

## Article

# An Integrated Epigenomic and Genomic View on Phyllodes and Phyllodes-Like Breast Tumors

Juergen Hench <sup>1</sup>, Tatjana Vlajnic <sup>1</sup>, Savas Deniz Soysal <sup>2,3</sup>, Ellen C Obermann <sup>4</sup>, Stephan Frank <sup>1</sup>, Simone Muenst <sup>1,\*</sup>

<sup>1</sup> Institute of Medical Genetics and Pathology, University Hospital Basel, Basel, Switzerland;

[juergen.hench@usb.ch](mailto:juergen.hench@usb.ch); [tatjana.vlajnic@usb.ch](mailto:tatjana.vlajnic@usb.ch); [stephan.frank@usb.ch](mailto:stephan.frank@usb.ch); [simone.muenst@usb.ch](mailto:simone.muenst@usb.ch)

<sup>2</sup> Visceral Surgery Research Laboratory, Clarunis, Department of Biomedicine, University of Basel, Basel, Switzerland; [savas.soysal@clarunis.ch](mailto:savas.soysal@clarunis.ch)

<sup>3</sup> Department of Surgery, Clarunis University Center for Gastrointestinal and Liver Diseases Basel, Basel, Switzerland;

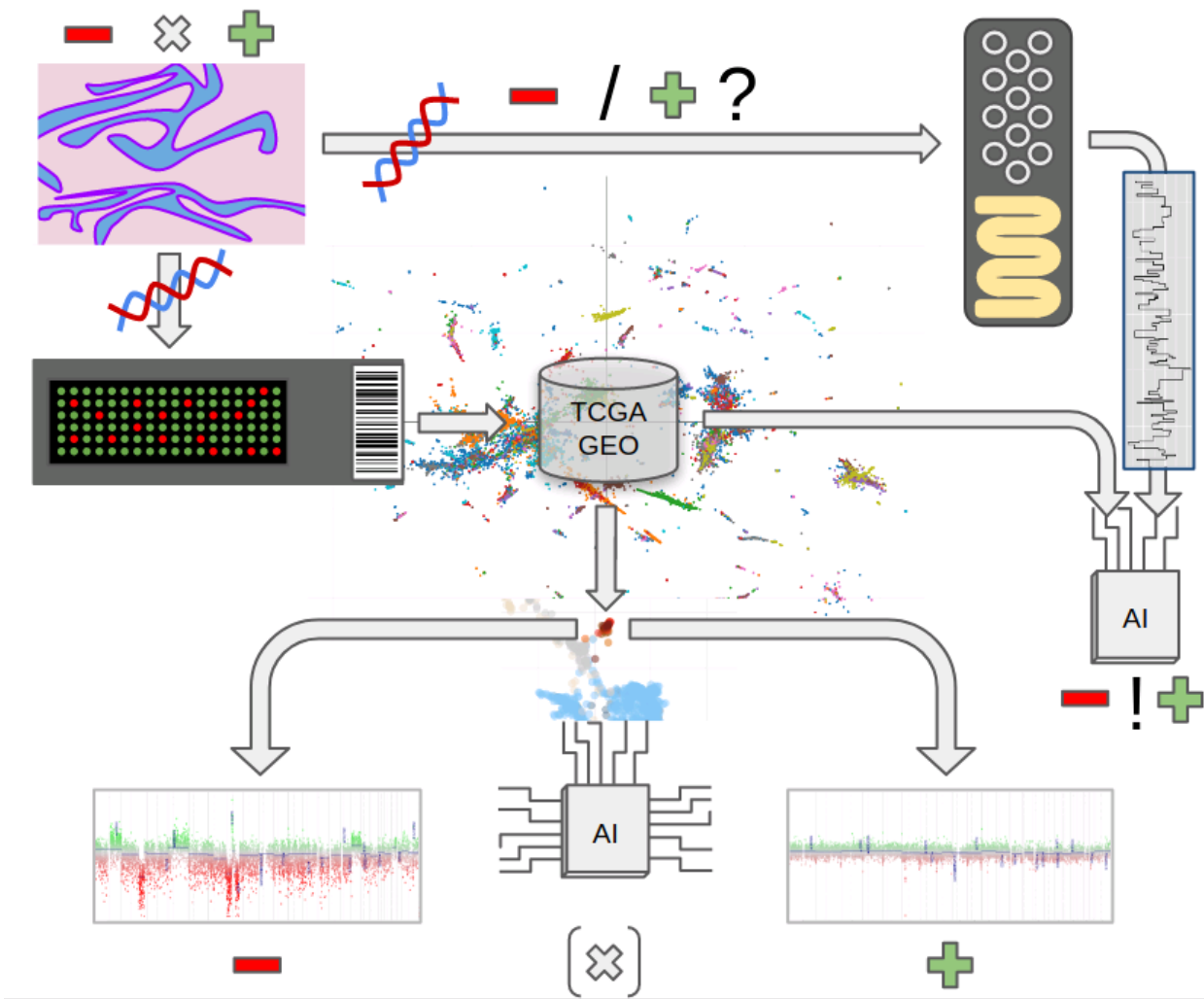
<sup>4</sup> Institute of Pathology, Cantonal Hospital Lucerne, Lucerne, Switzerland; [ellen.obermann@luks.ch](mailto:ellen.obermann@luks.ch)

\* Correspondence: [simone.muenst@usb.ch](mailto:simone.muenst@usb.ch) ; Tel.: 0041 61 265 25 25

**Simple Summary:** Fibroepithelial tumors of the breast represent a spectrum of mostly benign diseases. However, some of these tumors tend to recur and may even spread distantly to other body sites. Prediction of their biological behavior is currently morphology-centered. In this study we set out to answer the question whether their biologic behaviour might be reflected by specific DNA methylation and copy number profiles, both of which can be determined alongside each other in a diagnostic routine workflow through microarrays. We discovered that the fibroepithelial tumors seem to fall into two distinct copy number variant patterns, and that they are epigenetically related. Our study underlines the diagnostic usefulness of combined methylation/ copy number profiling in fibroepithelial breast tumors to predict clinical outcome.

