CXCR4 and adaptive anti-tumor immunity in the pancreas cancer TCGA cohort

Short Analysis Report

Piotr Tymoszuk, daas.tirol

2022-06-13

Contents

Analysis summary

- The primary goal of the analysis was to check co-expression of *CXCR4* with the gene expression readouts of immune infiltration in pancreatic cancer and, in particular, with cytotoxic T cells. The secondary tasks were to determine prognostic relevance of tumor *CXCR4* expression and changes in whole genome expression and signaling pathways in *CXCR4*$^{\text{high}}$ tumors as compared with *CXCR4*$^{\text{lo}}$ malignancies.

- A total of 147 primary pancreatic adenocarcinoma samples from the TCGA PAAD cohort were analyzed[1]. Estimates of Quantiseq immune infiltration and immune response scores for the TCGA PAAD RNA Seq samples were extracted from the TIMER 2.0 platform[2]. Gene members of the T cell exhaustion[3], cytotoxic signature[4], IFN-γ, expanded immune[5] and tumor inflammation signatures[6] were extracted from the respective literature and the signature scores calculated by GSVA[7].

- The methodology encompassed self-organizing map (SOM) participant clustering by the immune infiltration estimates[8,9], Pearson's correlation of *CXCR4* expression with estimates of immune cell infiltration and comparison of survival differences between *CXCR4*$^{\text{high}}$ and *CXCR4*$^{\text{low}}$ tumors using automatically optimized expression cutoffs[10,11], Kaplan-Meier method and Cox regression. Differentially regulated genes in *CXCR4*$^{\text{high}}$ vs *CXCR4*$^{\text{low}}$ tumors were identified by Benjamini-Hochberg-corrected two-tailed T test[12]. Differences in biological processes and signaling pathway modulation between the *CXCR4* expression strata were investigated with GO term and KEGG pathway enrichment procedures[13] and with the *SPIA* algorithm[14].

- By means of SOM and hierarchical clustering[8,9], three pancreatic tumor subsets were identified, which differed primarily in malignant cell content, cytotoxic T cells, regulatory CD4$^+$ T cells and TAM (tumor-associated macrophages) infiltration (**Figure 1A**, **Supplementary Figure S1** and **S2**). *CXCR4* was predominantly expressed in the malignancies characterized by high Treg, CD8$^+$ T cells and tumor-associate macrophage (TAM) abundance (**Figure 1**).

- Expression of *CXCR4* was significantly positively correlated with the values of the immune and microenvironment scores[2] and Quantiseq estimates of T regs, CD8$^+$ T cells and TAMs and was inversely correlated with the contents of non-leukocyte cells (**Figure 2**). This suggests that the chemokine receptor may be predominantly expressed by stroma cells and especially cytotoxic T cells, T regs and TAMs.

- Quite surprisingly, high *CXCR4* expression tended to be associated with better overall survival but was only weakly linked to tumor-related (TRS) and relapse-free survival (RFS) in Kaplan-Meier analysis (**Figure 3**).

- Comparison of the cancer immune signatures values[2–6] between the *CXCR4*$^{\text{high}}$ and *CXCR4*$^{\text{low}}$ tumors stratified by the optimal TRS cutoff (**Figure 3**) reveals higher expression of the immune, microenvironment, cytotoxicity and stroma score as well as expanded immune signature in the CXCR4 high expression strata (**Figure 4**). Investigation of the unique signature gene members points toward highly upregulated levels of chemokines (*CXCL9*, *CXCL10*, *CXCL13*), granzymes, T cell lineage markers (*CD3E*, *CD2*, *CD8A*) along with T cell exhaustion markers (*TIGIT*, *PDCD1*, *HAVCR2*) in *CXCR4*$^{\text{high}}$ neoplasms (**Figure 5**).

- Differential gene expression (**Figure 6**, **Supplementary Figure S3**) with subsequent pathway modulation and enrichment analysis[13,14] indicates higher activity of NK cell cytotoxicity, chemokine and focal adhesion pathways together with increased MAPK signaling in the *CXCR4*$^{\text{high}}$ than in the *CXCR4*$^{\text{low}}$ TCGA pancreatic tumors (**Figure 7**, **Supplementary Figure S4**).

- Collectively, the results suggest that *CXCR4* is likely expressed by tumor-infiltrating cytotoxcic CD8$^+$ cells, Tregs and macrophages. Unexpectedly, *CXCR4* gene does not seem to play a protective role or contribute to better survival in the TCGA PAAD cohort. Interestingly, *CXCR4* is co-expressed both with the typical pan T cell, CD8$^+$ T cell and cytotoxicity markers as well as some exhaustion genes.

The biological impact of *CXCR4* on T cell priming, anergy and exhaustion may be quite complicated (e.g. dependent of T cell subtype) and needs to be clarified by ex vivo experiments.

Methods

Study cohort, immune infiltration data, gene signatures

A total of 147 primary pancreatic tumor samples from the TCGA PAAD cohort were analyzed[1]. The normalized RNAseq (level 3) togehter with clinical data were fetched from the GDC repository using TCGA Assembler 2.0 (https://github.com/compgenome365/TCGA-Assembler-2/tree/master/TCGA-Assembler) as described[10].

Estimates of immune infiltration and immune response scores for the TCGA PAAD RNA Seq samples were extracted from the TIMER 2.0 platform[2] and, in case of multiple estimates of the same population by multiple algorithms, pooled with the gene set variation analysis (GSVA) algorithm[7] and termed further 'pooled multi-algorithm estimates'. Gene members of the T cell exhaustion[3], cytotoxic[4], IFN-γ, expanded immune[5] and tumor inflammation signatures[6] were extracted from the respective literature and the signature scores calculated by GSVA[7].

Data transformation and visualization

The study variables were transformed, analyzed and visualized with R version 4.0.5 with *tidyverse*[15,16], *cowplot*[17] and *survminer*[18] packages. RNAseq data in form of FKPM counts were transformed with the $log_2(FKPM + 1)$ function to improve normality.

Clustering

The samples were clustered using a two-step combined self-organizing map (SOM) and hierarchical clustering algorithm[8,9]. In the first step, samples were assigned to the nodes of hexagonal-topology SOM grids (samples: 7×7) with the Euclidean distance measure between the participants[8,9,19]. SOM assignment was accomplished with the tools provided by *kohonen* package[19] and home-developed wrappers (package *clustTools*, https://github.com/PiotrTymoszuk/clustTools). The SOM training process is visualized in **Supplementary Figure S1A**. In the second step, SOM nodes were subjected to hierarchical clustering with Euclidean distance measure. The optimal cluster number (k = 3) was determined by the bend of the within sum-of-squares and visual analysis of the dendrograms (**Supplementary Figure S1BC**). The hierarchical clustering was done with the base *hclust()* function and home-developed package *clusterTools* (https://github.com/PiotrTymoszuk/clustTools) for clustering quality control and visualization.

Differences of expression of immune estimates and *CXCR4* between the tumor sample clusters were compared with one-way ANOVA. P values were corrected for multiple comparisons with Benjamini-Hochberg method[12].

Co-expression analysis

Correlation of immune infiltration estimates values and *CXCR4* expression levels was done with Spearman correlation test. Significance ($\rho \neq 0$) was assessed with two-tailed T test. P values were corrected for multiple comparisons with Benjamini-Hochberg method[12]. Base R function *cor.test()* and the in-house-developed package *ExDA* were used in the correlation analysis (https://github.com/PiotrTymoszuk/ExDA).

Survival analysis

To compare differences in overall (OS), tumor-related survival (TRS) and relapse-free survival (RFS) between high and low expressors of *CXCR4* , the *CXCR4* expression levels were subjected to an optimal stratification procedure using an in-house developed iterative algorithm (development package *kmOptimizer*, https://

github.com/PiotrTymoszuk/kmOptimizer) described previously[10,11]. The minimal strata size was set to 20% of the cohort.

Differences in survival time between the low and high expressors were compared with Kaplan-Meier method and Mentel-Henszel test (package *survival* and *kmOptimizer*). P values were corrected for multiple comparisons with Benjamini-Hochberg method[12].

Differential expression of immune signatures, immune genes and whole-genome transcripts

Pancreatic tumors were stratified into *CXCR4*^{high} and *CXCR4*^{low} samples by the optimal cutoff for the difference in overall survival. Significance of differences in expression of immune signature values or expression of unique signature member genes between the *CXCR4* strata was assessed by two-tailed T test (development package *ExDA*, https://github.com/PiotrTymoszuk/ExDA). P values were corrected for multiple comparisons with Benjamini-Hochberg method[12].

