

Supplementary file for:

Metabolic heterogeneity in tumor cells impacts immunology in lung squamous cell carcinoma

Authors: Qian Wang^{a,*}, Na Sun^{a,*}, Chaoyang Zhang^a, Thomas Kunzke^a, Philipp Zens^{b,c}, Annette Feuchtinger^a, Sabina Berezowska^{b,d,#}, Axel Walch^{b,d,#}

Affiliations:

^a Research Unit Analytical Pathology, Helmholtz Zentrum München - German Research Center for Environmental Health, 85764 Neuherberg, Germany

^b Institute of Tissue Medicine and Pathology, University of Bern, 3008 Bern, Switzerland

^c Graduate School for Health Sciences, University of Bern, 3012 Bern, Switzerland

^d Department of Laboratory Medicine and Pathology, Institute of Pathology, Lausanne University Hospital and University of Lausanne, 1011 Lausanne, Switzerland

* Qian Wang and Na Sun share the first authorship.

Co-corresponding authors:

1. Prof. Axel Walch, Research Unit Analytical Pathology, Helmholtz Zentrum München, Ingolstädter Landstraße 1, 85764 Neuherberg, Germany; E-mail: axelkarl.walch@helmholtz-munich.de

2. Prof. Sabina Berezowska, Department of Laboratory Medicine and Pathology, Institute of Pathology, Lausanne University Hospital and University of Lausanne, Rue du Bugnon 25, Lausanne CH-1011, Switzerland; E-mail: sabina.berezowska@chuv.ch

Supplementary Materials and methods

Patient cohort and tissue samples

Tumor samples were obtained from 106 LUSC patients (**Table 1**) diagnosed at the Institute of Pathology, University of Bern, excluding cases with prior or concurrent LUSC in other organs. A tissue microarray (TMA) was constructed from formalin-fixed and paraffin-embedded (FFPE) tissue blocks [1]. The TMA, with digital annotation of scanned slides and automatic transferal of the punches was constructed from formalin-fixed and paraffin-embedded (FFPE) tissue blocks archived in the Institute of Pathology, University Bern, Switzerland. One representative tissue block was selected for each tumor after reviewing all available slides per case (H&E staining stained), and two tumor cores were randomly selected from the block for measurement and analysis. The histology of all cases was reevaluated according to the current World Health Organization guidelines for the diagnosis of LUSC [2]. All tumors were restaged according to the Union for International Cancer Control (UICC) 2017, 8th edition, tumor-node-metastasis (TNM) classification [3]. OS was defined as the time from resection to death of any cause. PFS was defined as the time from the first day of adjuvant therapy to the date of objective disease progression or death of any cause.

High mass resolution MALDI fourier-transform ion cyclotron resonance (FT-ICR) MSI

High mass resolution MALDI FT-ICR MSI was performed as previously described [4]. Briefly, the FFPE section (4 μm) was mounted onto indium tin oxide-coated glass slides (Bruker Daltonik, Bremen, Germany). The air-dried tissue section was spray-coated with 10 mg/mL 9-aminoacridine (9-AA) hydrochloride monohydrate matrix (Sigma-Aldrich, Munich, Germany) in methanol (70%) using the SunCollect™ sprayer (Sunchrom, Friedrichsdorf, Germany). Prior to matrix application, the FFPE tissue section was additionally incubated for 1 h at 70°C and deparaffinized in xylene (2 \times 8 min). Spray-coating of the matrix was performed in eight passes, utilizing 2 mm line distance and a spray velocity of 900 mm/min. Metabolite detection was performed in negative ion mode on a 7 T Solarix XR FT-ICR mass spectrometer (Bruker Daltonik). Mass spectra were acquired in the range of m/z 50-1,100 with a lateral resolution of 50 μm . Metabolites were annotated using HMDB Database (<http://www.hmdb.ca/>) with a mass tolerance of 4 ppm.

Immunohistochemistry (IHC)

Immunohistochemical staining for CD3, CD8, and PD-L1 was performed as previously described [1] on consecutive sections. Briefly, an automated immunostainer (Bond III, Leica Bio-systems, Muttenz, Switzerland) with anti-CD3 (Abcam Cambridge, UK, clone SP7, 1:400, RRID: AB_443425), anti-CD8 (Dako, clone C8/144B, 1:100, RRID: AB_2075537), and anti-PD-L1 (Cell Signaling Technology, clone E1L3N, 1:400, RRID: AB_2687655) was used. The counts of CD8+ and CD3+ TILs were determined using image analysis (Aperio Image Scope) and adjusted for core completeness. PD-L1 expression was assessed as the intensity of membranous staining by a pathologist (S. Berezowska).

