

Whole Genome Sequencing and single-cell transcriptomics identify *KMT2D* as a potential new driver for pituitary adenomas

Maxime Brunner^{1,4}, Jenny Meylan-Merlini¹, Maude Muriset¹, Sergey Oreshkov^{1,4},
Andrea Messina^{1,4}, Mahmoud Messerer^{2,4}, Roy Daniel^{2,4}, Ekkehard Hewer^{3,4}, Jean
Phillipe Brouland^{3,4}, Federico Santoni^{1,4}

¹ Department of Endocrinology, Diabetology and Metabolism, CHUV, Lausanne,
Switzerland

² Department of Neurosurgery, CHUV, Lausanne, Switzerland

³ University institute of Pathology, CHUV, Lausanne, Switzerland

⁴ Faculty of Biology and Medicine, Lausanne University, Lausanne, Switzerland

Correspondence to: federico.santoni@chuv.ch

Abstract

The pituitary gland is a main component of the endocrine system and a master controller of hormone production and secretion. Unlike neoplastic formation in other organs, Pituitary Neuroendocrine Tumors (PitNETs) are frequent in the population (16%) and, for unknown reasons, almost exclusively benign. So far, few genes have been identified as drivers for PitNETs, such as *GNAS* in somatotroph tumors and *USP8* in corticotroph tumors. Using whole genome sequencing, we uncover a potential novel driver, the histone methyltransferase *KMT2D*, in a patient in his 50s suffering from a mixed somato-lactotroph tumor. Coverage ratio between germline and tumor revealed extensive chromosomal alterations. Single-cell RNA sequencing of the tumor shows up-regulation of known tumorigenic pathways compared to a healthy reference, as well as a different immune infiltration profile compared to other PitNETs, more closely resembling the profile of carcinomas than adenomas. Genome-wide DNA methylation analysis identified 792 differentially methylated regions, including notable hypomethylation in the promoter of *SPON2*, an immune-related gene. Our results show that tumors considered as quiet and non-aggressive can share drivers, features, and epigenetic alterations with metastatic forms of cancer, raising questions about the biological mechanisms controlling their homeostasis.

Introduction

The pituitary gland, often referred to as the master regulator of the endocrine system, lies within the sella turcica at the base of the brain. It is composed of the adenohypophysis (anterior pituitary), arising from the oral ectoderm, and the neurohypophysis (posterior pituitary), originating from the neuroectoderm. The neurohypophysis consists primarily of axonal projections from the hypothalamus. Together with the hypothalamus, the pituitary gland plays a critical role in regulating essential physiological processes such as growth, puberty, metabolism, stress responses, reproduction, and lactation through hormone production and secretion.

Pituitary tumors, or Pituitary NeuroEndocrine Tumors (PitNETs), are the third most common type of intracranial tumors, accounting for approximately 15% of cases [1]. Although many PitNETs remain undiagnosed (with an autopsy prevalence of around 16%), the incidence of clinically relevant cases in the United States was estimated at 4.07 per 100,000 per year between 2012 and 2016 [2]. The vast majority of pituitary tumors are adenomas (95%), which can sometimes be invasive, invading neighboring tissues and having a higher risk of local recurrence as well as displaying metastatic lesions but only in 0.1% to 0.2% of cases [3]. Clinically, PitNETs can present with severe symptoms due to excessive hormone secretion, inhibition of specific hormone secretion, or signs and symptoms related to an expanding sellar mass, such as visual deficits and headaches.

The most common treatment for PitNETs is transsphenoidal surgery, followed by radiotherapy and drug control for certain subtypes [4]. Additionally, immunotherapy with PD-L1 has been proposed as an alternative treatment [5]. A recent study aimed to investigate the association between PD-L1 expression and the radiological and pathological behavior of PitNETs to determine its suitability as a target in relapse cases [6].

Classification of these tumors follows the latest WHO guidelines from 2022 [7], which distinguish tumors of the anterior and posterior lobes, as well as other hypothalamic tumors. Key features used to determine the type and subtype of a tumor include transcription factors, hormones, and other biomarkers, such as low-molecular-weight cytokeratin to determine the cell of origin. For instance, somatotroph tumors are divided into densely and sparsely granulated subtypes, based on the distribution of secretory granules within the tumor. Densely granulated tumors generally secrete more hormones than their sparsely granulated counterparts. These tumors are further characterized by staining for the transcription factor PIT1 (*POU1F1*, a lineage determinant) and GH (*GH1*), along with the glycoprotein hormone γ -subunit (*CGA*). Additionally, cytoke-

atin staining shows perinuclear patterns in densely granulated tumors and more than 70% fibrous bodies in sparsely granulated subtypes.

Despite significant advancements in the accurate classification of PitNETs, their etiology remains poorly understood. Genetic studies have highlighted the role of inherited germline mutations in genes such as *AIP* in Familial Isolated Pituitary Adenoma [8], *PRKARIA* in Carney Complex, *MEN1* in multiple endocrine neoplasia syndrome type 1 (MEN 1) [9], *CDKN1B* in MEN 4 [10], and *USP8* as an oncogene in sporadic corticotroph adenomas [11], as well as *GNAS* mutations in somatotroph adenomas [12].

In this report, we present a potential novel driver for pituitary neuroendocrine tumors: the histone-lysine N-methyltransferase 2D (*KMT2D*) gene. Germline pathogenic mutations in *KMT2D* are known to cause Kabuki syndrome, characterized by distinct facial features, growth delay, and mental retardation [13]. Somatic mutations in this gene have been implicated in various cancers, including breast cancer [14], lung squamous cell carcinoma (LUSC) [15], and B-cell-derived lymphoma [16], making it one of the most frequently altered genes in cancer. However, *KMT2D* has not previously been reported as a driver for pituitary neuroendocrine tumors.

Results

Patient information

In the context of a large collaborative research project, Pituitary Neuroendocrine Tumors (Pit-NETS) are fresh collected after surgery at the university hospital (CHUV) in Lausanne, Switzerland and processed for single-cell RNA sequencing (sc-RNAseq) and Whole Genome Sequencing (WGS) in our laboratory. Here, we present the case of a male patient in his 50s with a secreting mixed somatotroph macroadenoma causing acromegaly and elevated IGF-1 levels. Surgical intervention successfully alleviated acromegaly symptoms and normalized IGF-1 levels. Post-operative brain MRI confirmed complete tumor resection. The pathology report on the excised tissue mentions the loss of the usual acinar architecture on reticulin staining as well as cells having a densely granulated eosinophilic cytoplasm. No apparent sign of necrosis or mitosis. Cells express PIT1 (*POU1F1*), SF1, GH and PRL (Fig. 1). Proliferation index MIB1/Ki67 is around 1%. Very weak intensity staining with PD-L1 (*CD274*) for 40% of tumoral cells. The patient is referred as patient 16.

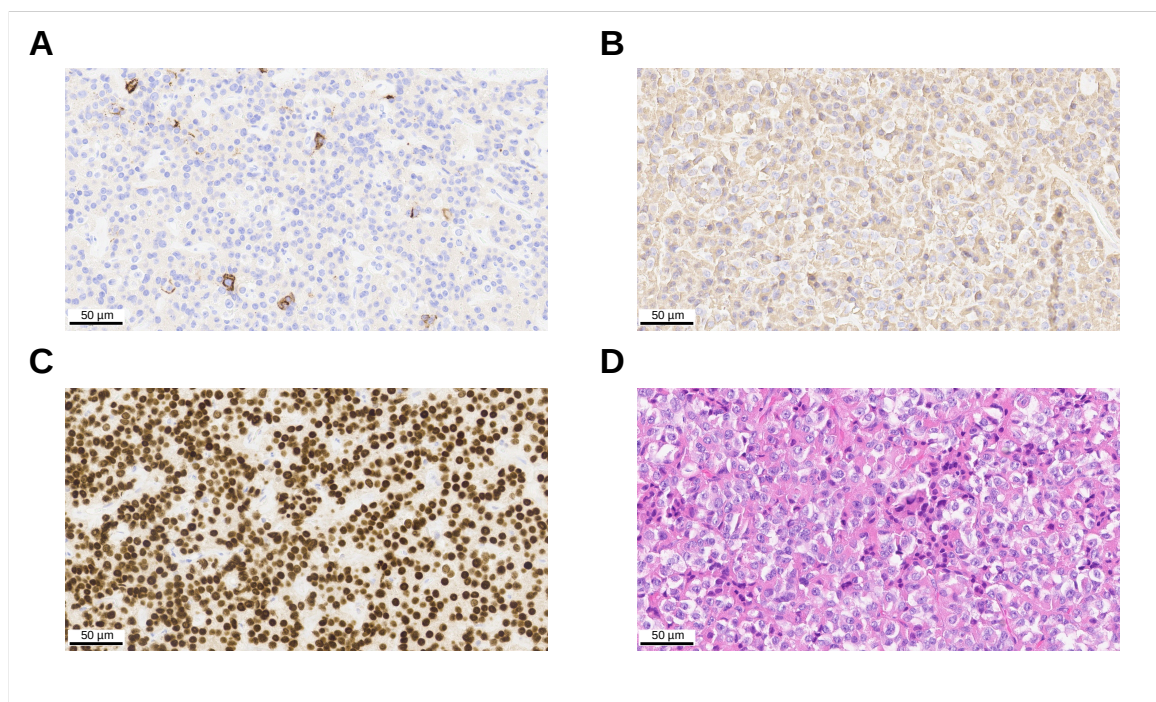


Figure 1: **A:** PRL staining. **B:** GH staining. **C:** POU1F1 staining. **D:** Hematoxylin and eosin staining.

Variant information

Variant analysis of matched germline (blood) and tumor WGS from the patient revealed 455 somatic single nucleotide mutations [cf. Methods], with the only plausible candidate to be a

driver in the gene *KMT2D*. The mutation is NM_003482:exon32:c.8047-2A>G, a splicing variant disrupting the acceptor site of exon 33. This variant is predicted as pathogenic by SpliceAI, ADA and RF [17, 18]. This variant is not reported in ClinVar and absent from gnomAD exome and genome databases. It is reported as rs2120513666 in dbSNP. Absent in germline, in the tumor tissue the variant presents with an allelic ratio of 50% indicating a tumor purity close to 100% in absence of structural variants in this region (Fig. 2A-B).

Structural alterations

Indeed differential coverage analysis [cf. Methods] between germline and tumor DNA showed extensive alteration in chromosomes 1, 4, 15, 16, 18 and 22 (full chromosomal deletions) as well as chromosomes 7, 9 and 13 (full chromosomal amplifications) but not in chromosome 12 (Fig. 2C).

Single-cell RNA seq

Tumor gene count matrix was generated by CellRanger (version 6.1.0) and processed with Seurat for quality control, normalization, dimensionality reduction, clustering and annotation [cf. Methods]. 2928 cells were retained after stringent quality control with a mean number of detected genes of 2031 and mean unique molecular identifiers (UMIs) of 7348 per cell. Initial clustering displayed 13 clusters (Fig. 3A) which after automatic and manual annotation were separated into tumor cells (3 clusters), immune cells (6 clusters) and structural cells (4 clusters). In depth analysis enabled us to separate the tumor clusters into *GHI* and *PRL* expressing cells, the immune cells into CD4 T cells, CD8 T cells and monocytes (with subcategories, cf. Immune cell analysis) and structural cells into stem cells, pericytes, fibroblasts and endothelial cells (Fig 3B).

Alterations at single-cell level

We validated the chromosomal alterations using Copykat in single cells (version 1.1.0). This software predicts the aneuploid or diploid status of each individual cell allowing for a clear separation of tumor cells versus normal cells. The resulting heatmap showed essentially the same pattern of altered chromosomes we observed in WGS with little differences mostly due to the increased coverage in WGS with respect to single-cell transcriptomics (Fig. 3C).

