

Genomic and transcriptomic analysis of pituitary adenomas reveals the impacts of copy number variations on gene expression and clinical prognosis among prolactin-secreting subtype

Yiyuan Chen¹, Hua Gao¹, Weiyan Xie¹, Jing Guo¹, Qiuyue Fang¹, Peng Zhao², Chunhui Liu², Haibo Zhu², Zhuang Wang³, Jichao Wang⁴, Songbai Gui^{2,5,6}, Yazhuo Zhang^{1,2,5,6}, Chuzhong Li^{1,2,5,6}

¹Department of Cell Biology, Beijing Neurosurgical Institute, Capital Medical University, Beijing 100070, China

²Department of Neurosurgery, Beijing Tiantan Hospital, Capital Medical University, Beijing 100070, China

³Department of Neurosurgery, The First Affiliated Hospital of Zhengzhou University, Zhengzhou 450052, China

⁴Department of Neurosurgery, People's Hospital of Xinjiang Uygur Autonomous Region, Xinjiang 830001, China

⁵China National Clinical Research Center for Neurological Diseases, Beijing 100070, China

⁶Brain Tumor Center, Beijing Institute for Brain Disorders, Beijing 100070, China

Correspondence to: Chuzhong Li; **email:** lichuzhong@ccmu.edu.cn

Keywords: pituitary adenoma, WGS, CNV, bromocriptine resistance

Received: August 8, 2020

Accepted: September 29, 2020

Published: December 19, 2020

Copyright: © 2020 Chen et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/3.0/) (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

Pituitary adenomas (PAs) are slow growing and benign primary intracranial tumors that often cause occupying effects or endocrine symptoms. PAs can be classified into various subtypes according to hormone secretion. Although widespread transcriptional alterations that cause aberrant hormone secretion have been characterized, the impact of genomic variations on transcriptional alterations is unclear due to the rare occurrence of single-nucleotide variations in PA. In this study, we performed whole-genome sequencing (WGS) on 76 PA samples across three clinical subtypes (PRL-PAs; GH-PAs, and NFPAs); transcriptome sequencing (RNA-seq) of 54 samples across these subtypes was also conducted. Nine normal pituitary tissues were used as controls. Common and subtype-specific transcriptional alterations in PAs were identified. Strikingly, widespread genomic copy number amplifications were discovered for PRL-PAs, which are causally involved in transcriptomic changes in this subtype. Moreover, we found that the high copy number variations (CNVs) in PRL-PA cause increased prolactin production, drug resistance and proliferative capacity, potentially through key genes with copy number amplification and transcriptional activation, such as *BCAT1*. This study provides insight into how genomic CNVs affect the transcriptome and clinical outcomes of PRL-PA and sheds light on the development of potential therapeutics for aberrantly activated targets.

INTRODUCTION

Pituitary adenomas (PAs) are primary intracranial tumors that occur in 0.1% of adults [1, 2]. PAs can be classified according to the presence of abnormal hormone secretion (functional PAs, 36%-54%; clinical nonfunctional PAs, NFPAs, 46%-64%) [3, 4]. Functional PAs can be further divided into subtypes according to hormone secretion

status, and prolactin-secreting PAs (PRL-PAs, 32%-51%) and growth hormone-secreting PAs (GH-PAs, 9%-11%) are the two most common subtypes accounting for a large percentage of functional PAs. Adrenocorticotropin-secreting PAs (ACTH-PAs, 3%-6%) and thyrotropin-secreting PAs (TSH-PAs, <1%) can cause obvious clinical symptoms, but their incidence is lower than that of PRL-PAs and GH-PAs [2, 5, 6]. Most gonadotropin

adenomas (GON-PAs) and pluri-hormonal PAs are NFPAs. The majority of pituitary tumors are benign and exhibit arrested cell cycle as well as aberrant growth factor signaling [7, 8].

Transcription level alterations have been linked to abnormal hormone secretion in different subtypes of PAs [9]. In ACTH-PAs, the proopiomelanocortin (POMC), T-Box transcription factor 19 (TBX19, TPIT) and epidermal growth factor receptor (EGFR) genes are reportedly upregulated [10]; aberrant splicing of estrogen-related receptor gamma (ESRRG) leads to stronger binding to POU class 1 homeobox 1 (POU1F1, PIT-1) and excessive prolactin secretion [11]. Transcriptomic and epigenomic approaches have been employed to illustrate divergent patterns of gene expression in several subtypes of PAs [12]. However, the global changes in gene expression in PAs are still under investigation due to the lack of normal pituitary tissue as a control.

Genomic analyses have identified genes with somatic single-nucleotide variations (SNVs) in PAs, e.g., Ubiquitin-specific peptidase 8 (USP8) in ACTH-PAs, G protein α_s (GNAS) in GH-PAs and splicing factor 3b subunit 1 (SF3B1) in PRL-PAs [13–16]. However, these genomic studies suggested that PAs are associated with low mutation burdens; in addition, the frequencies of the three recurrent mutations were quite low, indicating a limited impact of SNVs on widespread gene expression alterations in PAs [12]. The impact of other genomic alterations on transcriptomic changes, such as copy number variations, remains to be investigated.

In this study, we performed whole-genome sequencing (WGS) and transcriptomic sequencing (RNA-seq) on PRL-PAs, GH-PAs and NFPAs to identify subtype-specific genomic and transcriptomic alterations. Normal pituitary tissues were also used to identify common gene expression abnormalities in PAs. Common gene expression alterations were detected in PAs, including genes in neuronal pathways and growth pathways. Widespread and unrecognized genomic copy number amplifications were identified in PRL-PAs, contributing to specific transcriptional activation in numerous genes and worse clinical outcomes of PRL-PA patients.

RESULTS

Transcriptional landscape of pituitary tumors

We performed transcriptome sequencing (RNA-seq) on 54 PA samples (21 PRL-PAs, 11 GH-PAs, and 23 NFPAs) and 9 normal pituitary tissues (Supplementary Table 1). Hierarchical clustering showed that these

specimens clustered according to their clinical groups, suggesting widespread transcriptomic alterations in PAs and across PA subtypes and that transcriptional signatures can be used for molecular classification of PA subtypes. The similarity heatmap suggests that NFPAs exhibit more dramatic transcriptomic alterations (Figure 1A). Consistently, the number of DEGs between NFPAs and other groups (3800~4200) was significantly larger than the number of DEGs between other comparison groups (Supplementary Figures 1, 2). This result suggests that despite a lack of abnormal hormone secretion, NFPAs are characterized by more widespread gene expression alterations.

A total of 448 common DEGs were found to be shared across PA subtypes compared to normal pituitary tissue (Figure 1B). These genes are enriched with neuronal pathways, including neural active ligand-receptor binding and axon guidance (Figure 1C). The common DEGs are also enriched in five growth factor signaling pathways (Wnt, TGF- β , PI3K-AKT, Hippo, and STAT3-JAK), all of which have been linked to PA pathogenesis or therapy [17–21]. These genes commonly changed across PA subtypes are also involved in cytokine production pathways (Figure 1C), consistent with decreased CD8⁺ T cell infiltration in both functional PAs and NFPAs (Figure 1D, 1E) and suggesting suppression of the immune response as a common etiology of PAs.

Transcriptional signatures reveal altered pathways across PA subtypes

To identify transcriptional signatures associated with each PA subtype, 54 transcriptomic datasets were used to construct a weighted gene coexpression network [22] consisting of 62 modules with a size of 30~1500 genes (Figure 2A). Analysis of module trait association revealed modular alterations of gene expression in each subtype (Figure 2B and Supplementary Figure 3). The two abnormal hormone-secretion subtypes share some coexpression modules (e.g., MEturquoise, MEgreen and MEdarkgreen), and the module MEturquoise genes was enriched with pathways related to protein synthesis (Figure 2C, 2D), indicating that abnormal macromolecule biosynthesis may be the basis of aberrant hormone production and secretion. Several coexpression modules were significantly associated with the nonfunctional PA subtype, e.g., the Module purple showed enrichment of genes in the insulin signaling pathway (Figure 2E). As patients with hormone-abnormal PA usually develop insulin resistance and glucose abnormalities [23], activated insulin signaling in NFPAs might antagonize metabolic disorders, thus preventing excessive hormone secretion. The high expression of genes in subtype-correlated modules was confirmed by the heatmaps (Figure 2F).

Genomic copy number amplifications causally involved in gene transcriptional activation in PRL-PAs

Next, we attempted to explore genomic alterations underlying the observed transcriptomic changes in PA patients. WGS data from 76 PA samples and paired blood samples were analyzed. We noticed a similar low mutation burden in these PA samples, consistent with previous observations. Some previously reported SNVs associated with PAs, such as *GNAS* and 1/11 in GH-

PAs, were identified with a low frequency of occurrence (Supplementary Material 1). Strikingly, we found widespread and recurrent copy number variations, especially amplifications, in PRL-PAs. The high-CNV feature occurred in nearly half of the PRL-PA cases but rarely in other PA subtypes (Figure 3A). Specific genomic copy number amplification was not mentioned in a previous exome sequencing analysis of PAs, possibly because of the low number of PRL-PA patients involved.

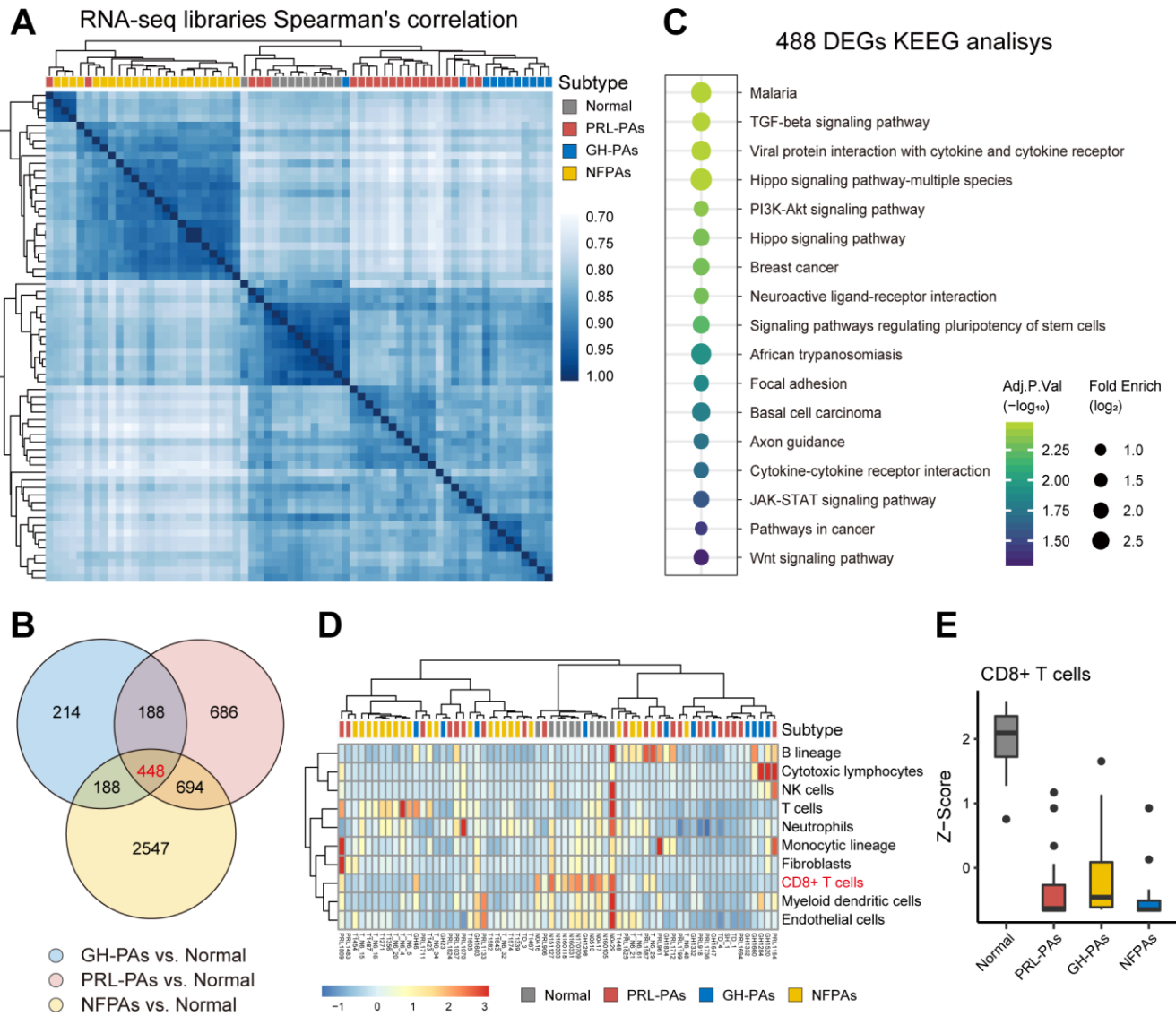


Figure 1. The transcriptional landscape of PAs. (A) Correlation heatmap of transcriptomic similarity among 21 PRL-PAs, 11 GH-PAs, 23 NFPAs and 9 normal pituitary tissues (Normal). Pituitary adenoma subtype is indicated by the color bar above the heatmap. (B) Venn diagram showing the intersection of DEGs among three subtypes of PAs vs. Normal. DEGs were identified by the R package DESeq2 under the cutoff of adjusted P value < 0.05. (C) KEGG pathway enrichment analysis of 448 overlapping DEGs, the dot plot shows pathways with an adjusted P value < 0.05. (D) The infiltration of eight subtypes of immune cell populations and two endothelial cell types in PA samples and normal pituitary tissues was evaluated using the expression levels of cell type specific markers using the MCP-counter [38]. The abundances of each cell types were normalized by z transformation across samples. (E) Boxplots of z score from Figure 1D showing the reduced infiltration of CD8⁺ T cells was reduced across PA subtypes compared to normal pituitary tissues.

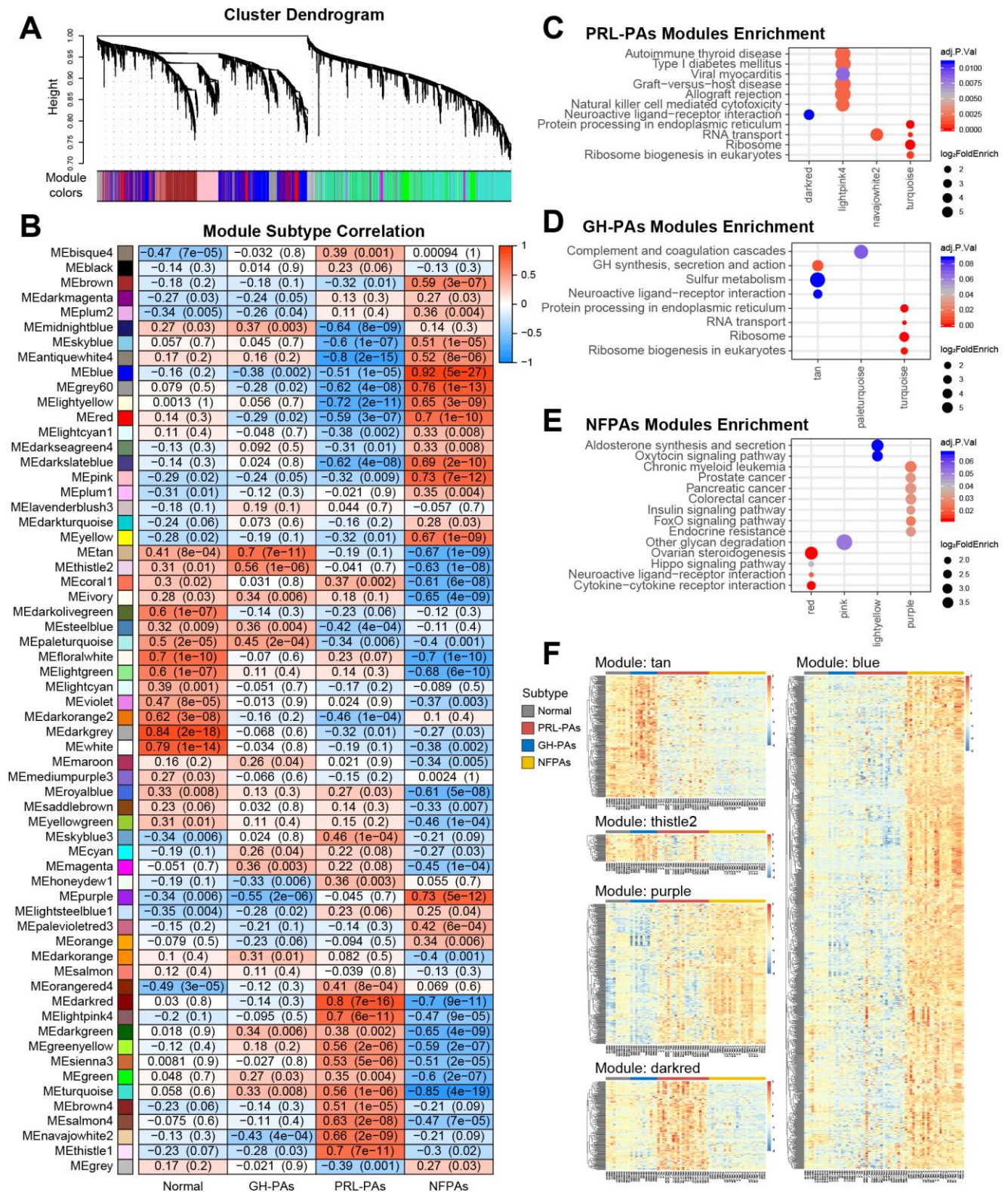


Figure 2. Weighted gene correlation network analysis (WGCNA) of transcriptome data reveals subtype specific coexpression modules. (A) WGCNA cluster dendrogram groups genes into distinct gene coexpression modules defined by dendrogram branch cutting. **(B)** Modules-subtype correlation heatmap of three PA subtypes and normal tissues. The cells in the heatmap were colored by the correlation between eigengene expression and each sample group, the correlation coefficients and P values were indicated in each cell. **(C–E)** KEGG pathway enrichment analysis of genes in the modules significantly associated with different PA subtypes. Adjusted P value < 0.05. **(F)** Expression profiles of genes in key modules associated with PA subtypes.

Table 1. Clinico-pathological characteristics of high- and low- CNV PRL-PAs.

Sample	Volume ¹	Invasive/non-Invasive	CNV ²	Drug resistance ³	Ki-67 IHC ⁴
P906	Large	Invasive	High	Resistance	8%
P918	Large	non-Invasive	High	NA	<1%
P1133	Large	non-Invasive	High	NA	2%
P1154	Giant	Invasive	High	NA	>3%
P1037	Giant	Invasive	High	Resistance	3%
P1408	Giant	non-Invasive	High	Resistance	<1%
P1100	Large	non-Invasive	High	Resistance	<1%
P1694	Large	Invasive	High	NA	5-8%
P1711	Large	non-Invasive	High	NA	NA
P1712	Giant	Invasive	High	Resistance	6%
P1736	Giant	Invasive	High	Resistance	8%
P1809	Large	Invasive	High	Resistance	7%
P25502	Large	Invasive	High	Sensitive	3-5%
P28607	Large	Invasive	High	Resistance	1-2%
P2890	Large	non-Invasive	High	NA	2-3%
P1824	Giant	Invasive	High	Sensitive	5-10%
P25505	Large	Invasive	High	Resistance	<1%
P1821	Large	non-Invasive	High	Resistance	NA
P1070	Large	non-Invasive	Low	Sensitive	2%
P961	Giant	Invasive	Low	NA	<1%
P1483	Giant	Invasive	Low	Sensitive	<1%
P1587	Large	Invasive	Low	NA	NA
P_N6_30	Large	non-Invasive	Low	Sensitive	<1%
P1199	Large	Invasive	Low	Resistance	<1%
P1144	Large	Invasive	Low	Sensitive	<1%
P1825	Giant	Invasive	Low	NA	1%
P640829	Large	Invasive	Low	NA	NA
P29115	Large	Invasive	Low	Sensitive	2-5%

¹ Tumor classification by volume, micro: <1 cm, large: 1-4 cm, giant: >4 cm.

² According to the hierarchical clustering of CNV signatures.

³ Drug (BCT) resistance, NA: no drug therapy was administered preoperatively.

⁴ Ki-67 IHC staining, NA: not enough tissue.

The genes located in copy number amplification regions significantly overlapped with the genes specifically upregulated in PRL-PAs compared to other subtypes (Figure 3B), implying a role of genomic CNV in shaping transcriptomic alterations in PRL-PAs. The 498 overlapping genes were enriched in pathways downstream of mTOR signaling, including the biosynthesis of amino acids, lysosome pathway, ribosome and AMPK signaling (Figure 3C). Cyclin-dependent kinase 6 (CDK6) was also transcriptionally activated by copy number amplification, which suggests the role of genomic copy number variation in the out-of-control cell cycle and tumor progression of PAs.

We further divided PRL-PA patients into high-CNV and low-CNV groups according to the hierarchical clustering of CNV signatures. Consistent with the increase in copy number, the expression fold change of

PRL-PA specific upregulated genes was significantly higher in the high-CNV group (Figure 3D), suggesting a causal role of CNVs in transcriptional changes in PRL-PA patients.

Genomic copy number amplifications cause high prolactin production and activation of genes downstream of the mTOR signaling pathway

To investigate the influence of genomic copy number amplification on the clinical outcomes of PRL-PA patients, we used an immunohistochemistry (IHC) staining approach to probe key genes under copy number amplifications in PRL-PAs. The copy number-amplified genes *BCAT1* and *MYC* were more abundant in the high-CNV group of patients (Figure 4A). To further analyze the pathological impacts of genomic copy number variation, we surveyed expression of prolactin in

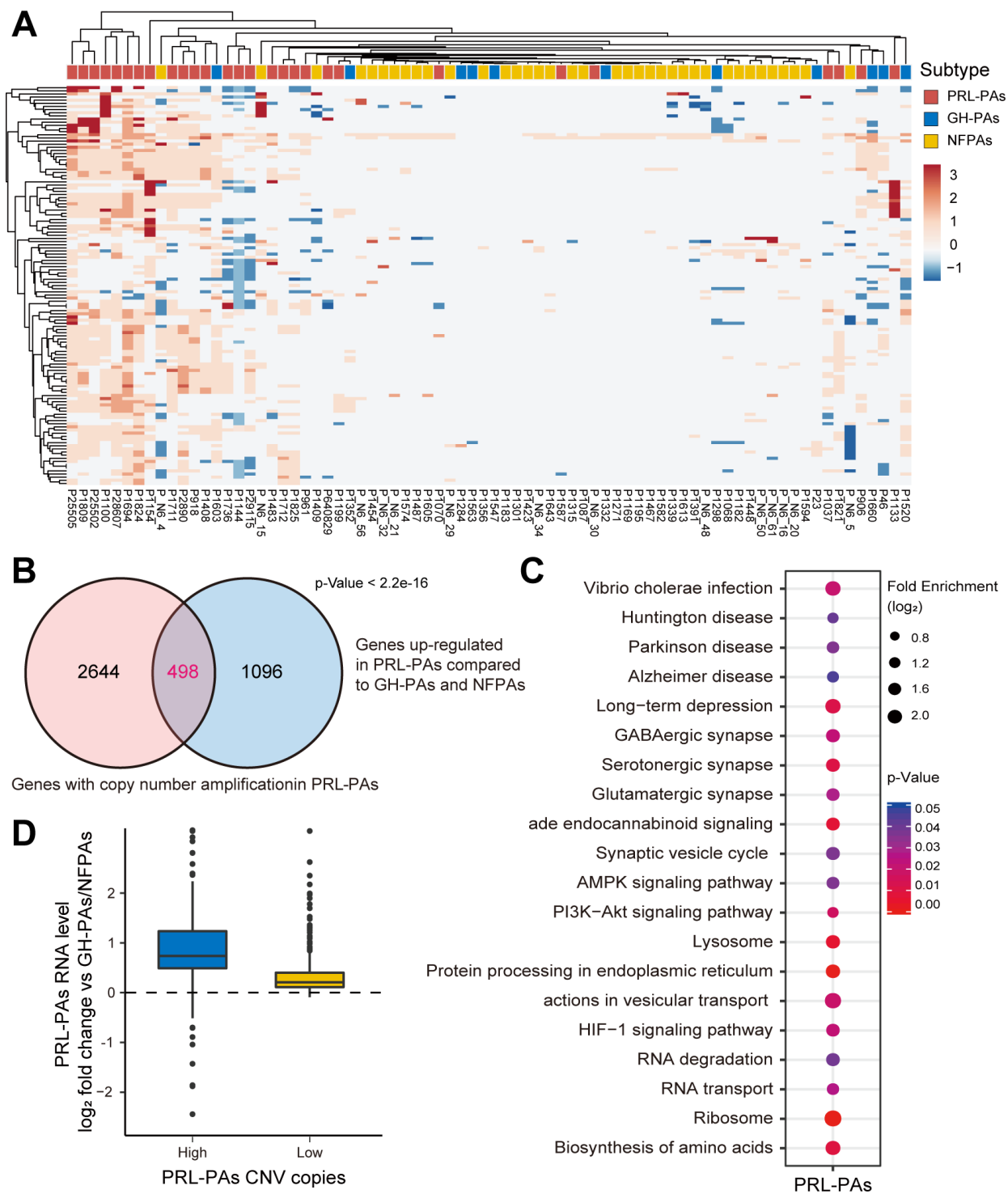


Figure 3. Copy number amplifications in PRL-PAs cause gene transcriptional activation. (A) The heatmap of CNV profiles across PA subtypes. The sample subtype was indicated by color bar above the heatmap, samples were grouped according to the similarity of CNV profiles. The heatmap was colored by CNVs, red indicates gain of copy number and blue indicates loss of copy number. Frequent copy number amplification was observed in PRL-PAs. (B) 498 genes with the copy number amplifications in PRL-PAs overlapped with up-regulated DEGs in the PRL-PAs compared to other subtypes. P value < 2.2e-16 by hypergeometric test. (C) KEGG pathway enrichment analysis of 498 up-regulated genes with both copy number amplifications and transcriptional up-regulation in PRL-PAs, the dot plot shows pathways with a P value < 0.05. (D) PRL-PAs samples were divided into two groups (high CNV and low CNV) according to the clustering results in (A). The boxplot shows the log₂ expression fold-changes of PRL-PAs specific up-regulation genes relative to GH-PAs/NFPAs in the high CNV group and low CNV group. The high CNV group exhibited transcriptional up-regulation (Median log₂ fold change ~0.75) while the low CNV group did not.

patients in the high- and low-CNV groups and found prolactin expression to be 4-fold higher in the former group (Figure 4B). ErbB signaling is a determinant of prolactin production, and the downstream transcription factor of ErbB signaling, MYC, also exhibited a 4-fold increase in the high-CNV group. These results suggest a

significant impact of genomic CNVs on the excessive production of prolactin in PRL-PA patients. The transcriptomic differences between the high- and low-CNV groups also involved numerous ribosome genes and branched-chain amino acid aminotransferase 1 (BCAT1), implicating activation of the mTOR signaling

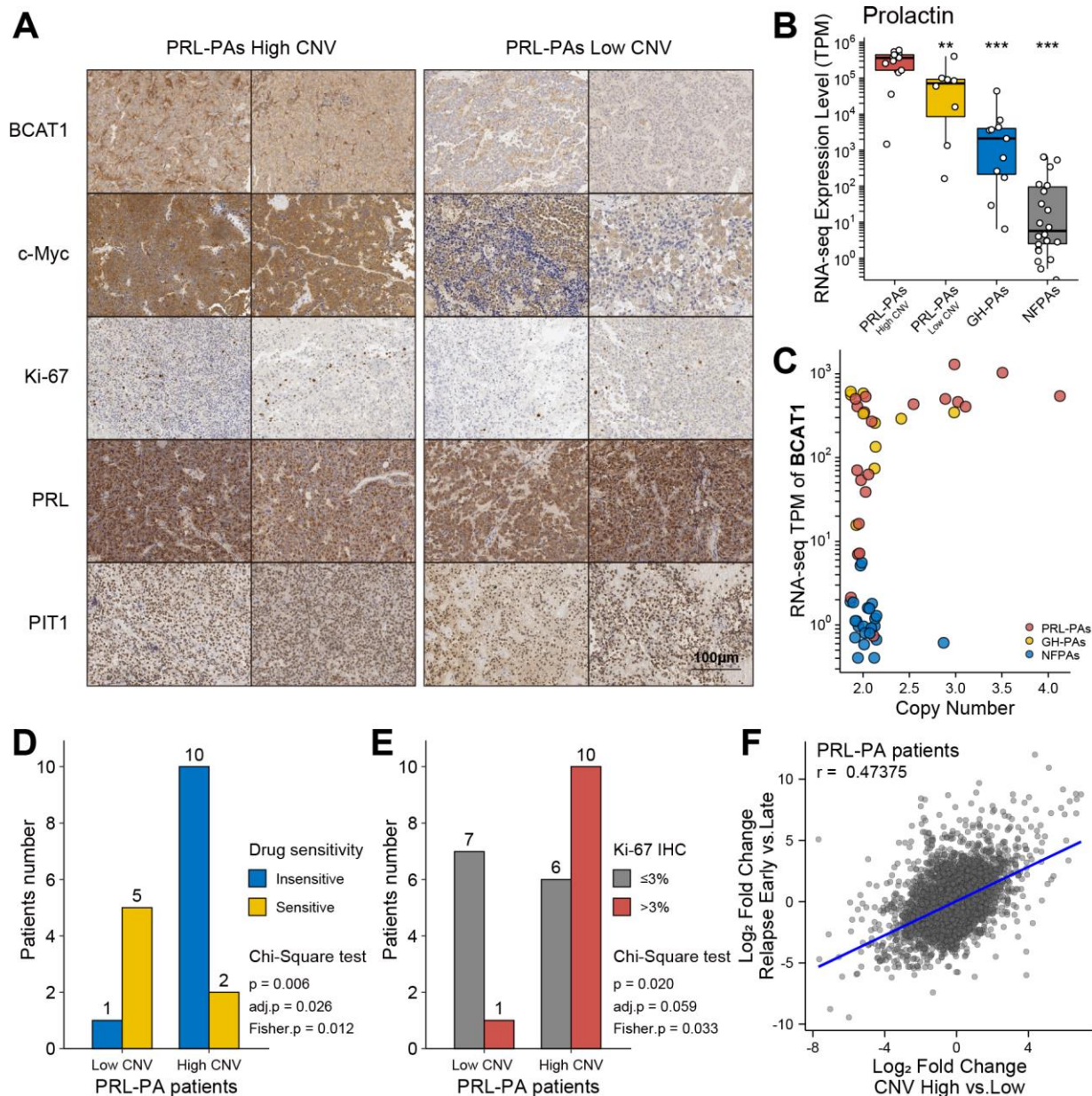


Figure 4. Clinical relevance of genomic copy number variation in PRL-PAs. (A) Immunohistochemistry of BCAT1, MYC, Ki-67, PRL, and PIT1 in high and low CNV PRL-PAs. $\times 100$ magnification, the scale bar = 100 μm . (B) Prolactin expression levels (TPM) in PRL-PAs with high CNV, PRL-PAs with low CNV, GH-PAs and NFPAs. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ (C) Copy number and expression level (TPM) of BCAT1 in different PA subtypes. Both copy number and expression level of BCAT1 were increased in PRL-PAs. (D) PRL-PA patients in the high CNV group more frequently developed drug resistance. The yellow bar indicates the number of patients that are sensitive to BCT treatment, while the blue bar indicates number of patients that are insensitive to the same treatment. P value = 0.026, Chi-Square test. (E) High CNV group in PRL-PAs exhibits a higher degree of malignancy. The PA malignancy was defined by the number of positive Ki-67 foci in IHC. P value = 0.059, Chi-Square test. (F) The Scatter plot shows positive correlation (spearman correlation coefficient: 0.47) of transcriptomic alterations in high CNV group relative to low CNV group (x-axis) and relapsed PAs relative to un-relapsed PAs (y-axis).

pathway in the high-CNV group (Figure 4C). BCAT1 undergoes correlated copy number amplifications and gene expression increments in PRL-PAs. Both BCAT1 and prolactin can activate the mTOR signaling pathway, which leads to excessive ribosome biogenesis.

Genomic copy number amplifications are associated with worse prognosis

Elevated Ki-67 IHC staining in the PRL-PA high-CNV group suggests a worse clinical prognosis (Figure 4A and Table 1). Dopamine agonists (bromocriptine, BCT) are the first choice for PRL-PA treatment (not applicable to patients with rapid visual loss and visual field defects), followed by surgery, which can reduce tumor size and prolactin levels [4, 24]. However, drug resistance arises in a subset of patients, as indicated by a maintained prolactin level or tumor size. The frequency of BCT resistance in patients in the high-CNV group (BCT resistance: 10/12 v.s. BCT sensitive: 1/6; $p = 0.006$, chi-square test) was significantly higher than that in the low-CNV group (Figure 4D). The Ki-67 label index marks aggressive behavior in pituitary tumors, and the high-CNV group (Ki-67 positive $\geq 3\%$: 10/16 vs. