INTRODUCTION
Combining cancer patient spatial transcriptomics (ST) and bulk RNA-Seq allows for a more nuanced understanding of the gene expression profiles. While large-scale sequencing projects have identified some NSCLC cancer drivers, they cannot fully address the complexity of the heterogeneous tissue composition present in patients. ST enables the measurement of gene expression in specific regions within a tissue sample, complementing and refining insights gained from bulk RNA-Seq.

METHODS
Collection of samples: Tumor tissues were globally collected using a standardized protocol, removing the ischemia time until freezing in liquid nitrogen. To ensure the quality of the samples, all tissues were Histofree-ED and Eosin stained and subjected to pathological QC. Samples need to be invasive, have a tumor content of > 30 % and Necrosis < 30 %. Normal tissues were processed in parallel and need to be free of tumor and representative regarding the tumor tissue to be included. Approximately 10 mg tissue were taken for nucleic acid extraction and protein lysate preparation each. To account for tumor heterogeneity, pathological QC was performed on two sections, before and after taking the analysis material. The tissues stayed frozen during the entire process.

RNA-Seq and Spatial Transcriptomics analyses were performed on one NSCLC patient, adenocarcinoma (ADC), stage IV. RNA-Seq differential expression was additionally done on 52 ADCs.

Nucleic Acid Extraction and Quality Assessment: Frozen tissue slices were manually cut with beta-mercaptoethanol containing sample buffer and homogenized using the Beadbug system. RNA was extracted using the Qagen AllPrep kit according to the manufacturer’s instructions. RNA concentration was quantified using Quibet RNA BR assay respectively. RNA quality was assessed using Agilent High-Sensitivity RNA ScreenTape kit respectively. RNA need to have a RIN > 4 or 280/260 > 1.8 to be selected for library preparation.

RNA-Seq: For whole transcriptome sequencing, RNA samples were depleted of the ribosomal RNA using the Ribo Zero kit (Illumina) and library preparation was performed using the Truseq Stranded Total RNA kit (Illumagen). The library preparation kit was used according to manufacturer’s instructions. Sequencing was performed as a Novaseq5000 system (Illumina).

Whole transcriptome sequencing datasets have >= 100 million total reads with less than 20% of the reads mapping to the rRNA according to Ensembl reference. Ribosomal depletion was performed to remove nuclear-RNA and miRNA. NGS data was aligned against Gencode genome assembly.

10xGenomics Visium Spatial Gene Expression tissue-seq libraries were prepared according to the 10x Visium Spatial Gene Expression library preparation protocol. The final library was sequenced at a concentration of 300 pM on a Novaseq® 6000 using a S1 v1.5 flowcell (Illumina) to reach a minimum sequencing depth of 25,000 reads per tissue covered spot. Sequencing data were normalized separately, using the Ensembl human genome (version 95). Expression data were normalized using TMM (trimmed mean of M-values) and further analyzed using the multi-dimensional scaling algorithm (UMAP). The final library was sequenced at a concentration of 300 pM on a Novaseq® 6000 using a S1 v1.5 flowcell (Illumina) to reach a minimum sequencing depth of 25,000 reads per tissue covered spot. Sequencing data were normalized separately, using the Ensembl human genome (version 95). Expression data were normalized using TMM (trimmed mean of M-values) and further analyzed using the multi-dimensional scaling algorithm (UMAP). The final library was sequenced at a concentration of 300 pM on a Novaseq® 6000 using a S1 v1.5 flowcell (Illumina) to reach a minimum sequencing depth of 25,000 reads per tissue covered spot.

RESULTS
Integration and clustering of 10xGenomics tumor and normal data into R

RESULTS
Spatially resolved data can provide insights into the identification and assignment of cancer-relevant genes that primarily show only minor regulation in bulk RNA-Seq data

CONCLUSION
A more refined in situ understanding of gene expression profiles in the NSCLC microenvironment, combined with large cohort data will help guide therapeutic target selection. The combination not only refines target expression (e.g. primary tumor or immune system) but also contains a strong indicator for target gene expression variability within patients’ tumors for improved precision medicine approaches.