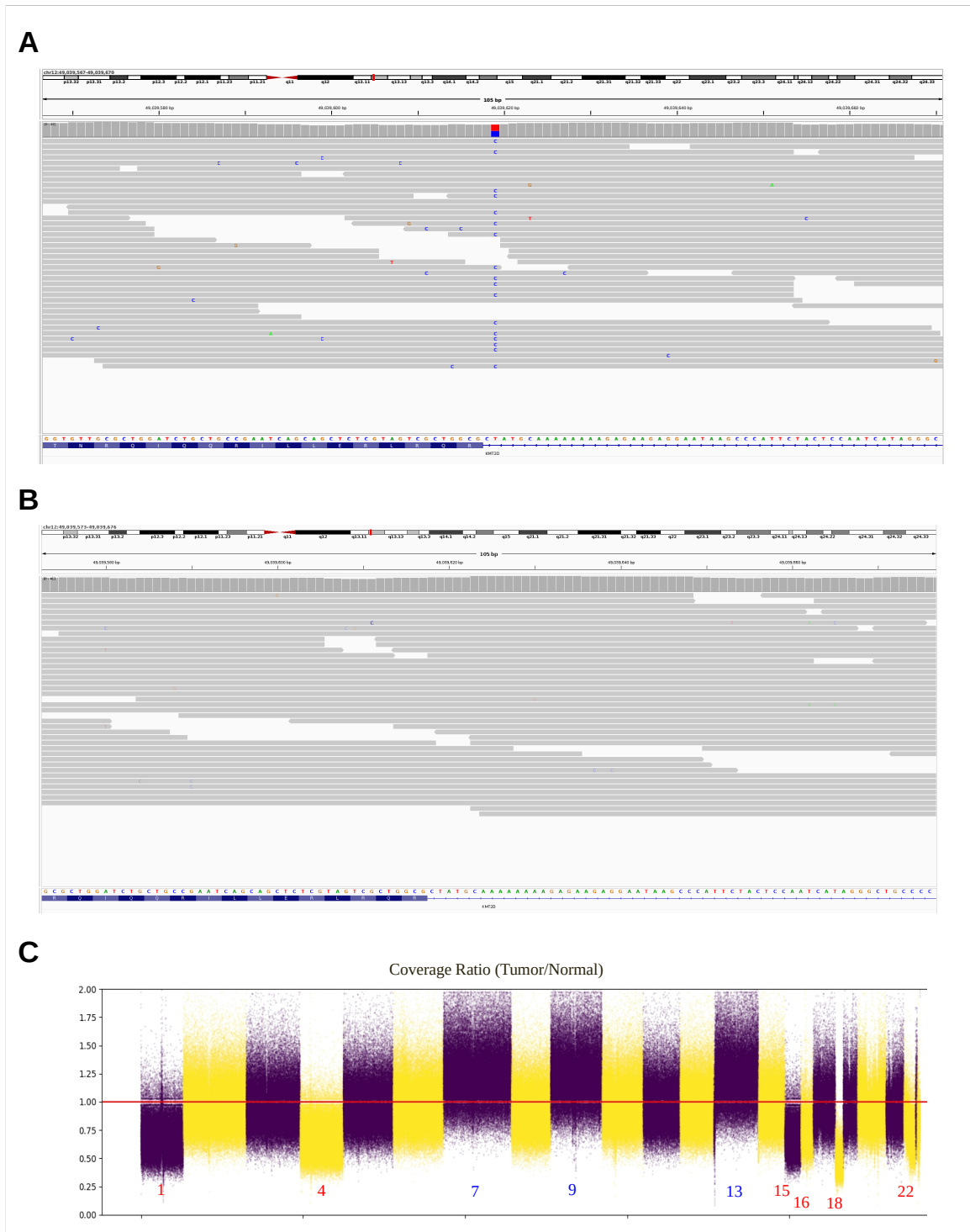


Figure 2: **A:** IGV output of Patient 16's tumor WGS coverage showing the somatic mutation. **B** IGV output of Patient 16's germline WGS coverage. **C:** \log_2 Matched germline and tumor coverage ratio.

***KMT2D* expression in PitNETs**

The fraction of cells expressing *KMT2D* was computed in all PitNET sc-RNA seq data sets and also in an integrated reference (somatotroph and lactotroph cells from the single-cell data set of 3 healthy anterior pituitary gland [19])(Fig. 3D). A consistently lower fraction was detected

in patient 16 compared to all the other tumors and reference indicating a lower transcript abundance due to the non-sense mediated decay of the aberrant spliced transcript.

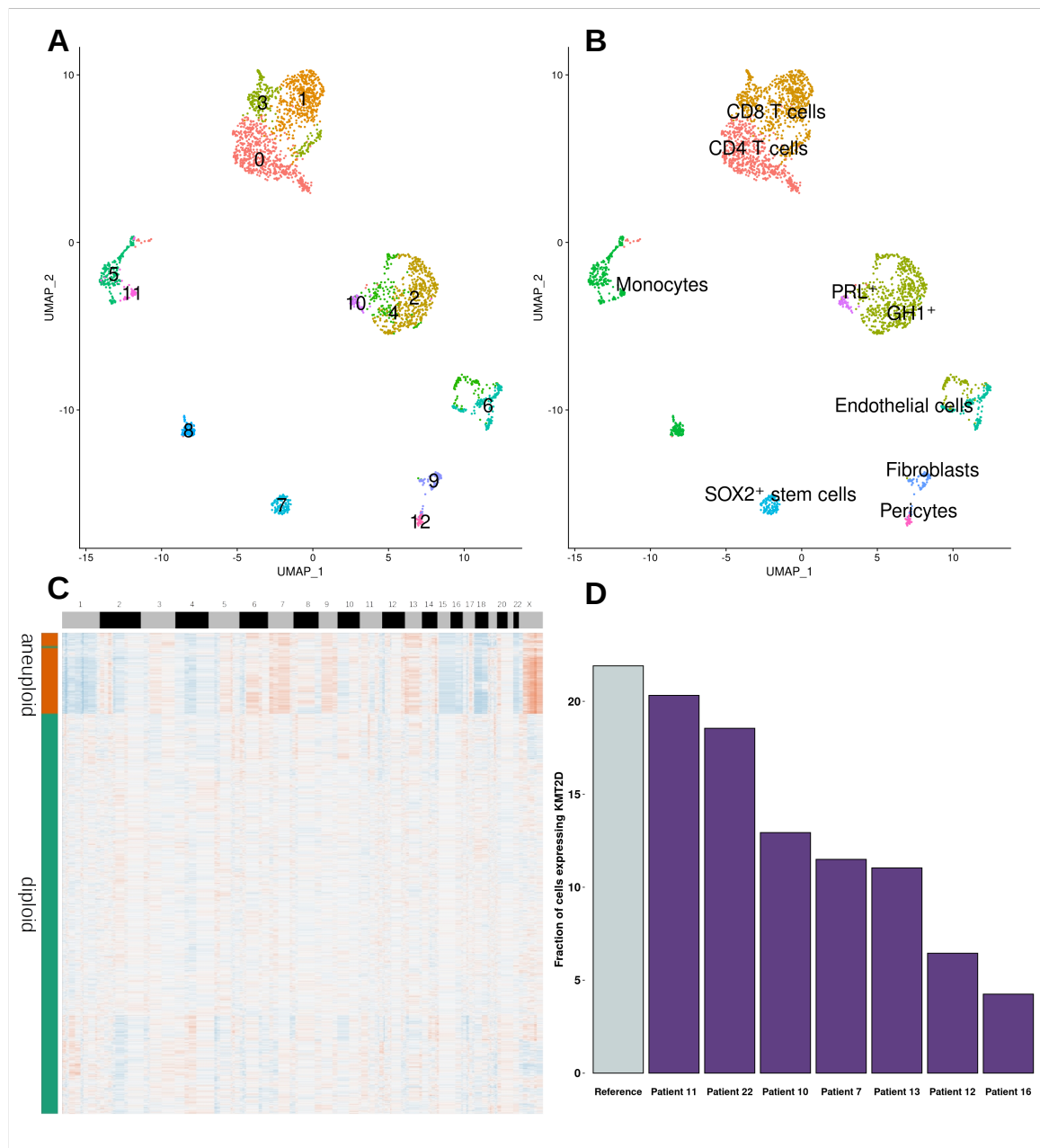


Figure 3: **A**: Initial clustering with 13 clusters. **B**: Annotated UMAP based on canonical markers and automated annotation pipeline. **C**: Copykat heatmap representing chromosomal alterations. Blue for deletion and orange for amplification. Y axis bar represent predicted aneuploid and diploid cells. **D**: Number of cells expressing *KMT2D* in different PitNET samples and the healthy reference.