Hematoxylin and eosin staining (H&E)

After MALDI MSI, the 9-AA matrix was removed from the tissue section surface using 70% ethanol (Carl Roth) for 5 minutes, followed by H&E staining of the very same tissue section. The H&E-stained tissue section was cover-slipped and scanned with an AxioScan.Z1 digital slide scanner (Carl Zeiss) equipped with a 20x magnification objective. The visualization and export of the images to TIFF was done with the software ZEN 2.3 blue edition (Carl Zeiss).

Immunofluorescence (IF)

Double staining of the TMA was performed using pan-cytokeratin (monoclonal mouse pan-cytokeratin plus [AE1/AE3+8/18], 1:75, catalogue no. CM162, Biocare Medical, US, RRID: AB_10582491) and vimentin (abcam, clone ab92547, 1:500, RRID: AB_10562134). Regions positive for pan-cytokeratin were defined as tumor. Regions negative for pan-cytokeratin but positive for vimentin were defined as stroma. Multi-images were exported as TIF files in assistance with tumor-only regions of interest (ROIs) annotation in QuPath (version 0.3.2).

Unsupervised pixel-wise k-means segmentation and metabolic tumor subpopulation determination

Using the segmentation tool in SCiLS Lab software (Version 2023a Pro, Bruker Daltonics), unsupervised *k*-means clustering was performed pixel-wisely containing overall resolved MS peaks within ROIs where only tumor regions of LUSC samples were annotated. Due to uncertainty regarding the extent of heterogeneity, the segmentation algorithm was run with *k* range from 2 to 15. The created segmentation maps were then used to identify the spatial distribution of clusters having similar spectra occur across the patients.

For further statistical analysis, we linked the survival data of patients to the presence of specific clusters and consequently determined the metabolic tumor subpopulations. In this process, a patient was assigned to a cluster if the cluster was sufficiently present in that patient's ROI (i.e., if the cluster contained a fraction of pixels above a particular threshold). A single patient could be assigned to more than one cluster if it contained significant tumor heterogeneity. The effect of choosing different thresholds on survival was investigated using Cox proportional hazards regression model as previously described [5]. Specifically, an iterative loop was created with thresholds ranging from 4% to 50%. A binary variable was created at each threshold by applying the threshold to the cluster ratio. Cox proportional hazards regression model was then built using the thresholded data. The fitness of the model at each *k* was assessed using the Akaike information criterion (AIC) involving patients overall survival; the model with the lowest AIC value was presumed to most closely fit the data [5]. The calculation was performed in Python (version 3.9). Applying the *k* value and threshold of pixels fraction corresponding to the lowest AIC value, patients were assigned to certain clusters, so called metabolic tumor subpopulations (MTSs).

Simpson's diversity index as heterogeneity score

The Simpson diversity index [6] measures the diversity and equals the probability that two randomly chosen pixels are from different types. The calculation formula in our study is defined as follows:

$$D = 1 - \sum_{i=1}^k p_i^2, \text{ where } p_i \text{ is the share of pixels in cluster } i \text{ and } k \text{ is the number of clusters.}$$

The formula is built considering the spatial locations of pixels for each patient in each cluster. Each pixel is located by x and y coordinates which were extracted from SCiLS Lab software (Version 2023a Pro, Bruker Daltonics). The within sum of squares of the coordinates x and y in each cluster is calculated for each patient. Summing up these within sum of squares for all clusters leads to the total within sum of squares of the patients. By using the within sum of squares, the spatial distribution of the pixels in one patient is measured, i.e. how heterogeneous and therefore far apart are the pixels within one cluster. The index can have values between 0 and 1. A value of 0 means that all the pixels from one patient are in one cluster. A higher value of the index, therefore, indicates higher diversity of the pixels in the different clusters for a patient. As the heterogeneity score for each patient, the Simpson index to measure the metabolic heterogeneity of the patients is calculated using the within sum of squares as entries p_i . The calculation was performed in R (version 4.2.2).