**Abstract:** Fibroepithelial lesions (FL) of the breast, in particular Phyllodes tumors (PT) and fibroadenomas, pose a significant diagnostic challenge. There are no generally accepted criteria that distinguish benign, borderline, malignant PT, and FA. Combined genome-wide DNA methylation and copy number variant (CNV) profiling is an emerging strategy to classify tumors. We compiled a series of patient-derived archival biopsy specimens reflecting the FL spectrum and histological mimickers including clinical follow-up data. DNA methylation and CNVs were determined by well-established microarrays. Comparison of the patterns with a pan-cancer dataset assembled from public resources including "The Cancer Genome Atlas" (TCGA) and "Gene Expression Omnibus" (GEO) suggests that FLs form a methylation class distinct from both control breast tissue as well as common breast cancers. Complex CNVs were enriched in clinically aggressive FLs. Subsequent fluorescence in situ hybridization (FISH) analysis detected respective aberrations in the neoplastic mesenchymal component of FLs only, confirming that the epithelial component is non-neoplastic. Of note, our approach could lead to the elimination of the diagnostically problematic category of borderline PT and allow for optimized prognostic patient stratification. Furthermore, the identified recurrent genomic aberrations such as 1q gains (including MDM4), CDKN2a/b deletions and EGFR amplifications may inform therapeutic decision-making.

**Keywords:** fibroepithelial breast lesions; phyllodes tumors; methylation analysis; copy number alterations; dimension reduction; unsupervised machine learning



## 1. Introduction

Phyllodes tumor (PT), a rare breast neoplasm, accounts for 0.3% to 1% of all breast tumors [1]. The World Health Organization (WHO) Classification (2019) currently divides PT into categories of benign (up to 75% of all PT), borderline, and malignant, based on a combination of several histologic features such as stromal cellularity, nuclear atypia, mitotic activity, stromal overgrowth, and delimitation of the tumor [1,2]. However, this morphology-based classification remains challenging, as there is considerable overlap between categories. Furthermore, diagnostic criteria cannot always be sufficiently appreciated on small biopsies. Given the resulting interobserver variability, the diagnosis of PT, and, in particular, the distinction between benign PT and fibroadenoma (FA) as well as between benign and borderline PT remains problematic in diagnostic routine [1]. Furthermore, the differentiation of malignant PT from metaplastic carcinoma or primary breast sarcoma is not straightforward either [3]. Studies have shown that the overall rate of concordantly diagnosed FA and benign PT lies between 40 and 60% [1]. Additionally, histological grading correlates with prognosis, but is not predictive of clinical behavior in all patients [1]. At the moment, no clinically applicable biomarkers exist, and pathogenesis as well as molecular background of PT remain largely unknown [1].

While benign PT have a low recurrence risk (10-17%), borderline and malignant PT tend to relapse in a significant proportion of patients (14-25% and 23-30%, respectively), justifying surgical excision with tumor-free margins of 10 mm [3]. Moreover, malignant PT metastasize in up to 29% [4], most commonly to the lungs and skeleton, invariably indicating a dismal prognosis [3]. Molecular characteristics that conclusively distinguish between FA, benign, borderline, and malignant PT as well as breast carcinomas and primary breast sarcomas would, therefore, satisfy an urgent, currently unmet clinical need in breast surgery.

In recent years, combined genome-wide DNA methylation and chromosomal copy number analysis by microarrays has gained considerable interest as a precise tool to classify benign and malignant tumors based on their individual, often lineage-reflecting methylation patterns [5–8]. Most prominently, the brain tumor methylation classifier has become a mainstay in neuropathological tumor diagnostics worldwide [5], and has already influenced several entity definitions in the 2016 WHO classification [9]. It has recently also been adopted for soft tissue tumors [10] and outperforms histology not only in precision but also diagnostic speed when applied to intraoperative cryo specimens, employing nanopore sequencers instead of microarrays [11,12]. The wealth of methylation data in public repositories allows unsupervised machine learning [13,14] approaches to cross-compare a single diagnostic case against thousands of other specimens [15–17]. As opposed to supervised machine learning-based static classifiers [5,7,10,18–20], unsupervised approaches are able to place data series extraordinarily rare tumors [21] in the context of a magnitude of neoplastic and non-neoplastic differentiation based on the raw data alone [12,17]. In addition to fine-tuned supervised machine learning [22], integrated interpretation of copy number alterations, genetic changes, and histology can significantly increase disease course prediction granularity [23,24].

Given their morphological and immunohistochemical characteristics, we hypothesized that PT and FA represent a fibroepithelial lesion (FEL) spectrum originating from similar or identical cells of origin affected by different initial genomic damage events. Similar observations have been made, e.g., in meningiomas [20,25] and primary brain tumors [5,26]. Methylation detection tools, in particular microarrays (Illumina bead chip arrays, Illumina) deliver both methylation signatures and genome-wide copy number profiles [6,11], providing a dual use for routine diagnostics. To test our hypothesis

and simultaneously generate backbone reference data to train machine learning systems, we employed the more comprehensive methylation array strategy.

## 2. Materials and Methods

### 2.1. Tissue collection

After identifying potential samples from FA, PT, breast carcinomas (BC), and primary breast sarcomas (BS) in the biobank at the Institute of Medical Genetics and Pathology, University Hospital Basel, a H&E stained cryo-section of the freshly frozen (FF) tissue was prepared, and diagnosis in each case was re-confirmed by a specialized breast pathologist (S.M.) by reviewing the frozen section, or, where available, FFPE slides. FF tissue of 37 samples (1 FA, 30 PT and 6 BS), diagnosed between 1990 and 2017, was included. Furthermore, FFPE specimens from 23 cases (2018-2020) from the archives of the Institute of Medical Genetics and Pathology, University Hospital Basel and the Institute of Pathology, Cantonal Hospital, Lucerne (10 FA, 12 PT and 1 metaplastic BC) were analyzed. The respective H&E stained sections were also reviewed and the diagnosis confirmed by an expert breast pathologist (S.M.). This study was approved by the Ethikkommission Nordwest- und Zentralschweiz (EKNZ, proposal number 2014-397 and PB\_2020-00071). The study was performed in accordance with the Declaration of Helsinki.

