A PCA: Bulk 3'RNA-seq

P40 ex vivo

P39 ex vivo

P41 ex vivo

- BA
- CL
- IMS
- NA

PC1

PC2

B tSNE: scRNA-seq

P41 ex vivo

P39 ex vivo

P40 ex vivo

- BA
- CL
- IMS
- NA

tSNE 1

tSNE 2

C Primary tumor   Recurrence

P10

PT area 1
PT area 2
PT area 3

R area 1
R area 2
R area 3

P03

PT area1
PT area 2
PT area 3

R area 1
R area 2
R area 3
R area 4

P08

PT area 1
PT area 2
PT area 3

R area 1
R area 2
R area 3

4 mm

D

PT area 3
PT area 2
R area 2
R area 1
PT area 1
P10

R area 2
R area 3
R area 1
PT area1
PT area 1
PT area 3
PT area 2
P03
R area 4
R area 3

R area 1
R area 2
P08
PT area 2
R area 3
PT area 1
PT area 3

PC1

PC2

Patient
- P03
- P08
- P10

Fill Color
- BA
- CL
- IMS

Text Color
- Primary (PT)
- Recurrence (R)

A

PC2 (13.5% variance explained)

PC1 (15.8% variance explained)

B

PC2 11 % variance explained

PC1 18 % variance explained

C

PC2 11 % variance explained

PC1 18 % variance explained

D DGE: R vs. PT

−log10(P.adjust)

log2(FC)

E DGE: CL vs. BA

−log10(P.adjust)

log2(FC)

F DGE: IMS vs. BA

−log10(P.adjust)

log2(FC)

G DGE: IMS vs. CL

−log10(P.adjust)

log2(FC)

H GSEA: R vs. PT

−log10(P.adjust FWER)

PD1 SIGNALING

TCA CYCLE AND RESPIRATORY
ELECTRON TRANSPORT

REGULATION OF HYPOXIA INDUCIBLE
FACTOR HIF BY OXYGEN

GLUCOSE
METABOLISM

RESPIRATORY ELECTRON
TRANSPORT

PYRUVATE METABOLISM AND
CITRIC ACID TCA CYCLE

GLYCOLYSIS

FORMATION OF TRANSCRIPTION
NER TC NER REPAIR COMPLEX

CITRIC ACID CYCLE
TCA CYCLE

NUCLEOTIDE
EXCISION REPAIR

PYRUVATE METABOLISM

DOUBLE STRAND
BREAK REPAIR

NES

I GSEA: CL.vs. BA

−log10(P.adjust FWER)

NES

J GSEA: IMS vs. BA

−log10(P.adjust FWER)

GLYCOLYSIS

REGULATION OF HYPOXIA
INDUCIBLE FACTOR HIF BY OXYGEN

RESPIRATORY ELECTRON TRANSPORT

PD1 SIGNALING

TCA CYCLE AND
RESPIRATORY ELECTRON TRANSPORT

GLUCOSE METABOLISM

NES

K GSEA: IMS vs. CL

−log10(P.adjust FWER)

DNA REPAIR

PD1 SIGNALING

NUCLEOTIDE EXCISION REPAIR

HOMOLOGOUS RECOMBINATION
REPAIR OF REPLICATION INDEPENDENT
DOUBLE STRAND BREAKS

DOUBLE STRAND BREAK REPAIR

FORMATION OF
TRANSCRIPTION COUPLED NER TC
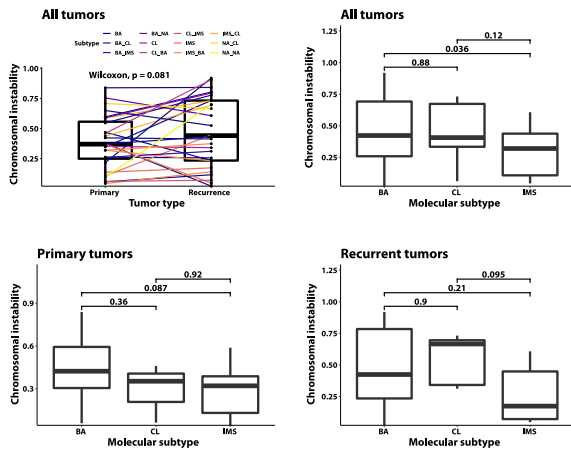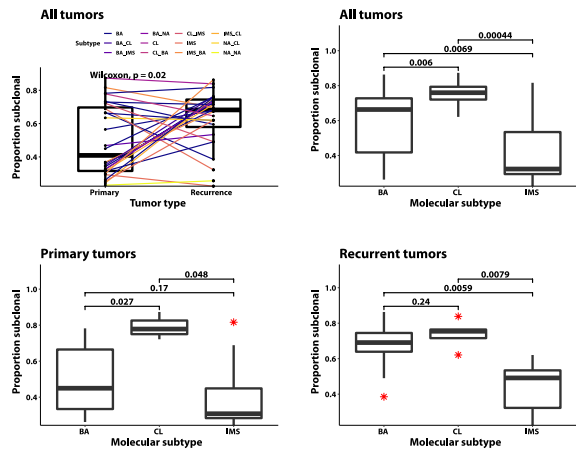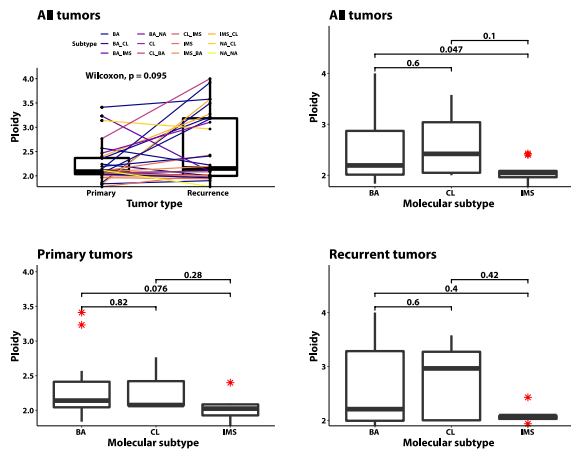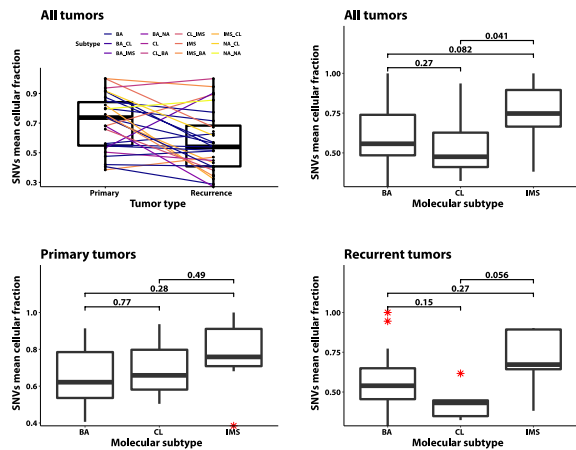NER REPAIR COMPLEX

GLYCOLYSIS

NES

+ HPV positive − HPV negative   ● BA; basal subtype   ● CL; classical subtype   ● IMS; inflamed-mesenchymal subtype

○ Other   ○ Oxygen Dependent Metabolism Various   ○ Glucose Metabolism Various   ○ PD1 Signaling   ○ DNA Repair Various

A Hypoxia | p < 1*10⁻⁴ **B** pEMT | p = 0.0112 **C** RadR | p = 0.0138 **D** TIS | p = 1*10⁻⁴

E Hypoxia | p < 1*10⁻⁴ **F** pEMT | p < 1*10⁻⁴ **G** RadR | p = 0.7575 **H** TIS | p < 1*10⁻⁴

I Hypoxia **J** pEMT **K** RadR **L** TIS

M Molecular Subtype **N** Subtype Change

A

Variants per sample
Median: 84.5

Observed variants

No. of samples

TP53 — 62%
FAT1 — 21%
PIK3CA — 18%
SEC16A — 18%
CCDC168 — 16%
PCLO — 16%
ANKRD17 — 14%
CSMD3 — 14%
KMT2D — 14%
LRP1B — 14%
NOTCH1 — 12%
SPATA31D1 — 12%
UBR4 — 12%
DMD — 11%
DNAH5 — 11%
FAT2 — 11%
KMT2B — 11%
NALCN — 11%
NAV2 — 11%
ADGRB3 — 9%

Patient ID
Tumor Type
Subtype

Legend (mutation types):
- Missense Mutation (green)
- In Frame Insertion (pink)
- Frame Shift Insertion (purple)
- Splice Site (orange)
- Nonsense Mutation (red)
- Multi Hit (black)

Patient ID: P01, P03, P04, P05, P09, P10, P12, P14, P17, P18, P19, P20, P21, P22, P23, P24, P25, P26, P27, P28, P29, P31, P32, P33, P34, P35, P36, P37

Tumor Type: Primary, Recurrence

Subtype: BA, CL, IMS, NA

B

Signature SBS2 Contribution

Kruskal–Wallis, p = 0.0099

0.036
0.0021
0.062
0.014

Contribution

Stage: I, II, III, IV, IVa

Title: Therapy-related transcriptional subtypes in matched primary and recurrent head and neck cancer

Authors: Peter Weber[1),2) ‡], Axel Künstner[3),4) ‡], Julia Hess[1),2) 5) ‡], Kristian Unger[1),2),5) ‡],

Sebastian Marschner[2),5)], Christian Idel[6)], Julika Ribbat-Idel[7)], Philipp Baumeister[8),2)], Olivier

Gires[8),2)], Christoph Walz[9)], Sibylle Rietzler[9)], Laura Valeanu[9)], Timm Herkommer[1)], Lisa

Kreutzer[1)], Olena Klymenko[5)], Guido Drexler[2),5)], Thomas Kirchner[9)], Cornelius Maihöfer[2),5)],

Ute Ganswindt[2),5),10)], Axel Walch[11)], Michael Sterr[12),13)], Heiko Lickert[12),13),14)], Martin Canis[8)],

Dirk Rades[15)], Sven Perner[7),16)], Mauricio Berriel Diaz[17)], Stefan Herzig[17)], Kirsten Lauber[2),5)],

Barbara Wollenberg[6),18) †], Hauke Busch[3),4)†], Claus Belka[2),5) †], Horst Zitzelsberger[1),2),5) †,*]

‡ have contributed equally

† shared senior authorship

* corresponding author

Affiliations:

1) Research Unit Radiation Cytogenetics, Helmholtz Zentrum München, Neuherberg, Germany

2) Clinical Cooperation Group "Personalized Radiotherapy in Head and Neck Cancer",

Helmholtz Zentrum München, Neuherberg, Germany

3) Luebeck Institute of Experimental Dermatology and Institute for Cardiogenetics, University of

Luebeck, Luebeck, Germany

4) University Cancer Center Schleswig-Holstein, University Hospital of Schleswig-Holstein, Campus Luebeck, Luebeck, Germany

5) Department of Radiation Oncology, University Hospital, LMU Munich, Munich, Germany

6) Department of Otorhinolaryngology, University of Luebeck, Luebeck, Germany

7) Institute of Pathology, University of Luebeck and University Hospital Schleswig-Holstein, Luebeck, Germany

8) Department of Otorhinolaryngology, Head and Neck Surgery, University Hospital, LMU Munich, Munich, Germany

9) Institute of Pathology, Faculty of Medicine, LMU Munich, Munich, Germany

10) Department of Therapeutic Radiology and Oncology, Innsbruck Medical University, Innsbruck, Austria

11) Research Unit Analytical Pathology, Helmholtz Zentrum München, Neuherberg, Germany

12) Institute of Diabetes and Regeneration Research, Helmholtz Diabetes Center, Helmholtz Zentrum München, Neuherberg, Germany

13) German Center for Diabetes Research (DZD), Neuherberg, Germany

14) Technical University of Munich, Munich, Germany

15) Department of Radiation Oncology, University of Luebeck, Luebeck, Germany

16) Pathology, Research Center Borstel, Leibniz Lung Center, Borstel, Germany

17) Institute of Diabetes and Cancer, Helmholtz Diabetes Center, Helmholtz Zentrum München, Neuherberg, Germany

18) Clinic of Otorhinolaryngology, Klinikum rechts der Isar, Technical University Munich, Munich, Germany

Corresponding author:

Prof. Dr. Horst Zitzelsberger

Research Unit Radiation Cytogenetics

Helmholtz Zentrum München

Ingolstädter Landstrasse 1

85764 Neuherberg

Germany

+49-89-3187-3421

Zitzelsberger@helmholtz-muenchen.de

Translational Relevance

Despite aggressive multi-modal treatment, patients suffering from advanced head and neck squamous cell carcinomas (HNSCC) are at high risk of recurrence as the clinical response to therapy depends both on the molecular characteristics of the individual tumors and on the extent of intratumor heterogeneity. A therapy-driven composition of tumor subpopulations has significant impact on therapy resistance and tumor recurrence. Here we present an in-depth analysis of molecular tumor dynamics in a unique cohort of primary and recurrent tumor pairs, which is of clinical importance in two ways: (i) it suggests transcriptome profiling of recurrences for therapeutic considerations and (ii) it identifies a basal tumor subtype as the most important molecular subgroup for alternative therapy approaches in recurrent tumors. Thus, transcriptome approaches should be included in molecular tumor boards and basal subtype-related processes should be considered for novel treatments options.