Whole-genome differentially expressed genes were identified by Benjamini-Hochberg-corrected T test[12] and 1.5-fold regulation cutoff (**Supplementary Table S2**).

GO and KEGG enrichment analysis, signaling pathway modulation analysis

Biological process Gene Ontology (GO) term and KEGG pathway enrichment analysis with the sets of genes significantly up- or down-regulated in *CXCR4*^{high} as compared with *CXCR4*^{low} tumors (**Supplementary Table S2**) was performed with *goana()* and *kegga()* functions, respectively, provided by *limma* package[13]. P values for significant enrichment were corrected for multiple comparisons by Benjamini-Hochberg method[12].

Signaling pathway modulation analysis employing the genes differentially expressed between the *CXCR4* expression strata (**Supplementary Table S2**) was conducted with *spia()* tool from *SPIA* package[14]. Significantly modulated pathways were defined by the aggregated, Benjamini-Hochberg-corrected enrichment/perturbation p value (pGFDR < 0.05)[12]. Magnitude of pathway modulation and activation/inhibition status is expressed as the total perturbation accumulation in the pathway (tA)[14]. Complete significant pathway modulation results are presented in **Supplementary Table S3**.

Data and code availability

The complete TCGA analysis pipeline is available at https://github.com/PiotrTymoszuk/CXCR4-TCGA.

Figures

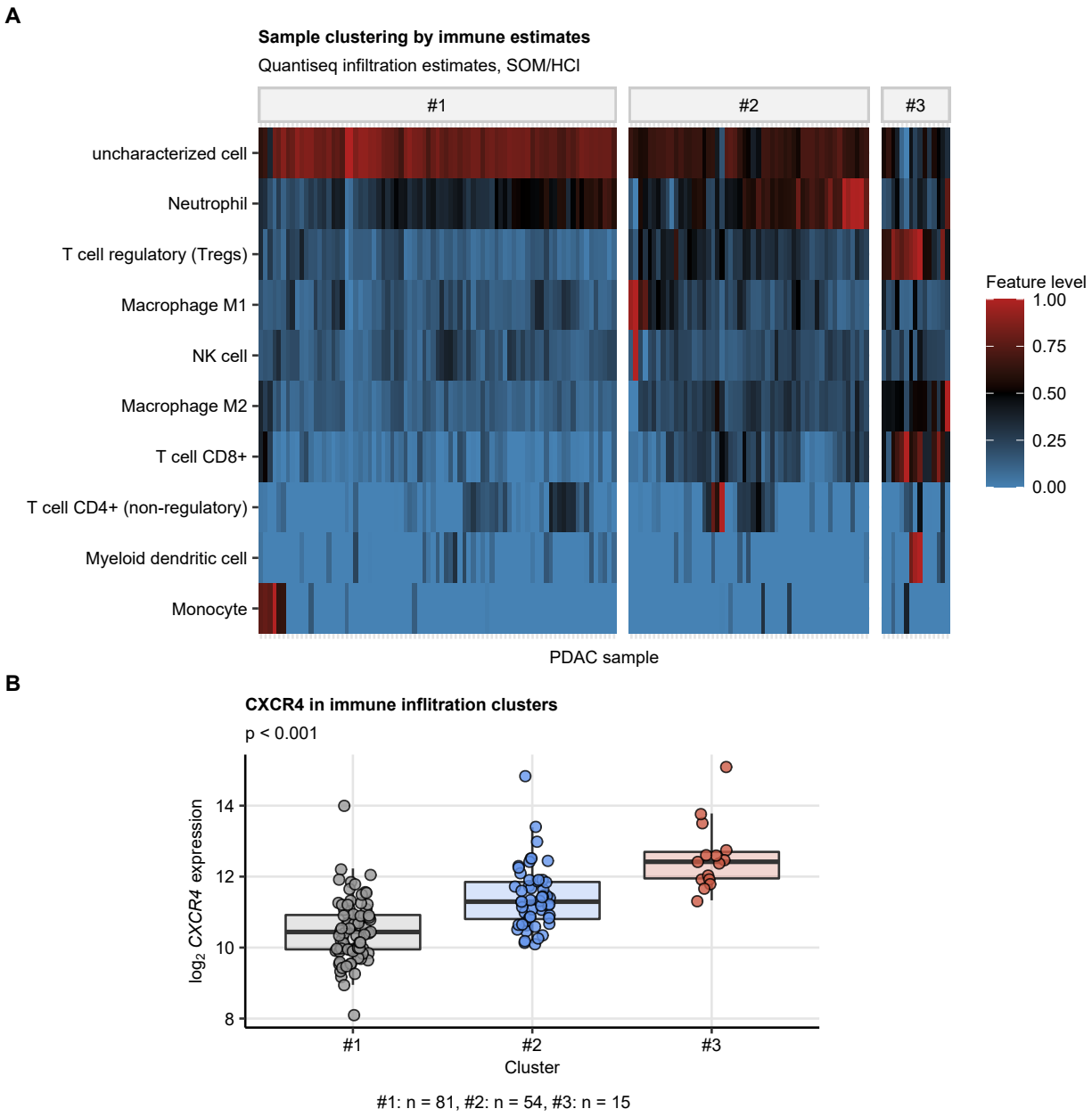

Figure 1: Predominant expression of *CXCR4* in stroma -, $CD8^+$ - and TAM-rich pancreatic tumors.

Figure 1. Predominant expression of *CXCR4* in stroma -, $CD8^+$ - and TAM-rich pancreatic tumors.

Tumor samples and pooled multi-algorithm immune infiltration estimates were clustered with a two step self-organizing map/hierarchical clustering procedure (**Supplementary Figure S1**). Differences in immune infiltration estimates and $\log_2$ *CXCR4* expression between the *Stroma-high*, *-intermediate* and *-low* tumor subsets were determined by one-way ANOVA and p values corrected for multiple testing with Benjamini-Hochberg method.

(A) Values of immune infiltration estimates in the *#1*, *#2* and *#3* tumor subsets presented as a heat map.

(B) $\log_2$ *CXCR4* expression in the tumor subsets. p value for the expression difference is shown in the plot heading. N numbers of samples assigned to the tumor subsets are provided under the plot.

