.81-NE)	1.38 (0.623-3.048)	0.4281
Avelumab plus BSC	>Median	Present	98	48	35.12 (26.05-NE)		
BSC alone	≤Median	Absent	18	13	13.68 (8.8-NE)	1.14 (0.622-2.071)	0.6788
BSC alone	≤Median	Present	89	60	16.07 (10.25-24.18)		
BSC alone	>Median	Absent	21	14	14.78 (11.5-NE)	1.14 (0.635-2.044)	0.6608
BSC alone	>Median	Present	86	58	17.81 (13.54-26.64)		

608

609 **Figure 1. Blood chromatin markers associated with JAV-Immuno in tumor and**

610 **interactions between TMB, treatment, and OS.**

611 **A.** Modulation of TMB interactions with treatment based on OS outcome by JAV-Immuno score.

612 Kaplan-Meier plots of high baseline TMB (≥median) vs low baseline TMB (<median) are shown

613 for subgroups defined by treatment (avelumab plus BSC or BSC) and baseline JAV-Immuno

614 score$^{hi/lo}$ (based on median split). A Cox proportional hazards model with no adjustment for

615 baseline covariates was used. A 2-sided Wald test was used to determine p values. **B.**

616 Discovery of CCMs using EpiSwitch platform. i. DNA from fixed and permeabilized nuclei was

617 digested with restriction enzymes and hybridized to an oligonucleotide array designed to detect

618 sequences brought in proximity by chromatin conformation. ii. Nested PCR assays were

619 designed for 150 sequences that could potentially differentiate between patients with JAV-

620 Immuno^{hi} and JAV-Immuno^{lo} tumors. iii. PCR product indicated whether a given chromatin

621 marker was present or absent. Predictive models based the presence or absence of chromatin

622 marker PCR products and built using methods such as XGBoost were used to narrow the list of

623 candidate markers to 25. **C.** Associations between blood chromatin markers and tumor JAV-

624 Immuno levels. Univariate associations are shown between JAV-Immuno score and the status

625 of 25 markers (present vs absent) detected by PCR assay across 3 template dilution levels (1✕,

626 2✕, and 4✕). The x-axis represents magnitude and direction of differences in mean JAV-Immuno

627 levels. A Wilcoxon rank-sum test comparing JAV-Immuno scores between the presence and

628 absence of each marker was performed at each dilution level. The y-axis represents the $\log_{10}$ of

629 2-sided p values based on the test. Markers and dilutions that were selected in the final model

630 are labeled. **D.** Correlations between individual genes covered by selected CCM markers and

631 JAV-Immuno signature scores in tumor are plotted as Pearson coefficients with log10(p value).

632 Genes with uncorrected p≤0.001 are labeled.

633 **BSC**, best supportive care; **CCM**, chromatin conformation marker; **HR**, hazard ratio; **NE**, not