### 2.2. Methylation and copy number analysis

The technology is based on a beadchip microarray (Infinium human methylation EPIC, by Illumina), consisting of a modified single nucleotide polymorphism (SNP) array to quantify DNA methylation. The current array covers approx. 850'000 CpG islands distributed across the entire genome. The procedure is well-established and part of our routine diagnostic practice: of each case, depending on biopsy size and tissue cellularity, 2 to 6 cryosections (70µm) or 7 to 15 FFPE sections (4µm) were used for DNA isolation (Promega Maxwell FFPE kit). DNA was quantified by absorption measurement (NanoDrop, Thermo Fisher, Waltham, MA, USA). After bisulfite-conversion and low-level amplification, the DNA was hybridized to bead chips which are then read on an appropriate scanner (typically iScan, service provided by Life&Brain, Bonn, Germany). The resulting data (IDAT format) were then preprocessed and normalized (SWAN), mapped to the genome and converted into beta values (which represent methylation state at each scanned site; all preprocessing via minfi) [27,28]. Top differentially methylated probes were determined by calculation of standard deviations across the entire dataset comprising >18'000 cases obtained from public resources including TCGA and Gene Expression Omnibus (GEO), as well as from in-house reference collections. The 75'000 probes with the highest standard deviations were selected. This filtered set of methylation beta values were then compared by uniform manifold approximation projection (UMAP) for dimension reduction as previously described [13,15]. This resulted in an unsupervised, bias-free grouping of samples sharing similar DNA methylation patterns, which often reflect individual (biological) entities 5. Of note, probes were not selected based on their annotation to specific genes. Copy number plots were generated with the conumee [29] in R. R 3.6.3 on Ubuntu Linux 18.04 (x86\_64) was used throughout this study.

### 2.3. Fluorescent in situ hybridization (FISH)

After deparaffinization and hydration of 3–4-µm-thick slides, sections were further processed for FISH according to our in-house protocol. In brief, the slides were pretreated automatically with the Leica Bond-III, then manually washed with water and dehydrated by 70%, 80%, and 100% ethanol. Subsequently, slides were incubated overnight with commercially available SPEC CDKN2A/CEN 9 and SPEC RB1/13q12 Dual Color Probe kits (ZytoVision, Bremerhaven, Germany) as well as LSI EGFR SpectrumOrange/CEP7 SpectrumGreen (Abbott, Chicago, IL, USA) probes.

### 3. Results

#### 3.1. Patient characteristics

In total, tissue of 41 PT was available for analysis. All patients with PT were female, and mean age at diagnosis was 51.7 years (range 14 to 86 years). All tumors were located in the breast, with the exception of one specimen, which was from a cerebellar metastasis of a malignant PT. Of the 41 analyzed tumors, 22 had an initial histologic diagnosis as benign, 8 as borderline and 10 as malignant PT, respectively. In one case from 1990, the diagnosis was "phyllodes tumor" without further specification. The initial diagnosis was confirmed in all cases by an experienced breast pathologist (S.M.).

#### 3.2. Patient outcome

Follow-up was available for 19 patients (mean follow-up time 75.8 months, range 6-219 months). Of the 9 patients with the diagnosis of a benign PT, 6 were disease free postoperatively, and three had recurrent disease: one patient after 45 months, with a second recurrence 19 months later, and the other two after 34 and 96 months, respectively. All three recurrent PT showed the same histology as the primary tumor. All 5 patients with a borderline PT remained disease free. Of the 5 patients with a malignant PT, 3 showed no evidence of disease, one patient presented with recurrent disease after 6 months, and one patient had a cerebellar metastasis after 36 months, with no further follow-up available after this event.

#### 3.3. DNA Methylation and copy number changes

We included a total of 18'537 methylome profiles, the majority of which were available through TCGA and GEO (Supplementary File PT20210527\_UMAP6\_all\_bVals\_top\_75000.xlsx). According to the annotation, these comprise 854 "breast cancer" (BC) and 97 "control breast" (CB) samples. While the majority of BC samples showed high-amplitude CNVs (769/854; 90%) and mostly clustered together with BCs from our cohort (Figure 1), a few cases (12/854; 1%) clustered with CB, likely representing BC samples with low tumor cell content as reflected by their low CNV amplitudes. No CB-annotated cases clustered with the high-amplitude CNV BCs. Out of high-amplitude BCs, a minority (94/769; 12%) showed an ERBB2 gene amplification; these cases did not form a separate cluster within BCs (Figure 1, Supplementary File PTplot.html)

In addition, 2 primary breast angiosarcomas and 1 metaplastic breast carcinoma were included to test whether they would fall into the respective reference data clusters. The angiosarcomas as well as the metaplastic carcinoma clustered in their respective groups, and no overlap with the PT was found (Figure 1, Supplementary File PTplot.html).

Interestingly, a fresh frozen sample of a 79 year-old patient, initially diagnosed as a malignant PT, showed a methylation profile consistent with diffuse large B-cell lymphoma (DLBCL). This case dated back to the pre-immunohistochemistry era, and retrospective immunohistochemical work-up of available FFPE tissue indeed confirmed DLBCL (CD20 positive and CD5 negative).

#### 3.4. Overlapping methylation patterns of phyllodes tumors and fibroadenomas

Interpretable methylation array data could be obtained for 38 of the 41 analyzed PT specimens while the remaining three samples clustered as "degraded DNA" array samples.

Additionally, 51 tumors, comprising 34 PT and all 17 FA, formed a methylation cluster in proximity to the BC and CB clusters, and distinct from adenocarcinomas of the breast (Figure 1). Likewise, the PT/FA methylation pattern differed from non-neoplastic breast tissue, but was more similar to the latter. This is reassuring, since the majority of neoplasms that do not carry driver alterations within strong epigenomic modifiers (e.g., IDH1/2, SMARCB1) largely retain epigenomic features of their precursor lineages 7



(<http://www.epidip.org>). The remaining 4 tumors histologically diagnosed as PT clustered elsewhere in proximity to mostly mesenchymal tumors (Supplementary Figure 1). The PT/FA cluster showed a slight trend to separate FA from PT, but overlap currently remains high and without clear distinction between PT histologically diagnosed as benign, borderline, or malignant (Figure 1 D).

### 3.5. Copy number alterations in phyllodes tumors and fibroadenomas

We visually classified copy number aberrations in a tumor type agnostic manner into four categories: flat, high-amplitude of CNV, low-amplitude of CNV, degraded/unclear. All but one PT histologically classified as malignant showed high-amplitude CNVs. We hence detected strong correlation between copy number alterations and morphological changes indicating malignancy. The remaining histologically malignant PT case showed low-amplitude CNVs with a gain of 1q and was from a 21 year-old patient for whom, unfortunately, no clinical follow-up was available.

