

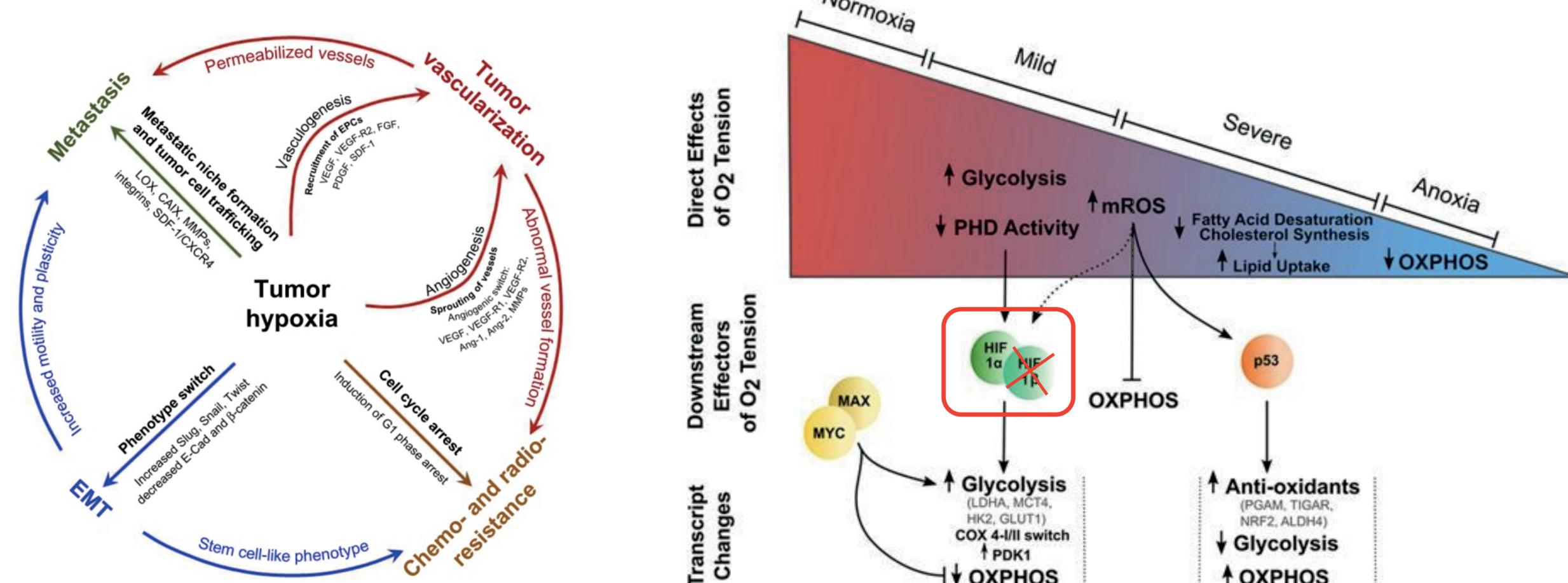
# Identification and Characterization of hypoxia-inducible Factor (HIF)-Dependent Alternative Splicing Events in Pancreatic Cancer

Philipp Markolin<sup>1\*</sup> & Natalie Davidson<sup>1\*</sup>, Christian K. Hirt<sup>1</sup>, Christophe Chabbert<sup>1</sup>, Nicola Zamboni<sup>3</sup>, Gerald Schwank<sup>1</sup>, Gunnar Rätsch<sup>2</sup> and Wilhelm Krek<sup>1</sup>

1. Institute of Molecular Health Sciences, Department of Biology, ETH Zurich, Switzerland 2. Biomedical Informatics Group, Department of Computer Science, ETH Zurich, Switzerland 3. Institute of Molecular Systems Biology, Department of Biology, ETH Zurich, Switzerland \* these authors contributed equally

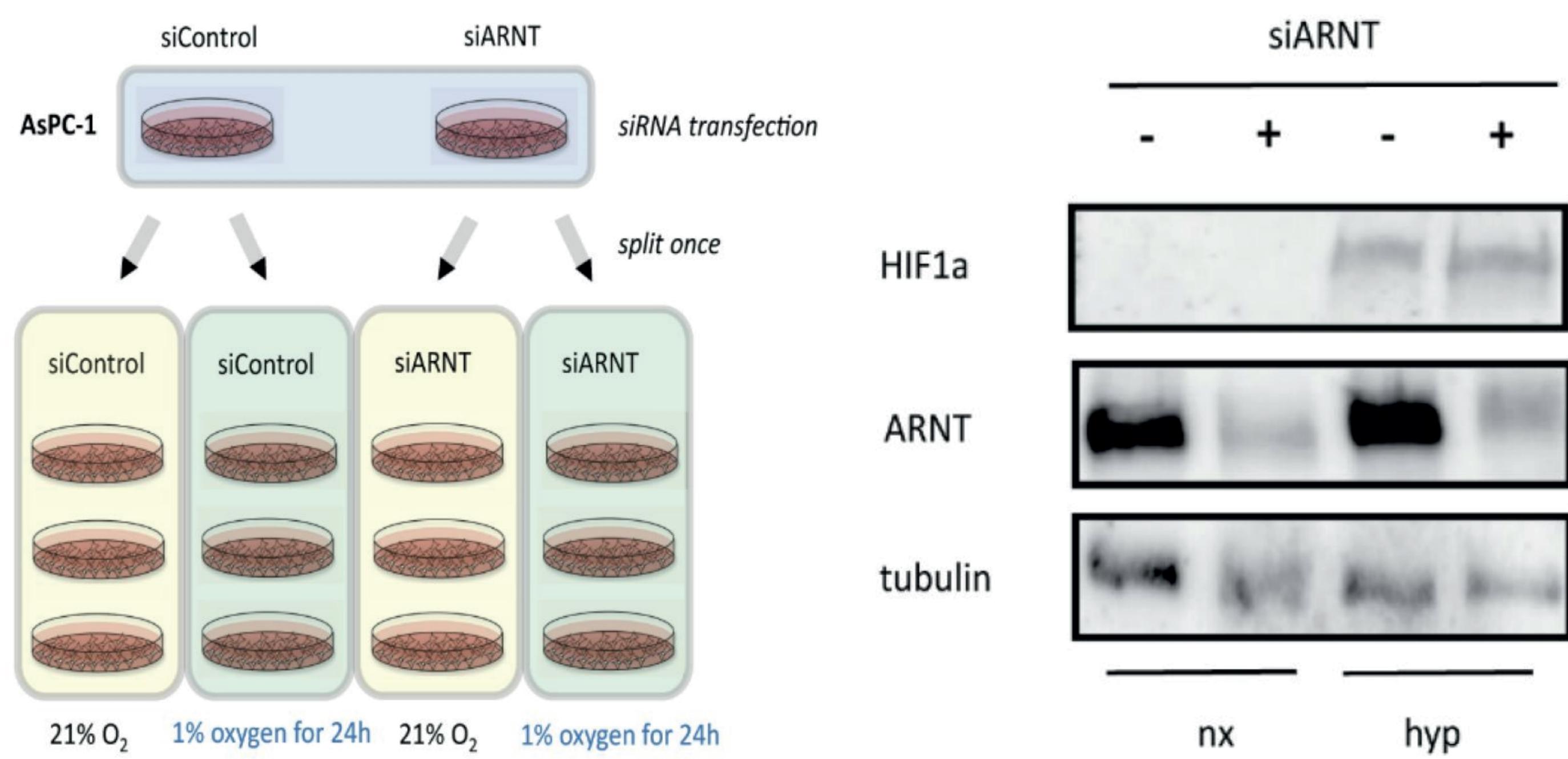
## Motivation

Hypoxia is prevalent in many tumors and associated with malignant tumor progression. Hypoxia mediates its effects primarily through hypoxia-inducible factors (HIFs), a family of heterodimeric transcription factors composed of distinct HIF1α and a shared HIF1β subunit. While the effects of HIFs on transcription programs have been extensively studied in the past, little is known about splicing. To determine such roles, we performed RNA-seq on human pancreatic cancer (PDAC) cells subjected to hypoxia in the absence or presence of HIF1β to define the HIF-dependency of any effects.