Abstract

Purpose:

The genetic relatedness between primary and recurrent head and neck squamous cell carcinomas (HNSCC) reflects the extent of heterogeneity and therapy-driven selection of tumor subpopulations. Yet, current treatment of recurrent HNSCC ignores the molecular characteristics of therapy-resistant tumor populations.

Experimental Design:

From 150 tumors, 74 primary HNSCCs were RNA-sequenced and 38 matched primary/recurrent tumor pairs were both, whole-exome and RNA-sequenced. Transcriptome analysis determined the predominant classical (CL), basal (BA) and inflamed-mesenchymal (IMS) transcriptional subtypes according to an established classification. Genomic alterations and clonal compositions of tumors were evaluated from whole-exome data.

Results:

While CL and IMS subtypes were more common in primary HNSCC with low recurrence rates, the BA subtype was more prevalent and stable in recurrent tumors. The BA subtype was associated with a transcriptional signature of PARTIAL EPITHELIAL-TO-MESENCHYMAL TRANSITION (P-EMT) and early recurrence. In 44% of matched cases, the dominant subtype changed from primary to recurrent tumors, preferably from IMS to BA or CL. Gene set enrichment analysis identified upregulation of HYPOXIA, P-EMT and RADIATION RESISTANCE signatures and downregulation of TUMOR INFLAMMATION in recurrences compared to index tumors. A relevant subset of primary/recurrent tumor pairs presented no evidence for a common clonal origin.

Conclusions:

Our study showed a high degree of genetic and transcriptional heterogeneity between primary/recurrent tumors, suggesting therapy-related selection of a transcriptional subtype with characteristics unfavorable for therapy. We conclude that therapy decisions should be based on genetic and transcriptional characteristics of recurrences rather than primary tumors to enable optimally tailored treatment strategies.

Word count: 6,190 words

Introduction

Head and neck squamous cell carcinomas (HNSCC) appear as a heterogenous disease in terms of etiology, phenotypes, biological, immunological, and clinical features (1). In line with this heterogeneity, molecular studies confirmed a molecular diversity on the genome, transcriptome, and immune landscapes of HNSCC (1-3). Etiological diversity is a major determinant of the heterogeneous HNSCC phenotype because of the two major entities (being either driven by human papilloma virus (HPV)-driven or induced by tobacco and alcohol abuse) that are now clearly defined in a new pathological classification system (4). Besides this inter-tumor heterogeneity also intra-tumor heterogeneity is a key feature of HNSCC and has been observed on different molecular levels (5-11). Deeper insights into tumor composition and cellular diversity were gained by the use of single-cell mRNA sequencing (scRNA-seq) (12). A recent scRNA-seq study on primary and metastatic HNSCC discovered partial epithelial-to-mesenchymal transition (p-EMT) as a key mechanism of nodal metastasis (9). However, the role of pEMT in the progression and recurrence of HNSCC after multi-modal treatment is so far unclear. A key component in multi-modal treatment is radiotherapy, which is significantly impacted by tumor hypoxia in terms of radiotherapy response (13). HNSCC cells can adapt to the hostile environment generated in hypoxic areas, thus promoting the emergence of more aggressive tumor phenotypes (14). Hypoxia contributes to radiation resistance of tumor cells, as does enhanced DNA repair, which in HNSCC is frequently associated with overexpression of the epidermal growth factor receptor (EGFR) and mutations in TP53 and CDKN2A (15). All these factors involved in radiation resistance support the development of tumor recurrence, which commonly occurs in advanced stage HNSCC – despite aggressive, multi-modal treatment comprising surgery and postoperative

7

radio(chemo)therapy. Local tumor relapses are consentaneously defined as tumors that develop within three years after the multi-modal treatment within the high-dose radiation field of treated tumors (16). A local relapse may originate from minimal residual disease or from surrounding preneoplastic cells, respectively, and thus may exhibit different levels of relationship to the primary tumor (17). In the latter case, the process of field cancerization, i.e. the occurrence of carcinogenic alterations across large areas of affected tissue in the absence of morphological change, may play an important role (18). Consequently, both the extent of genetic correspondence between primary and recurrent HNSCCs and the role of intra-tumor heterogeneity in tumor recurrence remain largely elusive. Cell populations contributing to local relapses were investigated in a surgical xenograft model of HNSCC of HNSCC which demonstrated how initially sparse clones propagated and eventually substituted dominating clones (19). These findings raise the question about genetic traits of such substituting clones in tumor recurrences. To answer these questions, genomic driver alterations play an important role. They are also a major focus in precision medicine of malignancies since they are used to target frequently aberrant molecular pathways in tumors (20). A randomized clinical trial compared molecularly targeted agents matched to specific molecular alterations *versus* conventional chemotherapy in refractory solid tumors, including HNSCC (21). The actionability of most molecular targets, based on the ESMO Scale of Actionability of molecular Targets (ESCAT) (22), was low in this SHIVA01 trial, highlighting the crucial importance of the type of molecular alteration beyond genomic driver genes and affected pathways (23). In the present study, we therefore aimed to identify similarities and differences on the exome and transcriptome levels in a panel of well-characterized matched pairs of primary and recurrent tumors from HPV-negative HNSCC patients. We hypothesized that

by performing a detailed matched-pair analysis, we could gain new insights into the development

of recurrence with possible implications for personalized treatment strategies.

Methods

HNSCC patient specimens

This observational study was designed to investigate matched pairs of primary tumors and recurrences in HNSCC patients after radiotherapy (RT) with curative intent. Formalin-fixed paraffin-embedded (FFPE) specimens of 150 HNSCCs - comprising 74 primary tumors and 38 matched primary/recurrent tumor pairs - were collected (resected tissues or pan-endoscopic biopsies; Ludwig-Maximilians-University (LMU) Clinics Munich and Luebeck University Clinics). This study was approved by the local ethics committees in Munich (EA 448-13/17-116) or Luebeck (16-277) and was carried out in accordance with the Declaration of Helsinki. HNSCC of the hypopharynx, larynx, oropharynx, or the oral cavity was histologically confirmed (**Table 1, Table S1**). Tumor stage was assessed using the UICC TNM Classification of Malignant Tumors, 7th edition. HPV-status was determined by p16^{INK4a} immunohistochemistry in combination with HPV DNA detection as described before (24).

HPV-negative primary/recurrent tumor pairs were included according to the following criteria: local or locoregional recurrence within the high-dose field after radio(chemo)therapy, tumor relapse within three years after initial diagnosis to exclude secondary cancers (and in the case of definitive radiotherapy treatment, at the earliest six months after therapy completion). 29 patients received initial surgical resection followed by adjuvant RT (median dose: 64 Gy; range 62.4-75.6 Gy), ten patients thereof underwent concomitant chemotherapy. Nine patients were treated by definitive radio(chemo)therapy (median dose: 70 Gy; range 69.96-80.5 Gy). Five patients underwent definitive or adjuvant RT followed by an interstitial HDBT-Boost to the tumor or tumor bed.

The independent LMU-KKG (LMU Munich, Clinical Cooperation Group "Personalized Radiotherapy in Head and Neck Cancer") cohort (n=74; **Table S1**) includes HNSCC patients with available follow-up data who received resection with curative intent followed by adjuvant radio(chemo)therapy (median dose 64 Gy) at the LMU Department of Radiation Oncology between 2008 and 2013 (25,26). In the case of close or positive microscopic resection margins and/or extracapsular extension, patients (n=49) received concurrent chemotherapy.

Tumor specimens were processed as serial sections (3 and 10 µm) from FFPE tumor tissue blocks, of which the first and last sections were stained with hematoxylin/eosin (HE). The HE-stained tissue sections were examined by a pathologist and the tumor area was defined. The annotated tumor area was dissected followed by simultaneous DNA and total RNA extraction as described in Supplementary Methods.

Primary *ex vivo* cell cultures were established from surgical material of three HPV-negative HNSCC patients (P39_ex_vivo, P40_ex_vivo, P41_ex_vivo; **Table S2**). Tissues were minced with a scalpel and incubated with Liberase[TM] (1 mg/ml in HBSS; Roche) for 1 hour at 37°C. The cell suspensions were treated as described in Supplementary Methods. Afterwards, the cells were seeded and incubated in a humified atmosphere at 37°C and 5% CO_2. The primary cells were cultivated as described in Supplementary Methods.

In this study, whole RNA, 3'mRNA and single-cell RNA sequencing was performed.

Whole RNA sequencing

For whole RNA sequencing, a protocol adapted for FFPE material was used. Prior to library generation, RNA integrity was assessed using Bioanalyzer systems and Agilent RNA 6000 Nano Reagents according to the manufacturer's protocols (Agilent Technologies, Inc., USA). The percentage of fragments >200 nucleotides (DV200) was determined followed by categorization

into low, medium, and good quality samples. Accordingly, RNA sequencing library preparation using the TruSeq Stranded Total RNA Library Prep Gold Kit (Illumina, Inc., USA) was adjusted for input quantity and RNA fragmentation time. After library preparation, library quality and quantity were evaluated using the Quanti-iT PicoGreen dsDNA Assay (Thermo Fisher Scientific, MA, USA) and the Bioanalyzer High Sensitivity DNA Analysis Kits (Agilent Technologies). Sequencing and data analyses were performed as described in Supplementary Methods.

3'mRNA sequencing

Different microscopically defined areas (n=19) of n=3 matched pairs of primary/recurrent FFPE tumor specimens and *ex vivo* primary cell cultures on four technical replicates of each bulk and subclone were subjected to 3'mRNA sequencing (3'RNA-seq). RNA isolation, quantification, and quality assessment were conducted as for whole RNA-seq. Sequencing libraries were generated using the QuantSeq 3'-RNA-Seq Library Prep Kit FWD for Illumina with dual indexing (Lexogen) according to the manufacturer's instructions for low quantity samples. After library quality assessment, sequencing was performed, and data analysis was carried out with R-packages as outlined in Supplementary Methods.

Single-cell RNA sequencing

Single-cell RNA sequencing on *ex vivo* cells of P39_ex_vivo, P40_ex_vivo and P41_ex_vivo was conducted on cell suspensions generated according to 10x Genomics Sample Preparation Demonstrated Protocol (10x Genomics, Inc. Pleasanton USA). They were microscopically inspected, counted and diluted to a specific target density. Sequencing libraries were constructed with Chromium Single Cell 3' Reagent Kits v2 (10x Genomics) and subsequent sequencing was performed as a pool of the three libraries on four channels of an Illumina HiSeq2000 machine. Processing of raw base calls, quality control and data analysis (Seurat 2.3.0) is detailed in

Supplementary methods. Classification of subtypes was performed by a custom R script that closely follows the method described by Puram *et al.* [9] and detailed in Supplementary Methods.

Bulk RNA sequencing data analysis

RNA sequencing data were processed as outlined in Supplementary Methods. Molecular subtyping was performed by the nearest centroid method (27). VST normalized expression values were first gene-wise scaled and median centered and then correlated to the centroid table obtained from Keck *et al.* (3). The subtype with highest Pearson correlation coefficient was assigned to each sample. For TCGA subtype calling, the same method was applied (a custom centroid table was computed from data of TCGA HNSCC samples). For pathway analyses, gene set enrichment approach (GSEA) was used with the java command line tool v3 (28,29).

Metagene activity scores were computed as median of gene-wise scaled and centered expression values. Thereby, for survival and radiation resistance scores, signature genes with negative associations were sign inverted. For comparisons primary *versus* recurrent tumors Wilcoxon tests with patient as pairing variable were used in inference testing. For subtype comparisons Kruskal-Wallis-Tests were computed with pairwise Wilcoxon Tests *post-hoc*. Univariable Cox proportional hazard models were computed and Kaplan-Meier analysis visualized with the R packages *survival* (30), and *survminer*. Statistical results from metagene activity scores are reported as nominal p-values.

Cibersortx deconvolution was used to determine the proportion of immune cells from the normalized read counts of all matched tumor pairs. In addition, tumor subclonality was reconstructed from aligned and recalibrated bam files using the HATCHet approach. Details are outlined in Supplementary Methods.

Whole Exome Sequencing

For whole exome sequencing (WES) DNA quality was assessed and samples were categorized into low, medium, and good quality according to Bioanalyzer profiles (details in Supplementary Methods). Sequencing library construction was adjusted according to quality measurements. Exome processing included quality checks, trimming, filtering and mapping to a human reference. The tools are outlined in Supplementary Methods. Single nucleotide variants and short insertions and deletions were called using *Mutect2*. Quality assessment for cross-sample contamination, filtering of somatic variants and variant annotation was performed as described in Supplementary Methods. Somatic signatures were estimated based on cophenetic correlation metric and signatures were compared to known SBS signatures.

Network diffusion and pathway analysis

After data processing (Supplementary Methods) resulting gene sets were plotted using the R package *ComplexHeatmap* (v 2.0.0) (31).