634 evaluable; **OS**, overall survival; **PCR**, polymerase chain reaction; **TMB**, tumor mutation burden.
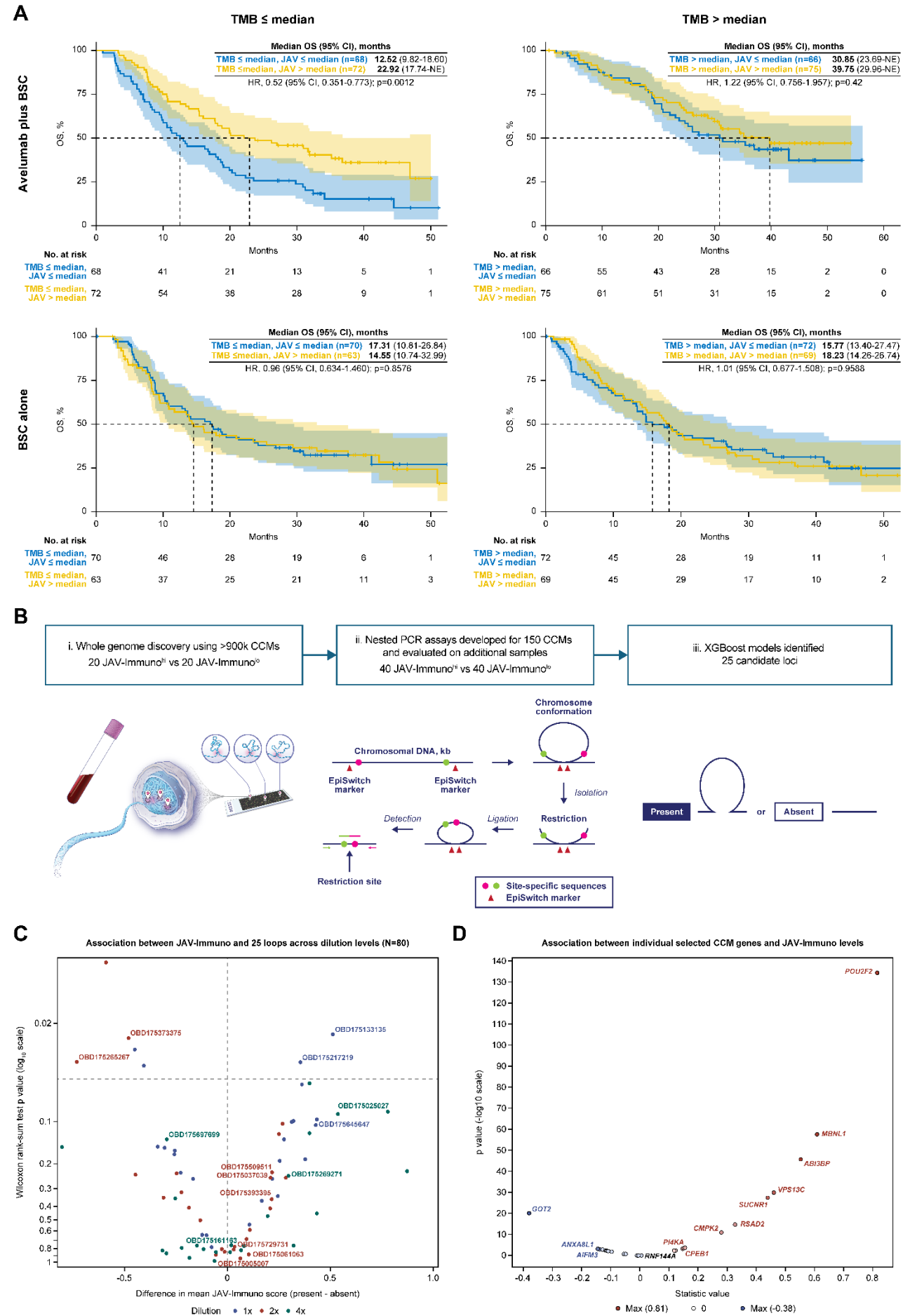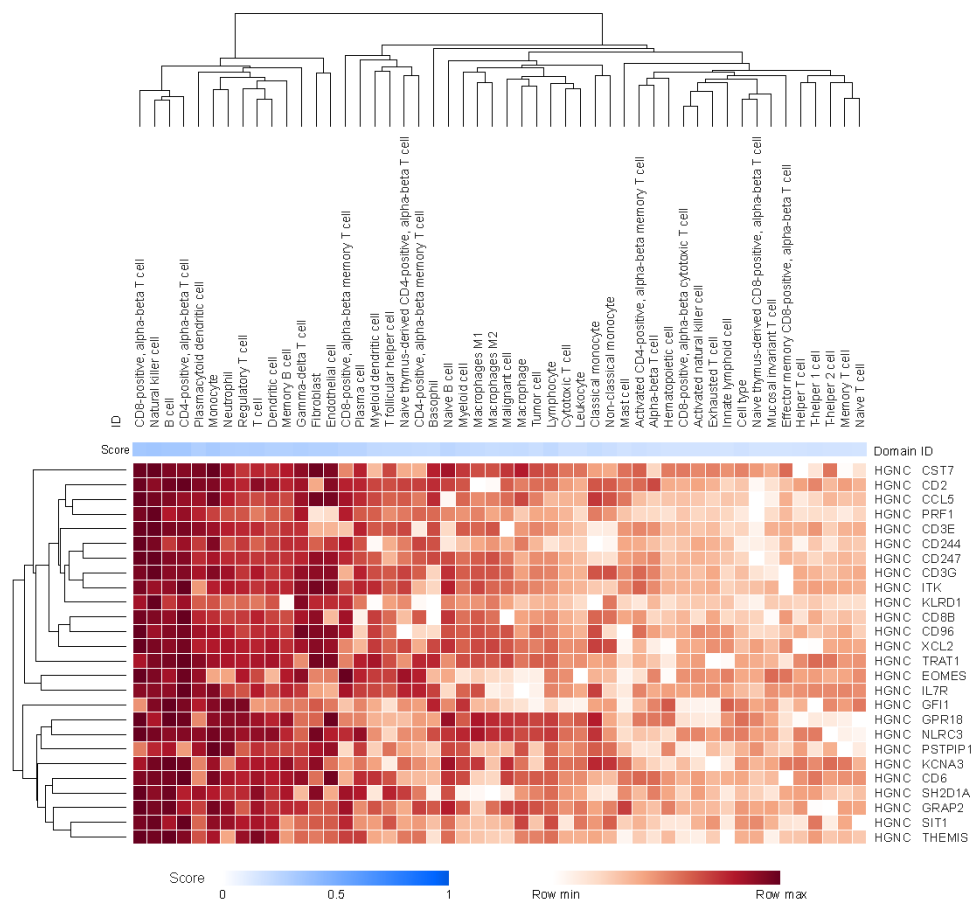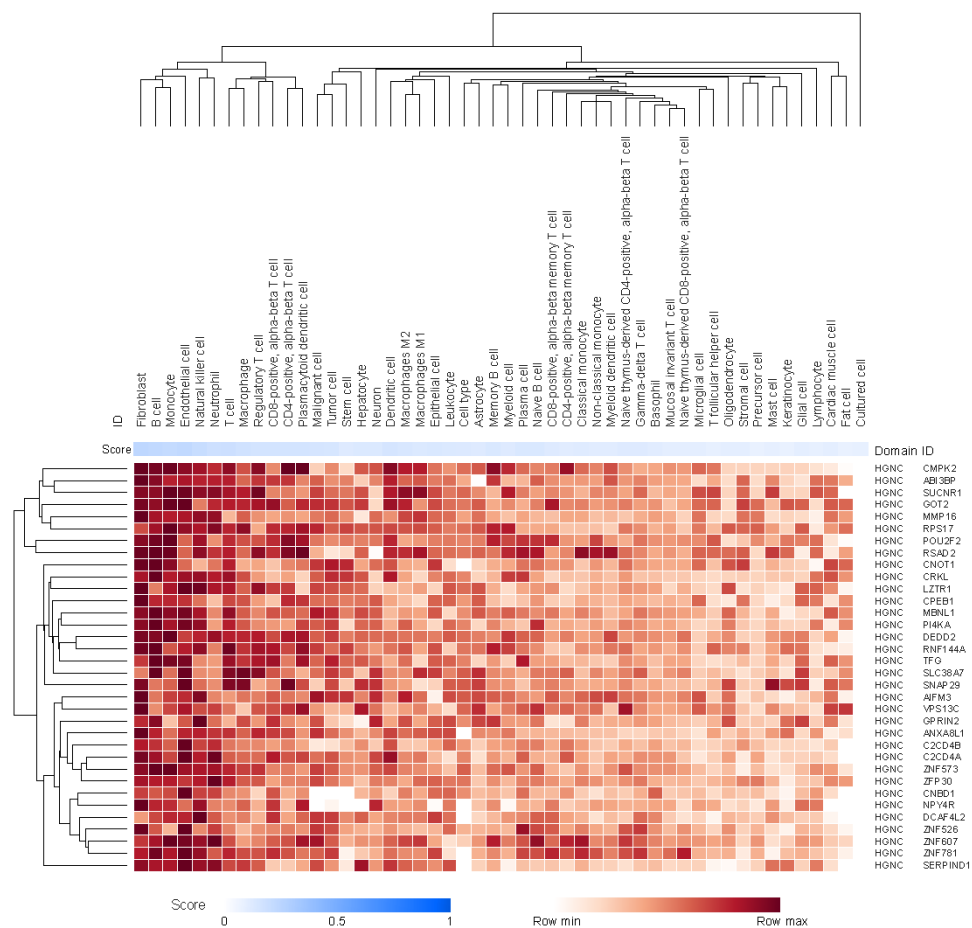
Figure 1