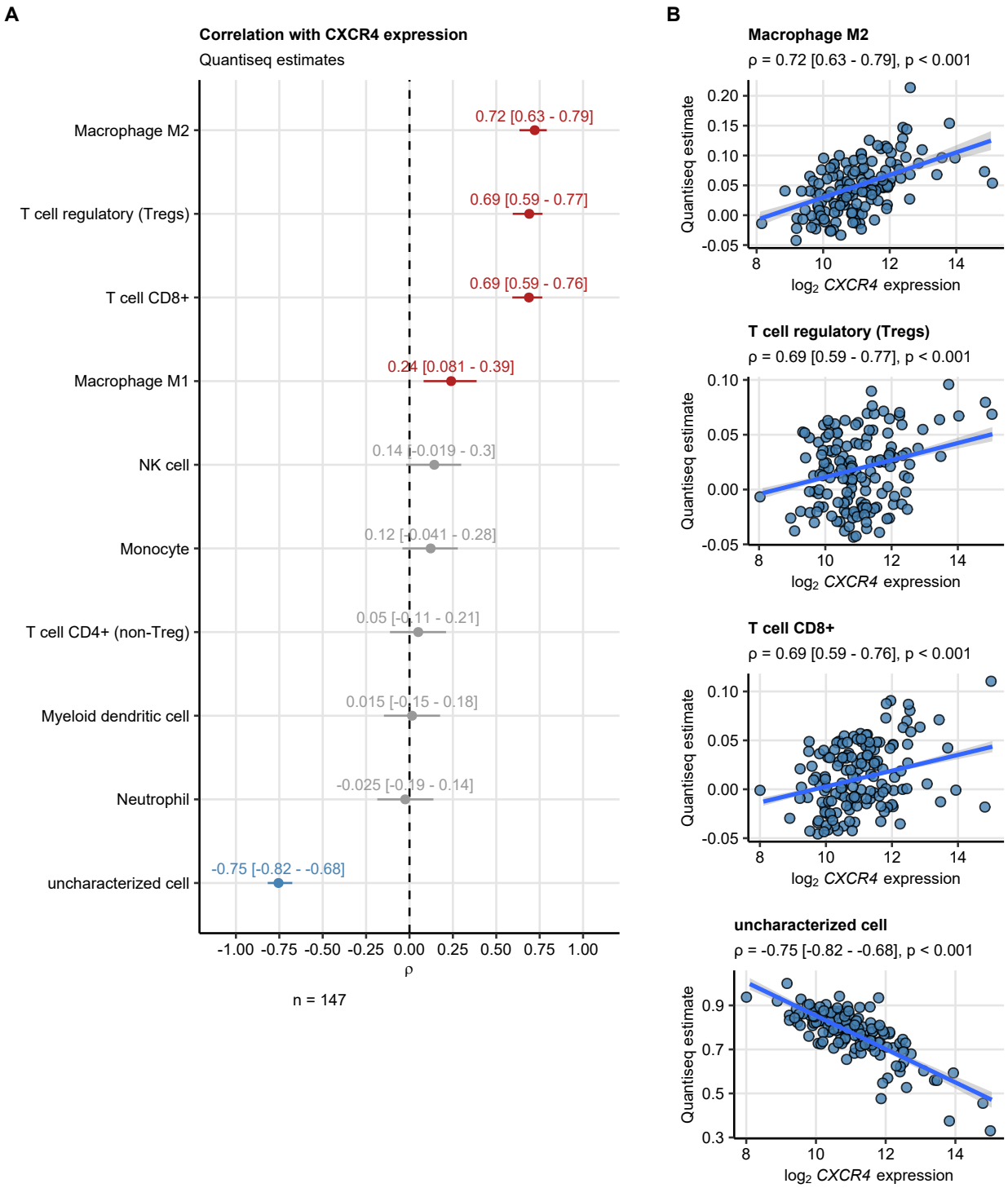

Figure 2: Correlation of the pooled multi-algorithm immune infiltration estimates with *CXCR4* expression.

Figure 2. Correlation of the pooled multi-algorithm immune infiltration estimates with *CXCR4* expression.

Pooled multi-algorithm estimates of immune cell features were correlated with log_2 expression levels of CXCR4 with Spearman correlation (**Supplementary Table S1**). P values were corrected for multiple comparisons with Benjamini-Hochberg method.

(A) ρ correlation coefficients with 95% confidence intervals are shown for the significant immune feature estimates. Red: positive correlation, blue: negative correlation. N number of observations is indicated below the plot.

(B) Plots of values of M2 macrophages, regulatory T cell, $CD8^+$ T cell and non-leukocyte uncharacterized cell infiltration estimates vs $\log_2$ *CXCR4* expression. Each point represents a single sample, blue lines represent fitted linear trends and gray ribbons represent 95% confidence intervals. Values of ρ with 95% confidence intervals and p values are shown in the plot captions.

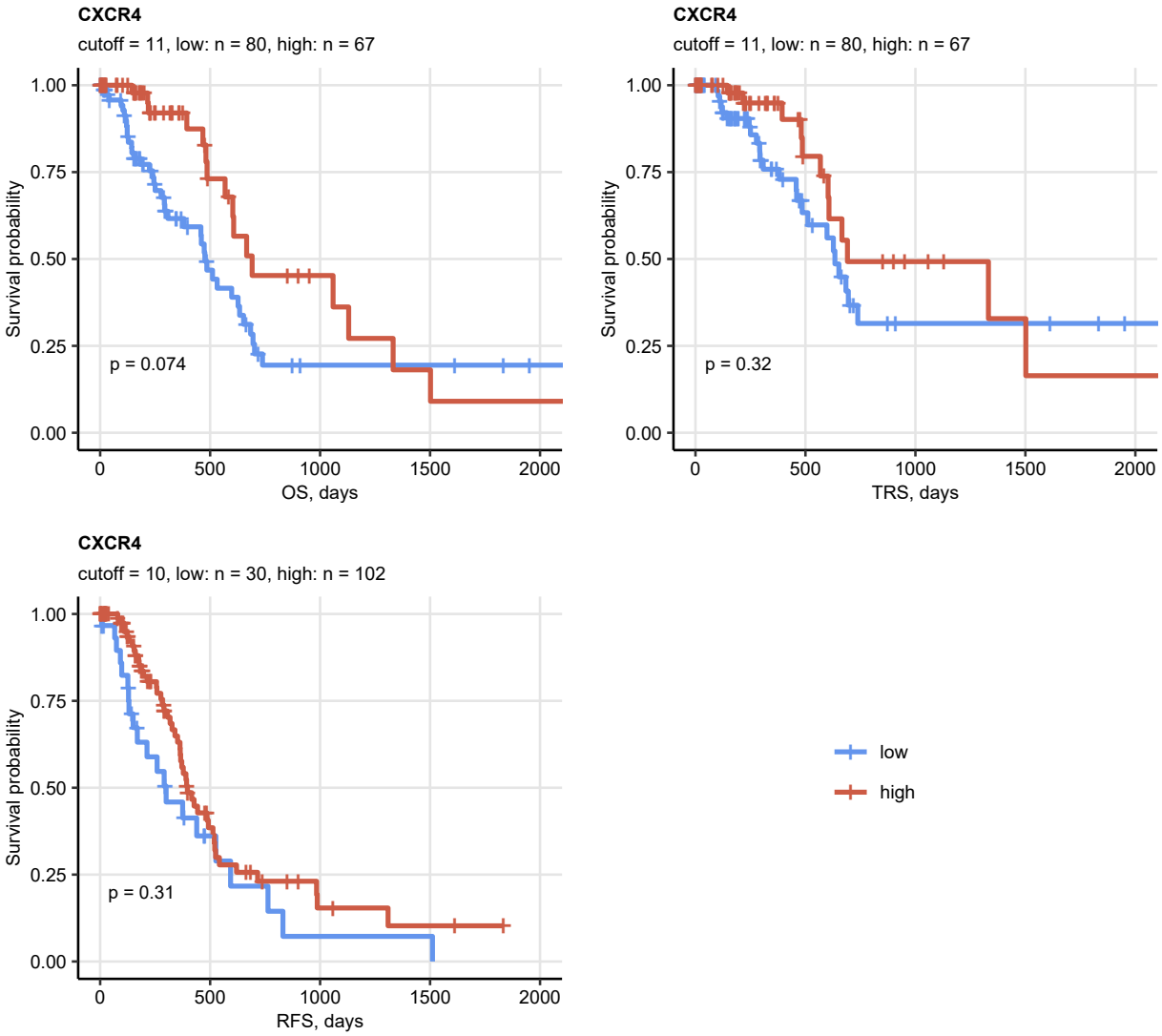

Figure 3: Differences in survival between $CXCR4^{high}$ and $CXCR4^{high}$ cancer patients.

Figure 3. Differences in survival between *CXCR4*$^{\text{high}}$ and *CXCR4*$^{\text{low}}$ cancer patients.

Tumor samples were stratified as high and low expressors for $\log_2$ *CXCR4* expression with automatically-determined optimal cutoffs. Differences in overall (OS), tumor-related (TRS) and relapse-free survival (RFS) were compared by Mentel-Henszel test corrected for multiple comparisons with Benjamini-Hochberg method. The fraction of survivors were presented in Kaplan-Meier plots. N numbers of observations in the *CXCR4* strata are indicated below the plots.

