Tumor eQTM screening reveals distinct CpG panels for deconvolving cancer immune signatures

**Supplementary figures**

Xiaoqing Yu\textsuperscript{1,2,†}, Ling Cen\textsuperscript{1,†}, Y. Ann Chen\textsuperscript{1}, Joseph Markowitz\textsuperscript{3,5}, Timothy I. Shaw\textsuperscript{1}, Kenneth Y. Tsai\textsuperscript{4}, Jose R. Conejo-Garcia\textsuperscript{2,5,*}, and Xuefeng Wang\textsuperscript{1,2*}

\textsuperscript{1}Department of Biostatistics and Bioinformatics, H. Lee Moffitt Cancer Center & Research Institute, Tampa, FL 33612, USA
\textsuperscript{2}Moffitt Cancer Center Immuno-Oncology Program, Tampa, FL 33612, USA
\textsuperscript{3}Department of Cutaneous Oncology, H. Lee Moffitt Cancer Center & Research Institute, Tampa, FL, USA
\textsuperscript{4}Department of Pathology, H. Lee Moffitt Cancer Center & Research Institute, Tampa, FL 33612, USA
\textsuperscript{5}Department of Immunology, H. Lee Moffitt Cancer Center & Research Institute, Tampa, FL 33612, USA

\* These authors contributed equally to this work.

\† To whom correspondence should be addressed. Tel: +1-813-745-6710 Fax: +1-813-745-6107 Email: xuefeng.wang@moffitt.org.
Supplementary Fig. S1. Characterization of eQTM-pos in cutaneous melanoma based on the TCGA SKCM cohort. (a) The two pie charts show the distribution of the melanoma eQTM-pos (with positive correlation between CpG and corresponding gene, i.e. $\rho > 0.3$) CpG sites, in terms of gene region and CpG island relationship, respectively. (b) Top 30 eQTM-pos CpGs ranked by correlation coefficients in melanoma. The orange bars indicate the magnitude of the positive correlations and the green bars show the magnitude of the positive correlations if there is eQTM-neg in the gene region.
Supplementary Fig. S2. Forest plot showing the hazard ratios of top prognostic eQTM CpG (for predicting patient overall survival) after adjusting for tumor purity and immune scores (Appendix to Fig. 1C).
Supplementary Fig. S3. Full Reactfoam plot of top prognostic eQTM genes showed they are highly enriched for immune related pathways. Reactfoam plot shows a high-level summary of the pathway enrichment analysis. The enriched pathways are shown in yellow shade. The zoomed-in plot for the main enriched pathways in Immune System is shown in Fig. 1D.
Supplementary Fig. S4. Fit criteria from best subset regression analysis shows that a 3-CpG panel reaches sufficient precision in predicting the CYT score in melanoma. The model was fit using seven candidate eQTM (CpGs listed in Supplementary Table S2) in CYT related genes. The plot was generated using the function ols_step_best_subset in R package "olsrr".
Supplementary Fig. S5. Volcano plots displaying the top predictive eQTM genes for APM and ISG.RS signatures. The x-axis is the average effect size (beta) in the final models (representing predictability) and y-axis indicate how many times that one eQTM was selected after 200 resampling training (representing reliability), same as in Fig. 3C-D.
Supplementary Fig. S6. The PCA loading plot for principal components 2 (x-axis) and 3 (y-axis) (supplementary to Fig. 4A). The PCA loading plot shows how strongly each gene correlates or influences each principal component.
Supplementary Fig. S7. Forest plot summarizing results from pan-cancer analysis of multivariable Cox models testing the prognostic value of the CYT-TCF7 subgroups. The hazard ratios are for the group II defined in the study (TCF7 low/CYT high). Supplementary to Fig. 4D.
Supplementary Fig. S8. Prognostic value of the TCF7 eQTM in the multivariable Cox models (Supplementary to Fig. 4E).

