CORRECTION

Open Access

Correction: Mesenchymal–epithelial transition in lymph node metastases of oral squamous cell carcinoma is accompanied by ZEB1 expression

Kai Horny[1,2], Christoph Sproll[3], Lukas Peiffer[1,2], Frauke Furtmann[1,2,4], Patricia Gerhardt[1,2], Jan Gravemeyer[1,2], Nikolas H. Stoecklein[5], Ivelina Spassova[1,2,4] and Jürgen C. Becker[1,2,4]*

Correction: Journal of Translational Medicine (2023) 21:267
https://doi.org/10.1186/s12967-023-04102-w

Following publication of the original article [1], we have been notified by the authors that the legends of supplementary material was not provided. Now legends are as follows:

Additional file 1: Figure S1. Histology of OSCC primary and metastatic tumors.

*Correspondence:
Jürgen C. Becker
j.becker@dkfz.de
[1] Translational Skin Cancer Research, German Cancer Consortium (DKTK), 45141 Essen, Germany
[2] German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany
[3] Department of Oral- and Maxillofacial Surgery, Medical Faculty, University Hospital of the Heinrich-Heine-University, Düsseldorf, Germany
[4] Department of Dermatology, University Medicine Essen, 45141 Essen, Germany
[5] Department of General, Visceral and Pediatric Surgery, Medical Faculty, University Hospital of the Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany

Additional file 2: Figure S2. Cell type identification by marker genes, automated reference-based annotation, differential expression and inferred CNVs.

Additional file 3: Figure S3. PEMT and epithelial differentiating gene expression signatures are comparable to previously published EMT signatures.

Additional file 4: Figure S4. Extended analysis of the tumor phenotype characterization for the lymph node metastasis of patient 1.

Additional file 5: Figure S5. Malignant phenotypes characterized across all analyzed patients.

Additional file 6: Figure S6. Inferred transcription factor activity might be biased by activator or repressor function.

Additional file 7: Figure S7. Cell type abundances across patients and tissue types.

Additional file 8: Figure S8. Bulk transcriptomes reveal the cellular composition of OSCC across tissue types.

Additional file 9: Figure S9. Characterization of OSCC-derived fibroblasts.

Horny *et al. Journal of Translational Medicine*   (2024) 22:277

Page 2 of 4

The legends of the supplementary files should be as follows:

Additional file 1: Figure S1. Histology of OSCC primary and metastatic tumors. Whole-slide image H&E staining for FFPE sections of OSCC samples. Scale bars depict 2 mm. High-resolution pictures are available through DOI:https://doi.org/10.6084/m9.figshare.20905837.v1.

Additional file 2: Figure S2. Cell type identification by marker genes, automated reference-based annotation, differential expression and inferred CNVs. (**A**) Expression of marker genes (x-axis) for each cell type (y-axis) in our cohort. Dots are colored by the average log-normalized gene expression and the dot size represents the percentage of cells with detected expression of the respective gene within the cell type. (**B**) UMAP of 41,284 cells from our cohort cells colored by SingleR annotations using the Monaco bulk RNA dataset on shared-nearest neighbor clusters with resolution 100. (**C**) Heatmap for scaled, log-normalized gene expression of all cell types from patient #1 and their top 10 DEGs (rows) against all other cells. DEGs are sorted from highest to lowest log2 foldchange. (**D**) Inferred CNVs across cells (rows) of different cell types without mitochondrial genes from patient #1. Columns show genes categorized in chromosomes and ordered by genome position; hence the size of the chromosome reflects the number of detected genes and not its nucleotide length. (**E**) Standard deviation of the log2 inferCNV values to the mean of non-malignant cells compared between non-malignant and malignant cells of patient #1.

Additional file 3: Figure S3. PEMT and epithelial differentiating gene expression signatures are comparable to previously published EMT signatures. (**A**) EMT hallmark gene set enrichment plot for log2 fold changes of pEMT cells against all other cells of lymph node metastasis from patient 1. Shown is the stepwise calculated enrichment score, black lines indicate genes present in the respective gene set. (**B**) Average log2 fold change of gene expression (x-axis) and differences in cellular fractions expressing the respective gene (y-axis) between pEMT and epithelial differentiated cell clusters. Labelled in red are genes with log2 foldchange below or above 1 that are included in the epithelial differentiation or pEMT signature, respectively, with top 10 genes named. The histogram on top shows the number of genes across the log2 fold change with in total 100 bins. (C, D) Average expression scores (y-axis) of the pEMT (**C**) and epithelial differentiation (**D**) signatures across tumor phenotypes from patient #1 depicted in Fig. 2A (x-axis) color-coded by these clusters.