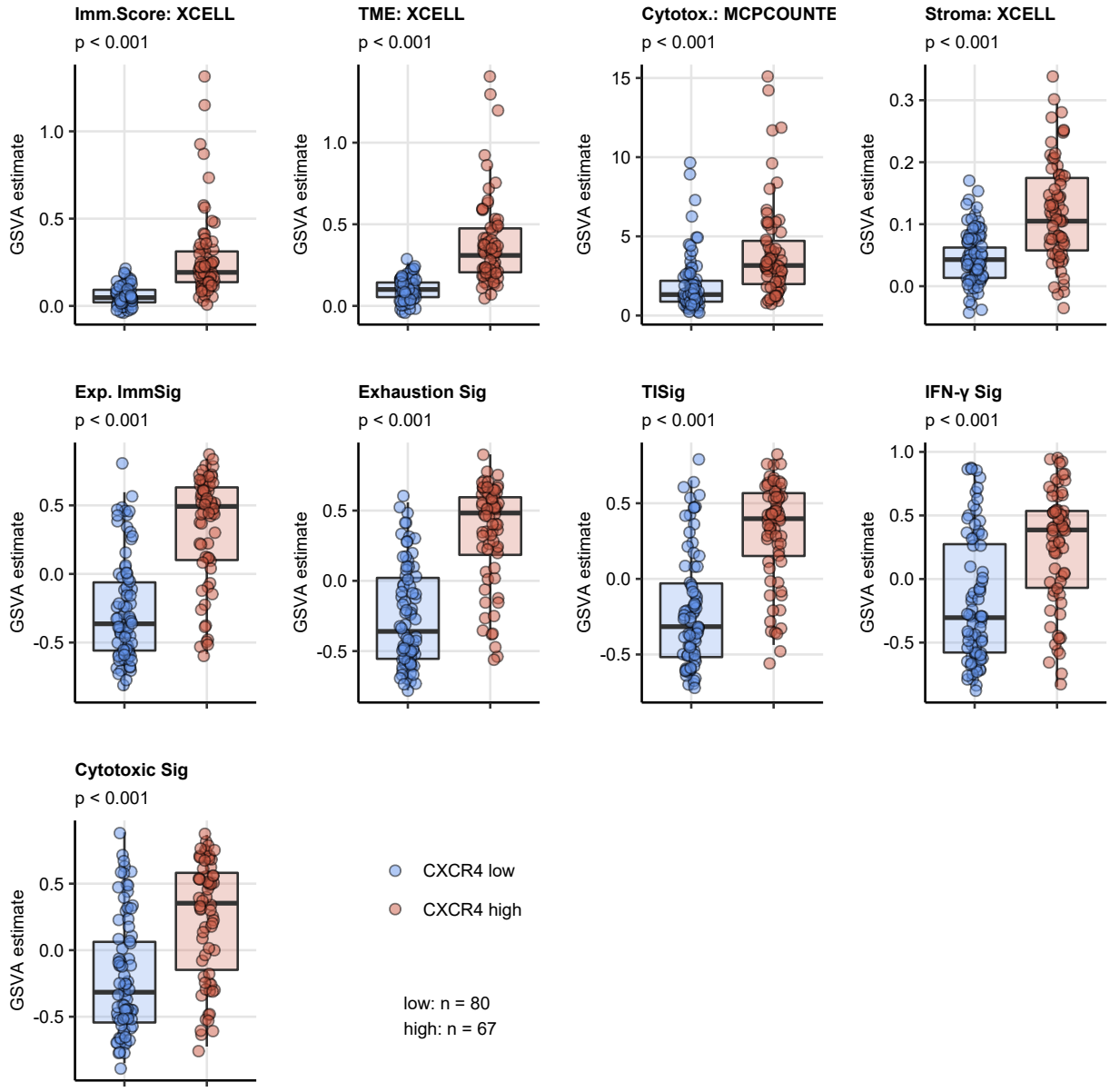

Figure 4: Differences in values of immune signatures between $CXCR4^{high}$ and $CXCR4^{high}$ tumors.

Figure 4. Differences in values of immune signatures between $CXCR4^{\text{high}}$ and $CXCR4^{\text{low}}$ tumors.

Tumor samples were stratified as $CXCR4^{\text{high}}$ vs $CXCR4^{\text{low}}$ expressors by the optimal cutoff for the difference in tumor-related survival. Differences in values of immune signatures (provided by TIMER: Immune Score, Microenvironment Score, Cytotoxicity Score and Stroma Score, provided by literature: Expanded Immune Signature, Exhaustion Signature, TISig: Tumor Inflammation Signature, IFN-γ Signature, Cytotoxic Signature) between the *CXCR4* strata were investigated with T test. Benjamini-Hochberg p values are presented in plot captions. N numbers of the strata samples are indicated next to the plots.

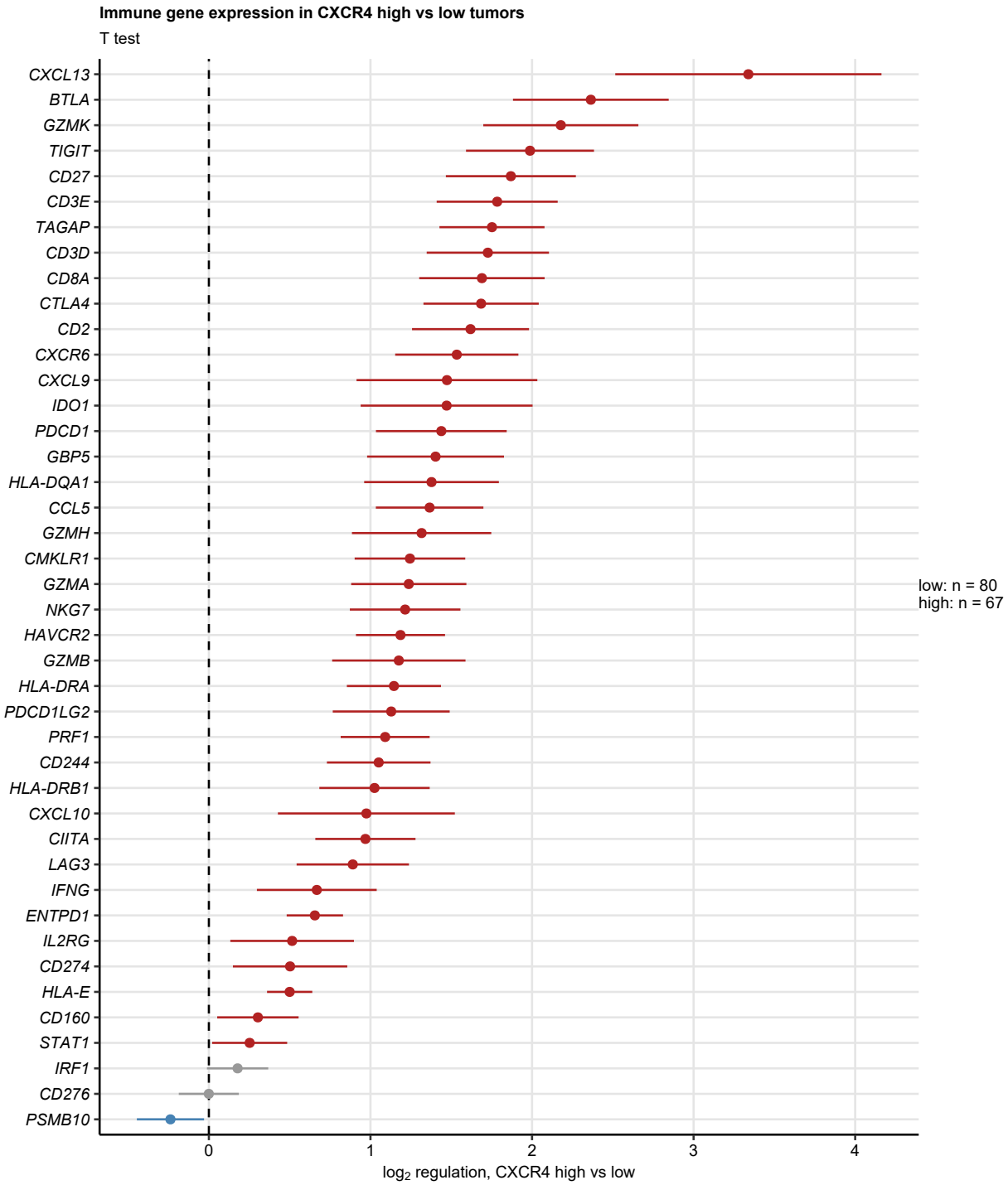

Figure 5: Differences in expression of unique gene members of immune signatures between $CXCR4^{high}$ and $CXCR4^{high}$ tumors.

Figure 5. Differences in expression of unique gene members of immune signatures between $CXCR4^{\text{high}}$ and $CXCR4^{\text{low}}$ tumors.

Tumor samples were stratified as $CXCR4^{\text{high}}$ vs $CXCR4^{\text{low}}$ expressors by the optimal cutoff for the dif-

ference in tumor-related survival. Differences in log_2 expression of the unique gene members of immune signatures (provided by TIMER: Immune Score, Microenvironment Score, Cytotoxicity Score and Stroma Score, provided by literature: Expanded Immune Signature, Exhaustion Signature, TISig: Tumor Inflammation Signature, IFN-γ Signature, Cytotoxic Signature) between the *CXCR4* strata were investigated with T test. $\log_2$ differences in expression between the $CXCR4^{\text{high}}$ vs $CXCR4^{\text{low}}$ samples with 95% confidence intervals are presented. N numbers of the strata samples are indicated next to the plot.