Differential gene expression analysis (DGEA)

We performed a differential gene expression analysis between the integrated reference and the tumor cells using a pseudobulk approach [cf. methods]. Hallmark gene sets (H) from the Broad Institute were selected to detect up- and down-regulated pathways. Results showed up-regulation

of multiple terms (Fig. 4A) which can be grouped into different categories; energy metabolism (Oxidative phosphorylation (NES: 4.18, adjusted P value: $1.26e^{-23}$), adipogenesis (NES: 2.39, adjusted P value: $1.1e^{-4}$), proliferation and cell growth (MYC targets (NES: 3.91, adj. $p = 6.67e^{-20}$), MTORC1 signaling (NES: 2.73, adj. $p = 2.16e^{-20}$), E2F targets (NES: 2.32, adj. $p = 5.59e^{-4}$), DNA damage and repair mechanisms (DNA repair (NES: 2.9, adj. $p = 9.61e^{-7}$), P53 pathway (NES: 2.4, adj. $p = 3.4e^{-4}$), Apoptosis (NES: 2.25, adj. $p = 1.51e^{-3}$)), stress response and survival (Unfolded protein response (NES: 2.41, adj. $p = 4.33e^{-4}$), ROS pathway (NES: 2.33, adj. $p = 7.21e^{-4}$)), immune response and inflammation (Interferon gamma (NES: 2.55, adj. $p = 1.02e^{-4}$) and alpha response (NES: 2.4, adj. $p = 6.19e^{-4}$), TNFA signaling (NES: 2.38, adj. $p = 4.33e^{-4}$)) and cell signaling pathways (UV Response up (NES: 2.69, adj. $p = 3.23e^{-5}$). Down-regulated pathways are KRAS signaling down (NES:-2.17, adj. $p = 2.09e^{-3}$) and UV Response down (NES: -2.11, adjusted P value: $2.89e^{-3}$). Interestingly, the GO term "KRAS signaling down" indicates a reduced expression of of known KRAS downregulated genes. Indeed, KRAS appears up-regulated in the tumor (log2FC: 1.47, adjusted P value: $3.34e^{-35}$).

Immune cells analysis

While exploring the general composition of patient 16's sc-RNA seq, we noticed a peculiar distribution of cell-types compared to the other PitNETs (i.e. increased number of immune cells proportional to the tumor cells). In this specific single cell sample there are more than 60% immune cells while generally in PitNETs, the proportion is around 5% (Fig. 4B). To determine the composition of immune cell subcategories, we used Azimuth (<https://app.azimuth.hubmapconsortium.org>) for label transfer (i.e. automatic annotation) was used. It allows for the prediction of the most likely cell type compared to a comprehensive list of peripheral blood mononuclear cells (PBMCs) [20]. Predicted cell-types were effector memory CD8 T cells (37.6%), central memory CD4 T cells (26.9%), CD14 monocytes (15.9%), effector memory CD4 T cells (6.8%), central memory CD8 T cells (4.8%) and different, low proportion immune cell-types that we grouped in an "other" category (7.9%). Then, we performed the same analysis for the other PitNETs as comparison and results showed a lower fraction of monocytes, a higher fraction of central memory CD4 T cells as well as the presence of effector memory CD4 T cells (absent in all others). In addition, there is a general increase in cytotoxic activity (both effector and central memory CD8 T cells). Lastly, there is a complete absence of hematopoietic stem and progenitor cells (HSPCs - present in all others) (Fig. 4C).

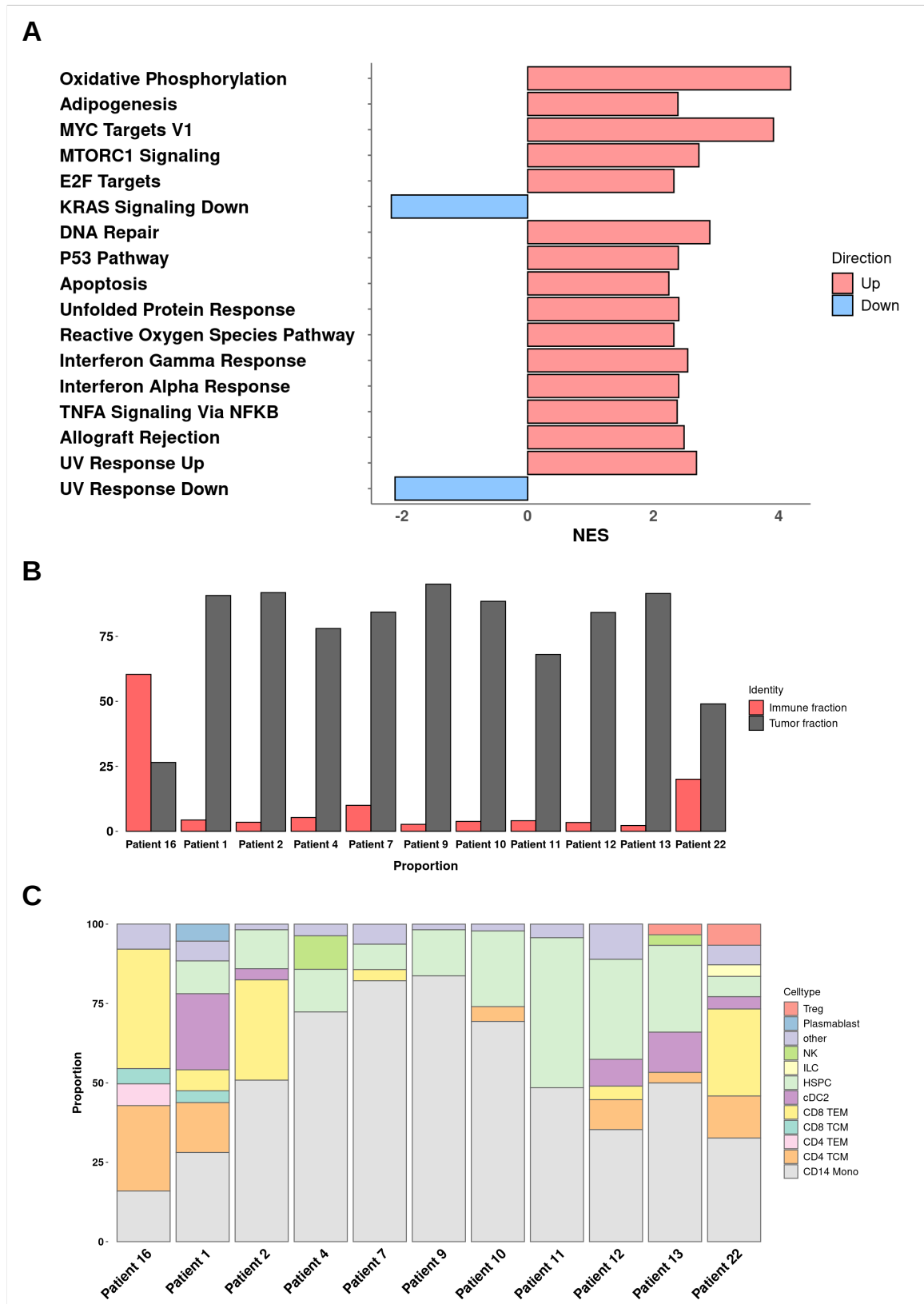


Figure 4: **A**: Differential gene expression analysis results showing HALLMARKS up- and down-regulated between patient 16 and reference. **B**: Tumor and immune cells proportion in PitNET samples. **C**: Immune cells composition in the PitNETs predicted by Azimuth.