(**E**) Heatmap of correlation coefficients of GSVA scores between 91 EMP-related signatures of malignant cells, derived from the EMTome database and selected publications (9, 10, 14). On the right side, the correlation coefficients between GSVA scores of EMT signatures from the EMTome database and of epithelial differentiation and pEMT signatures from patient #1 (right) are shown and next to it, annotated as"EXPR_perc", is the fraction of genes with non-zero expression and the size of the respective EMT signature in log10 scale with the respective number next to it. Rows and columns are hierarchically clustered using a spearman correlation distance $(1-cor(x,y))$ and ward.D2 method.

Additional file 4: Figure S4. Extended analysis of the tumor phenotype characterization for the lymph node metastasis of patient #1. (**A**) Top 5 enriched gene sets from log2 foldchanges of respective tumor phenotypes by normalized enrichment scores (x-axis). Gene sets of respective phenotypes are sorted from highest to lowest enrichment. Bars are colored by the negative decadic logarithm of the Benjamini–Hochberg adjusted p-value (padj). (**B**) First two PCs of OSCC cells with all six principal curves that are derived from trajectory inference. Graph on top visualizes the relationship between malignant phenotypes, described by the principal curves forming a branching trajectory. Cells responding to environmental conditions form their own branch, indicating that the strong reactive response determines their predominant phenotype. (**C**) Inferred CNVs across tumor cells of patient #1 (rows) for all chromosomes (columns). Columns show genes categorized in chromosomes and ordered by genome position; hence the size of the chromosome reflects the number of detected genes and not its nucleotide length. Mitochondrial genes were excluded. (**D**, **E**) Inferred CNVs across tumor cells (rows) of chromosome 1 (**D**) and chromosome 17 (**E**) showing genes (columns) ordered by genome position. The signal on chromosome 1 is located on a genomic position on which S100 genes are accumulating and the signal on chromosome 17 on a location with accumulation of cytokeratins; most of these genes are highly expressed in the more epithelial differentiated cells.

Additional file 5: Figure S5. Malignant phenotypes characterized across all analyzed patients. (**A**) Number of cells (y-axis) for each library (x-axis) showing the cells that are used for 10×Genomics scRNAseq (light blue) and all recovered, i.e., detected, cells after sequencing (blue). Based on manufacturers information a recovery rate around 50% is expected. (**B**) Heatmap for scaled, log-normalized gene expression of tumor cells (columns) split by respective phenotype depicted in Fig. 3C and the top

Horny *et al. Journal of Translational Medicine*   (2024) 22:277

Page 3 of 4

10 DEGs (rows) of the respective phenotype against all other tumor cells. DEGs are sorted from highest to lowest log2 foldchange and row sections are ordered the same as column section. On bottom, the respective patient and localization is annotated for each cell. (**C**) Top 5 enriched gene sets from log2 foldchanges of respective tumor phenotypes by normalized enrichment scores (x-axis). Gene sets of respective phenotypes are sorted from highest to lowest enrichment. Bars are colored by the negative decadic logarithm of the Benjamini–Hochberg adjusted p-value (padj). (**D**) UMAP of OSCC cells as depicted in Fig. 3C with PCs corrected for patient-specific effects using harmony. Cells are annotated according to their patient id. (**E**) Inferred CNVs across EMP-related OSCC cells from patient HN01 (rows) for all chromosomes (columns). Cells split by their EMP phenotype do not show any differences in their inferred CNVs pattern. Columns show genes categorized in chromosomes and ordered by genome position; hence the size of the chromosome reflects the number of detected genes and not its nucleotide length. Mitochondrial genes were excluded. (**F**) UMAPs of malignant cells from all respective patients. Cells are annotated SNN clusters and renamed according to the predominant phenotype. (**G**) Same plot as depicted in Fig. 3F with the names of all patient-specific clusters as shown in E followed be the patient id.