Survival analyses

To define the high and low metabolic heterogeneity, we iterate over each heterogeneity score as cutoffs to evaluate the association between cutoff-optimized patients' metabolic heterogeneity levels and their OS. With a minimum p-value of 0.029 in the log-rank test, the cutoff of 0.436 was chosen to distinguish between low and high metabolic heterogeneity levels. Following the separation, the difference in PFS was evaluated between high and low metabolic heterogeneity levels. Among MTSs, Kaplan-Meier survival curve analyses involving both OS and PFS were performed based on AIC-thresholded determination. Multivariate Cox proportional hazards model was performed to evaluate the effect of metabolic heterogeneity levels and tumor subpopulations on survival, respectively together with UICC stage, sex, age, tumor size, distant metastases, and primary resection status. To manage the possibility that some patients belonged to more than one MTS, an extended Cox regression model [7] was incorporated into algorithms to evaluate the statistical differences with the Wald test. Kaplan-Meier survival curve analyses for TILs were performed with the R packages survival v3.5-3 and survminer v0.4.9.

Bioinformatics

Metabolic pathway enrichment analysis and associated genes identification were achieved by mapping correlated metabolites to KEGG pathway library via MetaboAnalyst v5.0 (<https://www.metaboanalyst.ca/>). The results of pathway enrichment were evaluated by Fisher's exact test and visualized the pathways with at least 2 hits. The association of genes with CD8+ TIL immune infiltrating level by tumor purity adjustment and the survival impact of genes were investigated integrated TCGA-LUSC cohort via TIMER [8]. The correlation model between genes with TILs immune infiltrating levels was built using LASSO regression.

The associations of metabolites to immune processes were investigated by the Immunometabolic Atlas (IMA), a public web application and library of R functions [9]. The action and metabolism pathways of anti-cancer drugs as well as the resistance information were obtained from Small Molecule Pathway Database 2.0 (SMPDB, <https://www.smpdb.ca/>) and DRESIS [10] (<https://idrblab.org/dresis/>). All the networks were built in Cytoscape (v3.8.2).