Methylation profile in PitNETs

To investigate the epigenetic landscape of pituitary tumors, we conducted a genome-wide DNA methylation analysis across our cohort of PitNETs. We focused our analysis on the 1000 most variable methylation probes to identify significant patterns. The boxplots illustrate the methylation levels of these probes across all tumor samples in our cohort (Fig. 5). Our analysis confirmed the previously reported hypomethylation in POU1F1-PIT1 lineage [21], which includes somatotroph, lactotroph, and thyrotroph tumors. These tumors, including Patient 16's somatotroph tumor, consistently exhibited lower overall methylation levels across the most variable probes compared to other tumor types. Notably, despite the presence of the *KMT2D* mutation and extensive chromosomal alterations in Patient 16's tumor, we did not detect significant DNA methylation differences between this tumor and other somatotroph tumors.

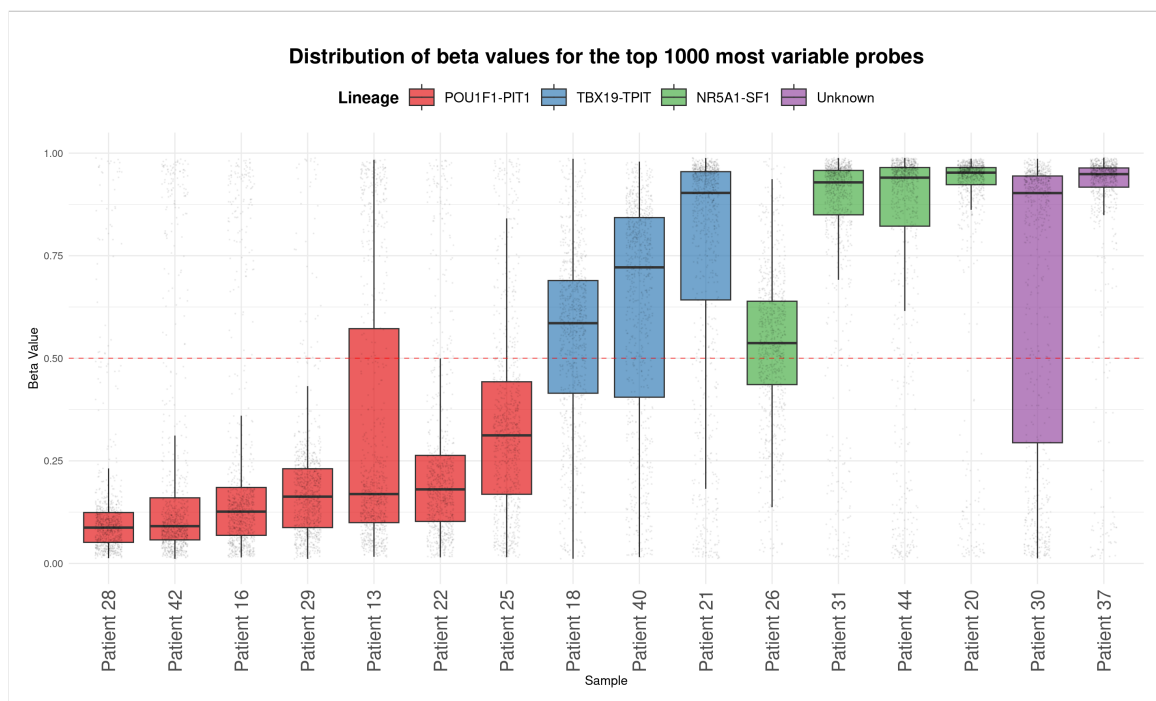


Figure 5: Boxplot representing beta values of the top 1000 most variable probes in 16 PitNETs. POU1F1 lineage includes somato- and lactotroph tumors. TBX19 lineage produces corticotroph tumors and NR5A1 gonadotroph tumors (Non-functioning).

Discussion

In this case report, we present a patient in his 50s with a secreting mixed somato-lactotroph tumor of the pituitary gland exhibiting extensive chromosomal alterations. Variant calling from whole-genome sequencing (WGS) data revealed a splicing variant in *KMT2D*. This histone methyltransferase enzyme is associated with active gene transcription and chromatin regulation. In normal conditions, it acts as a tumor suppressor gene, preventing uncontrolled cell growth and division [22]. The loss of function of *KMT2D* disrupts histone methylation, leading to dysregulation of DNA damage repair and gene expression [23]. This disruption increases chromosomal instability, which can potentially lead to severe forms of cancer [24–26]. However, it has not been previously reported as a driver for pituitary neuroendocrine tumors (PitNETs). Subsequent single-cell RNA sequencing (scRNA-seq) analyses of the tumor revealed unusual results compared to other PitNETs. Firstly, a reduced number of *KMT2D* transcripts were detected compared to a healthy reference for the action of the nonsense mediated decay (NMD) pathway on the aberrantly spliced transcript. Differential gene expression (DGE) analysis showed up-regulation of several pathways associated with tumor maintenance and progression, such as oxidative phosphorylation, DNA repair mechanisms, inflammation and immune response, and KRAS signaling. This suggests similarities between this adenoma and more aggressive carcinomas [ref]. The extensive chromosomal alterations observed in this patient are the consequence of a significant chromosomal instability, which is often a hallmark of cancer progression. Chromosomal instability can lead to an increased mutation rate and the activation of oncogenic pathways. In this case, the up-regulation of KRAS related pathways is noteworthy. KRAS is a well-known oncogene, and its activation can drive tumorigenesis through various mechanisms, including promoting cell proliferation and survival, metabolic reprogramming, and evasion of immune surveillance. However, unlike in lung cancer where up-regulated KRAS disrupts the circadian gene *PER2* leading to increased activation of glycolytic pathways [27], this disruption is not observed in our patient's tumor. This difference may be tissue-specific, indicating that the role of KRAS activation varies across different tumor types. Immune cell analysis between patient 16 and other PitNETs highlighted an inverse ratio (more immune cells than tumor cells), indicating an increased involvement of the tumor microenvironment. Moreover, immune infiltration was different, with increased immune surveillance and cytotoxic activity (higher numbers of CD4 and CD8 T cells). We speculate that this might counteract a possible transition to malignancy and that the tumor microenvironment, including immune cell infiltration, plays a