For the histologically classified FAs and benign PT, all tumors showed low-amplitude CNVs or flat copy number profiles, with the exception of 3 cases (2 benign PT and 1 FA) featuring high-amplitude, malignant-looking CNV profiles with either CDKN2a/b deletion (2 benign PT) or MDM4 gain (1 FA). Importantly, the histology of these 3 cases was concordant with the initial diagnosis. Of note, one histologically benign PT with CDKN2a/b deletion showed an increased proliferation rate as well as strong expression of p53, both of which have been linked to malignancy in PT 19; this patient was initially resected with a very close resection margin (<1mm), and developed recurrent disease after 46 months, and then again after 19 months. The recurrent PT were both again excised with clear margins, and featured benign histology in both instances. Further follow-up of the patient is not available. For the second benign PT with CDKN2a/b deletion, no recurrence has been recorded up to now. The FA with MDM4 gain was histologically unremarkable, and no recurrent disease was recorded. Interestingly, of the 4 PT histologically diagnosed as borderline, two showed high- and the other two low-amplitude CNV profiles, indicating that borderline PT do not seem to represent a distinct biological entity, but instead may be classified as either benign PT with flat copy number profile, or malignant PT with high-amplitude CNVs.

CNV profile analysis also revealed recurrent genomic aberrations. 17 PT showed a gain of 1q (including MDM4), while 6 PT featured a CDKN2a/b deletion, of which 4 were diagnosed as malignant (see above). Notably, of 9 PT with 1q gain and available follow-up, only one recurred after 6 months, while both PT with CDKN2a/b deletion and follow-up developed recurrent disease, in one case recurring twice over a period of 65 months.

One of the malignant PT showed a Rb1 deletion, and one patient with 2 unilateral PT (one classified as malignant, and one as PT NOS/borderline in histology) harbored a potentially targetable epidermal growth factor receptor (EGFR) gene amplification in both tumors (Figure 3). In this patient, the malignant PT additionally showed Rb1 deletion, and the borderline PT showed an additional MDM4 amplification on top of the EGFR amplification. This suggests a joined tumor origin with divergent genetic aberrations, which in one case led to a malignant PT. Table 1 summarizes the identified CNVs.

FISH analysis of PT cases with genomic aberrations was able to verify the CDKN2a deletions in 3 tumors, as well as the Rb1 deletion in 2 tumors and the EGFR amplification in 2 tumors. In the remaining 3 cases with CDKN2a deletion, probe hybridization failed, probably due to the age of the tissue. Importantly, all examined deletions and amplifications were only present in the so-called stroma, and not in the adjacent epithelial cells, as illustrated in Figure 3.

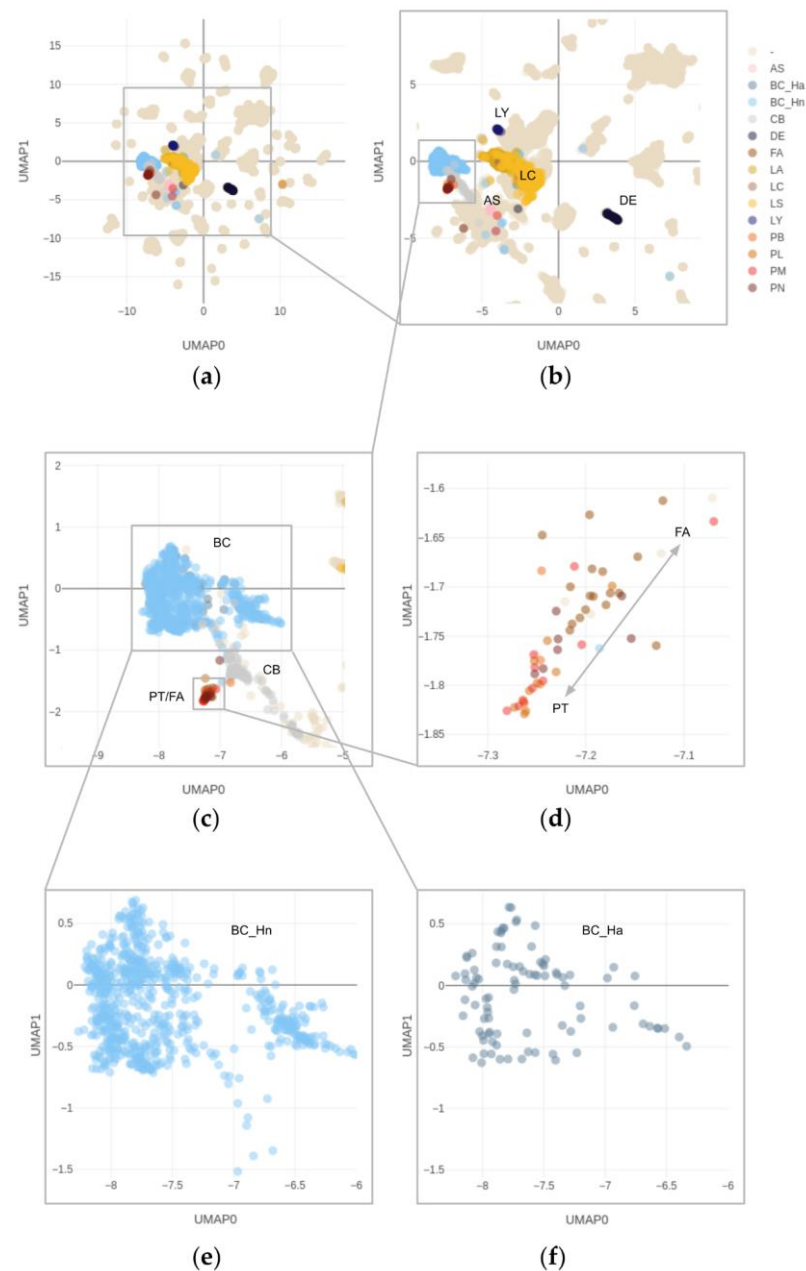
### 3.6. Proof-of-concept experiment using nanopore sequencing

Nanopore same-day diagnostics requires native tumor DNA [11,11] which was available for some archival specimens. To demonstrate the technical validity of nanopore sequencing as an ultra-fast alternative to microarrays, we ran an aliquot of a histologically