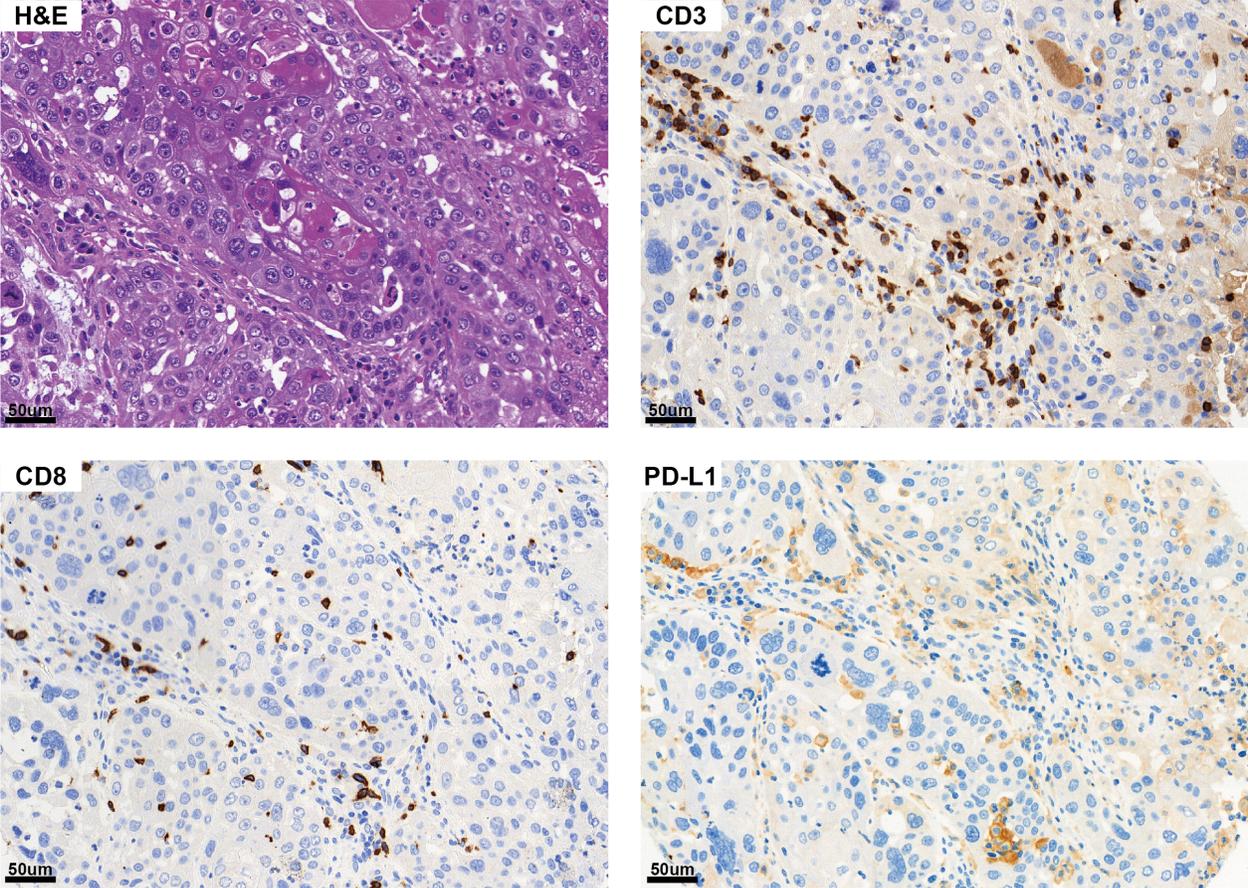
Statistics

The significance of differences in clinicopathological characteristics between high and low heterogeneity levels was evaluated by Chi-squared test or Fisher's exact test. Kruskal–Wallis test was used to calculate the differences of Simpson's diversity index distributions among MTSs. A p-value < 0.05 was considered statistically significant in all results.

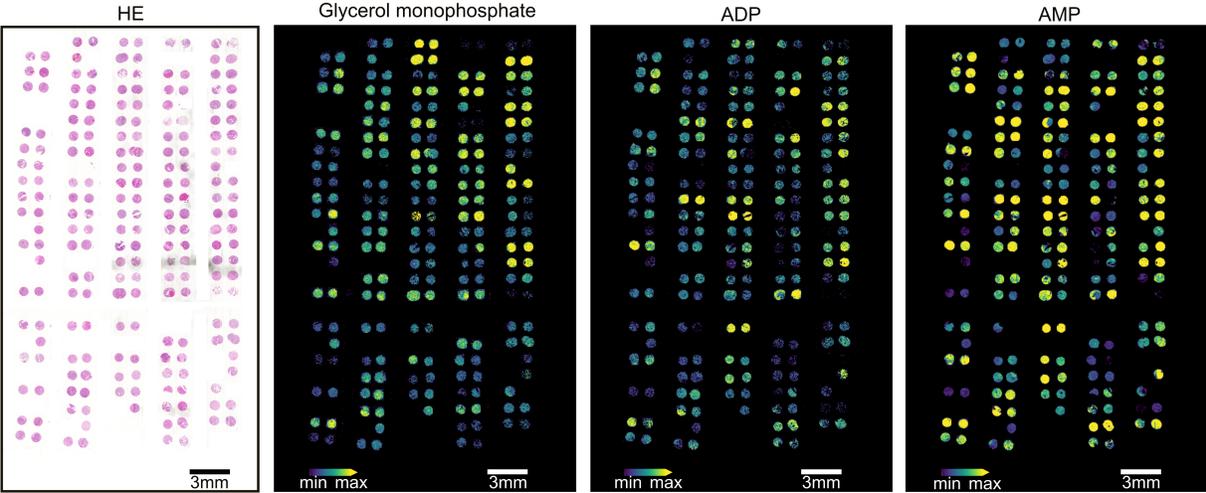
References

1. Keller, M.D., et al., *Adverse prognostic value of PD-L1 expression in primary resected pulmonary squamous cell carcinomas and paired mediastinal lymph node metastases*. *Mod Pathol*, 2018. **31**(1): p. 101-110.
2. Board, W.C.o.T.E., *Thoracic Tumours: WHO Classification of Tumours*. 5 ed. Vol. 5. 2021: International Agency for Research on Cancer (IARC) (UN).
3. Nepl, C., et al., *Comparison of the 7th and 8th Edition of the UICC/AJCC TNM Staging System in Primary Resected Squamous Cell Carcinomas of the Lung-A Single Center Analysis of 354 Cases*. *Front Med (Lausanne)*, 2019. **6**: p. 196.
4. Ly, A., et al., *High-mass-resolution MALDI mass spectrometry imaging of metabolites from formalin-fixed paraffin-embedded tissue*. *Nat Protoc*, 2016. **11**(8): p. 1428-43.
5. Balluff, B., et al., *De novo discovery of phenotypic intratumour heterogeneity using imaging mass spectrometry*. *J Pathol*, 2015. **235**(1): p. 3-13.
6. Simpson, E.H., *Measurement of Diversity*. *Nature*, 1949. **163**(4148): p. 688-688.
7. Andersen, P.K. and R.D. Gill, *Cox's Regression Model for Counting Processes: A Large Sample Study*. *The Annals of Statistics*, 1982. **10**(4): p. 1100-1120.
8. Li, T., et al., *TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells*. *Cancer Res*, 2017. **77**(21): p. e108-e110.
9. Maas, P., et al., *The Immunometabolic Atlas: A tool for design and interpretation of metabolomics studies in immunology*. *PLoS One*, 2022. **17**(5): p. e0268408.
10. Sun, X., et al., *DRESIS: the first comprehensive landscape of drug resistance information*. *Nucleic Acids Res*, 2023. **51**(D1): p. D1263-d1275.