Additional file 6: Figure S6. Inferred transcription factor activity might be biased by activator or repressor function. (**A**) Distribution of the mean activity of all cells from patient #1 for all transcription factors split by repressor, ambiguous and activators. Repressors and activators are defined based on more than 90% of the target genes being either repressed or upregulated, transcription factors with less than 90% for both are in the ambiguous class. (**B**) Distribution of the fraction of cells within a respective cell cluster with a transcription factor activity of greater than 0. The clusters include all cell types and malignant cell clusters from patient #1 split by activators, ambiguous and repressors. Clusters with high fraction of cells with activity greater than 0 indicate an active transcription factor, which is more prominent across repressors than for activators.

Additional file 7: Figure S7. Cell type abundances across patients and tissues. (**A**) Relative fractions (x-axis) of cell types (y-axis) across different patients (left) or tissue types (middle) with the absolute number of cells per cell type (right), colored by cell types from Fig. 5C. "NA" denotes cells that could not be demultiplexed from hashed samples and hence could not be assigned to a tissue type. (**B**) Relative fractions (x-axis) of cell types (y-axis) across different patients (left) with absolute

numbers per cell type (right). (**C**) UMAP of 41,284 cells based on OSCC scRNAseq data from our cohort and colored by patients. (**D**) UMAP of 21,037 cells based on CD45-negative and HPV-negative primary HNSCC from Kürten et al. and colored by patients.

Additional file 8: Figure S8. Bulk transcriptomes reveal the cellular composition of OSCC across tissue types. (**A**) Fractions of cell types (x-axis) across all samples (y-axis) including primary tumors (PT), metastatic lymph nodes (MET) and tumor-free lymph nodes (LN) for cells detected by scRNAseq (left panel) or deconvoluted from bulk transcriptome analysis (right panel). (**B**) Pie charts showing the average fraction of cell types across samples from each tissue type, derived from scRNAseq data (top) and bulk transcriptome deconvolution (bottom) and colored by cell type. cDCs: conventional dendritic cells; pDCs: plasmacytoid dendritic cells; RBCs: red blood cells; ECs: endothelial cells.

Additional file 9: Figure S9. Characterization of OSCC-derived fibroblasts. (**A**) Heatmap of scaled, log-normalized expression of the top 5 differentially expressed genes (DEGs) (rows) for fibroblasts and pericytes (columns) split by their respective phenotype. DEGs are sorted from highest towards lowest log2 foldchange and row sections are ordered like column sections. (**B**) Normalized enrichment scores (NES) of top 5 enriched gene sets for each fibroblast phenotype. Gene sets are sorted from highest to lowest NES and the bar chart is colored by negative decadic logarithm of Benjamini–Hochberg adjusted p-values (padj). (**C**) Scaled, log-normalized expression of collagens (COL) (rows) across fibroblasts and pericytes split by respective phenotypes (columns). Rows are clustered by their similarity using the Euclidean distance and ward.D2 method. (**D**) Selected genes (y-axis) expressed across phenotypes (x-axis). Dots are colored by averaged log-normalized gene expression and dot size represents the percentage of cells expressed in this phenotype, i.e., cells with more than 1 unique molecular identifier (UMI) detected in the respective gene. (**E–H**) Analog to A-D for fibroblasts and pericytes from Kürten et al. dataset. (**I**) Composition of phenotypes across tissue types in pie charts (top) and across samples as bar chart (bottom). Pie charts show the average fraction of phenotypes across fibroblasts and pericytes for each tissue type. The bar chart shows the fraction of phenotypes (x-axis) across samples (y-axis) on the left with the absolute abundance of cells on the right side, colored by tissue type. (**J**) Similar plot as in I for the Kürten et al. dataset. As all samples represent primary tumors, they were summarized in one pie chart.

Also, although description for "Additional file 10: Table S1" and "Additional file 11: Materials and Methods" are correct, the uploaded files have been swapped and uploaded in PDF format rather than CSV format that should be the case.

The original article [1] was updated.

Published online: 14 March 2024

Reference
1. Horny K, Sproll C, Peiffer L, Furtmann F, Gerhardt P, Gravemeyer J, Stoecklein NH, Spassova I, Becker JC. Mesenchymal–epithelial transition in lymph node metastases of oral squamous cell carcinoma is accompanied by ZEB1 expression. J Transl Med. 2023;21:267. https://doi.org/10.1186/s12967-023-04102-w.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.