(A) Prognostic significance of TCF7 eQTM (based on methylation value of cg259477408) across 33 TCGA cancer types. The estimated hazard ratios are based on the Cox regression models for each individual cancer type after adjusting for tumor purity, and ssGSEA-based immune scores.

(B) The forest plot summarizes the results from the multivariable Cox model adjusting for sex, grade, tumor purity, and immune scores. The hazard ratio from the group II remains significant after the adjustment.
Supplementary Fig. S9. Stratified eQTM results by primary vs metastatic tumor types in TCGA.

(A) Pie charts showing the distribution of negative and positive eQTMs stratified by tumor types.
(B) The UpSet plots show the number of shared eQTMs identified using the primary and metastatic tumors only.
(C) Scatter plots comparing the correlation of eQTMs identified based two tumor types.
(D) Scatterplots and correlations (Pearson) between seven cis-eQTM CpGs and tumor CYT score in metastatic melanoma tumors.
(E) Scatterplots and correlations (Pearson) between seven cis-eQTM CpGs and tumor CYT score in primary melanoma tumors.
(F) Scatter plot of cg259477408 (eQTM in TCF7) methylation value vs gene expression (log2 TPM) value of TCF7 and PDCD1.
(G) Survival analysis of TCGA SKCM patients stratified according to the CYT-TCF joint eQTM signatures based on metastatic tumors only.
**Supplementary Fig. S10. Basic clinical characteristics of patients in each melanoma patient subgroup defined in Fig. 4D.**

<table>
<thead>
<tr>
<th>Variable</th>
<th>I (N = 7)</th>
<th>II (N = 8)</th>
<th>III (N = 8)</th>
<th>IV (N = 7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FEMALE</td>
<td>2 (29%)</td>
<td>2 (25%)</td>
<td>3 (38%)</td>
<td>2 (29%)</td>
</tr>
<tr>
<td>MALE</td>
<td>5 (71%)</td>
<td>6 (75%)</td>
<td>5 (62%)</td>
<td>5 (71%)</td>
</tr>
<tr>
<td>Grade</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>5 (71%)</td>
<td>4 (50%)</td>
<td>4 (50%)</td>
<td>2 (29%)</td>
</tr>
<tr>
<td>IV</td>
<td>2 (29%)</td>
<td>4 (50%)</td>
<td>4 (50%)</td>
<td>5 (71%)</td>
</tr>
<tr>
<td>Survival Status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alive</td>
<td>5 (71%)</td>
<td>8 (100%)</td>
<td>3 (38%)</td>
<td>2 (29%)</td>
</tr>
<tr>
<td>Dead</td>
<td>2 (29%)</td>
<td>5 (62%)</td>
<td>5 (71%)</td>
<td></td>
</tr>
<tr>
<td>Tissue</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metastatic</td>
<td>7 (100%)</td>
<td>8 (100%)</td>
<td>4 (50%)</td>
<td>7 (100%)</td>
</tr>
<tr>
<td>Primary</td>
<td></td>
<td></td>
<td>4 (50%)</td>
<td></td>
</tr>
</tbody>
</table>

\(^1\text{n (%)}\)
Supplementary Fig. S11. Comparisons of eQTM correlations with partial correlations adjusted for tumor purity.

(A) The UpSet plots show the number of shared eQTMs (positive eQTMs on the left panel, and negative ones on the right panel) identified using the unadjusted CpG-GE correlation and tumor purity adjusted correlation.

(B) Scatter plots comparing the unadjusted correlation (X axis) with tumor purity adjusted correlation coefficients (Y axis) on eQTMs identified based on the adjusted-correlation method.

(C) Scatter plots comparing the unadjusted correlation (X axis) with tumor purity adjusted correlation coefficients (Y axis) on eQTMs identified based on the unadjusted-correlation method.
Supplementary Fig. S12. Heatmap representation of all TCGA melanoma samples using the selected immunosuppressive genes. The top panel shows the gene expression and the bottom panel shows the CpG methylation values from the corresponding eQTM.