Genomic copy number and genomic biomarkers

For genomic copy number analysis of primary/recurrent tumor pairs Mutect2 variant calling was run including the "genotype-germline-sites" and "genotype-pon-sites" flags in order to record germline records existent in the normal samples and pool of normal samples as required by PureCN (version 1.22.2). Details of genomic copy number analysis are outlined in Supplementary Methods and include the determination of ploidy, cellularity, extent of subclonality and chromosomal instability for each tumor sample.

Assessment of clonality of matched primary and recurrent tumor pairs

Analysis of clonality was performed with logR-ratios calculated from GC-normalized read coverages using the Bioconductor Clonality R package. Pairs of genomic copy number profiles were considered independent if the calculated odds-ratio of independence was >1.

Data availability statement

Processed bam files from next-generation whole-exome sequencing have been deposited in the European Genome Phenome Archive (EGAS00001004941). RNA sequencing and single-cell RNA sequencing data have been deposited at Gene Expression Omnibus under GSE173855 (RNA-seq), GSE186461 (RNA-seq tumor areas) and GSE173964 (scRNA-seq), respectively. Access to all codes used in this study can be obtained from the Github repository (https://github.com/ZytoHMGU/HNSCC-GenTrans-PrimRelapse) upon request.

Results

To elucidate the extent of genetic relationship in pairs of primary and recurrent HPV-negative HNSCC we combined an exome and a transcriptome analysis of patient and *ex vivo* samples. In a first step, we characterized three primary *ex vivo* HNSCC cultures (**Table S2**) by both bulk 3'RNA-seq and scRNA-seq, which revealed an overt degree of heterogeneity of transcriptional subtypes within each tumor. Transcriptional subtyping of different tumor areas from RNA-seq in three primary/recurrent tumor pairs revealed a subtype heterogeneity within selected samples. This prompted us to hypothesize that tumor recurrence upon radio(chemo)therapy is related to tumor cell subpopulations that have advantageous molecular phenotypes. For a more in-depth examination of therapy-related molecular phenotypes in HNSCC recurrences, we focused on patients who developed a recurrence in the high-dose irradiation field within three years of initial diagnosis after definitive or adjuvant (chemo)radiation therapy (**Table 1**). For comparison, we also investigated the independent LMU-KKG cohort (n=74) of HNSCC patients with clinical follow-up data on recurrences by means of whole transcriptome sequencing (**Table S1**).

Transcriptomic heterogeneity of primary ex vivo HNSCC cultures and different tumor areas in HNSCC specimens

We generated scRNA-seq profiles from primary *ex vivo HNSCC* cultures of three patients (**Table S2**; P39_ex_vivo 5,536 cells; P40_ex_vivo 3,055 cells; P41_ex_vivo 5,388 cells) encompassing transcriptomes for in total 14,267 high quality cells. For transcriptional subtyping of each transcriptome, centroid genes as published by Keck *et al.* (3) were used. The assignment of basal (BA), classical (CL) and inflamed-mesenchymal (IMS) subtypes revealed a mixture of subtypes (P39_ex_vivo: BA 1,490 cells, CL 81 cells, IMS 403 cells; P40_ex_vivo: BA 512 cells, CL 1,712

cells, IMS 32 cells; patient P41_ex_vivo: BA 236 cells, CL 589 cells, IMS 62 cells; apart from non-assignable cells for each case) in all *ex vivo* cultures. The transcriptional subtypes of each case were determined by bulk 3'RNA sequencing of five isolated single-cell subclones and the bulk culture. Diverse subtypes in the subclones were confirmed and highly concordant findings for predominant subtypes in single-cell and bulk 3'RNA sequencing were obtained (**Figures 1A and 1B**). Hence, bulk mRNA sequencing was able to identify the dominating subtype in each case. In addition, RNA-seq analysis of different tumor areas (n=19) from three primary/recurrent tumor pairs revealed heterogeneous transcriptional subtypes in one primary and two recurrent tumors confirming at the same time the concept of predominant transcriptional subtypes (**Figures 1C and 1D**). These hypothesis-generating approaches rationalized a systematic bulk mRNA sequencing study on two FFPE tumor cohorts for which scRNA-seq analyses were not feasible.

Changes of predominant transcriptional subtypes between primary and recurrent tumors

To examine therapy-related transcriptional subtypes, we assessed the dominating subtypes in tumor pairs (n=38) of primary HNSCCs and local or loco-regional recurrences. For comparison and a more stable subtype calling, we included an independent cohort of n=74 primary HNSCCs with known recurrence status. First, an unsupervised clustering of all samples and between matched tumor pairs based on a principal component analysis (PCA) could not discern any specific subgroups with regard to general technical parameters or clinical characteristics of different cohorts (**Figures S1A, S1B**). Next, we performed a transcriptional subtype calling according to subtypes defined by Keck *et al.* (3) (malignant cells only: BA, CL, IMS) or TCGA-defined subtypes (1) (basal - BA, atypical - AT, classical – CL, mesenchymal – MS related to cancer-associated fibroblasts). Based on the median expression profiles of centroid genes, the MS subtype

centroid (TCGA) shared similarity with both BA subtypes, whilst the IMS centroid (Keck *et al.*) was similar to the AT centroid of TCGA. (**Figure S1C**). Hierarchical cluster analysis of with expressions of the "Keck subtype" centroid genes sorted samples mainly according to the assigned predominant subtypes (**Figure S2**). PCA of dominating subtypes according to Keck *et al.* (3) showed a clear clustering of individual subtypes and of HPV-related subtypes for all tumors of the reference cohort (**Figure 2A**). Furthermore, a PCA on the primary and recurrent tumors indicated the correlation or likewise dissimilarity between the pairs according to their two-dimensional separation along the first two principal components (**Figure 2B**). After exclusion of four outlier tumor pairs due to low quality, 19 primary tumors maintained the initial subtype in the recurrent tumors, while 15 tumors changed the predominant transcriptomic subtype between primary and relapsed tumor (**Figure 2C, Table 1**). Interestingly, the majority of these subtype switchers changed to the BA subtype (60%; 9/15), four cases to the CL subtype (27%; 4/15), and only two cases to the IMS subtype (13%; 2/15). "BA" was the most common subtype in primary tumors (50%; 17/34), was maintained in 94% of recurrences and selected in 60% of non-BA primary tumors in recurrences. These findings strongly suggested a preferential association of the BA subtype with tumor recurrence. The prevalence of individual subtypes was further investigated in the independent LMU-KKG HNSCC cohort (n=74 primary tumors; **Table S1**). This cohort with a recurrence rate of only 14% (10/74) showed a different subtype prevalence (BA: 24.4%, CL: 37.8%, IMS: 37.8%) compared to the cohort of primary/recurrent tumor pairs.

To explore the transcriptional diversity between primary tumors and recurrences and between individual transcriptional subtypes, we performed differential gene expression (DGE) and gene set enrichment (GSEA) analyses identifying deregulated cancer-related gene sets and pathways (**Figure 2D-G; Tables S3-10**). Gene sets contributing to metabolic functions (GLUCOSE and

OXYGEN METABOLISM) and DNA REPAIR processes (NUCLEOTIDE EXCISION REPAIR and DOUBLE-STRAND BREAK REPAIR) were enriched in relapsed tumors, whereas PD1 SIGNALING was diminished. GSEA between CL and BA subtypes revealed less distinct differences which were mainly associated with repair and metabolic processes (**Figure 2H-K**), whereas the IMS subtype showed enriched PD1 SIGNALING and diminished GLUCOSE METABOLISM in comparison to the BA subtype. Compared to the CL subtype, the IMS subtype also exhibited enriched PD1 SIGNALING, yet diminished DNA repair processes (NER, HR, DSB REPAIR). GSEA enrichment maps summarize the major molecular characteristics between primary/recurrent tumors and between the BA subtype and the other subtypes (**Figure S3**).

Tumor recurrence and transcriptional subtypes associated with prognosis-related gene sets

On the basis of differentially enriched gene sets between primary and relapsed tumors, we obtained deregulated pathways and cellular functions in tumor recurrences and individual transcriptional subtypes. Next, we selected published sets of metagenes from prognostic signatures that are related to the GSEA related functions and important prognostic features: HYPOXIA, RADIATION RESISTANCE, P-EMT, EGFR, SURVIVAL, TUMOR INFLAMMATION SIGNATURE (TIS) (9,32-35). Signature expression scores were calculated and subjected to differential analysis between primary and recurrent tumors in matched-pair comparisons and between groups defined by transcriptional subtypes. In recurrent tumors, HYPOXIA and P-EMT were upregulated and the expression of RADIATION RESISTANCE was elevated compared to primary tumors. In contrast, TIS expression scores were significantly decreased in recurrences (paired Wilcox test, $p<0.05$) (**Figure 3A-D**). Besides the activity change in prognostic signatures during tumor relapse, we also examined the interdependence between transcriptional subtypes and these signatures. HYPOXIA, P-EMT and EGFR

showed the strongest expression in the BA subtype, whereas the IMS subtype had the highest and the BA subtype the lowest SURVIVAL expression score (Kruskal test, p<0.05). At the same time, the IMS subtype showed the highest and the CL subtype the lowest expression score for TUMOR INFLAMMATION (Kruskal test, p<0.05) (**Figures 3E-H, S4**). Time to local/locoregional recurrence reflected the impact of transcriptional subtypes and prognostic signatures on the latency time of tumor relapse. High expression scores of HYPOXIA, P-EMT and RADIATION RESISTANCE were associated with a significantly shorter latency time to recurrence (pairwise Wilcox test, p<0.05) (**Figure 3I-L**). The BA subtype showed the shortest latency time for relapse development (pairwise Wilcox test, p<0.05) (**Figure 3M**). Interestingly, the tumors considered as subtype switchers had a significantly prolonged time to local/loco-regional recurrence compared to non-switchers (pairwise Wilcox test, p<0.05) suggesting complex processes of transcriptional reprogramming and/or subclone selection after therapy (**Figure 3N**). At the same time, the initially determined transcriptional distance between primary and recurrent tumors (**Figure 2B**) had no significant impact on latency time (**Figure S4**).

We also observed a significantly reduced TIS expression in the CL subtype regardless of primary and recurrent disease status (**Figure S5A-B**). This effect was even more pronounced, if only recurrent tumors were compared (**Figure 3H**). Thus, the CL subtype showed the lowest TIS expression score. We further investigated, if this observation was also reflected by other immunological parameters, such as the immune checkpoint marker PD-L1 and infiltrating immune cell populations. PD-L1 staining and derived tumor proportional score (TPS) revealed no significant differences between primary and recurrent tumors, however, confirmed a significantly reduced PD-L1 expression associated with the CL subtype (**Figure S5C-D**). Deconvolving 29 infiltrating immune cell populations from RNA-seq counts showed for some of them different

frequencies between primary and recurrent tumors. Considering only immune cell populations that correlated either with TUMOR INFLAMMATION expression or PD-L1 TPS scores resulted in a panel of ten immune cell types, including dendritic cells, monocytes, B cells and lymphocytes of which dendritic cells (mDC, pDC) correlated with both scores. Similar to TIS and TPS scores, dendritic cells were significantly reduced in CL subtypes. On the other hand, natural killer (NK) cells and plasmablasts were significantly reduced in BA subtypes. With regard to primary/recurrent tumor comparisons, NK cells and plasmablasts were strongly reduced in recurrences which is in line with the prevalence of BA subtypes in recurrent tumors (**Figure S6A-E**).

A high degree of diversity at exome DNA level between tumor pairs

Genetic diversity was also investigated by WES of primary/recurrent tumor pairs (n=28). The tumor samples had 92 (±97, median 85±51) somatic mutations in average. In total, 19 genes showed a mutation frequency >10%. TP53 was the most commonly mutated gene in 62% of samples (**Figure 4A, Table S11**). Differences in mutational signatures among covariates were assessed. Different tumor stages showed a significant difference for the COSMIC Single Base Signature (SBS) SBS2 (**Figure 4B**). The summary statistics demonstrated missense mutations as the most common variant class and single nucleotide variants (SNV) as the most frequent variant type (Figures S7).

We used a pathway enrichment against the Molecular Signature database (MSigDB) hallmark gene sets from the results of a network propagation approach (36) on the somatic mutations to elucidate the molecular effect of the patients' mutatomes. The resulting interaction networks revealed three clusters of up- and downregulated gene sets in all tumor samples, but no specifically deregulated gene sets in neither the primary nor the recurrent tumors (**Figure S8A**). Compared to the HNSCC TCGA cohort, the paired tumor samples of this study revealed a slightly lower mutational burden

not reaching statistical significance (**Figure S8B**). A comparison of SNVs of the top 20 genes between primary and recurrent tumors (**Figure 4A**) showed on average an overlap of 70.9% (s.d. ± 33.31%) mutations between paired tumors. Thus, an unexpectedly low degree of similarity for tumor gene variants between primary tumors and recurrences was observed. With further consideration of various quality controls (**Tables S12 and S13**) due to the usage of FFPE materials we are confident that the WES data are valid and better suited for variant calling than RNA-seq data (**Figure S9**).