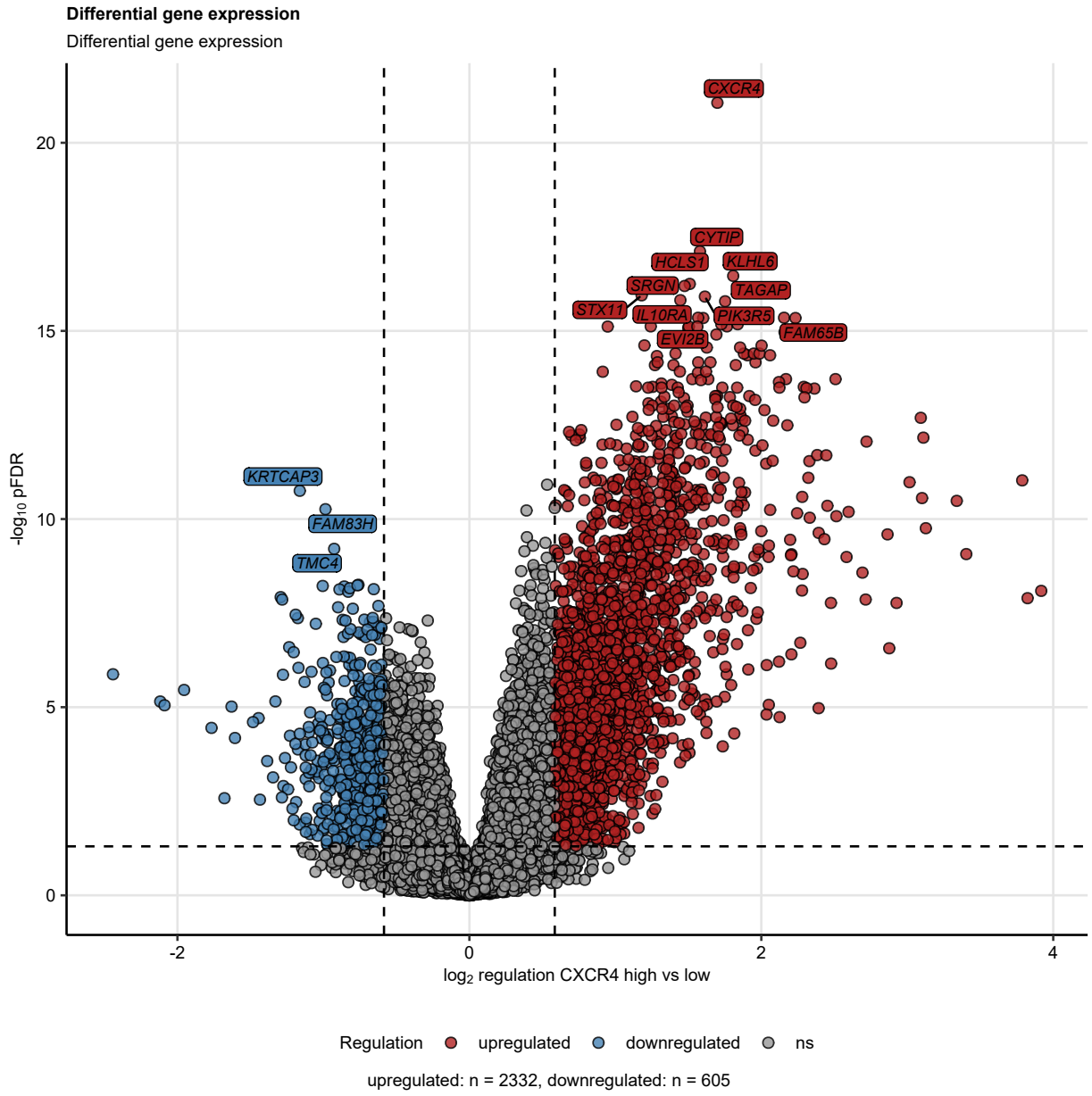

Figure 6: Differences in gene expression between $CXCR4^{high}$ and $CXCR4^{high}$ tumors.

Figure 6. Differences in gene expression between $CXCR4^{\text{high}}$ and $CXCR4^{\text{low}}$ tumors.

Tumor samples were stratified as $CXCR4^{\text{high}}$ vs $CXCR4^{\text{low}}$ expressors (low: n = 80, high: n = 67) by the optimal cutoff for the difference in tumor-related survival. Differentially expressed genes were determined by Benjamini-Hochberg-corrected T test and 1.5-fold regulation cutoff (**Supplementary Table S2**). log_2 differences in expression between the $CXCR4^{\text{high}}$ and $CXCR4^{\text{low}}$ tumors and Benjamini-Hochberg-corrected p values (pFDR) for each gene are presented in a volcano plot. Numbers of differentially expressed genes are indicated under the plot.

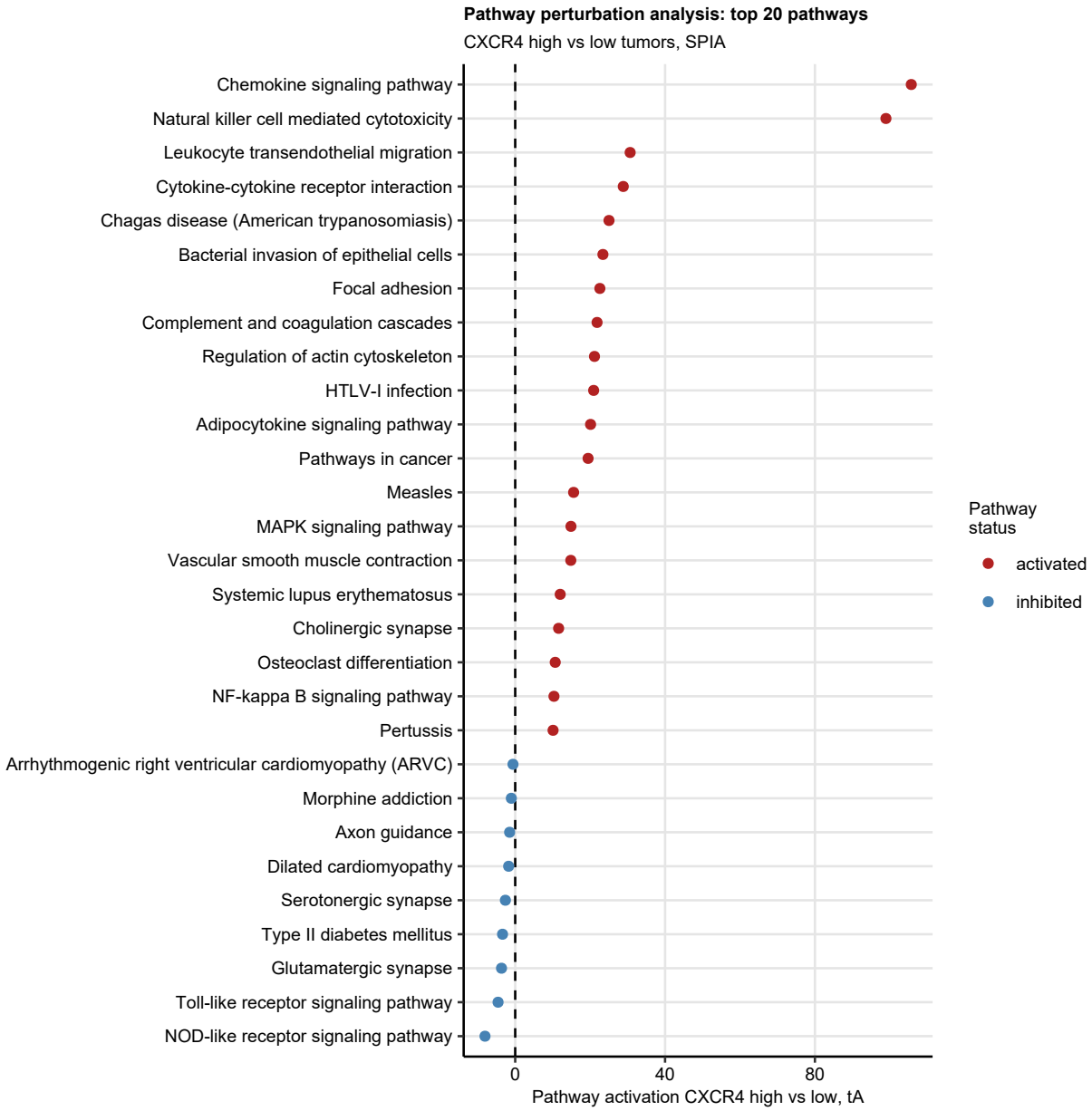

Figure 7: Differences in signaling pathway modulation between $CXCR4^{high}$ and $CXCR4^{high}$ tumors.