A

TMB ≤ median

Avelumab plus BSC

Median OS (95% CI), months
TMB ≤ median, JAV ≤ median (n=68) 12.52 (9.82-18.60)
TMB ≤median, JAV > median (n=72) 22.92 (17.74-NE)
HR, 0.52 (95% CI, 0.351-0.773); p=0.0012

No. at risk
| TMB ≤ median, JAV ≤ median | 68 | 41 | 21 | 13 | 5 | 1 |
| TMB ≤ median, JAV > median | 72 | 54 | 38 | 28 | 9 | 1 |

TMB > median

Median OS (95% CI), months
TMB > median, JAV ≤ median (n=66) 30.85 (23.69-NE)
TMB > median, JAV > median (n=75) 39.75 (29.96-NE)
HR, 1.22 (95% CI, 0.756-1.957); p=0.42

No. at risk
| TMB > median, JAV ≤ median | 66 | 55 | 43 | 28 | 15 | 2 | 0 |
| TMB > median, JAV > median | 75 | 61 | 51 | 31 | 15 | 2 | 0 |

BSC alone

Median OS (95% CI), months
TMB ≤ median, JAV ≤ median (n=70) 17.31 (10.81-26.84)
TMB ≤median, JAV > median (n=63) 14.55 (10.74-32.99)
HR, 0.96 (95% CI, 0.634-1.460); p=0.8576

No. at risk
| TMB ≤ median, JAV ≤ median | 70 | 46 | 28 | 19 | 6 | 1 |
| TMB ≤ median, JAV > median | 63 | 37 | 25 | 21 | 11 | 3 |

Median OS (95% CI), months
TMB > median, JAV ≤ median (n=72) 15.77 (13.40-27.47)
TMB > median, JAV > median (n=69) 18.23 (14.26-26.74)
HR, 1.01 (95% CI, 0.677-1.508): p=0.9588

No. at risk
| TMB > median, JAV ≤ median | 72 | 45 | 28 | 19 | 11 | 1 |
| TMB > median, JAV > median | 69 | 45 | 29 | 17 | 10 | 2 |