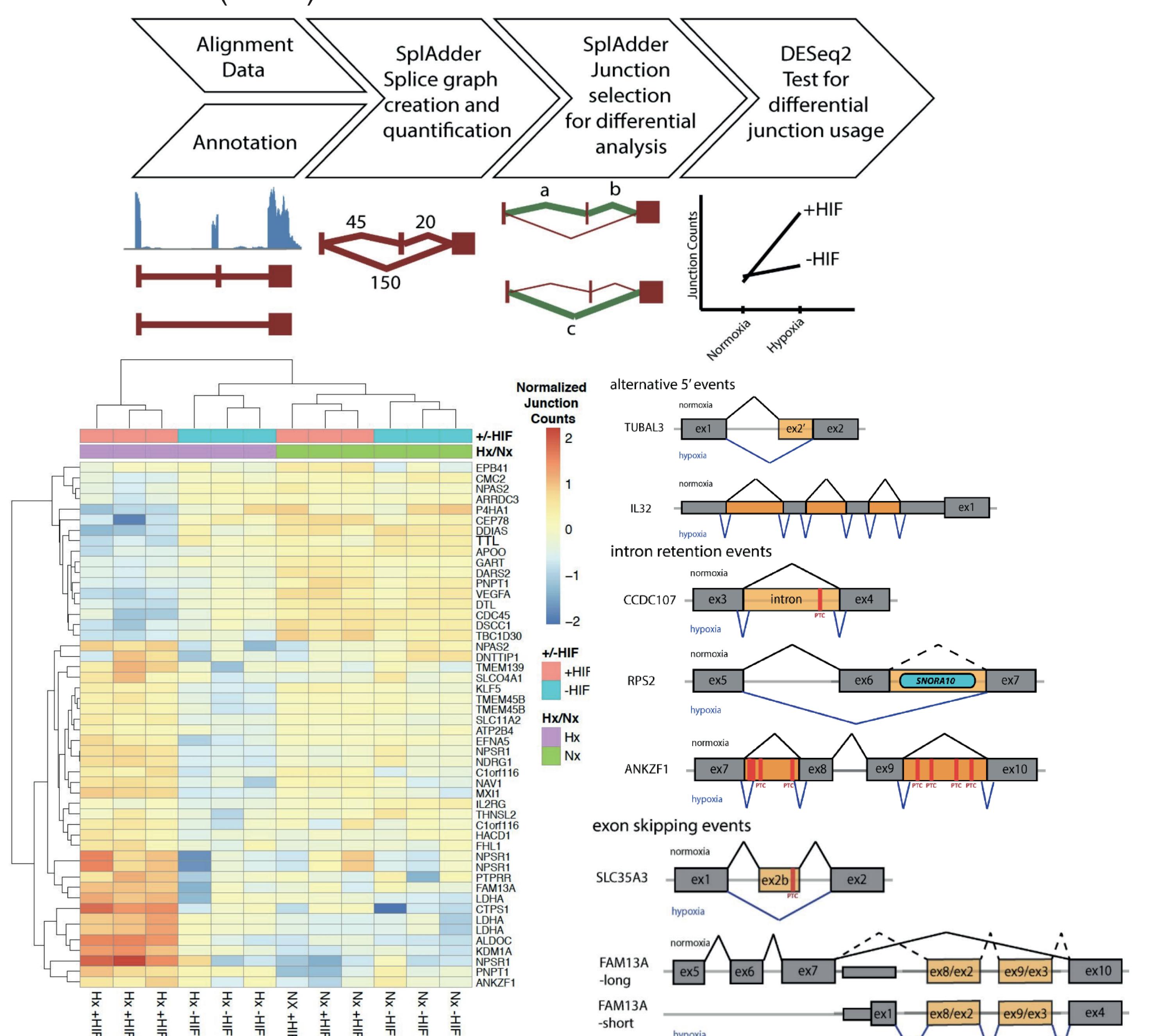
## Experimental Design

To understand the downstream splicing effects of HIF, we designed an experiment with four conditions: Hypoxic+/-HIF, Normoxic+/-HIF. Each condition was done in triplicate; an overview of the design is pictured below. HIF's effects were silenced through silencing HIF1β (ARNT), HIF1α's dimerization partner. The silencing of HIF1β can be seen in the western blot pictured below.



## HIF-Dependent Splice Event Detection

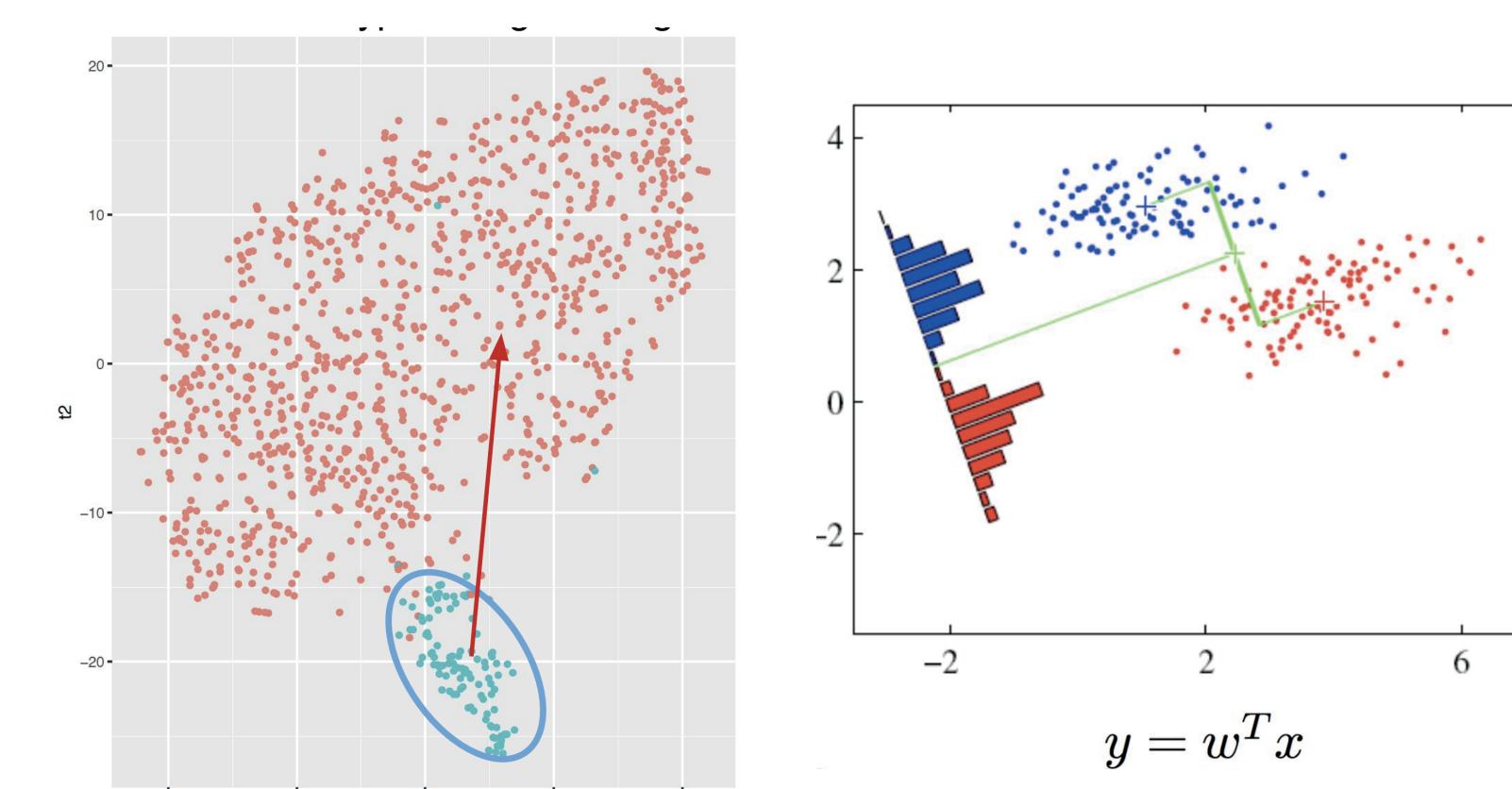
An overview of our computational analysis is pictured below. To identify significant events, we used DESeq2 to test the interaction between oxygen and ARNT state using junction counts. We identified 538 HIF-dependent events (FDR<15%), of which 38 events have percent-spliced-in change ( $\Delta\text{PSI}$ ) > 0.05. These events are depicted in the heatmap and gene-body diagrams below. We experimentally validated events using multiple PDAC cell lines and patient-derived PDAC organoids. More than half (22/38) were confirmed in TCGA.