Figure 7. Differences in signaling pathway modulation between $CXCR4^{\text{high}}$ and $CXCR4^{\text{low}}$ tumors.

Differential signaling pathway modulation in $CXCR4^{\text{high}}$ and $CXCR4^{\text{low}}$ expression strata (low: n = 80, high: n = 67) was investigated with the *SPIA* algorithm (**Supplementary Table S3**). Significantly modulated pathways were identified by the value of combined (enrichment and perturbation) p value corrected for multiple comparisons with Benjamini-Hochberg method (pGFDR < 0.05). Estimates of the total pathway perturbation (tA) in the $CXCR4^{\text{high}}$ vs $CXCR4^{\text{low}}$ tumors are shown for the significantly modulated pathways.

Supplementary Figures

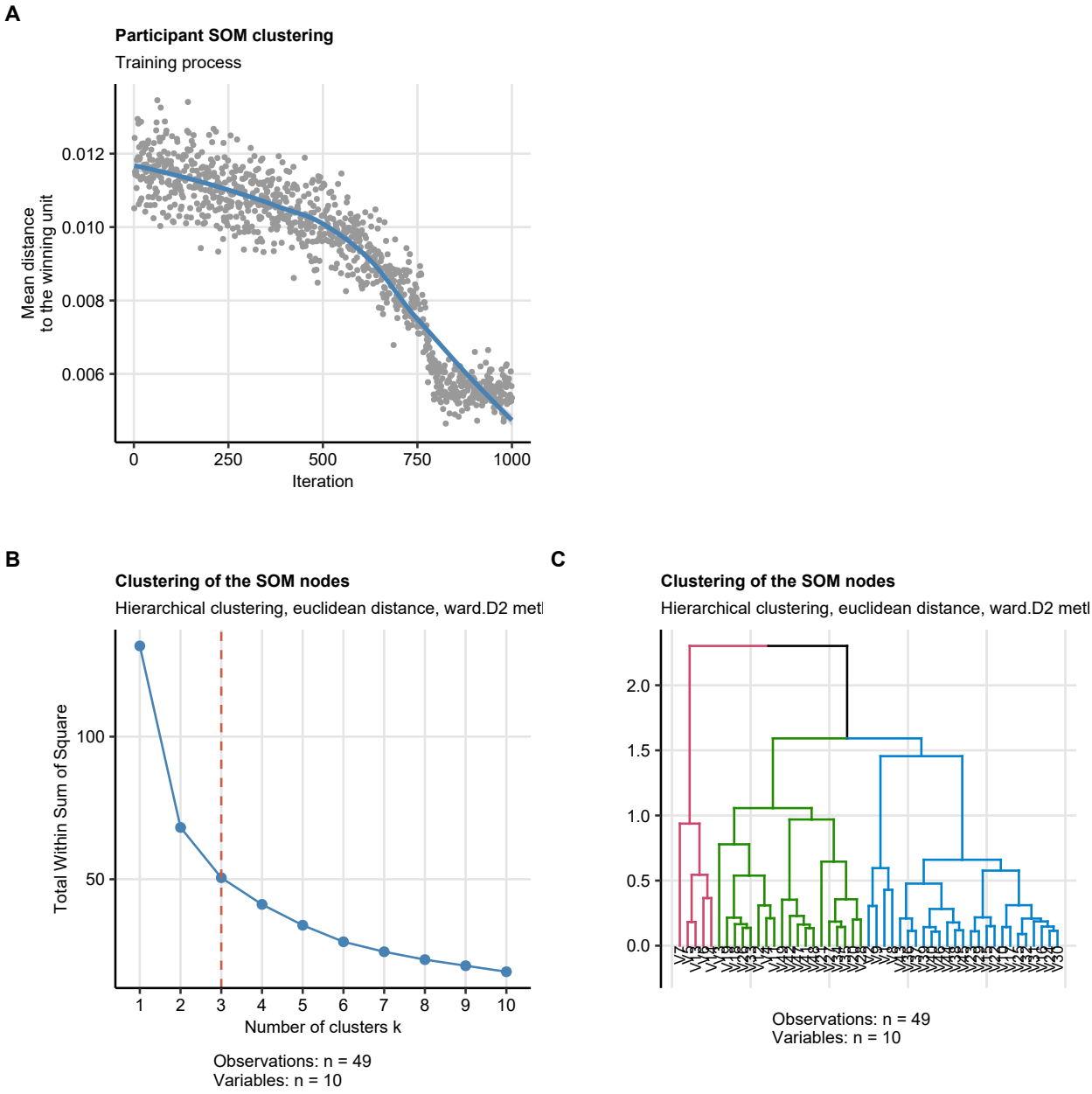

Figure S1: Definition of tumor sample and immune infiltration feature clusters.

Supplementary Figure S1. Definition of tumor sample and immune infiltration feature clusters.

Tumor samples and pooled multi-algorithm immune infiltration estimates were clustered with a two step, self-organizing map (SOM, samples: 7 × 7 hexagonal grid, Euclidean distance) and k-means clustering (Euclidean distance) procedure. The optimal cluster number was determined by the bend of the total within-cluster sum of squares (WSS) curve.

(A) Progress of the SOM training procedure visualized as the drop of the mean distance to the winning unit with the algorithm iterations.

(B) Determination of the optimal cluster number in hierarchical clustering of the SOM nodes by finding the bend of the within-cluster WSS curve.

(C) Histogram of the SOM node clustering structure.

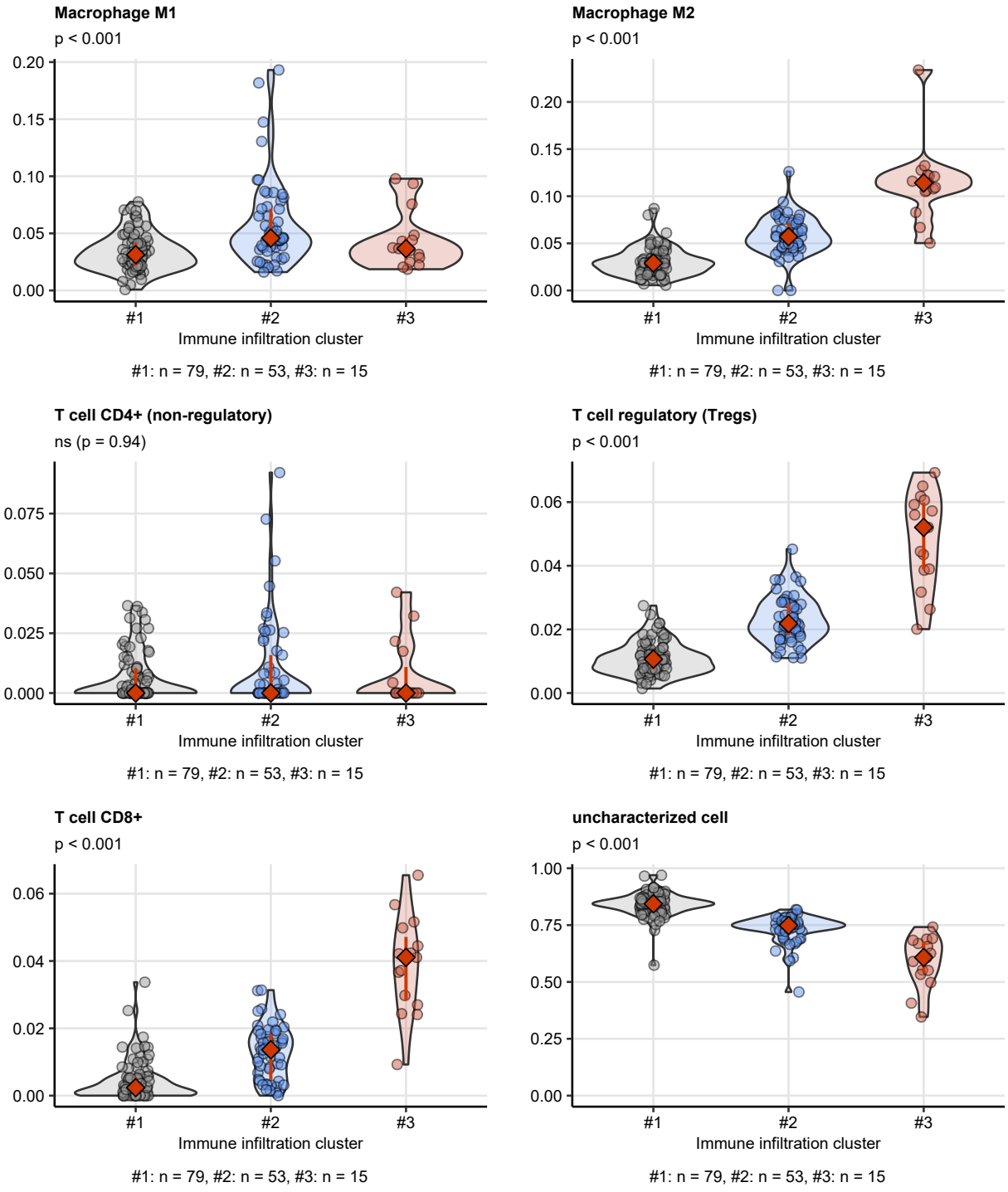

Figure S2: Levels of tumor-associated macrophages, T cells and uncharacterized, bona-fide malignant cells in the immune infiltration clusters.