B

i. Whole genome discovery using >900k CCMs
20 JAV-Immuno^{hi} vs 20 JAV-Immuno^{lo}

ii. Nested PCR assays developed for 150 CCMs and evaluated on additional samples
40 JAV-Immuno^{hi} vs 40 JAV-Immuno^{lo}

iii. XGBoost models identified 25 candidate loci

Chromosome conformation
Chromosomal DNA, kb
EpiSwitch marker
EpiSwitch marker
Isolation
Restriction
Ligation
Detection
Restriction site
Present or Absent
Site-specific sequences
EpiSwitch marker

C

Association between JAV-Immuno and 25 loops across dilution levels (N=80)

Dilution: 1x 2x 4x

D

Association between individual selected CCM genes and JAV-Immuno levels

Max (0.81) 0 Max (-0.38)

686 **Figure 2. Cell types and functions of genes covered by the selected EpiSwitch CCM.** The

687 heat maps display aggregated associations between genes (rows) and cell types (columns in **A**

688 and **B**) or functions (columns in **C** and **D**) across public data represented in the Data4Cure

689 CURIE knowledge graph. The association score summarizes data and literature evidence from

690 publicly available datasets and publications linking biomedical entities. Hierarchical clustering

691 was applied to associations to generate the heat map. Association between cell types and

692 genes in the JAV-Immuno signature (**A**) and the selected CCM set (**B**). Only genes with mean

693 scores of ≥0.1 for at least 4 cell types are included. T cells and NK cells are strongly associated

694 with JAV-Immuno, whereas genes in the selected CCMs display more diverse associations,

695 including with endothelial cells, fibroblasts, B cells, and monocytes, as well as NK and T cells.

696 Comparison of Gene Ontology biological processes associated with JAV-Immuno (**C**) and the

697 selected CCM set (**D**). JAV-Immuno displays strong connections with T-cell processes, whereas

698 selected CCM genes are associated with more diverse processes.

699 **CCM**, chromatin conformation marker; **NK**, natural killer.