significant role in preventing the adenoma from progressing to carcinoma. Immunohistochemical analysis revealed light PD-L1 expression in approximately 40% of tumor cells in this pituitary neuroendocrine tumor. While PD-L1 expression is often associated with immune evasion in various cancer types, its significance in PitNETs appears to be more nuanced. Our review of other pathology reports indicates that PD-L1 expression levels in PitNETs can be highly variable, with even some indolent adenomas showing weak expression in up to 80% of cells. Thus, PD-L1 expression alone may not be a reliable indicator of tumor aggressiveness or invasiveness in this context. However, when considered alongside other findings—such as the *KMT2D* mutation, extensive chromosomal alterations detected by WGS, and the unusually diverse immune cell infiltrate observed in single-cell analysis—it contributes to a complex picture of tumor biology. While these collective findings don't necessarily indicate malignancy, they highlight the tumor's unique characteristics and emphasize the need for comprehensive, multi-modal analysis in tumor characterization. Our methylation analysis further supports this complex picture of tumor biology. We analyzed the 1000 most variable methylation probes across our cohort of pituitary tumors. Interestingly, we observed that somatotroph tumors, including Patient 16's tumor, consistently showed lower overall methylation profiles compared to the other tumor types. This hypomethylation pattern in somatotroph tumors suggests a distinct epigenetic landscape that may contribute to their unique cellular characteristics and behavior. Notably, despite the presence of the *KMT2D* mutation and extensive chromosomal alterations in Patient 16's tumor, we did not find significant methylation differences between this tumor and other somatotroph tumors. This suggests that the *KMT2D* mutation in this context may not dramatically alter the overall methylation profile beyond what is typically observed in somatotroph tumors. Interestingly, although epigenetic alterations, along with strong drivers as *KMT2D* and significant genetic and transcriptomic changes may contribute to bridge the gap between pituitary adenomas and carcinomas, PitNETs, as in the presented case, appear resistant to malignant transitions. This suggests that the interplay between genetic mutations, epigenetic regulation, and the tumor microenvironment, in the context of the particular anatomical location outside but yet close to the brain, may be implicated in the maintenance of PitNETs as adenomas. Understanding this peculiar tumor behavior may offer new avenues for targeted therapies or diagnostic approaches in the future.

Competing interests

The Authors declare no competing financial or non-financial interests

Data availability

Single cell processed data are available in NCBI GEO under the accession number XXX.

Author contributions

F.S designed and supervised the study, M.B., J.M. and Maude.M. performed the experiments, A.M. helped with the experiments, M.B. analysed the single cell and methylation data, S.O. and F.S. analysed the genetic data, Mahmoud.M. and R.D. provided the tumors and clinical data, H.E. and J.P.B. analysed the tissues and wrote the clinical reports, M.B. and F.S. wrote the manuscript. All authors contributed to the manuscript.

Funding

This work was supported by the Swiss National Science Foundation (310030_185292), Horizon2020 (847941) and Novartis Foundation for medical-biological research (18A052) to F.S.

Methods

Tissue collection and dissociation

PitNETs are resected at the university hospital in Lausanne, Switzerland. After the operation, the department of pathology takes half of the fresh tumor to perform routine analyses and the clinical report. The other part is dedicated to this project. To dissociate the cells we are taking advantage of Worthington papain dissociation kit. In short, tumor is first minced to break the strongest adherence between tissue and cells. Then pieces are mixed with the digestion mix (EBSS, papain, DNase and collagenase) and heat up at 37°C for 30 minutes in a shaker set to 60 RPM. Every 10 minutes, the liquid is passed through a 10 mL pipette to further separate the cells from the tissues. At the end of the incubation, remaining visible parts are mechanically dissociated with a 200µl pipette tip set on a 1000µl tip and pressed to the side of the tube to ensure maximum dissociation thank to laminar flow. Next, sample is filtered on a 40µm cell strainer and centrifuged for 5 minutes at 300g. The pellet is resuspended in 500µl diluted 1/10 ovomucoid (trypsin inhibitor) to stop the digestion process. Next steps involve gradient centrifugation in concentrated ovomucoid for 6 minutes at 70g to get rid of aggregates and dead cells. Pellet is then resuspended in 600µl of cold HBSS and filtered a second time on a 40µm cell strainer. Cells are counted before the second gradient centrifugation in order to resuspend them in the desire volume of HBSS to hit a range of 800-1200 cells per µl.

Single-cell capture, cDNA library preparation, and sequencing

For this project, we used the 10X Chromium workflow. A Chromium Next GEM Chip G (10X Genomics) is loaded with approximately 6'000 cells and sequencing libraries are prepared strictly following the manufacturer's recommendations (manual CG000204 revD). Briefly, an emulsion encapsulating single cells, reverse transcription reagents, and cell barcoding oligonucleotides are generated. After the reverse transcription step, the emulsion is broken, and double-stranded cDNA are generated and amplified for 12 cycles in a bulk reaction. The cDNA is fragmented, a P7 sequencing adaptor is ligated, and a 3' gene expression library generated by PCR amplification for 12-14 cycles depending on the initial cDNA amount. Libraries were quantified using a fluorimetric method (Q-Bit), and their quality was assessed on a Fragment Analyzer (Agilent Technologies). Sequencing was performed with DNBSEQ™ technology in BGI Europe (Denmark). Primary data processing and assembly was performed with the Cell Ranger Gene Expression pipeline (version 6.1.0, 10X Genomics) with GRCh38 (refdata-gex-GRCh38-2020-A)

as transcriptome reference.

DNA extraction and WGS

Extraction of genomic DNA is done with AllPrep DNA/RNA Mini Kit (Qiagen) following standard protocol. Tumor and matched blood samples have been whole genome sequenced by BGI Poland with PE150 30X/90Gb on DNBSEQ machines. Alignment and mapping has been performed by a custom pipeline based on Sentieon v202308.03 [28] with GRCh38 human genome reference. For somatic variant discovery TNScope algorithm from Sentieon [29] was used in a tumor-normal matching sample mode. From resulting variant set any variant not passing any TNScope filter was removed as potentially of non-tumor origin.. Annotation of resulting vcf files with tumor variants has been performed with Annovar annotation tool using dbSNP v150, SpliceAI, gnomad exonic v 2.11 and gnomad genomic v 3.0 databases [30].