Supplementary Figure S2. Levels of tumor-associated macrophages, T cells and uncharacterized, bona-fide malignant cells in the immune infiltration clusters.

Levels of M1 and M2 tumor-associated macrophages, $CD4^+$ non-regulatory and regulatory T cells, cytotoxic $CD8^+$ T cells and uncharacterized bona-fide malignant cells estimated by the QuantTIseq algorithm were

compared between the immune infiltration clusters of the TCGA cohort (**Figure 1** and **Supplementary Figure S1**). Statistical significance was assessed by Kruskal-Wallis test corrected for multiple testing with Benjamini-Hochberg method. P values are indicated in the plot captions. Estimate levels are presented in violin plots. Points represent single tumor samples, orange diamonds and whiskers code for medians with interquantile ranges. Number of samples assigned to the clusters are indicated under the plots.

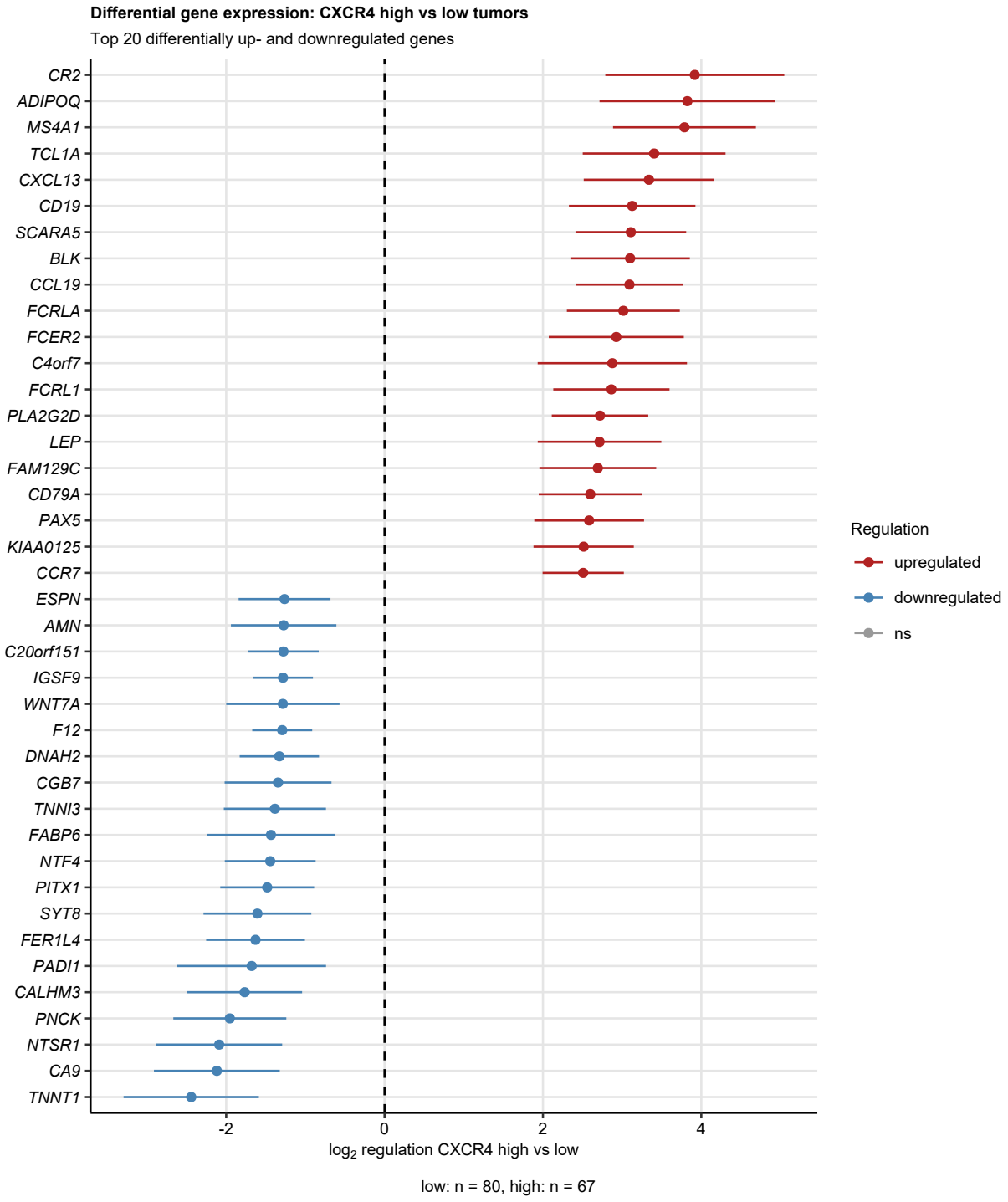

Figure S3: Top 20 strongest up- and downregulated genes in $CXCR4^{high}$ vs $CXCR4^{high}$ tumors.

Supplementary Figure S3. Top 20 strongest up- and downregulated genes in $CXCR4^{\text{high}}$ vs $CXCR4^{\text{low}}$ tumors.

Tumor samples were stratified as $CXCR4^{\text{high}}$ vs $CXCR4^{\text{low}}$ expressors by the optimal cutoff for the difference in tumor-related survival. Differentially expressed genes were determined by Benjamini-Hochberg-corrected

T test and 1.5-fold regulation cutoff (**Supplementary Table S2**). log_2 differences in expression between the $CXCR4^{high}$ and $CXCR4^{low}$ tumors with 95% confidence intervals for the top 20 strongest up- and downregulated genes are presented. N number of observations is indicated below the plot.