The WES data were used to determine genomic copy number variations (CNV) of primary/relapsed tumor pairs (n=28). Unsupervised hierarchical clustering of CNVs revealed three main clusters which were neither associated with primary/recurrent tumor types nor with tumor site, sex, or transcriptional subtype (**Figure 5A**) or the primary/recurrence status of tumors. 29 amplified copy number regions on chromosomes 3, 12 and 20, covering 167 genes were associated with the CL subtype and one deleted region covering the gene CROCC gene on chromosome 1 was associated with the IMS subtype (**Figure S10, Tables S14-15**). Analysis of clonal relatedness of primary and recurrent tumors revealed that 6/28 (21%) primary/recurrent tumor pairs are likely to be genetically unrelated (odds ratio >1), while the others appeared to be clonally related (79%, odds ratio <1; **Figure S11, Table S16**).

Exome-derived SNV and CNV data were used to explore the extent of chromosomal instability, subclonality, ploidy and cellularity of each tumor sample. An overview of the most frequent CNVs is depicted in **Figure 5A**. Matched pairs of primary/recurrent tumors were composed of up to five tumor subclones. A clonal expansion was observed in 16/28 tumor pairs indicating a potential clonal selection of therapy-related subclones (**Figure 5B**). Chromosomal instability scores were not significantly different between primary tumors (mean: 0.38), and recurrences (0.48), but were

significantly lower in IMS tumors (mean: 0.28) compared to the other subtypes (mean: 0.46, p=0.028; **Figures 5C**, **S12C**). As reflected by the proportion of subclonal variants within all identified alterations, there was an increased level of subclonality in recurrences (mean: 74%) compared to primary tumors (mean: 70%, p=0.02). Along the transcriptional subtypes, CL exhibited the highest (mean: 76%), BA an intermediate (mean: 59%) and IMS the lowest level of subclonality (mean: 42%, p<0.06; **Figures 5D**, **S12A**). Ploidy was not significantly different between primary (mean: 2.25) and recurrent tumors (mean: 2.55) but was significantly increased in BA/CL-subtypes (mean: 2.50) compared to the IMS subtype (mean: 2.1, p=0.033; **Figures 5E**, **S12D**). The mean size of the cellular fractions estimated to carry SNV/CNV alterations was increased in primary (mean: 70%) compared to recurrent tumors (mean: 57%, p=0.01) and increased in IMS tumors (mean: 74%) compared to BA/CL tumors (mean: 59%, p=0.042; **Figures 5F** and **S12B**). Cellularity (i.e., the proportion of tumor cells in the analyzed bulk) was determined as another genomic biomarker for each tumor. Primary tumors had lower levels of cellularity (mean: 28%) compared to recurrent tumors (mean: 42%, p= 0.0015), which was the lowest for the IMS subtype showed the lowest mean value of cellularity (mean: 22%), followed by the BA subtype (mean: 34%) and CL subtype (mean: 60%, p<0.01). A comparison of the degree of cellularity estimated from the exome data with those determined from the histological hematoxylin/eosin-stained slides indicated a moderate (Pearson correlation: 0.59) but statistically significant agreement (p<0.001; **Figures S12E-F**).

Thus, CNVs confirmed genetic heterogeneity between tumor pairs and disclosed a clear link between transcriptional subtypes, CNVs, and derived subclones. In this sense, the more similar CL and BA subtypes both showed a disparity to the IMS subtype.

Discussion

Genetic diversity across and within tumors represents a major challenge for cancer therapy affecting cancer pathways and phenotypic heterogeneity (37). As a consequence, diverse cancer cell populations evolve during tumor progression and/or in response to treatment (38,39). Tumor cell dynamics are poorly understood in the therapy-related tumor recurrence of HNSCC. Thus, a longitudinal analysis of matched samples of primary and relapsed tumors is of paramount clinical importance. However, such cohorts are rare and sample sizes are usually small. Here, we established and analyzed the hitherto largest sample cohort of primary HNSCCs and corresponding recurrences that emerged upon surgery and radio(chemo)therapy. We used NGS to investigate the genetic diversity and dynamics at the genomic and transcriptomic level.

First, we used RNA sequencing of HNSCC patient samples to systematically determine dominating transcriptional subtypes. We classified each individual HPV-negative tumor sample as BA, CL or IMS subtype according to Keck *et al.* (3) and compared this classification approach with TCGA-derived subtypes (9) in our whole dataset. Hereinafter, we focused our analyses on the tumor cell-specific subtypes according to Keck *et al.* ("Keck subtypes") (3). The predominant subtypes of our HNSCC cohorts revealed a complex picture for subtype prevalence, diversity, and longitudinal persistence. (i) The CL and IMS subtypes were more frequent in primary HNSCC with a low recurrence rate, while the BA subtype was enriched in primary/recurrent tumor pairs. (ii) The BA subtype was stable between primary and recurrent tumors and associated with the expression of P-EMT (9) and EGFR (40) gene signatures and early recurrence. (iii) 44% of primary tumors changed the predominant subtype in the recurrence, preferably switching from IMS to BA or CL subtypes, respectively. (iv) The CL subtype of recurrent tumors showed a very low expression of a TUMOR INFLAMMATION signature, suggesting mechanisms of immune escape in

24

these recurrences. This was supported by significantly reduced intra-tumoral immune cell populations (dendritic cells) and PD-L1 scores in CL subtype tumors pointing to a strong link between DC cell infiltration and PD-L1 tumor expression in this particular subtype. Differential expression of immune markers was already reported in individual "Keck subtypes" correlating PD-L1 expression with an inflamed phenotype consistent with the IMS subtype (3). Simultaneously, a lack of immune markers for the BA subtype was observed. In our study, the co-occurrence of a low TUMOR INFLAMMATION signature and the CL subtype in recurrent tumors was very pronounced suggesting a combinatorial use of both markers for the prediction of tumor recurrence probability. Also, the reduced NK cell and plasmablast infiltration in the BA subtype is a striking finding of this study. A correlation of the BA subtype with enhanced P-EMT and EGFR signatures indicates a preferred route of tumor recurrence via a BA subtype/P-EMT axis. The P-EMT program includes upregulation of mesenchymal and restriction of epithelial genes, resembling EMT processes, yet in the absence of EMT-typical transcription factors. It is considered as a continuum of states recapitulating aspects of EMT and being linked to invasive properties, such as local invasion and nodal metastasis (9,41-43). Importantly, in HNSCC P-EMT has been exclusively associated to metastasis so far, yet its putative role in tumor recurrence is unknown. Also, EGFR was enhanced in BA subtypes which points to an EGF-mediated EMT in HNSCC (44). In our dataset an elevated P-EMT signature in the BA subtype was correlated with early tumor relapse. This was also noticed for HYPOXIA and RADIATION RESISTANCE signatures, while it remains unclear whether this is also related to the elevated P-EMT program or rather independent thereof. Additionally, recurrences had reduced PD1 SIGNALING and upregulated DNA REPAIR and METABOLISM PROCESSES compared to primary tumors. Since DNA repair capacity is strongly linked to radiation resistance, this was a conclusive finding in line with tumor relapse.

Interestingly, several DNA repair pathways were upregulated, including the gene and the associated base excision repair pathway. NTHL1 plays a pivotal role in cellular sensitivity towards clinical agents including ionizing radiation that induce double strand breaks and thereby has a strong influence on radiation sensitivity (45).

A main finding of our study was the frequent switch of the predominant subtype between primary and recurrent tumors. Therapeutic pressure apparently selects a therapy-resistant phenotype governed by a specific transcriptional subtype. It can be assumed that either intra-tumor heterogeneity of pre-existing subclones or transcriptional reprogramming are key drivers of such therapy-related phenotypes, however, this could not conclusively clarified in this study. Along these lines, it has been previously reported for glioblastoma that several transcriptional subtypes exist in most tumor samples and that subtype changes occur between primary recurrent tumors in about 50% of cases (46). Of note, also in the present study subtype switchers were observed in approximately 50% of the patients. This process was associated either with reduced expression of ANTI-TUMOR IMMUNITY response (47) or with an enhanced expression of P-EMT, suggesting increased tumor aggressiveness (9). A prolonged time to tumor relapse in patients with subtype changes points to time-consuming processes within tumor cells resulting in subtype-related resistance.

Here, we also generated a mutational landscape of matched primary/recurrent tumor pairs revealing dissimilarities for somatic mutation patterns. Various quality checks were performed on the FFPE-derived SNVs to exclude technical artefacts and to confirm that normal, primary tumor and recurrent tissues were from the same source. The observed genomic heterogeneity between tumor pairs is in line with the transcriptomic findings and with recent publications reporting that five out of ten primary and relapsed tumor pairs were not genetically related to each other (5,6).

With regard to driver mutations between primary and relapsed tumors in our cohort the best match was observed for TP53 with concordant results in 16/ 28 (57%) tumor pairs. In case of NOTCH1, only 2 /28 (7%) primary tumor/recurrent pairs showed agreement for this driver mutant. Previously, genes have been reported that were mutated in metastatic or recurrent tumors, but not in the corresponding primary tumors (6). This includes discoidin domain receptor tyrosine kinase 2 (DDR2/TKT) (6), which is mutated in about 2.2% of HNSCC cases according to cBioPortal for Cancer Genomics (https://www.cbioportal.org). This gene was neither mutated in our filtered set of somatic mutations nor in the unfiltered data. Another recent study on the comparison of HPV-positive primary and corresponding recurrent HNSCCs similarly observed that the HPV-related mutational landscape of primary tumors changed in recurrences to landscape features of HPV-unrelated HNSCCs (48). Looking at genomic disparities between tumor pairs or different tumor groups, we found a significant difference between different tumor stages for the SBS2 signature which is attributed to the activity of the AID/APOBEC family of cytidine deaminases. Members of these cytidine deaminases are recognized as a major cause for genomic mutations in many cancers (https://cancer.sanger.ac.uk/cosmic/signatures/SBS/). The increased prevalence of this signature in tumors of advanced stages is in line with a higher mutational burden in advanced HNSCCs.

A substantial proportion of primary/recurrent tumor pairs (21%) also significantly changed their CNV profiles. Genetically unrelated tumor pairs have also been found before in the study by deRoest et al. (5) who showed a higher proportion (80%) of unrelated tumor pairs, though in a smaller number of tumor pairs (n=10) than in the present study. The authors proposed CNV profiles as markers of genetic relationship due to their huge variations between independent tumors and between primary/recurrent tumor pairs. We observed this partially loose genetic relationship

between the tumor pairs was also on the transcriptomic level. The dominating transcriptional subtypes indicated a genuine development of recurrences based on clonal selection and evolutionary processes in tumors which was also postulated from a study on cellular populations in local recurrences from a surgical HNSCC xenograft model (19). Our additional analysis of clonal compositions of matched pairs of primary tumors and recurrences revealed up to five tumor subclones in some bulk samples. Indeed, the majority of tumor pairs showed signs of clonal expansion suggesting a possible clonal selection process during radiotherapy in those cases. However, in the remaining monoclonal cases molecular reprogramming might play a more important role. In our data, the estimated extent of subclonality was increased in recurrences, and the estimated average cellular fractions with subclonal alterations were smaller in recurrences compared to primary tumors, thus supporting this assumption. The cancer hallmark 'genomic instability', besides structural and small nucleotide variants, also manifests in the extent of CNVs. In our data, the estimate of chromosomal instability is reflected by the proportion of the genome affected by CNVs which was more pronounced in CL and BA subtypes compared to IMS. This suggests that the transcriptional subtypes are associated with – and presumably a consequence of - the extent of genomic instability. We further estimated tumor cellularity from the exome data, which were increased in recurrent (all biopsied) compared to primary tumors (n=4 biopsied, n=24 resected) – presumably due to differences in tissue sampling. However, we also observed a strikingly reduced cellularity of the IMS tumors, which is probably caused by immune-cell infiltrations in these tumors. In general, cellularity estimated from the WES data and that determined from histological slides were in good agreement which supports plausibility of the data. With regard to single CNVs none was associated with primary/recurrence status. 29 amplified regions, most of them clustering within the same regions on chromosomes 3, 12 and 20, were

associated with the CL subtype. The low extent of statistical association points to pronounced inter-tumor heterogeneity at the genomic level.

In summary, our study demonstrates genetic diversity and dynamics at the genomic and transcriptomic level in a cohort of paired primary and recurrent HNSCCs. Although this is the largest published cohort of tumor pairs explored in depth so far, there are certainly a number of limitations. Firstly, the sample size is small even though patients from two different clinics were included. Secondly, the biomaterial was very limited and only available as FFPE tissues, which provides DNA and RNA in limited yields and quality. Therefore, we performed serial sections and intense quality control measurements at multiple levels. Thirdly, variations in therapeutic treatments between patients required a strict definition for the selection of tumor recurrences related to the high-dose field of radiotherapy. Fourthly, somatic mutations reported previously might have been missed because a strict filtering scheme towards somatic mutations was applied to avoid any technical artifacts (sequencing depth of at least 40 reads, variant allele frequency of 10%, filtering for rare mutations).