References: 1. Kahles, Andre, et al. "SplAdder: identification, quantification and testing of alternative splicing events from RNA-Seq data." Bioinformatics 32.12 (2016): 1840-1847. 2. Love, Michael, Simon Anders, and Wolfgang Huber. "Differential analysis of count data—the DESeq2 package." Genome Biol 15 (2014): 550. 3. Yuan Tang, Gao Tao and Xiao Nan (2015). dml: Distance Metric Learning in R. R package version 1.1.0. <https://CRAN.R-project.org/package=dml>. 4. Eales, K. L., K. E. R. Hollinshead, and D. A. Tennant. "Hypoxia and metabolic adaptation of cancer cells." Oncogenesis 5.1 (2016): e190. 5. The role of hypoxia in cancer progression, angiogenesis, metastasis, and resistance to therapy

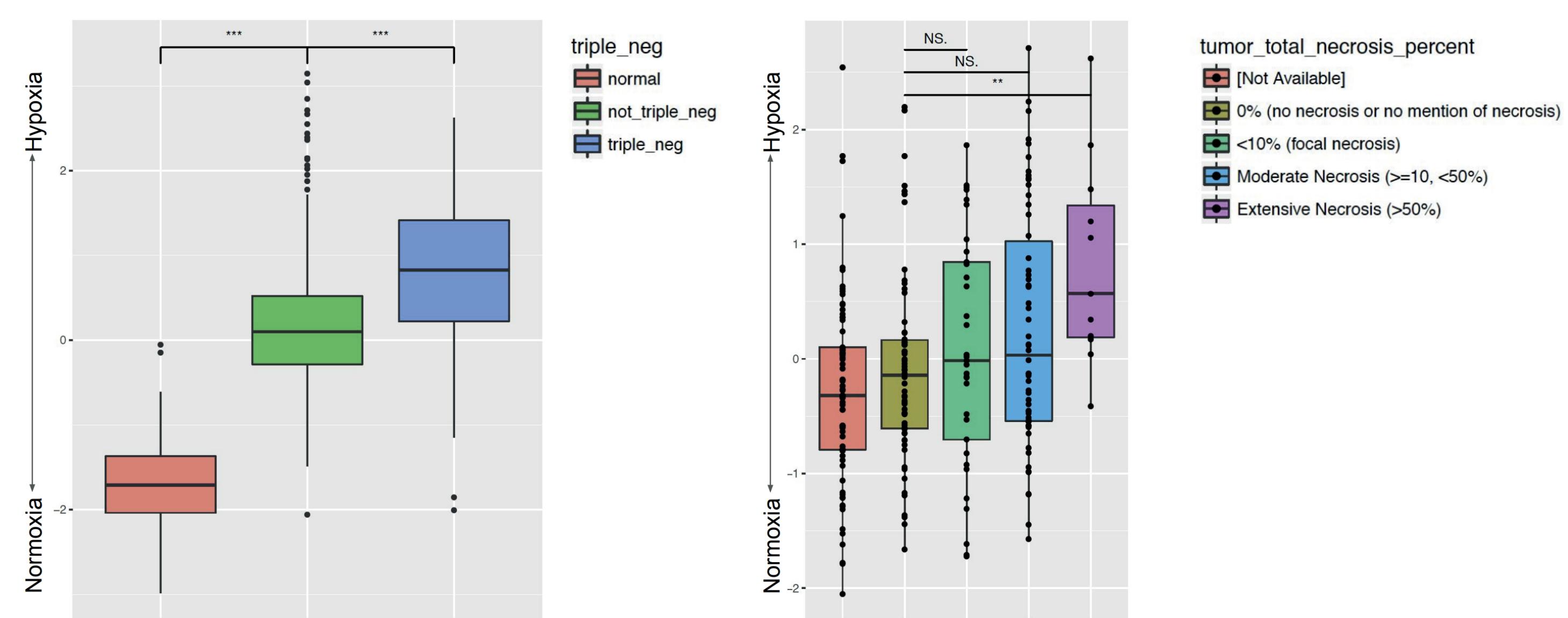
## Identifying Hypoxia in TCGA

Next, we explored if the splice events we identified were seen in other hypoxic contexts, namely the TCGA cohort. In order to do this we first have to estimate the hypoxic state of a specific sample. This was done using gene expression values for 32 previously identified key hypoxia genes to estimate a hypoxia score for each sample in TCGA. Shown below is a t-SNE plot of breast cancer (pink) and adjacent normal samples (blue) using the 32 genes. A weighting of the features was estimated using linear discriminant analysis to maximize the distance between the tumor and normal sample. The projected y value is the hypoxia score.