Figure 2

A

B

700

C
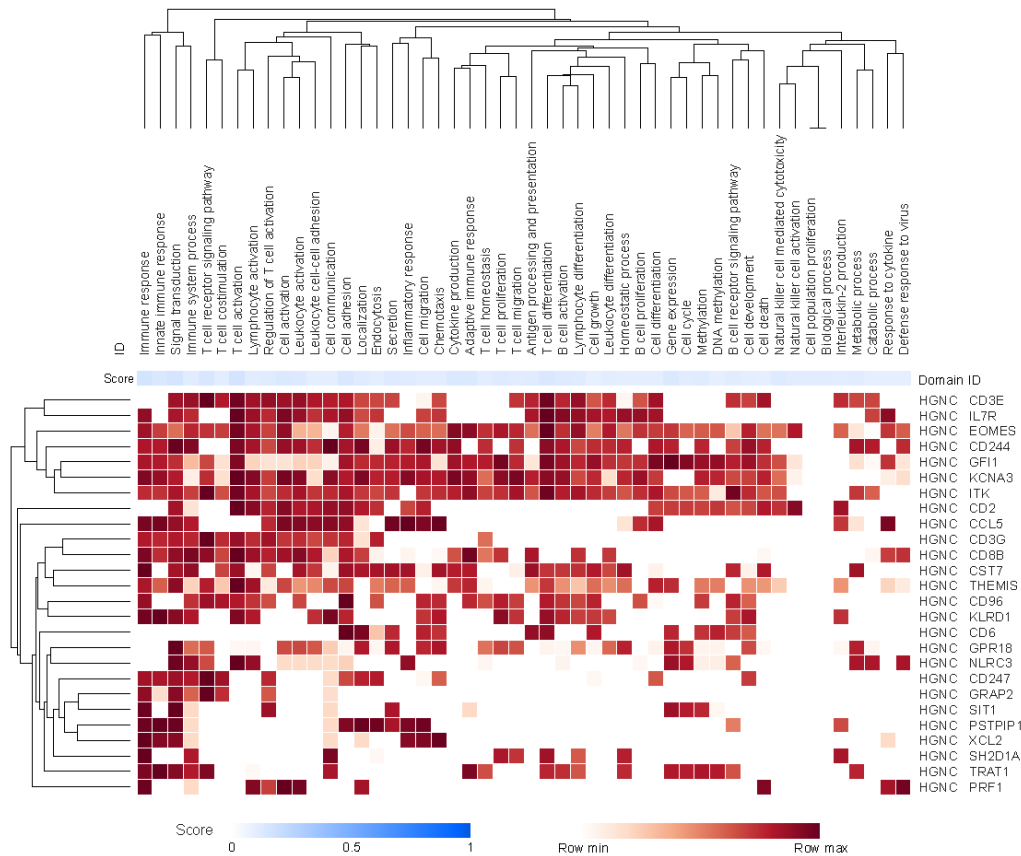
D
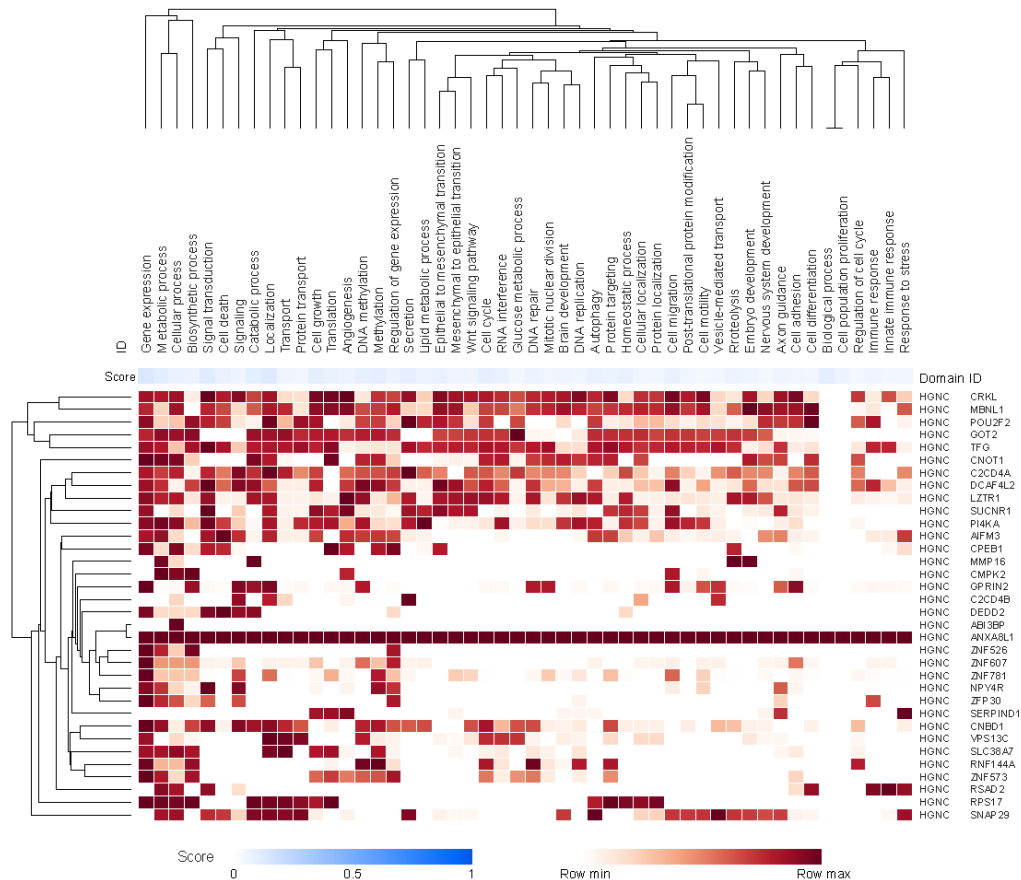
Figure 3. Spatial transcriptomic analysis in bladder cancer sections of marker set genes compared with JAV-Immuno genes. A. i. Hematoxylin and eosin stain of an index bladder cancer section (A1-1) showing the position of lymphoid aggregates annotated by a pathologist. ii. Spatial indexing transcriptomic data (Visium) from the same specimen processed by BayesSpace. Eight regions identified by BayesSpace were manually merged into 4 functional regions and annotated by majority cell type and histology. **B.** i. The A1-1 bladder specimen was annotated according to average expression of a published TLS signature [27] per spot. ii. Average expression level of the TLS signature over the 4 main regions. iii. Heat map