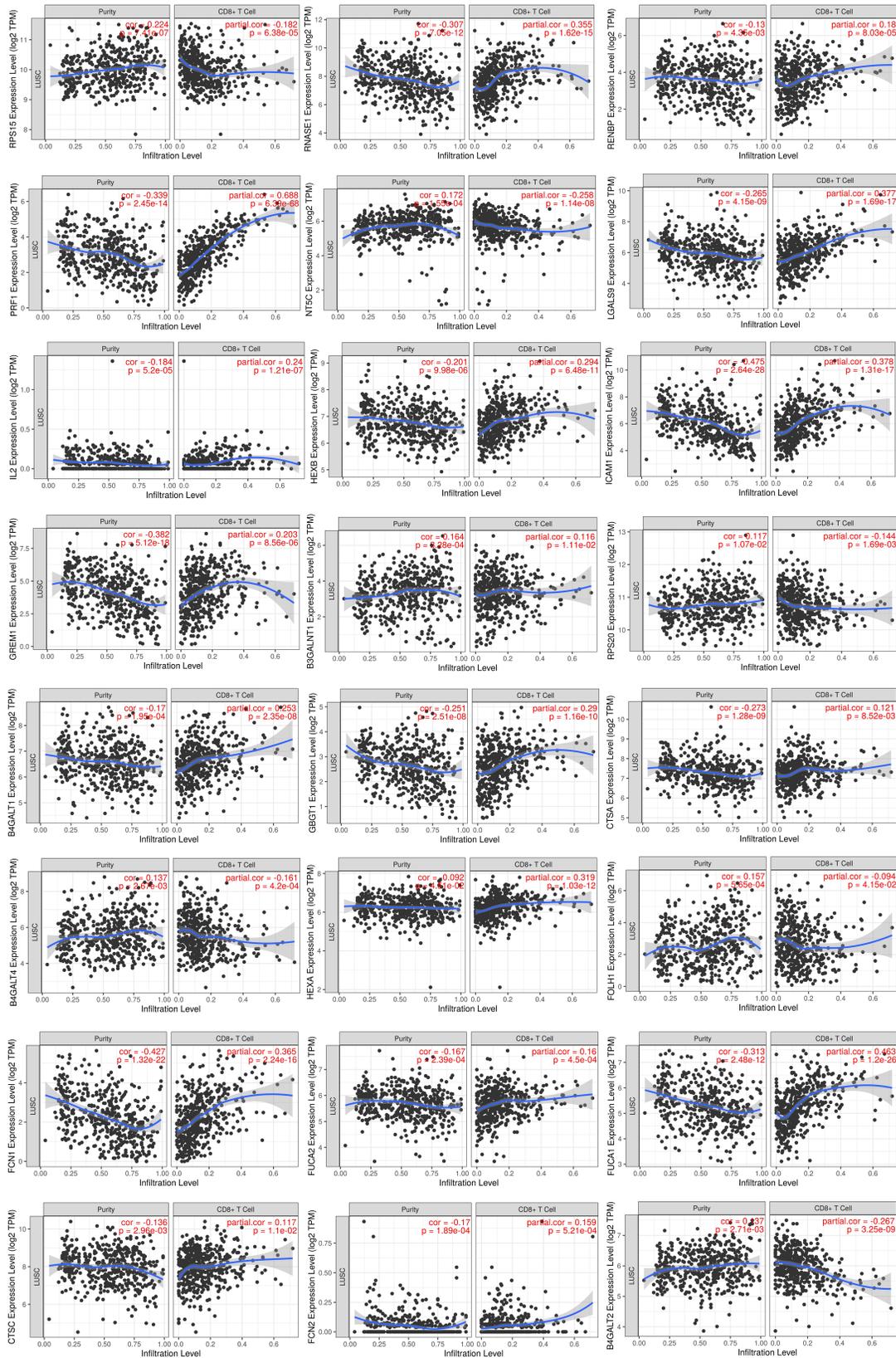
Supplementary Fig. 1A. High resolution IHC images of immunological markers.



Supplementary Fig. 1B. Metabolites distribution.



Supplementary Fig. 2. CD8+ TIL associated genes in validation LUSC cohort.



Supplementary Fig. 3. Distribution of the certain TILs-correlated metabolites in each MTS.