---

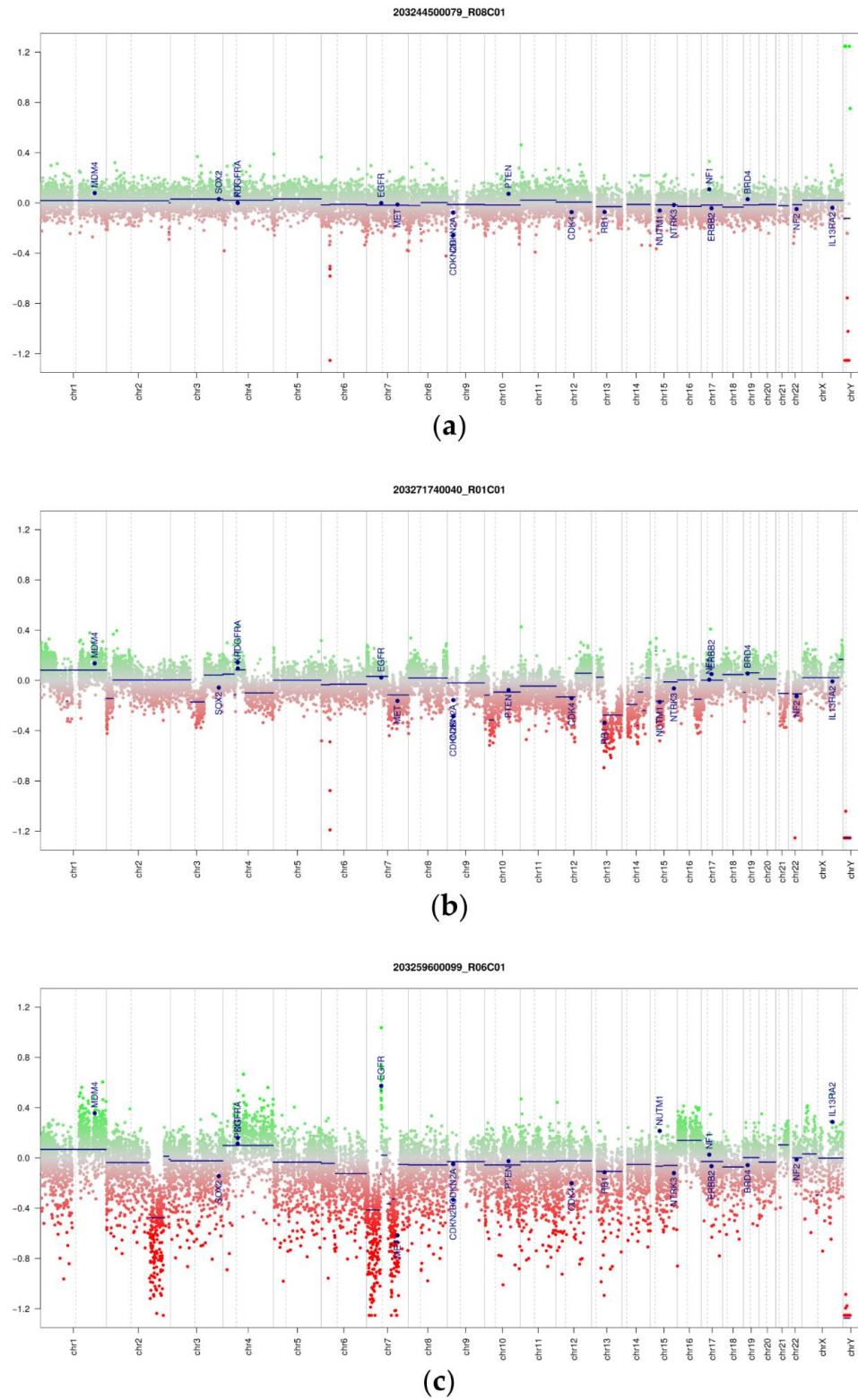
malignant PT on NanoDiP [12]. Empirically determined run settings applied in daily brain tumor routine were applied without modification, in particular 150 megabases of high-quality reads. These were achieved in 2h 5min run time after which NanoDiP terminated the sequencing procedure. All run details are included in the supplementary material (S5). The resulting UMAP [13] plot based on 4488 CpG sites identified in 142994 reads shows the same pattern as the one generated by array data alone (Figure 1), placing FLs close to breast control tissue samples, which in turn are located in proximity to invasive breast cancers. The copy number plot from nanopore read alignment recapitulates the respective aberrations found in the microarray analysis (supplementary material S5).

## 3.2. Figures, Tables and Schemes

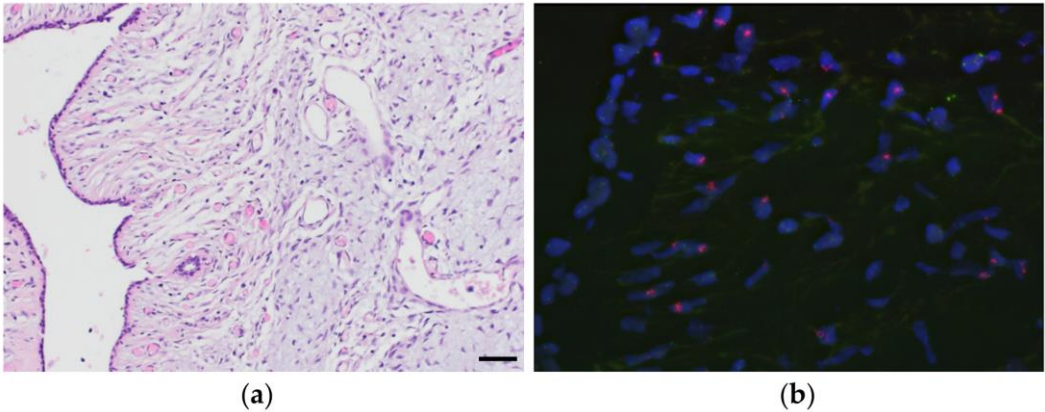


**Figure 1.** UMAP plot of the combined TCGA, GEO, and inhouse data reference set alongside phyllodes tumor specimens, zoomed-in view. The specimens are annotated as follows: BC\_Ha = breast adenocarcinoma, Her-2 amplified; BC\_Hn = breast adenocarcinoma, Her-2 not amplified; CB = control breast; DE = degraded DNA; FA = fibroadenoma; LA = lung adenocarcinoma; LC = lung cancer, NOS; LS = lung squamous cell carcinoma; LY = diffuse large B cell lymphoma; PB = PT borderline; PL = PT benign; PM = PT malignant; PN = PT, NOS. Note that phyllodes tumors cluster in the vicinity to reference control breast tissue (CB) but form a distinct methylation class. An interactive (zoomable, annotated) plot can be found in the supplementary file (PTplot.html). In addition, the plot coordinates are provided with Sentrax ID annotation in XLSX (MS Excel) and RDA (R 3.6.3) format.





**Figure 2.** Genome-wide copy number variation profiles of 3 phyllodes tumors. a) benign phyllodes tumor. b) malignant phyllodes tumor. c) Some phyllodes tumors, here a malignant form, show a potentially targetable EGFR gene amplification. The remaining copy number profiles of PT and FA can be accessed on <http://www.epidip.org> by searching for the respective Sentrax IDs.