Our reported key findings (genetically unrelated tumor pairs, subtype switchers, prevalence of the BA subtype in recurrences) raise the question about clinical translation. Most importantly, the observed heterogeneity within the same patient suggests that therapeutic decisions should be taken from molecular profiles of the recurrent, not the primary tumor. Additionally, HNSCC of the BA subtype may demand for harsher treatment and/or require a tighter clinical follow-up schedule. Also, alternative therapy of the BA subtype is a great challenge and of utmost importance since further dose escalation is not possible in recurrences, and immune checkpoint inhibition therapy is successful in only 20% of cases (49). Comparative GSEA maps between primary/relapsed tumors and between the BA and the other subtypes provided first hints for a therapeutic targeting

in BA subtype-driven recurrences. Finally, the major role of transcriptomic reprogramming in recurrent HNSCC far outweighs the information from genome analysis, the latter being the focus of current molecular tumor boards standard operating procedures. Transcriptomics for clinical use in cancer diagnostics are currently rare and neither part of clinical guidelines (50) nor covered by health insurance. Therefore, as a consequence and perspective of our study, a stratification of recurrences based on RNA sequencing and transcriptional subtype classification of recurrent HNSCC represents a future clinical need.

References:

1.   Cancer Genome Atlas Network. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* 2015;**517**(7536):576-82 doi 10.1038/nature14129.

2.   Mandal R, Senbabaoglu Y, Desrichard A, Havel JJ, Dalin MG, Riaz N, *et al.* The head and neck cancer immune landscape and its immunotherapeutic implications. *JCI Insight* 2016;**1**(17):e89829 doi 10.1172/jci.insight.89829.

3.   Keck MK, Zuo Z, Khattri A, Stricker TP, Brown CD, Imanguli M, *et al.* Integrative analysis of head and neck cancer identifies two biologically distinct HPV and three non-HPV subtypes. *Clin Cancer Res* 2015;**21**(4):870-81 doi 10.1158/1078-0432.CCR-14-2481.

4.   O'Sullivan B, Huang SH, Su J, Garden AS, Sturgis EM, Dahlstrom K, *et al.* Development and validation of a staging system for HPV-related oropharyngeal cancer by the International Collaboration on Oropharyngeal cancer Network for Staging (ICON-S): a multicentre cohort study. *Lancet Oncol* 2016;**17**(4):440-51 doi 10.1016/S1470-2045(15)00560-4.

5.   de Roest RH, Mes SW, Poell JB, Brink A, van de Wiel MA, Bloemena E, *et al.* Molecular Characterization of Locally Relapsed Head and Neck Cancer after Concomitant Chemoradiotherapy. *Clin Cancer Res* 2019;**25**(23):7256-65 doi 10.1158/1078-0432.CCR-19-0628.

6.   Hedberg ML, Goh G, Chiosea SI, Bauman JE, Freilino ML, Zeng Y, *et al.* Genetic landscape of metastatic and recurrent head and neck squamous cell carcinoma. *J Clin Invest* 2016;**126**(1):169-80 doi 10.1172/JCI82066.

7.   McGranahan N, Swanton C. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. *Cell* 2017;**168**(4):613-28 doi 10.1016/j.cell.2017.01.018.

8.   Mroz EA, Rocco JW. MATH, a novel measure of intratumor genetic heterogeneity, is high in poor-outcome classes of head and neck squamous cell carcinoma. *Oral Oncol* 2013;**49**(3):211-5 doi 10.1016/j.oraloncology.2012.09.007.

9.   Puram SV, Tirosh I, Parikh AS, Patel AP, Yizhak K, Gillespie S, *et al.* Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. *Cell* 2017;**171**(7):1611-24 e24 doi 10.1016/j.cell.2017.10.044.

10.   Rasmussen JH, Lelkaitis G, Hakansson K, Vogelius IR, Johannesen HH, Fischer BM, *et al.* Intratumor heterogeneity of PD-L1 expression in head and neck squamous cell carcinoma. *Br J Cancer* 2019;**120**(10):1003-6 doi 10.1038/s41416-019-0449-y.

11.   Huang C, Chen L, Savage SR, Eguez RV, Dou Y, Li Y, *et al.* Proteogenomic insights into the biology and treatment of HPV-negative head and neck squamous cell carcinoma. *Cancer Cell* 2021 doi 10.1016/j.ccell.2020.12.007.

12.   Tanay A, Regev A. Scaling single-cell genomics from phenomenology to mechanism. *Nature* 2017;**541**(7637):331-8 doi 10.1038/nature21350.

13.   Nordsmark M, Bentzen SM, Rudat V, Brizel D, Lartigau E, Stadler P, *et al.* Prognostic value of tumor oxygenation in 397 head and neck tumors after primary radiation therapy. An international multi-center study. *Radiother Oncol* 2005;**77**(1):18-24 doi 10.1016/j.radonc.2005.06.038.

14.   Harris AL. Hypoxia--a key regulatory factor in tumour growth. *Nat Rev Cancer* 2002;**2**(1):38-47 doi 10.1038/nrc704.

15.   Hutchinson MND, Mierzwa M, D'Silva NJ. Radiation resistance in head and neck squamous cell carcinoma: dire need for an appropriate sensitizer. *Oncogene* 2020;**39**(18):3638-49 doi 10.1038/s41388-020-1250-3.

16.   Braakhuis BJ, Tabor MP, Leemans CR, van der Waal I, Snow GB, Brakenhoff RH. Second primary tumors and field cancerization in oral and oropharyngeal cancer: molecular techniques provide new insights and definitions. *Head Neck* 2002;**24**(2):198-206 doi 10.1002/hed.10042.

17.   van Houten VM, Leemans CR, Kummer JA, Dijkstra J, Kuik DJ, van den Brekel MW, *et al.* Molecular diagnosis of surgical margins and local recurrence in head and neck cancer patients: a prospective study. *Clin Cancer Res* 2004;**10**(11):3614-20 doi 10.1158/1078-0432.CCR-03-0631.

18.   Curtius K, Wright NA, Graham TA. An evolutionary perspective on field cancerization. *Nat Rev Cancer* 2018;**18**(1):19-32 doi 10.1038/nrc.2017.102.

19.   Roh V, Abramowski P, Hiou-Feige A, Cornils K, Rivals JP, Zougman A*, et al.* Cellular Barcoding Identifies Clonal Substitution as a Hallmark of Local Recurrence in a Surgical Model of Head and Neck Squamous Cell Carcinoma. *Cell Rep* 2018;**25**(8):2208-22 e7 doi 10.1016/j.celrep.2018.10.090.

20.   Rodon J, Soria JC, Berger R, Miller WH, Rubin E, Kugel A*, et al.* Genomic and transcriptomic profiling expands precision cancer medicine: the WINTHER trial. *Nat Med* 2019;**25**(5):751-8 doi 10.1038/s41591-019-0424-4.

21.   Le Tourneau C, Delord JP, Goncalves A, Gavoille C, Dubot C, Isambert N*, et al.* Molecularly targeted therapy based on tumour molecular profiling versus conventional therapy for advanced cancer (SHIVA): a multicentre, open-label, proof-of-concept, randomised, controlled phase 2 trial. *Lancet Oncol* 2015;**16**(13):1324-34 doi 10.1016/S1470-2045(15)00188-6.

22.   Mateo J, Chakravarty D, Dienstmann R, Jezdic S, Gonzalez-Perez A, Lopez-Bigas N*, et al.* A framework to rank genomic alterations as targets for cancer precision medicine: the ESMO Scale for Clinical Actionability of molecular Targets (ESCAT). *Ann Oncol* 2018;**29**(9):1895-902 doi 10.1093/annonc/mdy263.

23.   Moreira A, Masliah-Planchon J, Callens C, Vacher S, Lecerf C, Frelaut M*, et al.* Efficacy of molecularly targeted agents given in the randomised trial SHIVA01 according to the ESMO Scale for Clinical Actionability of molecular Targets. *Eur J Cancer* 2019;**121**:202-9 doi 10.1016/j.ejca.2019.09.001.

24.   Wintergerst L, Selmansberger M, Maihoefer C, Schuttrumpf L, Walch A, Wilke C*, et al.* A prognostic mRNA expression signature of four 16q24.3 genes in radio(chemo)therapy-treated head and neck squamous cell carcinoma (HNSCC). *Mol Oncol* 2018;**12**(12):2085-101 doi 10.1002/1878-0261.12388.

25.   Hess J, Unger K, Maihoefer C, Schuttrumpf L, Wintergerst L, Heider T*, et al.* A Five-MicroRNA Signature Predicts Survival and Disease Control of Patients with Head and Neck Cancer Negative for HPV Infection. *Clin Cancer Res* 2019;**25**(5):1505-16 doi 10.1158/1078-0432.CCR-18-0776.

26.   Maihoefer C, Schuttrumpf L, Macht C, Pflugradt U, Hess J, Schneider L*, et al.* Postoperative (chemo) radiation in patients with squamous cell cancers of the head and neck - clinical results from the cohort of the clinical cooperation group "Personalized Radiotherapy in Head and Neck Cancer". *Radiat Oncol* 2018;**13**(1):123 doi 10.1186/s13014-018-1067-1.

27.   Wilkerson MD, Yin X, Hoadley KA, Liu Y, Hayward MC, Cabanski CR*, et al.* Lung squamous cell carcinoma mRNA expression subtypes are reproducible, clinically important, and correspond to normal cell types. *Clin Cancer Res* 2010;**16**(19):4864-75 doi 10.1158/1078-0432.CCR-10-0199.

28.   Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J*, et al.* PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* 2003;**34**(3):267-73 doi 10.1038/ng1180.

29.   Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA*, et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005;**102**(43):15545-50 doi 10.1073/pnas.0506580102.

30.   Therneau TM, Grambsch PM. Modeling Survival Data: Extending the Cox Model. Springer, New York; 2000.

31.   Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 2016;**32**(18):2847-9 doi 10.1093/bioinformatics/btw313.

32.   Ayers M, Lunceford J, Nebozhyn M, Murphy E, Loboda A, Kaufman DR*, et al.* IFN-gamma-related mRNA profile predicts clinical response to PD-1 blockade. *J Clin Invest* 2017;**127**(8):2930-40 doi 10.1172/JCI91190.

33.   Chen J, Fu G, Chen Y, Zhu G, Wang Z. Gene-expression signature predicts survival benefit from postoperative chemoradiotherapy in head and neck squamous cell carcinoma. *Oncol Lett* 2018;**16**(2):2565-78 doi 10.3892/ol.2018.8964.

34. Foy JP, Bazire L, Ortiz-Cuaran S, Deneuve S, Kielbassa J, Thomas E, *et al.* A 13-gene expression-based radioresistance score highlights the heterogeneity in the response to radiation therapy across HPV-negative HNSCC molecular subtypes. *BMC Med* 2017;**15**(1):165 doi 10.1186/s12916-017-0929-y.

35. Winter SC, Buffa FM, Silva P, Miller C, Valentine HR, Turley H, *et al.* Relation of a hypoxia metagene derived from head and neck cancer to prognosis of multiple cancers. *Cancer Res* 2007;**67**(7):3441-9 doi 10.1158/0008-5472.CAN-06-3322.

36. Cowen L, Ideker T, Raphael BJ, Sharan R. Network propagation: a universal amplifier of genetic associations. *Nat Rev Genet* 2017;**18**(9):551-62 doi 10.1038/nrg.2017.38.

37. Burrell RA, McGranahan N, Bartek J, Swanton C. The causes and consequences of genetic heterogeneity in cancer evolution. *Nature* 2013;**501**(7467):338-45 doi 10.1038/nature12625.

38. Ding L, Ley TJ, Larson DE, Miller CA, Koboldt DC, Welch JS, *et al.* Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* 2012;**481**(7382):506-10 doi 10.1038/nature10738.

39. Klein CA. Selection and adaptation during metastatic cancer progression. *Nature* 2013;**501**(7467):365-72 doi 10.1038/nature12628.

40. Bossi P, Bergamini C, Siano M, Cossu Rocca M, Sponghini AP, Favales F, *et al.* Functional Genomics Uncover the Biology behind the Responsiveness of Head and Neck Squamous Cell Cancer Patients to Cetuximab. *Clin Cancer Res* 2016;**22**(15):3961-70 doi 10.1158/1078-0432.CCR-15-2547.

41. Lambert AW, Pattabiraman DR, Weinberg RA. Emerging Biological Principles of Metastasis. *Cell* 2017;**168**(4):670-91 doi 10.1016/j.cell.2016.11.037.

42. Nieto MA, Huang RY, Jackson RA, Thiery JP. Emt: 2016. *Cell* 2016;**166**(1):21-45 doi 10.1016/j.cell.2016.06.028.

43. Baumeister P, Hollmann A, Kitz J, Afthonidou A, Simon F, Shakhtour J, *et al.* High Expression of EpCAM and Sox2 is a Positive Prognosticator of Clinical Outcome for Head and Neck Carcinoma. *Sci Rep* 2018;**8**(1):14582 doi 10.1038/s41598-018-32178-8.

44. Pan M, Schinke H, Luxenburger E, Kranz G, Shakhtour J, Libl D, *et al.* EpCAM ectodomain EpEX is a ligand of EGFR that counteracts EGF-mediated epithelial-mesenchymal transition through modulation of phospho-ERK1/2 in head and neck cancers. *PLoS Biol* 2018;**16**(9):e2006624 doi 10.1371/journal.pbio.2006624.