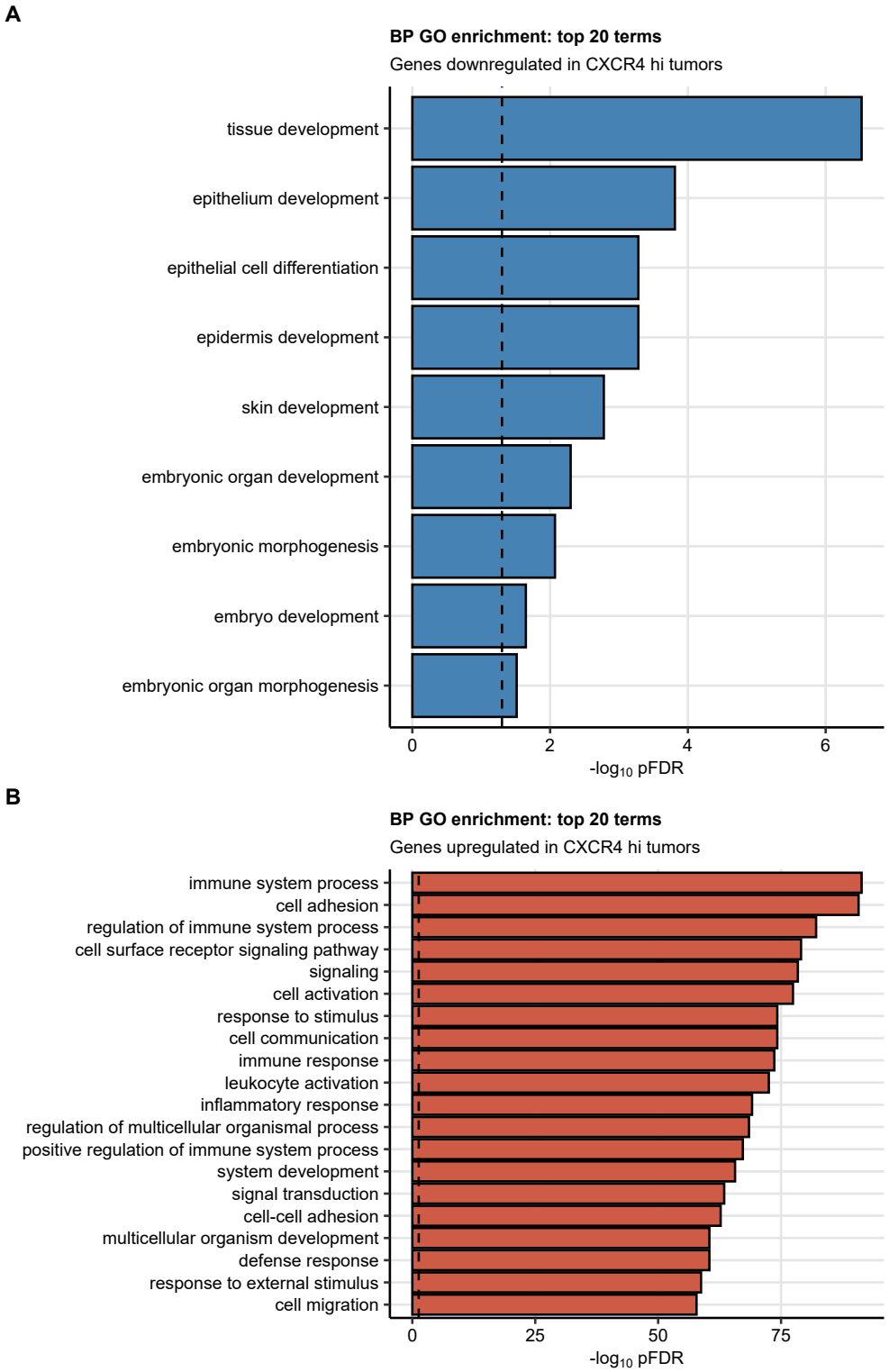

Figure S4: Top 20 significantly enriched GO terms within the sets of the genes up- and downregulated in $CXCR4^{high}$ vs $CXCR4^{high}$ tumors.

Supplementary Figure S4. Top 20 significantly enriched GO terms within the sets of the genes up- and downregulated in $CXCR4^{\text{high}}$ vs $CXCR4^{\text{low}}$ tumors.

Biological process gene ontology (BP GO) term enrichment analysis for the gene sets differentially regulated in *CXCR4*$^{\text{high}}$ vs *CXCR4*$^{\text{low}}$ tumors ((low: n = 80, high: n = 67)) was performed with the *goana()* tool from *limma* package. Benjamini-Hochberg-corrected enrichment p values (pFDR) for the top 20 most significantly enriched BP GO terms and KEGG pathways are presented in bar plots. Note: for the downregulated gene set no significant no significant GO term and no significant KEGG pathway enrichment was detected.

Supplementary Tables

Table S1: **Pearson correlation of immune features with *CXCR4* expression in pancreatic cancer.** Estimates of immune cell features were obtained for the TCGA RNA seq samples with the algorithms supported by TIMER and correlated with $\log_2$ expression levels of *CXCR4* with Pearson correlation. Rho: correlation coefficient with 95% confidence intervals, pFDR: p value corrected for multiple comparisons with Benjamini-Hochberg method.
The table is available as an Excel file.

—

—

—

Table S2: **Differentially expressed genes in *CXCR4* expression strata.** Tumor samples were stratified as $CXCR4^{hi}$ and $CXCR4^{low}$ by the optimal cutoff for the difference in tumor-related survival. Differentially expressed genes were determined by Benjamini-Hochberg-corrected T test and 1.5-fold regulation cutoff. $\log_2$ Regulation: $\log_2$ difference in expression between the $CXCR4^{hi}$ and $CXCR4^{low}$ samples, Lower/Upper CI: 95% confidence interval for the $\log_2$ expression difference, pFDR: pFDR: p value corrected for multiple comparisons with Benjamini-Hochberg method.
The table is available as an Excel file.

—

—

—

Table S3: **Differential signaling pathway modulation in *CXCR4* expression strata investigated with SPIA algorithm.** Pathway: KEGG pathway name, Modulation tA: the observed total perturbation accumulation in the pathway, pGFDR: global, combined (enrichment and perturbation) p value corrected for multiple comparisons with Benjamini-Hochberg method.
The table is available as an Excel file.

—

—

—

References

1. Raphael BJ, Hruban RH, Aguirre AJ, et al. Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. *Cancer Cell*. 2017;32(2):185-203.e13. doi:10.1016/j.ccell.2017.07.007
2. Li T, Fu J, Zeng Z, et al. TIMER2.0 for analysis of tumor-infiltrating immune cells. *Nucleic Acids Research*. 2020;48(W1):W509-W514. doi:10.1093/NAR/GKAA407
3. Woroniecka K, Chongsathidkiet P, Rhodin K, et al. T-cell exhaustion signatures vary with tumor type and are severe in glioblastoma. *Clinical Cancer Research*. 2018;24(17):4175-4186. doi:10.1158/1078-0432.CCR-17-1846
4. Yan K, Lu Y, Yan Z, Wang Y. 9-Gene Signature Correlated With CD8+ T Cell Infiltration Activated by IFN-γ: A Biomarker of Immune Checkpoint Therapy Response in Melanoma. *Frontiers in Immunology*. 2021;12:2223. doi:10.3389/fimmu.2021.622563
5. Ayers M, Lunceford J, Nebozhyn M, et al. IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade. *Journal of Clinical Investigation*. 2017;127(8):2930-2940. doi:10.1172/JCI91190
6. Danaher P, Warren S, Lu R, et al. Pan-cancer adaptive immune resistance as defined by the Tumor Inflammation Signature (TIS): Results from The Cancer Genome Atlas (TCGA). *Journal for ImmunoTherapy of Cancer*. 2018;6(1):1-17. doi:10.1186/s40425-018-0367-1
7. Hänzelmann S, Castelo R, Guinney J. GSVA: Gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinformatics*. 2013;14(1):7. doi:10.1186/1471-2105-14-7
8. Kohonen T. *Self-Organizing Maps*. Vol 30. Springer Berlin Heidelberg; 1995. doi:10.1007/978-3-642-97610-0
9. Vesanto J, Alhoniemi E. Clustering of the self-organizing map. *IEEE Transactions on Neural Networks*. 2000;11(3):586-600. doi:10.1109/72.846731
10. Kocher F, Tymoszuk P, Amann A, et al. Deregulated glutamate to pro-collagen conversion is associated with adverse outcome in lung cancer and may be targeted by renin-angiotensin-aldosterone system (RAS) inhibition. *Lung Cancer*. 2021;159:84-95. doi:10.1016/j.lungcan.2021.06.020
11. Eigentler A, Tymoszuk P, Zwick J, et al. The impact of Cand1 in prostate cancer. *Cancers*. 2020;12(2). doi:10.3390/cancers12020428
12. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*. 1995;57(1):289-300. doi:10.1111/j.2517-6161.1995.tb02031.x
13. Young MD, Wakefield MJ, Smyth GK, Oshlack A. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biology*. 2010;11(2):R14. doi:10.1186/gb-2010-11-2-r14
14. Tarca AL, Draghici S, Khatri P, et al. A novel signaling pathway impact analysis. *Bioinformatics*. 2009;25(1):75-82. doi:10.1093/bioinformatics/btn577
15. Wickham H, Averick M, Bryan J, et al. Welcome to the Tidyverse. *Journal of Open Source Software*. 2019;4(43):1686. doi:10.21105/joss.01686
16. Wickham Hadley. *ggplot2: Elegant Graphics for Data Analysis*. 1st ed. Springer-Verlag; 2016. https://ggplot2.tidyverse.org
17. Wilke CO. *Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures*. 1st ed. O'Reilly Media; 2019:389.
18. Kassambara A, Kosinski M, Biecek P. survminer: Drawing Survival Curves using 'ggplot2'. Published online 2016. https://cran.r-project.org/package=survminer
19. Wehrens R, Kruisselbrink J. Flexible self-organizing maps in kohonen 3.0. *Journal of Statistical Software*. 2018;87(7):1-18. doi:10.18637/jss.v087.i07