**Figure 3.** a) H&E of a borderline PT (magnification 200x) and b) corresponding FISH image (magnification 400x), which shows an EGFR gene amplification (red) in the stromal cells, but not the adjacent benign epithelium (top right). Green is the centromere probe for chromosome 7.

**Table 1.** This is a table. Tables should be placed in the main text near to the first time they are cited.

Age at diagnosis	Histological Diagnosis	Methylation Category	Follow-up (months)	Recurrent Disease	CNV aberrations	Confirmed by FISH
50	malignant PT	PHYT_MAL	6	yes	1 gain	
64	PT	PHYT_NOS		NA	1q gain	
59	benign PT	PHYT_NOS	81	no	1q gain	
65	benign PT	PHYT_BOR		NA	1q gain	
65	malignant PT	PHYT_BOR	30	no	1q gain	
72	malignant PT	PHYT_MAL	91	no	1q gain	
21	malignant PT	PHYT_MAL		NA	1q gain	
66	malignant PT	PHYT_NOS	48	no	1q gain	
50	malignant PT	PHYT_MAL	135	no	1q gain	
50	malignant PT	PHYT_MAL	122	no	1q gain	
83	malignant PT	PHYT_NOS	60	no	1q gain	
40	benign PT	PHYT_NOS	15	no	1q gain	
54	benign PT	PHYT_BEN			1q gain	
51	malignant PT	PHYT_MAL		NA	1q gain, EGFR amp., RB1 del.	RB1 del., EGFR ampl.
51	benign or borderline PT	PHYT_NOS		NA	1q gain , MDM4 amplification, EGFR amp.	EGFR ampl.
69	benign PT	PHYT_BEN		NA	1q/MDM4 gain	
60	benign PT	PHYT_BEN		NA	CDKN2a/b deletion	

50	benign PT	PHYT_MAL		NA	CDKN2a/ b deletion	
48	malignant PT	PHYT_MAL	36	yes	CDKN2a/ b deletion	CDKN2a/b deletion
42	benign PT	PHYT_BEN		NA	MDM4 gain	
46	benign PT	PHYT_BEN		NA	MDM4 gain	
82	borderline PT	PHYT_NOS		NA	MDM4 gain, CDKN2a/ b deletion	
36	benign PT	PHYT_BEN	65	yes	MDM4 gain, CDKN2a/ b deletion	CDKN2a/b deletion
38	FA	BR_FAD			MDM4 gain, malignan t-looking CNV	
64	benign PT	PHYT_MAL	204	no	RB1 deletion	RB1 deletion
83	malignant PT	PHYT_MAL		NA	susp. 1q gain (bad DNA)	

#### 4. Discussion

Our combined methylation and copy number analysis revealed that PT do indeed represent a biologically distinct group of breast neoplasms, and are part of a spectrum between normal breast tissue and invasive breast cancer; within the PT group, tumors form a gradient from benign (closely resembling benign breast tissue) to malignant. While the methylation profiles of PT and FA converge in a cluster distinct from BC and normal breast tissue, their copy number profiles prompts for a separation of the FA/PT tumor class into malignant (high CNVs) and benign (flat copy number profiles or few CNVs) forms. This suggests that tumors histologically categorized as “borderline” may not represent a distinct biological entity, but instead separate into benign and malignant PT as revealed by combined methylation and copy number analysis. Our molecular approach could thus be used to discriminate benign from malignant PT, especially in the diagnostically difficult borderline category, thereby aiding clinical patient management. Omitting the twilight category of “borderline” PT is likely to not only streamline the diagnostic process but may also contribute to an optimized diagnostic and prognostic patient stratification. This may help clinicians as well as patients to more confidently plan for potential revision surgery and follow-up, since borderline tumors with a flat CNV profile most likely follow a benign clinical course requiring no additional treatment, whereas borderline tumors with a high CNV profiles are potentially malignant, requiring wide excision as well as close clinical follow-up.

Of note, confirmation of our results on the NanoDiP platform confirm the feasibility of our approach and underline the straightforward clinical applicability of our FL reference data collection.

Our analysis also revealed recurrent CNV aberrations such as 1q gains, CDKN2a/b deletions and MDM4 gains (Figure 2). Of note, these CNV aberrations were verified by FISH. Using a single microarray-based technique, or, alternatively, nanopore-based sequencing [11], both copy number and methylation profiles are obtained and evaluated simultaneously.

Interestingly, out of all PT with 1q gain and available follow-up (n=9), only one recurred, while both PT with CDKN2a/b deletions developed recurrent disease. Truncating as well as non-synonymous CDKN2a mutations as well as homozygous CDKN2a deletions are known to occur in recurrent PT with histologically benign, borderline, and malignant characteristics [20,21]. These findings suggest that loss of CDKN2a gene function might underlie (or contribute to) PT recurrence, independent of histological grade, and indicate that CDKN2a analysis could be useful to identify patients at risk for recurrent disease.

One patient had two distinct PT at presentation (one borderline and one malignant histologically), both of which showed EGFR amplifications, suggesting that these represent two clonally related but histologically distinct tumors. Detection of EGFR amplification by FISH has been described in up to 16% of PT, and has been associated with tumor progression [30]. Unfortunately, no clinical follow-up is available for our patient.

Finally, both EGFR amplifications and CDKN2a/b deletions represent potentially targetable gene aberrations. With new anti-EGFR therapies and CDK4/6 inhibitors entering clinical practice, identification of these alterations may become part of the routine molecular diagnostic work-up of these tumors

## 5. Conclusions

The distinct methylation and CNV signatures of the histological PT/FA spectrum not only allows the diagnostic discrimination of PT from histological mimics such as sarcomas or carcinomas, but also enables the distinction between benign and malignant PT. These findings may potentially eliminate the need of a borderline category, paving the way to an optimized prognostic patient stratification and clinical management. Verification of copy number aberrations through FISH confirms that the stroma (as opposed to epithelial cells) represents the neoplastic component in PT. As recurrent genomic aberrations such as EGFR amplification and CDKN2a/b deletion may represent therapeutic targets, their diagnostic identification could impact clinical management of recurrent or metastatic PT patients.

Lastly, while having analyzed only a single case so far with our diagnostic same-day nanopore sequencing [12] that our institution routinely applies in brain tumor diagnostics, we demonstrate the immediate clinical applicability of our FL reference data collection which we make publicly available alongside this manuscript.