45. Limpose KL, Trego KS, Li Z, Leung SW, Sarker AH, Shah JA, *et al.* Overexpression of the base excision repair NTHL1 glycosylase causes genomic instability and early cellular hallmarks of cancer. *Nucleic Acids Res* 2018;**46**(9):4515-32 doi 10.1093/nar/gky162.

46. Klughammer J, Kiesel B, Roetzer T, Fortelny N, Nemc A, Nenning KH, *et al.* The DNA methylation landscape of glioblastoma disease progression shows extensive heterogeneity in time and space. *Nat Med* 2018;**24**(10):1611-24 doi 10.1038/s41591-018-0156-x.

47. Watermann C, Pasternack H, Idel C, Ribbat-Idel J, Bragelmann J, Kuppler P, *et al.* Recurrent HNSCC Harbor an Immunosuppressive Tumor Immune Microenvironment Suggesting Successful Tumor Immune Evasion. *Clin Cancer Res* 2020 doi 10.1158/1078-0432.CCR-20-0197.

48. Harbison RA, Kubik M, Konnick EQ, Zhang Q, Lee SG, Park H, *et al.* The mutational landscape of recurrent versus nonrecurrent human papillomavirus-related oropharyngeal cancer. *JCI Insight* 2018;**3**(14) doi 10.1172/jci.insight.99327.

49. Burtness B, Harrington KJ, Greil R, Soulieres D, Tahara M, de Castro G, Jr., *et al.* Pembrolizumab alone or with chemotherapy versus cetuximab with chemotherapy for recurrent or metastatic squamous cell carcinoma of the head and neck (KEYNOTE-048): a randomised, open-label, phase 3 study. *Lancet* 2019;**394**(10212):1915-28 doi 10.1016/S0140-6736(19)32591-7.

50. Carlson RW, Larsen JK, McClure J, Fitzgerald CL, Venook AP, Benson AB, 3rd, *et al.* International adaptations of NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw* 2014;**12**(5):643-8 doi 10.6004/jnccn.2014.0068.

Author contributions: CB, BW, HZ designed the study; JH, SM, CI, CM, PB, MC established and annotated the cohort; UG, SM, CM, CB, DR, collected and evaluated treatment schemes; JR-I, SR, LV, CW, AW performed histopathological evaluation and tissue selection; KU, AK performed DNA and copy number profiling; JH, PW performed low-coverage whole-genome and RNA-seq analysis; PW, KU, AK performed data analysis; HB, OG, KL, TH, LK, MBD, SH contributed to the data analysis; OK, GD, KL performed primary cultures of HNSCC samples; MS, HL, KU, PW performed single cell RNA-seq analysis; BW, TK, SP, CI, DR, CB contributed tumor samples and clinical data; HZ, KL, PW, OG, KU, AK, HB, JH wrote the manuscript with contributions of all authors.

Table 1: Patient data of tumor pairs (primary and recurrent tumors)

Patient			Primary Tumor							Recurrent Tumor			
Patient ID	Age [1]	Sex [2]	Site [3]	TNM stage [4]	UICC stage [5]	Radiotherapy dose (Gy) /treatment [6]	Chemotherapy	Predominant Transcriptional Subtype [7]	PD-L1 [8]	Type [9]	Time to Relapse (days) [10]	Predominant Transcriptional Subtype [7]	PD-L1 [8]
P01	48	m	OC	pT2N1M0	III	64/adj. EBRT	no	IMS	75	LR	201	BA	75
P02	68	f	HP	pT1N2bM0	IVa	64/adj. EBRT	Cisplatin/5-FU	IMS	3	LRR	215	IMS	20
P03	51	m	OC	pT3N0M0	III	64/adj. EBRT	no	BA	25	LR	204	CL	10
P04	60	m	OP	cT4N0M0	IVa	70/adj. EBRT	5-FU/MMC	BA	15	LR	416	BA	20
P05	58	f	OC	pT2N0M0	II	64/adj. EBRT	no	BA	15	LRR	315	BA	5
P06	60	m	LA	cT4N1M0	IVa	70/def. EBRT	5-FU/MMC	IMS	30	LR	779	CL	1
P07	55	m	OP	cT4N2bM0	IVa	69.96/def. EBRT	Cisplatin weekly	BA	10	LR	254	BA	0
P08	47	m	OC	pT2N1M0	III	64/adj. EBRT	no	BA	20	LR	199	BA	20
P09	69	f	HP	cT3N1M0	III	69.96/def. EBRT	Cetuximab mono	IMS	10	LR	461	IMS	20
P10	47	m	OC	pT2N0M0	II	64/adj. EBRT	no	IMS	35	LR	236	BA	10
P11	64	m	OP	cT4N2bM0	IVa	70/def. EBRT	MMC mono	IMS	25	LR	399	BA	15
P12	60	m	HP	pT4N2cM0	IVa	64/adj. EBRT	Cisplatin/5-FU	CL	0	LRR	323	IMS	25
P13	55	m	OC	pT4a/N2cM0	IVa	64/adj. EBRT	Cisplatin/5-FU	BA	20	LRR	303	BA	85
P14	60	m	OP	pT3N2bM0	IVa	64/adj. EBRT	Cisplatin/5-FU	BA	20	LR	233	BA	25
P15	53	f	OP	pT1N1M0	III	64/adj. EBRT	Cisplatin/5-FU	IMS	85	LRR	492	BA	10
P16	66	m	OP	cT4N2bM0	IVa	70/def. EBRT	Cetuximab mono	BA	10	LR	360	BA	20
P17	59	m	LA	pT2N0M0	II	73.5/adj. EBRT	no	IMS	1	LR	439	BA	0
P18	58	m	OP	pT4aN0M0	IVa	67.2/adj. EBRT	no	BA	2	LR	237	BA	33
P19	63	f	OC	pT2N2cM0	IVa	67.8/adj. EBRT	no	BA	33	LR	112	BA	63
P20	53	m	LA	pT2N0M0	II	62.4/adj. HDBT	no	BA	100	LR	97	BA	30
P21	71	m	OP	pT4aN2cM0	IVa	70.29/adj. EBRT	Cisplatin/Carboplatin weekly	BA	0	LRR	231	BA	0
P22	48	m	OC	pT3N2cM0	IVa	75.6/adj. EBRT	no	IMS	0	LRR	362	CL	0
P23	68	f	OP	pT2N2cM0	IVa	70.4/adj. EBRT/HDBT	no	IMS	67	LRR	333	IMS	67
P24	66	m	LA	pT1N2bM0	IVa	64.8/adj. EBRT	Cetuximab mono	IMS	1	LR	370	CL	0
P25	54	m	LA	pT2N2aM0	IVa	69.63/adj. EBRT	Cisplatin weekly	CL	0	LR	590	CL	0
P26	70	m	OC	pT4aN1M0	IVa	69.3/adj. EBRT	no	BA	67	LRR	163	BA	10
P27	78	m	HP	pT3N2bM0	IVa	75.5/adj. EBRT/HDBT	no	IMS	83	LR	359	IMS	0
P28	55	m	OC	cT3N2bM0	IVa	73.95/def. EBRT	Docetaxel/Cisplatin	IMS	3	LR	1146	BA	40
P29	59	m	OP	pT3N1M0	III	69.96/adj. EBRT	no	BA	93	LR	368	BA	70
P30	47	f	OP	cT4aN2cM0	IVa	80.5/def. EBRT/HDBT	Cisplatin weekly	IMS	0	LR	415	BA	0
P31	61	m	OP	pT4aN0M0	IVa	66/adj. EBRT	Cisplatin weekly	IMS	87	LR	787	BA	90
P32	47	m	OP	pT1N2bM0	IVa	68.7/adj. EBRT/HDBT	no	BA	1	LR	274	BA	0
P33	55	m	OC	cT1N2bM0	IVa	74/def. EBRT/HDBT	Cisplatin weekly	BA	16	LR	374	IMS	0
P34	58	f	LA	pT3N2bM0	IVa	64/adj. EBRT	Cisplatin/5-FU	CL	0	LR	147	BA	0
P35	87	m	LA	pT4aN0M0	IVa	64/adj. EBRT	no	NA	NA	LR	671	CL	NA
P36	55	m	LA	pT3N0M0	III	64/adj. EBRT	no	BA	NA	LR	138	NA	NA
P37	74	m	OP	pT2N0M0	II	64/adj. EBRT	no	NA	NA	LRR	239	NA	NA
P38	63	m	OC	pT2N2bM0	IVa	64/adj. EBRT	MMC mono	NA	NA	LR	182	BA	NA

[1] years; [2] m=male, f=female; [3] LA=larynx, OP=oropharynx, OC=oral cavity, HP=hypopharynx; [4] TNM according to UICC TNM Classification of Malignant Tumors, 7th edition; [5] UICC stage [6] EBRT=external beam radiation therapy, HDBT=high-dose-rate brachytherapy, adj.=adjuvant, def.=definitive; [7] Subtype according to Keck et al. (2015), IMS=inflamed mesenchymal, BA=basal, CL=classical; [8] Tumor proportion score (TPS) %; [9] LR=local recurrence, LRR=loco-regional recurrence; [10] Time to LR or LRR relapse in days calculated from the start of radiotherapy treatment.
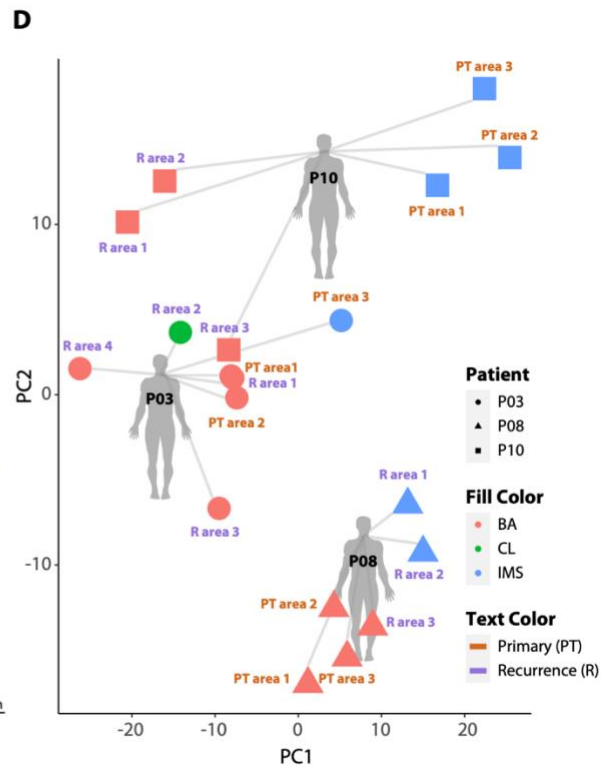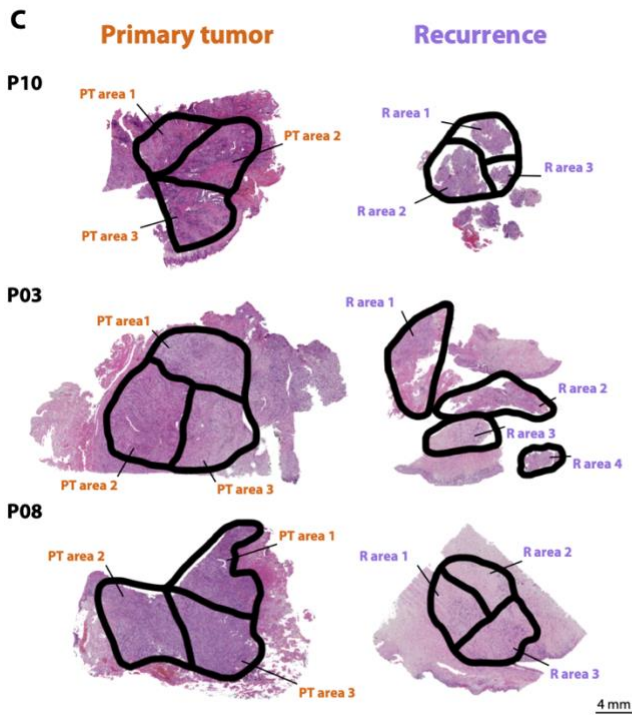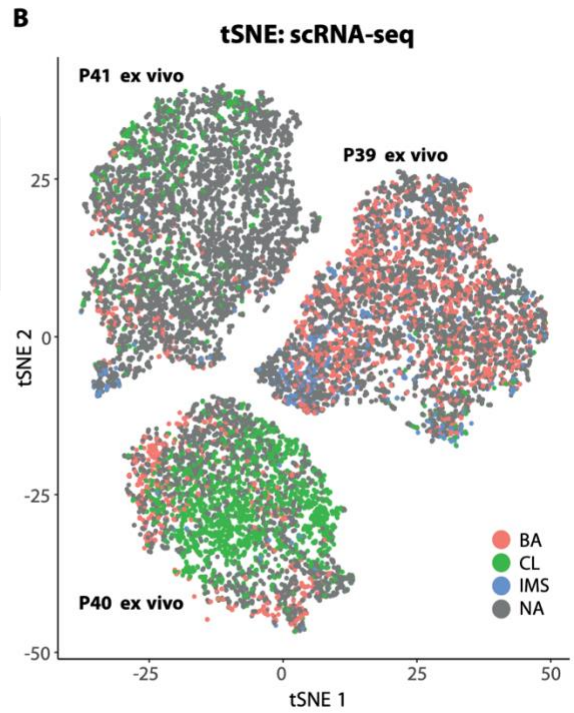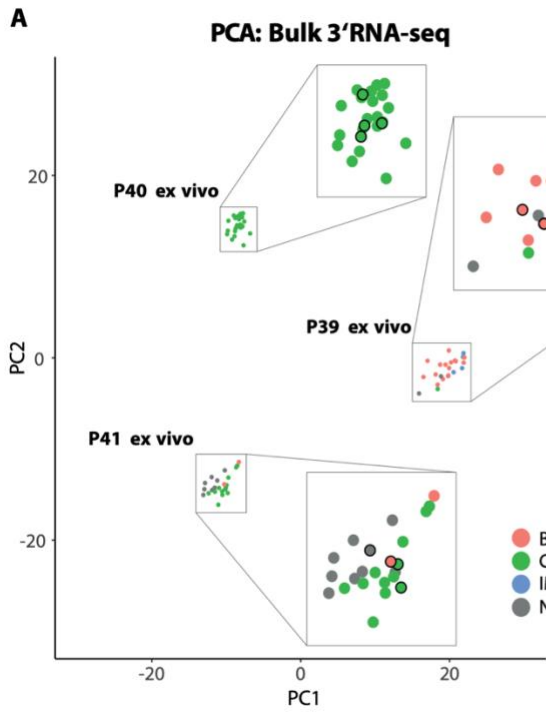
A PCA: Bulk 3'RNA-seq