depicting individual gene expression across the 4 regions. Color represents aggregation by mean of values normalized by z-score. Note that most genes in the signature are expressed predominantly in the lymphoid aggregate region. **C.** i. A1-1 specimen annotated according to average expression of JAV-Immuno per spot. ii. Average expression level of JAV-Immuno over the 4 main regions. iii. Heat map depicting individual gene expression across the 4 regions. Note that most genes in the JAV-Immuno signature are also expressed predominantly in the lymphoid aggregate region. iv. Concordance between JAV-Immuno genes localized to lymphoid aggregates in the A1-1 bladder specimen and TLS-high regions in the bladder/renal/breast cancer collection. **D.** i. Average expression level of the CCM model genes over the 4 regions. ii. Heat map depicting individual gene expression across clusters based on distribution of expression between regions. Approximately half of the marker-set genes are expressed in lymphoid aggregates. iii. The A1-1 bladder specimen annotated according to average expression of the CCM model genes per sp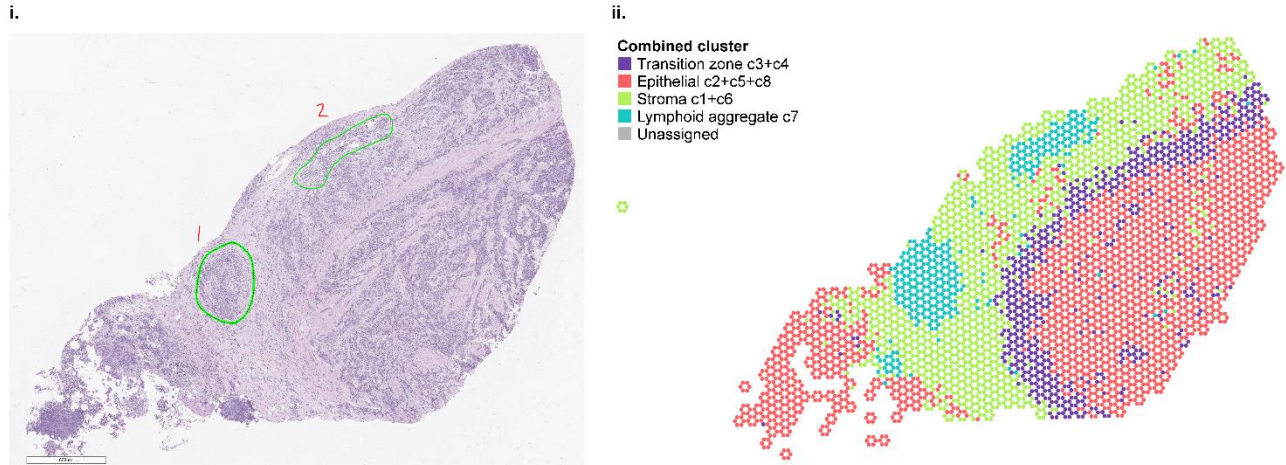ot. Separate spatial images are shown for the epithelial and lymphoid regions. iv. Concordance between genes localized to lymphoid aggregates in the A1-1 bladder specimen and TLS-h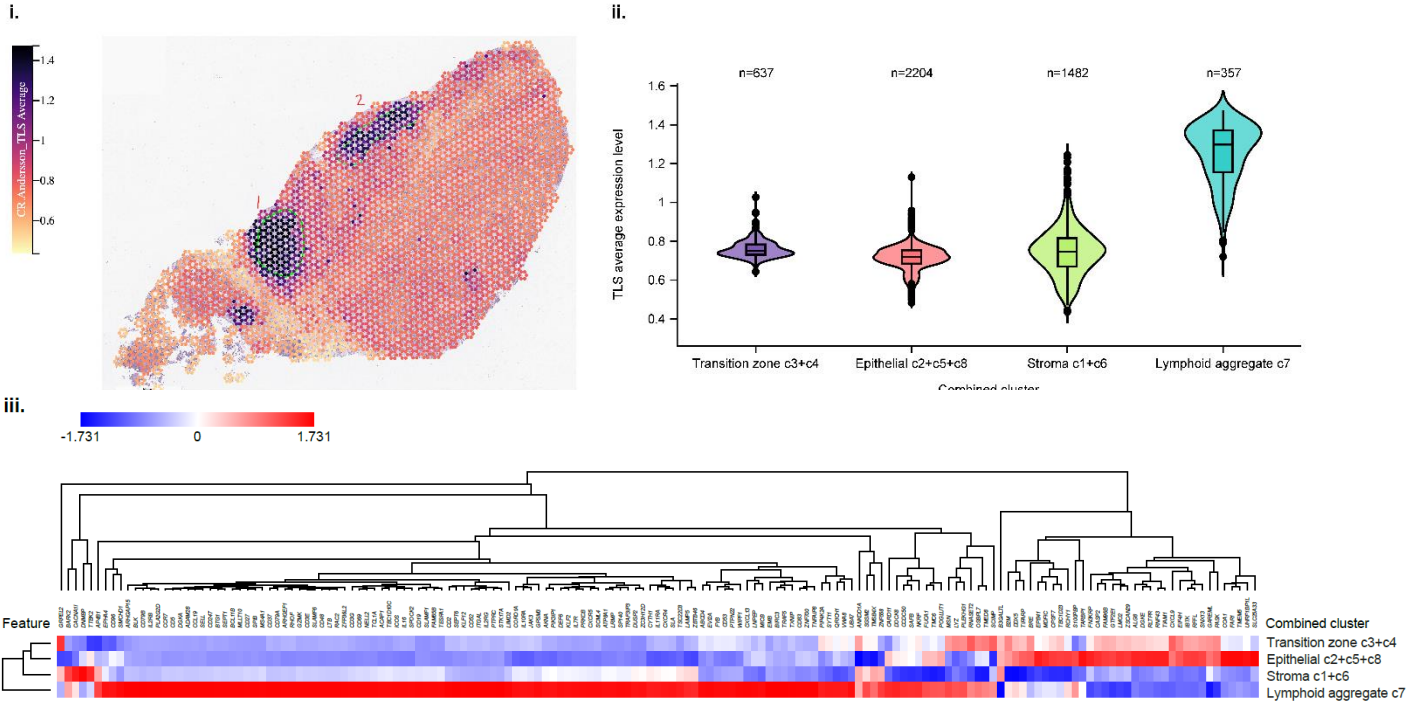igh regions in the bladder/renal/breast cancer collection. **CCM**, chromatin conformation marker; **TLS**, tertiary lymphoid structures.
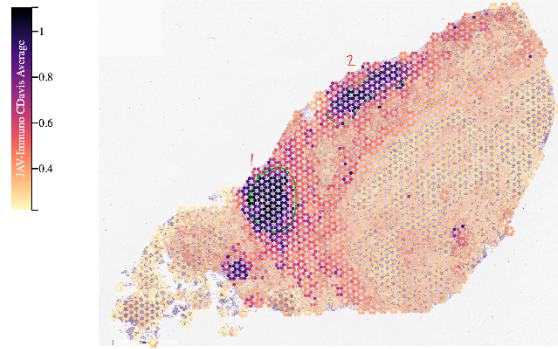
Figure 3
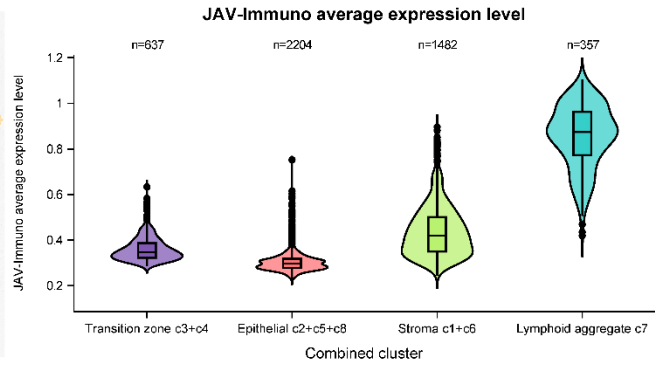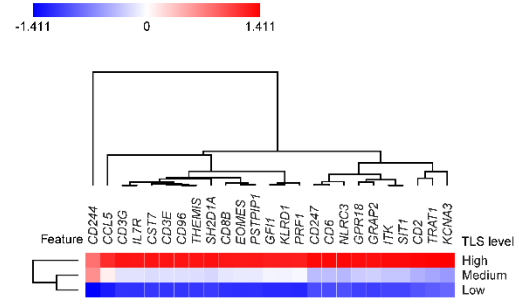
A