## Supplementary Materials:

*File S1: PTplot.html.* Interactive UMAP plot with annotation of the methylation classes mentioned in this manuscript. Requires an up-to-date web browser; tested with Chrome and Firefox. For reference on how to zoom, search, hide, etc. refer to the plotly user manual.

*File S2: PT20210527\_UMAP6\_all\_bVals\_top\_75000\_cgList.rds.rda.* CpG list in Illumina annotation in R 3.6.3 RDA format, refer to Illumina's documentation concerning the genomic positions of the probes.

*File S3: PT20210527\_UMAP6\_all\_bVals\_top\_75000.xlsx.* Coordinates of the UMAP plot in MS XLSX format, annotated with Sentrix IDs.

---

*File S4: PT20210527\_UMAP6\_all\_bVals\_top\_75000.rda.* Coordinates of the UMAP plot in R 3.6.3 RDA format, annotated with Sentrix IDs.

*File S5: NanoDiP\_PT.zip.* Report and nanopore run characteristics of the proof-of-concept experiment (PDF and HTML files with plots).

**Author Contributions:** S.M. and J.H. performed the study concept and design; S.M. collected and reviewed all patient samples, and wrote part of the manuscript; J.H. developed software, provided acquisition, analysis and interpretation of data, and wrote part of the manuscript; E.O. provided patient samples including diagnostic review and revised the manuscript; S.D.S. provided material support and revised the manuscript; T.V. and F.S. provided reviewed the data and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Foundation Domarena and the Parrotia Foundation.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of northwest and central Switzerland (EKNZ) (proposal number 2014-397 and PB\_2020-00071).

**Informed Consent Statement:** Patient consent was waived by the Ethics Committee (EKNZ) due to inclusion of archival samples, with consecutive loss of follow-up of patients.

**Data Availability Statement:** The datasets generated and/or analyzed during the current study are available in the GEO repository, <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE179458>. Furthermore, all data can be viewed in a processed manner on <http://www.epidip.org>. Individual cases can be identified by searching for their Sentrix ID.

**Acknowledgments:** We thank Bruno Grilli and Julian Pollinger for expertly performing FISH analyses.

**Conflicts of Interest:** S.M. is part of a digital advisory board for Roche. The other authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.



## References

1. Zhang, Y.; Kleer, C.G. Phyllodes Tumor of the Breast: Histopathologic Features, Differential Diagnosis, and Molecular/Genetic Updates. *Arch. Pathol. Lab. Med.* **2016**, *140*, 665–671, doi:10.5858/arpa.2016-0042-RA.
2. *Breast Tumours*; Lokuhetty, D., White, V.A., Watanabe, R., Cree, I.A., WHO, International Agency for Research on Cancer, Eds.; World Health Organization Classification of tumours; 5th ed.; International Agency for Research on Cancer: Lyon, 2019; ISBN 978-92-832-4500-1.
3. Tan, B.Y.; Acs, G.; Apple, S.K.; Badve, S.; Bleiweiss, I.J.; Brogi, E.; Calvo, J.P.; Dabbs, D.J.; Ellis, I.O.; Eusebi, V.; et al. Phyllodes Tumours of the Breast: A Consensus Review. *Histopathology* **2016**, *68*, 5–21, doi:10.1111/his.12876.
4. Abdalla, H.M.; Sakr, M.A. Predictive Factors of Local Recurrence and Survival Following Primary Surgical Treatment of Phyllodes Tumors of the Breast. *J. Egypt. Natl. Cancer Inst.* **2006**, *18*, 125–133.
5. Capper, D.; Jones, D.T.W.; Sill, M.; Hovestadt, V.; Schrimpf, D.; Sturm, D.; Koelsche, C.; Sahm, F.; Chavez, L.; Reuss, D.E.; et al. DNA Methylation-Based Classification of Central Nervous System Tumours. *Nature* **2018**, *555*, 469–474, doi:10.1038/nature26000.
6. Euskirchen, P.; Bielle, F.; Labreche, K.; Kloosterman, W.P.; Rosenberg, S.; Daniau, M.; Schmitt, C.; Masliah-Planchon, J.; Bourdeaut, F.; Dehais, C.; et al. Same-Day Genomic and Epigenomic Diagnosis of Brain Tumors Using Real-Time Nanopore Sequencing. *Acta Neuropathol. (Berl.)* **2017**, *134*, 691–703, doi:10.1007/s00401-017-1743-5.
7. Moran, S.; Martínez-Cardús, A.; Sayols, S.; Musulén, E.; Balañá, C.; Estival-Gonzalez, A.; Moutinho, C.; Heyn, H.; Diaz-Lagares, A.; de Moura, M.C.; et al. Epigenetic Profiling to Classify Cancer of Unknown Primary: A Multicentre, Retrospective Analysis. *Lancet Oncol.* **2016**, *17*, 1386–1395.
8. Djirackor, L.; Halldorsson, S.; Niehusmann, P.; Leske, H.; Capper, D.; Kuschel, L.P.; Pahnke, J.; Due-Tønnessen, B.J.; Langmoen, I.A.; Sandberg, C.J.; et al. Intraoperative DNA Methylation Classification of Brain Tumors Impacts Neurosurgical Strategy. *Neuro-Oncol. Adv.* **2021**, vdab149, doi:10.1093/noajnl/vdab149.
9. World Health Organization *WHO Classification of Tumours of the Central Nervous System*; Louis, D.N., Ohgaki, H., Wiestler, O.D., Cavenee, W.K., Eds.; World Health Organization classification of tumours; Revised 4th edition.; International Agency for Research on Cancer: Lyon, 2016; ISBN 978-92-832-4492-9.
10. Koelsche, C.; Schrimpf, D.; Stichel, D.; Sill, M.; Sahm, F.; Reuss, D.E.; Blattner, M.; Worst, B.; Heilig, C.E.; Beck, K.; et al. Sarcoma Classification by DNA Methylation Profiling. *Nat. Commun.* **2021**, *12*, 498, doi:10.1038/s41467-020-20603-4.
11. Kuschel, L.P.; Hench, J.; Frank, S.; Hench, I.B.; Girard, E.; Blanluet, M.; Masliah-Planchon, J.; Misch, M.; Onken, J.; Czabanka, M.; et al. *Robust Methylation-Based Classification of Brain Tumors Using Nanopore Sequencing*; Oncology, 2021;
12. Hench, J.; Hultschig, C. NanoDiP - Nanopore Digital Pathology Available online: <https://github.com/neuropathbasel/nanodip> (accessed on 26 November 2021).
13. McInnes, L.; Healy, J.; Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *ArXiv180203426 Cs Stat* **2018**.
14. van der Maaten, L.; Hinton, G. Visualizing Data Using T-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.
15. Haefliger, S.; Tzankov, A.; Frank, S.; Bihl, M.; Vallejo, A.; Stebler, J.; Hench, J. NUT Midline Carcinomas and Their Differentials by a Single Molecular Profiling Method: A New Promising Diagnostic Strategy Illustrated by a Case Report. *Virchows Arch. Int. J. Pathol.* **2020**, doi:10.1007/s00428-020-02869-7.
16. Saleh, C.; Jaszczuk, P.; Hund-Georgiadis, M.; Frank, S.; Cordier, D.; Hench, I.B.; Todea, A.; Wasilewski, A.; Wilmes, S.; Grigioni, G.; et al. Differentiation of Rare Brain Tumors through Unsupervised Machine Learning: Clinical Significance of in-Depth Methylation and Copy Number Profiling Illustrated through an Unusual Case of IDH Wildtype Glioblastoma. *Clin. Neuropathol.* **2020**, doi:10.5414/NP301305.
17. Hench, J.; Frank, S. EpiDiP Server Available online: <http://www.epidip.org>.
18. Jurmeister, P.; Schöler, A.; Arnold, A.; Klauschen, F.; Lenze, D.; Hummel, M.; Schweizer, L.; Bläker, H.; Pfitzner, B.M.; Mamlouk, S.; et al. DNA Methylation Profiling Reliably Distinguishes Pulmonary Enteric Adenocarcinoma from Metastatic Colorectal Cancer. *Mod. Pathol.* **2019**, doi:10.1038/s41379-019-0207-y.
19. Jurmeister, P.; Bockmayr, M.; Seegerer, P.; Bockmayr, T.; Treue, D.; Montavon, G.; Vollbrecht, C.; Arnold, A.; Teichmann, D.; Bressen, K.; et al. Machine Learning Analysis of DNA Methylation Profiles Distinguishes Primary Lung Squamous Cell Carcinomas from Head and Neck Metastases. *Sci. Transl. Med.* **2019**, *11*, eaaw8513, doi:10.1126/scitranslmed.aaw8513.