P40 ex vivo

P39 ex vivo

P41 ex vivo

BA
CL
IMS
NA

PC2

PC1

B tSNE: scRNA-seq

P41 ex vivo

P39 ex vivo

P40 ex vivo

BA
CL
IMS
NA

tSNE 2

tSNE 1

C

Primary tumor **Recurrence**

P10

PT area 1
PT area 2
PT area 3

R area 1
R area 3
R area 2

P03

PT area1
PT area 2
PT area 3

R area 1
R area 2
R area 3
R area 4

P08

PT area 1
PT area 2
PT area 3

R area 1
R area 2
R area 3

4 mm

D

PT area 3
PT area 2
PT area 1

R area 2
R area 1

P10

R area 2
R area 3
PT area 3

R area 4
PT area1
R area 1
PT area 2

P03

R area 3

R area 1
R area 2

P08

PT area 2
R area 3
PT area 1
PT area 3

PC2

PC1

Patient
P03
P08
P10

Fill Color
BA
CL
IMS

Text Color
Primary (PT)
Recurrence (R)

Figure 1: Transcriptomic tumor heterogeneity of three primary HNSCC *ex vivo* cultures and different tumor areas of matched primary and recurrent HNSCC tissue specimens. 3'RNA and single-cell mRNA sequencing (scRNA-seq) were performed on *ex vivo* cultures of three HNSCC cases (patients P39_ex_vivo, P40_ex_vivo and P41_ex_vivo). To assess possible heterogeneity within tumors we used the subtype defining genes for dimensionality reduction and visualized the first two components. Thus, spatial distance of symbols in (**A,B,D**) corresponds to difference in characteristics of the transcriptional subtype. 3'RNA sequencing of bulk mRNAs (**A**; PCA, principal component analysis; PC1, PC2: Principal component 1 or 2) revealed similar results than scRNA-seq (**B**; tSNE plot for individual tumor cells) for the predominant transcriptional subtype in each case (color code: red for BA=basal, green for CL=classical, blue for IMS=inflamed-mesenchymal subtypes, grey for NA=not applicable). **(A)** Individual patient derived cultures are shown magnified on the right, with the bulk samples outlined in black for distinction from the derived subclones. Both sequencing approaches revealed a distinct heterogeneity of transcriptional subtypes either for different subclones of the bulk analysis (**A**) or individual tumor cells (**B**). Different tumor areas (n=19) from three selected matched primary and recurrent HNSCC tissue specimens were microscopically defined. Neither intra- (between PT and R) nor inter-tumoral (between the individual defined areas) histomorphological heterogeneity was visible **(C).** 3'RNA-seq analysis revealed (**D**; PCA) heterogeneity of assigned transcriptional subtypes in different tumor areas of one FFPE tumor block for one primary tumor and two recurrent tumor specimens along with a stability of the predominant subtype.
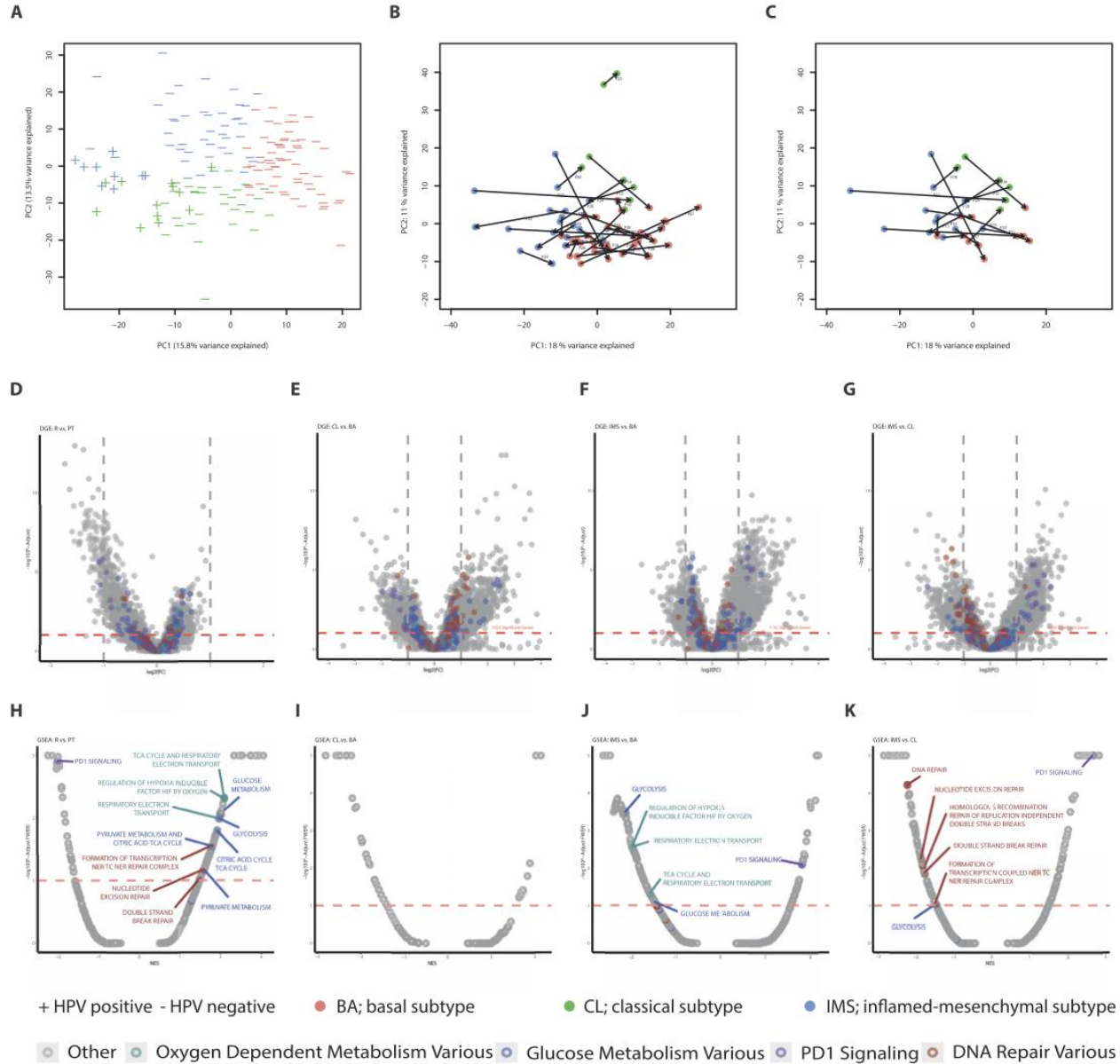
Figure 2: Transcriptional subtypes of primary and recurrent tumors. mRNA sequencing was performed on matched primary/recurrent tumor pairs and on several cohorts of primary tumors (n=126 patients/160 samples). (**A**). Principal component analysis (PCA) of predominant transcriptional subtypes for all samples, showing a clustering of individual subtypes (BA=basal n=66 samples, CL=classical n=42 samples, IMS=inflamed-mesenchymal n=52 samples) and of