i.

ii.

Combined cluster
- Transition zone c3+c4
- Epithelial c2+c5+c8
- Stroma c1+c6
- Lymphoid aggregate c7
- Unassigned

B

i.

ii.

iii.

C

i.

ii.

JAV-Immuno average expression level

iii.

iv.

D

i.

ii.

iii.

Prominent in epithelial
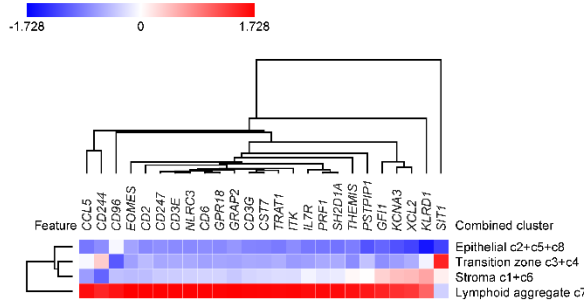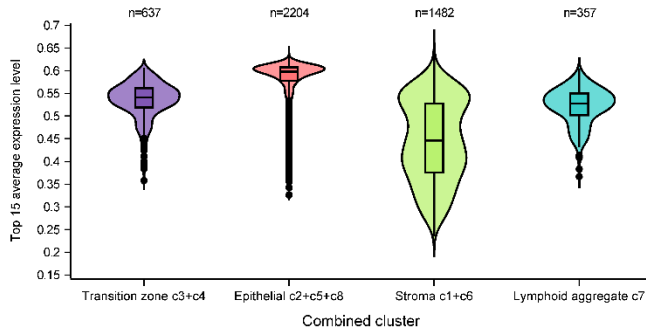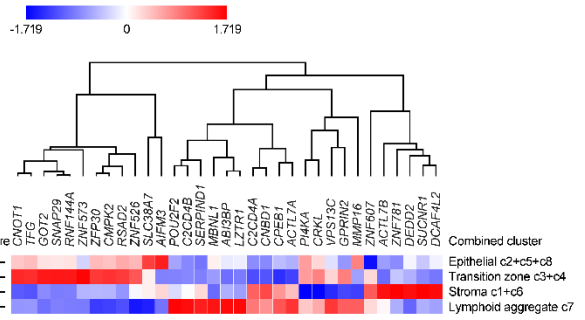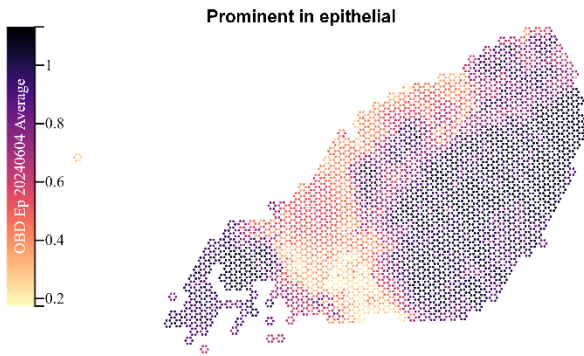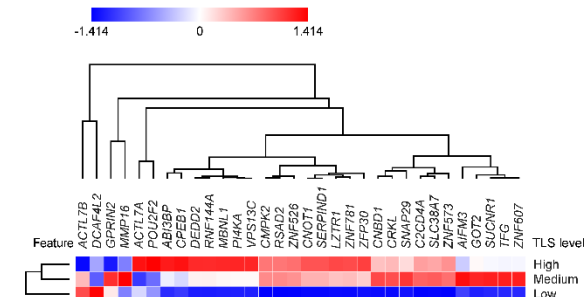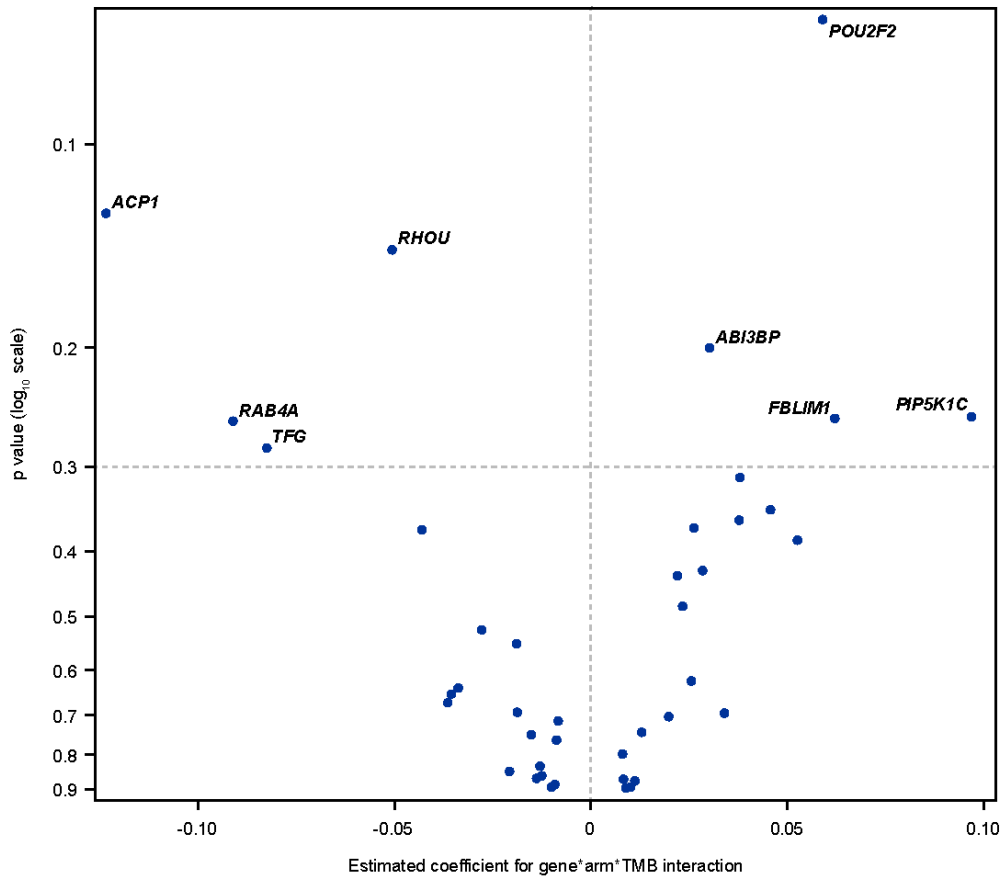
Prominent in lymphoid aggregate

iv.

81
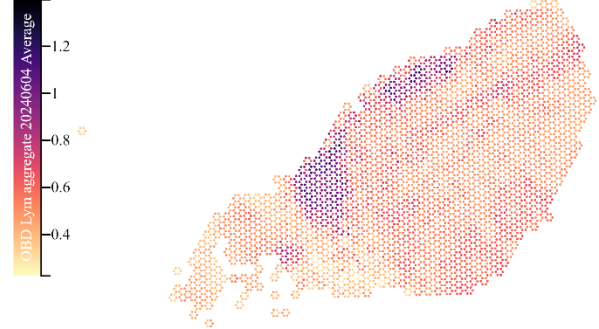
818 **Figure 4. Modulation of interactions between TMB and OS by *POU2F2*. A.** Three-way

819 interactions between POU2F2 expression in tumor, TMB, and treatment on OS outcome.

820 Continuous measures of TMB and gene expression ($\log_2$ TPM) were used in the analysis.

821 Genes that were expressed in ≤50% of samples or had a coefficient of variation of ≤5 for

822 expression were excluded from the analysis. The estimated coefficients for the 3-way interaction

823 and the p value on the y-axis were calculated based on a Cox model containing the

824 gene*TMB*treatment interaction term and all the related lower-order terms. Note that *POU2F2*

825 showed the strongest monogenic effect. **B.** Presence of the blood marker encompassing

826 *POU2F2* is associated with reduced expression of *POU2F2*, downstream target genes, and

827 gene expression signatures characteristic of TLS. **OS**, overall survival; **TLS**, tertiary lymphoid

828 structures; **TMB**, tumor mutation burden; **TPM**, transcript count per million.

829

830

831

832

833

834

835

836

837

838

839

840

841

842

843

Figure 4

A

B