20. Sahm, F.; Schrimpf, D.; Stichel, D.; Jones, D.T.; Hielscher, T.; Schefzyk, S.; Okonechnikov, K.; Koelsche, C.; Reuss, D.E.; Capper, D.; et al. DNA Methylation-Based Classification and Grading System for Meningioma: A Multicentre, Retrospective Analysis. *Lancet Oncol.* **2017**, *18*, 682–694.
21. Schweizer, L.; Thierfelder, F.; Thomas, C.; Soschinski, P.; Kim, H.-Y.; Jödicke, R.; Woltering, N.; Förster, A.; Teichmann, D.; Siewert, C.; et al. Molecular Characterisation of Sporadic Endolymphatic Sac Tumours and Comparison to von Hippel-Lindau Disease-Related Tumours. *Neuropathol. Appl. Neurobiol.* **2021**, *47*, 756–767, doi:10.1111/nan.12741.
22. Maros, M.E.; Capper, D.; Jones, D.T.W.; Hovestadt, V.; von Deimling, A.; Pfister, S.M.; Benner, A.; Zucknick, M.; Sill, M. Machine Learning Workflows to Estimate Class Probabilities for Precision Cancer Diagnostics on DNA Methylation Microarray Data. *Nat. Protoc.* **2020**, *15*, 479–512, doi:10.1038/s41596-019-0251-6.
23. Capper, D.; Stichel, D.; Sahm, F.; Jones, D.T.W.; Schrimpf, D.; Sill, M.; Schmid, S.; Hovestadt, V.; Reuss, D.E.; Koelsche, C.; et al. Practical Implementation of DNA Methylation and Copy-Number-Based CNS Tumor Diagnostics: The Heidelberg Experience. *Acta Neuropathol. (Berl.)* **2018**, *136*, 181–210, doi:10.1007/s00401-018-1879-y.
24. Maas, S.L.N.; Stichel, D.; Hielscher, T.; Sievers, P.; Berghoff, A.S.; Schrimpf, D.; Sill, M.; Euskirchen, P.; Blume, C.; Patel, A.; et al. Integrated Molecular-Morphologic Meningioma Classification: A Multicenter Retrospective Analysis, Retrospectively and Prospectively Validated. *J. Clin. Oncol.* **2021**, *39*, 3839–3852, doi:10.1200/JCO.21.00784.
25. Paramasivam, N.; Hübschmann, D.; Toprak, U.H.; Ishaque, N.; Neidert, M.; Schrimpf, D.; Stichel, D.; Reuss, D.; Sievers, P.; Reinhardt, A.; et al. Mutational Patterns and Regulatory Networks in Epigenetic Subgroups of Meningioma. *Acta Neuropathol. (Berl.)* **2019**, doi:10.1007/s00401-019-02008-w.
26. Perez, E.; Capper, D. Invited Review: DNA Methylation-based Classification of Paediatric Brain Tumours. *Neuropathol. Appl. Neurobiol.* **2020**, *46*, 28–47, doi:10.1111/nan.12598.
27. Fortin, J.-P.; Triche, T.J.; Hansen, K.D. Preprocessing, Normalization and Integration of the Illumina HumanMethylationEPIC Array with Minfi. *Bioinformatics* **2016**, btw691, doi:10.1093/bioinformatics/btw691.
28. Maksimovic, J.; Gordon, L.; Oshlack, A. SWAN: Subset-Quantile Within Array Normalization for Illumina Infinium HumanMethylation450 BeadChips. *Genome Biol.* **2012**, *13*, R44, doi:10.1186/gb-2012-13-6-r44.
29. Volker Hovestadt, M.Z. *Conumee*; Bioconductor, 2017;
30. Kersting, C.; Kuijper, A.; Schmidt, H.; Packeisen, J.; Liedtke, C.; Tidow, N.; Gustmann, C.; Hinrichs, B.; Wülfing, P.; Tio, J.; et al. Amplifications of the Epidermal Growth Factor Receptor Gene (Egfr) Are Common in Phyllodes Tumors of the Breast and Are Associated with Tumor Progression. *Lab. Investig. J. Tech. Methods Pathol.* **2006**, *86*, 54–61, doi:10.1038/labinvest.3700358.