**eLife’s transparent reporting form**

We encourage authors to provide detailed information within their submission to facilitate the interpretation and replication of experiments. Authors can upload supporting documentation to indicate the use of appropriate reporting guidelines for health-related research (see [EQUATOR Network](https://www.equator-network.org/), life science research (see the [BioSharing Information Resource](https://www.ebi.ac.uk/biorepository/)), or the [ARRIVE guidelines](https:// ARRIVEguidelines.org/) for reporting work involving animal research. Where applicable, authors should refer to any relevant reporting standards documents in this form.

If you have any questions, please consult our [Journal Policies](https://elifesciences.org/journal-policies) and/or contact us: editorial@elifesciences.org.

**Sample-size estimation**

- You should state whether an appropriate sample size was computed when the study was being designed
- You should state the statistical method of sample size computation and any required assumptions
- If no explicit power analysis was used, you should describe how you decided what sample (replicate) size (number) to use

Please outline where this information can be found within the submission (e.g., sections or figure legends), or explain why this information doesn’t apply to your submission:

For the data shown in Figure 1, Figure 1- figure supplement 1, and Figure 2- figure supplement 2 we chose to have 8 mice per group (and a total of 4 groups). From 3 groups we were able to isolate tumor samples and distant organs from all 8 mice and in one group of 7 mice. The information on the sample-size can be found in the figure legends for Figure 1, Figure 1- figure supplement 1, and Figure 2- figure supplement 2.

For the data shown in figure 7 we had 4 groups. In each group there were from 37-74 embryos depending on the experimental conditions. The exact number of fish embryos in each group used to quantify the area and circularity of the primary tumor as well as the number of disseminated cells are indicated in the figure 7 legend.

**Replicates**

- You should report how often each experiment was performed
- You should include a definition of biological versus technical replication
- The data obtained should be provided and sufficient information should be provided to indicate the number of independent biological and/or technical replicates
- If you encountered any outliers, you should describe how these were handled
- Criteria for exclusion/inclusion of data should be clearly stated
- High-throughput sequence data should be uploaded before submission, with a private link for reviewers provided (these are available from both GEO and ArrayExpress)

Please outline where this information can be found within the submission (e.g., sections or figure legends), or explain why this information doesn’t apply to your submission:
For each in vitro experiment (e.g. western blotting, immunofluorescence, RT-qPCR) the required information can be found in the respective figure legend and Materials and methods section.

The link for the high-throughput RNA-seq data as well as the GEO accession number for reviewers can be found in the Materials and methods section.

**Statistical reporting**

- Statistical analysis methods should be described and justified
- Raw data should be presented in figures whenever informative to do so (typically when N per group is less than 10)
- For each experiment, you should identify the statistical tests used, exact values of N, definitions of center, methods of multiple test correction, and dispersion and precision measures (e.g., mean, median, SD, SEM, confidence intervals; and, for the major substantive results, a measure of effect size (e.g., Pearson's r, Cohen's d)
- Report exact p-values wherever possible alongside the summary statistics and 95% confidence intervals. These should be reported for all key questions and not only when the p-value is less than 0.05.

Please outline where this information can be found within the submission (e.g., sections or figure legends), or explain why this information doesn’t apply to your submission:

| This information can be found in the Materials and methods (Subtitle: Statistical analysis) and in the appropriate figure legends. |

(For large datasets, or papers with a very large number of statistical tests, you may upload a single table file with tests, Ns, etc., with reference to sections in the manuscript.)

**Group allocation**

- Indicate how samples were allocated into experimental groups (in the case of clinical studies, please specify allocation to treatment method); if randomization was used, please also state if restricted randomization was applied
- Indicate if masking was used during group allocation, data collection and/or data analysis

Please outline where this information can be found within the submission (e.g., sections or figure legends), or explain why this information doesn’t apply to your submission:

| N/A |

**Additional data files (“source data”)**

- We encourage you to upload relevant additional data files, such as numerical data that are represented as a graph in a figure, or as a summary table
- Where provided, these should be in the most useful format, and they can be uploaded as “Source data” files linked to a main figure or table
- Include model definition files including the full list of parameters used
- Include code used for data analysis (e.g., R, MatLab)
- Avoid stating that data files are “available upon request”
Please indicate the figures or tables for which source data files have been provided:

We included the full size confocal images as a Source Datasets 1-4 for Figure 2e, Figure 4 a, c, and e, Figure 5 a, d and f, Figure 4 – figure supplement 1 b, Figure 4 – figure supplement 2 c, Figure 5 – figure supplement 1 b and c, Figure 5- figure supplement 2 a, b, d and e. The related Source dataset number is indicated in the figure legends.

The differentially expressed genes in the RNA-seq analysis were obtained by using DESeq2 Bioconductor package. Generally applicable gene set enrichment (GAGE) Bioconductor package was used for pathway analysis, and KEGG pathway maps were rendered with Pathview (https://pathview.uncc.edu). Morpheus (https://software.broadinstitute.org/morpheus/) was used to generate the gene heatmap shown in Figure 2 a, b and c.


We included two Supplementary Excel data files containing: i) more than 2-fold significantly altered genes found in RNA-seq analysis (Supplementary file 1, related to figure 2) and ii) all genes significantly up- and downregulated in the RNA-seq analysis and listed for both cell lines for the selected pathways (Supplementary file 2, related to figure 2).