## Validation of Hypoxia Score in TCGA

Across the known hypoxic cancer types, pancreatic, breast, and lung adenocarcinoma, we were able to distinguish tumor from normal samples using only the hypoxia score with a test-set precision-recall AUC and ROC AUC greater than 70%. In order to ensure that we are truly identifying a hypoxia signal and not only differences between tumor normal, we compared our score to analogous clinical markers of hypoxia: necrosis in sarcoma samples and triple negative status in breast cancer. We find that these clinical markers are significantly correlated with the hypoxic score, as pictured below.



## Correlation of Splice Events with TCGA Hypoxia Score

Using the hypoxic score across all samples in TCGA, we correlated the score with the PSI values we identified as HIF-specific in our experiment using the model displayed below. We find significant correlations for the two events that were extensively validated, SLC35A3 (exon\_skip\_926) and FAM13A (exon\_skip\_13265), as well as several others within the TCGA cohort as pictured in the heatmap below. Blue squares signify a significant ( $p\text{-adj} < 0.05$ ) negative correlation between PSI and hypoxia score, whereas red signifies a positive correlation. We believe that this indicates that these events are not specific to our experiment, but can be generalized into other hypoxic environments.

$$\text{PSI} \sim \beta_0 + \beta_{\text{Exp}} \cdot \text{Exp} + \beta_{\text{Gender}} \cdot \text{Gender} + \beta_{\text{HypoxiaScore}} \cdot \text{HypoxiaScore} + \varepsilon \quad \varepsilon \sim \text{Gaussian}(\mu, \sigma^2)$$