HPV[+]-related subtypes (CL n=12 samples, IMS n=9 samples). (**B**). PCA of predominant transcriptional subtypes of primary and recurrent tumor pairs (n=34 patients/68 samples) indicating either low or high correspondence which is apparent from the bar lengths linking primary and recurrent tumor pairs. (**C**). Same as (B), but depicting only the 15 tumor pairs changing the transcriptional subtype between primary tumor and recurrence. 19 tumor pairs maintained the initial subtype (subtypes between primary and relapsed tumor: BA-BA n=14, CL-CL n=1, IMS-IMS n=4; IMS-BA n=8, IMS-CL n=3, CL-BA n=1, BA-CL n=1, CL-IMS n=1). (**D, H**) Differential gene expression (DGE) analysis and gene set enrichment analysis (GSEA) between primary and recurrent tumor pairs (n= 34 patients/68 samples). DGE (volcano plot) shows differentially expressed genes between primary and recurrent tumors (**D**), GSEA reveals a diminished PD1 SIGNALING as well as enriched DNA REPAIR and METABOLIC processes in relapses (**H**). (**E, I**) Differentially expressed genes between CL and BA subtypes (volcano plot, **E**), gene sets mainly for DNA REPAIR and METABOLIC processes are differentially expressed between CL and BA subtypes (GSEA, **I**). (**F, J**) DGE shows differentially expressed genes between IMS and BA subtypes (volcano plot, **F**), GSEA reveals enriched PD1 SIGNALING and diminished GLUCOSE METABOLISM in IMS (**J**). (**G, K**) Differentially expressed genes between IMS and CL subtypes (volcano plot, **G**), showing an enriched PD1 SIGNALING and diminished DNA REPAIR processes in IMS (GSEA, **K**). NES: enrichment score normalized to mean enrichment of random samples of the same size; log2(FC): log2 fold change of gene expression; DGE and GSEA are represented by plots (**D** – **K**) in which significantly different expression values are indicated above the red dashed line representing an adjusted p-value (FDR) <0.1. Log2-fold changes of gene expressions <2 and >2 are indicated by perpendicular dashed lines within the volcano plots (**D** – **G**)
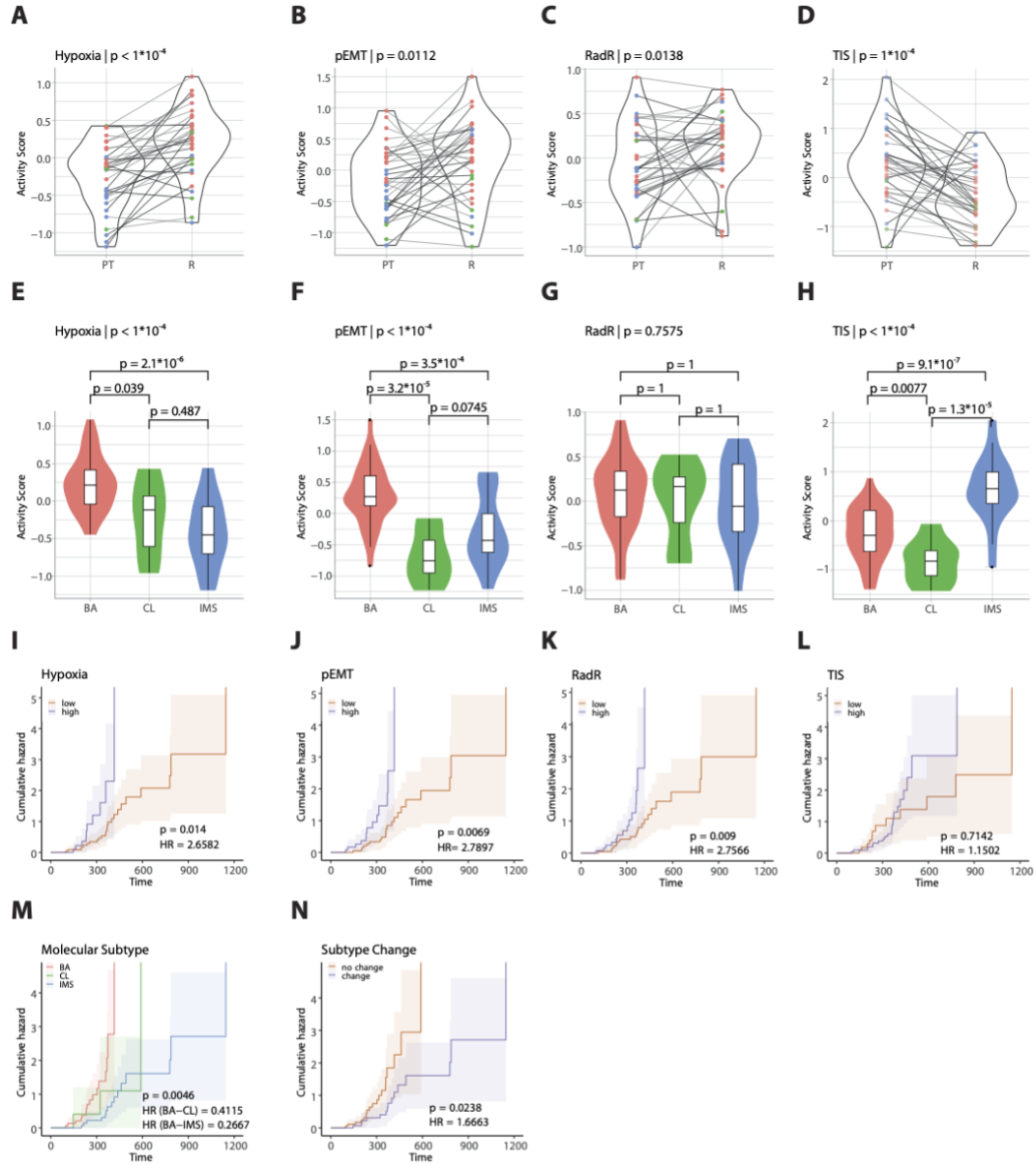
Figure 3: Association of transcriptional subtypes with gene signatures. Pairwise comparisons between primary and recurrent tumor pairs were performed for activity scores of prognostic signatures for hypoxia (n=34 patients/68 samples) (**A**) HYPOXIA, (**B**) P-EMT, (**C**) RADIATION RESISTANCE, (**D**) TUMOR INFLAMMATION. P-values < 0.05 indicate significantly different activity scores between tumor pairs. The activity scores of the same prognostic signatures were compared between basal (BA), classical (CL) and inflamed-mesenchymal (IMS) subtypes; significant differences between individual subtypes were indicated by p < 0.05 (**E-H**). Kaplan-Meier analysis depicting the development of tumor relapses dependent on time (days) and activity scores for each prognostic signature (**I-L**). Kaplan-Meier analysis (loco-regional recurrence) also demonstrates significantly different latencies for transcriptional subtypes (n=34 primary tumors; BA n=16, CL n=3, IMS n=15) (**M**) and a longer latency for subtype switcher (primary tumors with subtype change n=15; without subtype change n=19) (**N**).
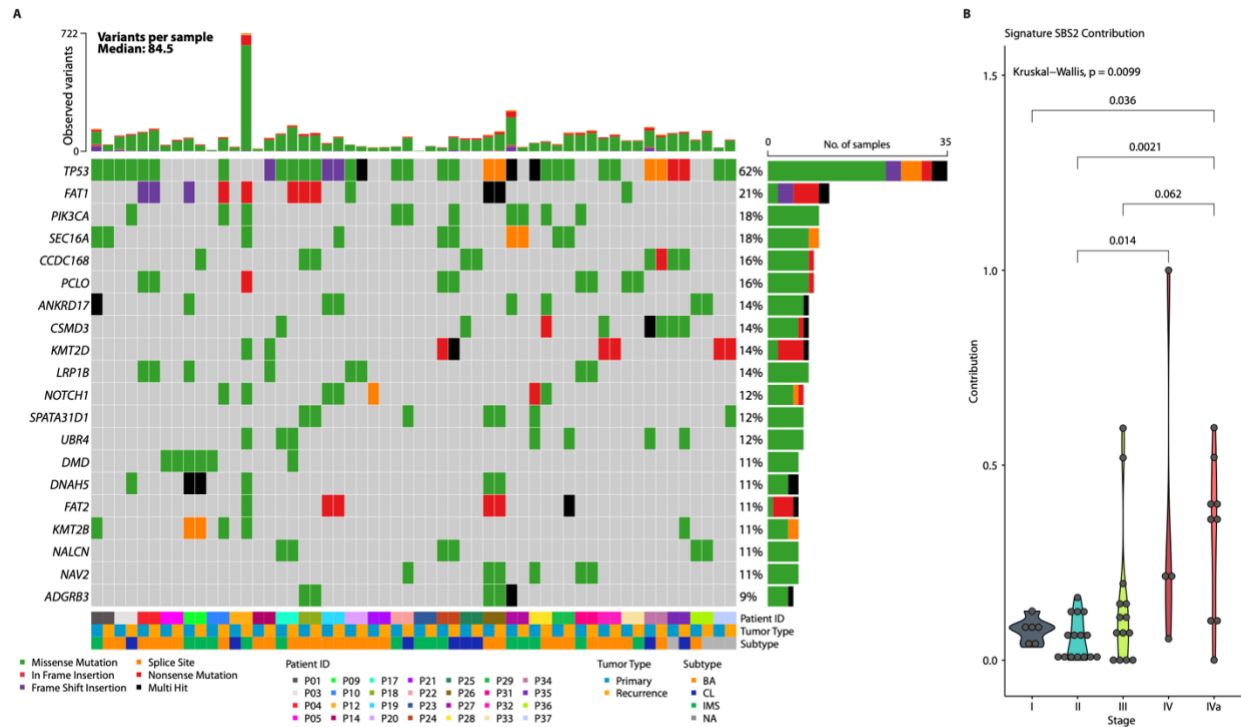
Figure 4: Landscape of genomic mutations, mutational signatures, and heterogeneity among samples. (**A**) Genomic mutations of HNSCC for matched tumor and recurrence pairs and number of variants per sample. Each column represents a sample and the most frequent DNA mutations after removing the top 20 frequently mutated genes in public exomes. Primary tumor and corresponding recurrence are shown next to each other for each sample. Additionally, subtype information inferred from RNA-seq data is shown for samples where RNA-seq data was available. Color codes at the bottom indicate patient-IDs, tumor type (index tumor and recurrence per patient next to each other), variant classification and transcriptional subtype. (**B**) Single Base Substitution (SBS) signatures in primary and recurrent tumor pairs for SBS2. The SBS2 contribution in UICC tumor stages I – IV is indicated.
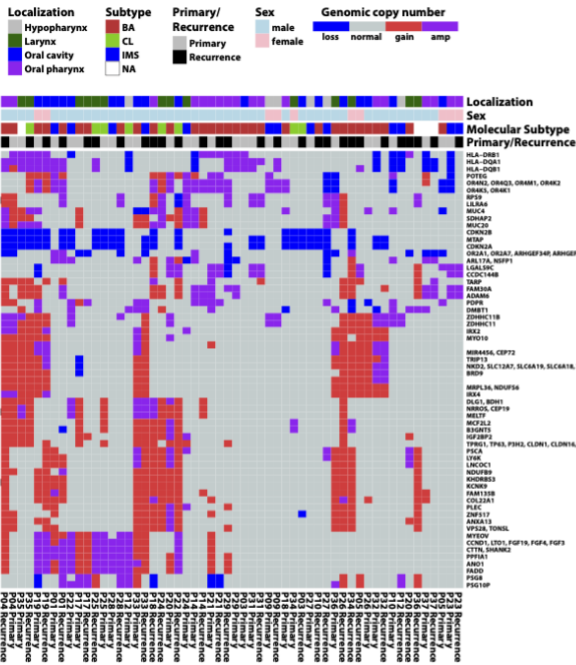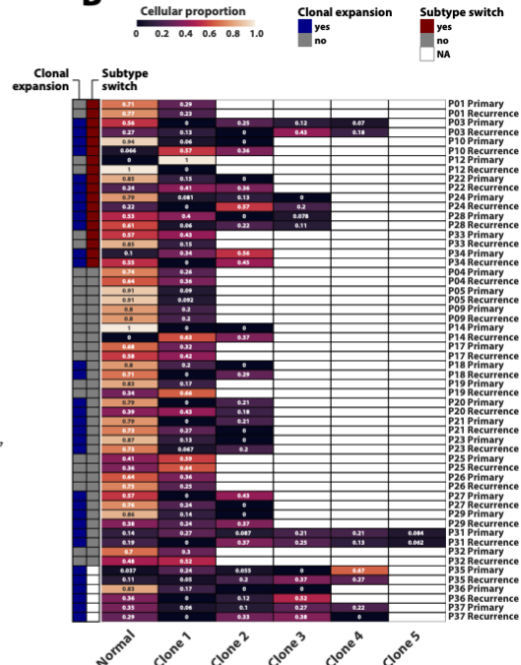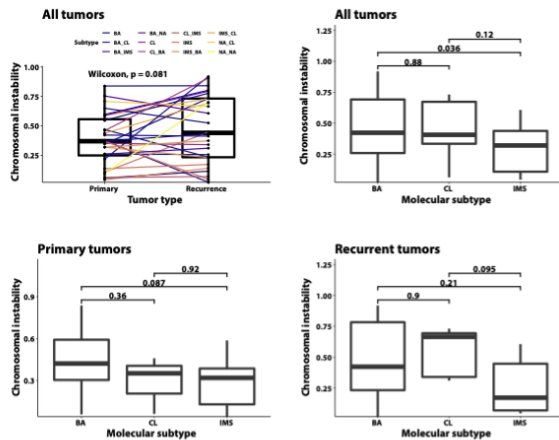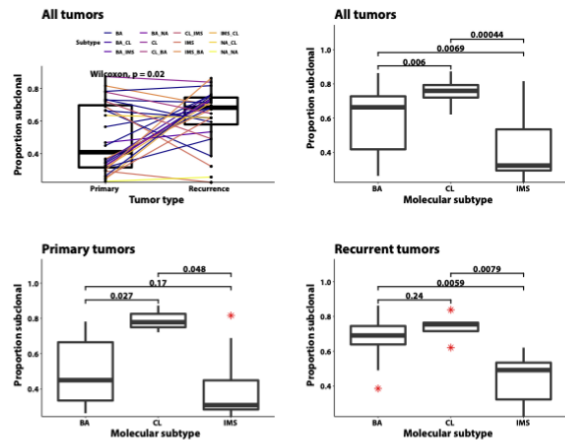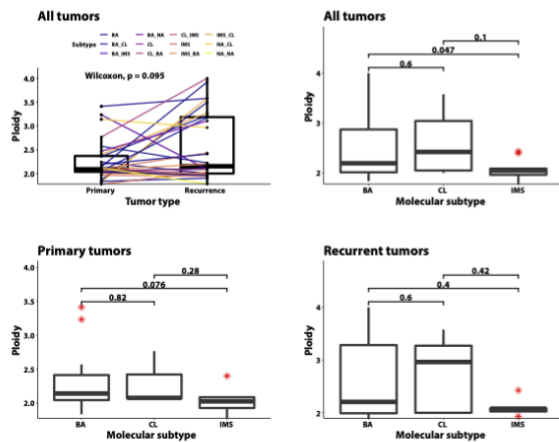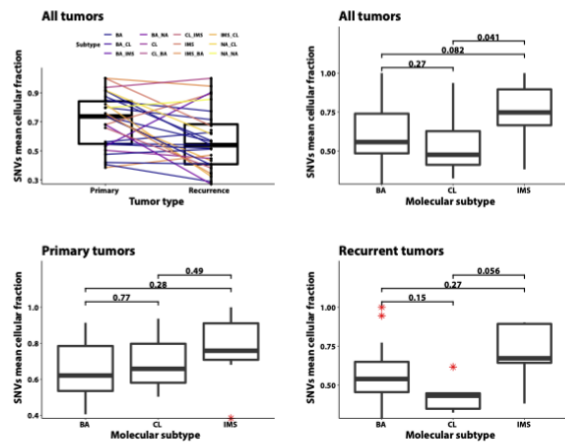
Figure 5: Genomic copy number variations (CNV) in primary and recurrent tumors and analysis of heterogeneity. (**A**) Heatmap and hierarchical cluster analysis of CNVs detected from whole exome sequencing data in 28 primary/recurrent tumor pairs. CNVs in different copy number regions consisting of one or more genes were indicated (blue: deletion, red: amplification, purple: focal amplifications). The main clusters after unsupervised hierarchical clustering were not associated with tumor site, sex, transcriptional subtype or primary/recurrent tumor type. (**B**) Heatmap with clonal compositions of matched pairs of primary tumors and recurrences. Given are the proportions of normal cells and up to five tumor subclones of the bulks. For each tumor-recurrence pair, the presence of clonal expansion probability is noted in addition to transcriptional subtype change. (**C**) Chromosomal instability for all primary and recurrent tumors according to the transcriptional subtypes. Chromosomal instability remains similar between primary tumors and recurrences, but is significantly decreased in IMS subtype. (**D**) Mean proportions for all primary and recurrent tumors according to the transcriptional subtypes. Recurrences show a higher mean proportion of subclonal variants compared to primary tumors with CL and IMS having the highest and lowest proportion of subclonal variants. (**E**) Ploidy for all primary and recurrent tumors according to the molecular subtypes. Ploidy remain unchanged between primary and recurrent tumors but are lowest in IMS subtype. (**F**) Size of the cellular fractions for all primary and recurrent tumors according to the molecular subtypes. The sizes are opposite to the proportions of subclones and was lowest in recurrences and highest in the IMS subtype.