Match germline tumor coverage ratio

Analysis of structural variations has been performed through tumor vs matched germline (blood) coverage ratio. Germline single nucleotide variants with allelic ratio = 0.5 and quality score > 400 have been matched with overlapping tumor variants and the log₂ ratio of the normalized coverages per variant reported.

Sample processing

Data processing was done with R (version 4.2.1). For each samples, the same processing is used for quality control and downstream analyses. Quality control is performed using SingleCellExperiment (version 1.26.0), DropletUtils (version 1.24.0), scuttle (version 1.14.0), AnnotationDbi (version 1.66.0) and scDbFinder (version 1.18.0). Shortly, data object is created and genes are mapped to chromosomes using EnsDb.Hsapiens.v86 (version 2.99.0). Next, cells expressing less than 200 genes are removed from the analysis. Then to account for outliers in mitochondrial gene expression, isOutlier is used with default parameters (Median Absolute Deviation = 3). Afterward, to detect and remove putative doublets, scDbFinder with dbr.sd = 1 is used. Finally, the main processing was done with Seurat [31] (version 5.0.3) using high quality cells as input. Data normalization is performed with SCTransform() followed by linear dimension reduction (RunPCA(), default parameter), non-linear dimension reduction (RunUMAP(), with dims = 1:20, neighbors detection with Share Nearest Neighbor graph (SNN graph, FindNeighbors() with dims same as above) and finally cluster detection (FindClusters() with default parameters).

First round of cell-type assignment was done by i. extracting markers per cluster (FindMarkers(), default parameters) and comparing them with current knowledge, ii. literature reviewing.

Integration and differential gene expression analysis

Prior to perform differential gene expression analysis, integration of 3 data sets of healthy adult pituitary gland to create a unified data set is performed with Seurat. Briefly, individual data set previously processed with the above described workflow are merged into a single object. Then, object layers are integrated with IntegrateLayers() using RPCAIntegration as integration method. Rest of the processing is the same.

Differential gene expressions analysis is then performed in a pseudobulk because of an excess of false positive in the default Seurat workflow [32]. Thus we used DESeq2 (version 1.44.0) package for bulk DGEA. Briefly, each data set (i.e. the integrated reference and patient 16) were split into four pseudo-replicates with an equivalent number of cells and all clusters were pulled together and considered as one single cluster. Then, DESeq2 with default parameter was applied. Genes not present in a least 20% of the cells were removed and differentially expressed genes with adjusted p-value of < 0.01 and \log_2 fold change > 1 & < -1 are kept.

Finally, to perform gene ontology, fgsea [33] (version 1.30) was used with H: hallmark gene sets and C5: ontology gene sets from GSEA MSigDB collections from the Broad Institute (<https://www.gsea-msigdb.org/gsea/msigdb/collections.jsp>). Number of permutation (nPerm-Simple) was set to 10000 and pathways with an adjusted p-value < 0.05 are kept.

Azimuth label transfer

For the immune cell analysis, we annotated them using Azimuth webtool and the Human PBMCs reference (<https://app.azimuth.hubmapconsortium.org/app/human-pbmc>). Azimuth is a set of single-cell RNA and single-cell ATAC seq references that can be used to annotate cells of a query data set. For the PBMCs reference, it contains 3 layers of increased cell-type specificity (layer 1 is more general compared to layer 2 and so on). For our analysis, we used layer 2 and plot the data using ggplot2 (version 3.5.1).

Methylation analysis

The methylation assay was performed by the IGE3 genomics platform of Geneva University in Switzerland using Illumina Infinium Methylation Assay. In silico processing was done with the Sesame R package (1.23.8) using default parameters.

Ethical approval

The study has been performed in accordance with the Declaration of Helsinki and approved by the Commission cantonale d'éthique de la recherche sur l'être humain (Vaud) ref. 2019-02033.

All patients have signed the informed consent.

References

- [1] Shereen Ezzat et al. "The prevalence of pituitary adenomas: A systematic review". In: *Cancer* 101 (3 Aug. 2004), pp. 613–619. ISSN: 0008543X. DOI: 10.1002/cncr.20412.
- [2] Quinn T Ostrom et al. "CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2012–2016". In: *Neuro-Oncology* 21.Supplement₅ (Nov. 2019), pp. v1–v100. ISSN: 1522-8517. DOI: 10.1093/neuonc/noz150. eprint: https://academic.oup.com/neuro-oncology/article-pdf/21/Supplement_5/v1/30340077/noz150.pdf. URL: <https://doi.org/10.1093/neuonc/noz150>.
- [3] Dieter K. Lüdecke et al. "Pathohistological classification of pituitary tumors: 10 years of experience with the German Pituitary Tumor Registry". In: *European Journal of Endocrinology* 156 (2 Feb. 2007), pp. 203–216. ISSN: 08044643. DOI: 10.1530/eje.1.02326.
- [4] Mark E. Molitch. "Diagnosis and treatment of pituitary adenomas: A review". In: *JAMA - Journal of the American Medical Association* 317 (5 Feb. 2017), pp. 516–524. ISSN: 15383598. DOI: 10.1001/jama.2016.19699.
- [5] Yong Hwy Kim and Jung Hee Kim. "Transcriptome analysis identifies an attenuated local immune response in invasive nonfunctioning pituitary adenomas". In: *Endocrinology and Metabolism* 34 (3 2019), pp. 314–322. ISSN: 20935978. DOI: 10.3803/EnM.2019.34.3.314.
- [6] Giulia Cossu et al. "PD-L1 Expression in Pituitary Neuroendocrine Tumors/Pituitary Adenomas". In: *Cancers* 15 (18 Sept. 2023). ISSN: 20726694. DOI: 10.3390/cancers15184471.
- [7] Sylvia L. Asa et al. *Overview of the 2022 WHO Classification of Pituitary Tumors*. Mar. 2022. DOI: 10.1007/s12022-022-09703-7.
- [8] Albert Beckers et al. *Familial isolated pituitary adenomas (FIPA) and the Pituitary adenoma predisposition due to mutations in the aryl hydrocarbon receptor interacting protein (AIP) gene*. 2013. DOI: 10.1210/er.2012-1013.

- [9] Crystal D.C. Kamilaris and Constantine A. Stratakis. *Multiple endocrine neoplasia type 1 (MEN1): An update and the significance of early genetic and clinical diagnosis*. 2019. DOI: 10.3389/fendo.2019.00339.
- [10] Raitis Peculis et al. "Whole exome sequencing reveals novel risk genes of pituitary neuroendocrine tumors". In: *PLoS ONE* 17 (8 July Aug. 2022). ISSN: 19326203. DOI: 10.1371/journal.pone.0265306.
- [11] Zeng Yi Ma et al. "Recurrent gain-of-function USP8 mutations in Cushing's disease". In: *Cell Research* 25 (3 Mar. 2015), pp. 306–317. ISSN: 17487838. DOI: 10.1038/cr.2015.20.
- [12] Cristina L Ronchi et al. "Landscape of somatic mutations in sporadic GH-secreting pituitary adenomas". In: *European Journal of Endocrinology* 174.3 (Mar. 2016), pp. 363–372. ISSN: 0804-4643. DOI: 10.1530/EJE-15-1064. eprint: <https://academic.oup.com/ejendo/article-pdf/174/3/363/45518417/363.pdf>. URL: <https://doi.org/10.1530/EJE-15-1064>.
- [13] Yi Rou Wang et al. "Kabuki syndrome: review of the clinical features, diagnosis and epigenetic mechanisms". In: *World Journal of Pediatrics* 15 (6 Dec. 2019), pp. 528–535. ISSN: 18670687. DOI: 10.1007/s12519-019-00309-4.
- [14] Emily Tinsley et al. *KMT2C and KMT2D aberrations in breast cancer*. June 2024. DOI: 10.1016/j.trecan.2024.02.003.
- [15] Yuanwang Pan et al. "KMT2D deficiency drives lung squamous cell carcinoma and hypersensitivity to RTK-RAS inhibition". In: *Cancer Cell* 41 (1 Jan. 2023), 88–105.e8. ISSN: 18783686. DOI: 10.1016/j.ccell.2022.11.015.
- [16] Jie Li et al. "Loss of CREBBP and KMT2D cooperate to accelerate lymphomagenesis and shape the lymphoma immune microenvironment". In: *Nature Communications* 15 (1 Dec. 2024). ISSN: 20411723. DOI: 10.1038/s41467-024-47012-1.
- [17] Kishore Jaganathan et al. "Predicting Splicing from Primary Sequence with Deep Learning". In: *Cell* 176 (3 Jan. 2019), 535–548.e24. ISSN: 10974172. DOI: 10.1016/j.cell.2018.12.015.
- [18] Xueqiu Jian, Eric Boerwinkle, and Xiaoming Liu. "In silico prediction of splice-altering single nucleotide variants in the human genome". In: *Nucleic Acids Research* 42.22 (Nov. 2014), pp. 13534–13544. ISSN: 0305-1048. DOI: 10.1093/nar/gku1206. eprint: <https://academic.oup.com/nar/article-pdf/42/22/13534/17424019/gku1206.pdf>. URL: <https://doi.org/10.1093/nar/gku1206>.

- [19] Qilin Zhang et al. "Single-cell sequencing identifies differentiation-related markers for molecular classification and recurrence prediction of PitNET". In: *Cell Reports Medicine* 4 (2 Feb. 2023). ISSN: 26663791. DOI: 10.1016/j.xcrm.2023.100934.
- [20] Yuhan Hao et al. "Integrated analysis of multimodal single-cell data". In: *Cell* 184 (13 June 2021), 3573–3587.e29. ISSN: 10974172. DOI: 10.1016/j.cell.2021.04.048.
- [21] Mario Neou et al. "Pangenomic Classification of Pituitary Neuroendocrine Tumors". In: *Cancer Cell* 37 (1 Jan. 2020), 123–134.e5. ISSN: 18783686. DOI: 10.1016/j.ccell.2019.11.002.
- [22] Shilpa S Dhar and Min Gyu Lee. *Cancer-epigenetic function of the histone methyltransferase KMT2D and therapeutic opportunities for the treatment of KMT2D-deficient tumors*. 2021. URL: www.oncotarget.com.
- [23] Dong Xu Wang et al. *Mutation status of the KMT2 family associated with immune checkpoint inhibitors (ICIs) therapy and implicating diverse tumor microenvironments*. Dec. 2024. DOI: 10.1186/s12943-023-01930-8.
- [24] Jing Xu et al. "KMT2D Deficiency Promotes Myeloid Leukemias which Is Vulnerable to Ribosome Biogenesis Inhibition". In: *Advanced Science* 10.19 (2023), p. 2206098. DOI: <https://doi.org/10.1002/advs.202206098>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/advs.202206098>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/advs.202206098>.
- [25] Shidong Lv et al. "Histone methyltransferase KMT2D sustains prostate carcinogenesis and metastasis via epigenetically activating LIFR and KLF4". In: *Oncogene* 37 (10 Mar. 2018), pp. 1354–1368. ISSN: 14765594. DOI: 10.1038/s41388-017-0026-x.
- [26] R. Tyler Hillman et al. "KMT2D/MLL2 inactivation is associated with recurrence in adult-type granulosa cell tumors of the ovary". In: *Nature Communications* 9 (1 Dec. 2018). ISSN: 20411723. DOI: 10.1038/s41467-018-04950-x.
- [27] Hunain Alam et al. "KMT2D Deficiency Impairs Super-Enhancers to Confer a Glycolytic Vulnerability in Lung Cancer". In: *Cancer Cell* 37 (4 Apr. 2020), 599–617.e7. ISSN: 18783686. DOI: 10.1016/j.ccell.2020.03.005.
- [28] Donald Freed et al. "The Sentieon Genomics Tools - A fast and accurate solution to variant calling from next-generation sequence data". In: *bioRxiv* (2017). DOI: 10.1101/115717. eprint: <https://www.biorxiv.org/content/early/2017/05/12/115717.full.pdf>. URL: <https://www.biorxiv.org/content/early/2017/05/12/115717>.

- [29] Donald Freed, Renke Pan, and Rafael Aldana. “TNscope: Accurate Detection of Somatic Mutations with Haplotype-based Variant Candidate Detection and Machine Learning Filtering”. In: *bioRxiv* (2018). DOI: 10.1101/250647. eprint: <https://www.biorxiv.org/content/early/2018/01/19/250647.full.pdf>. URL: <https://www.biorxiv.org/content/early/2018/01/19/250647>.
- [30] Siwei Chen et al. “A genomic mutational constraint map using variation in 76,156 human genomes”. In: *Nature* 625 (7993 Jan. 2024), pp. 92–100. ISSN: 14764687. DOI: 10.1038/s41586-023-06045-0.
- [31] Yuhan Hao et al. “Dictionary learning for integrative, multimodal and scalable single-cell analysis”. In: *Nature Biotechnology* 42 (2 Feb. 2024), pp. 293–304. ISSN: 15461696. DOI: 10.1038/s41587-023-01767-y.
- [32] Jordan W. Squair et al. “Confronting false discoveries in single-cell differential expression”. In: *Nature Communications* 12 (1 Dec. 2021). ISSN: 20411723. DOI: 10.1038/s41467-021-25960-2.
- [33] Gennady Korotkevich et al. “Fast gene set enrichment analysis”. In: *bioRxiv* (2021). DOI: 10.1101/060012. eprint: <https://www.biorxiv.org/content/early/2021/02/01/060012.full.pdf>. URL: <https://www.biorxiv.org/content/early/2021/02/01/060012>.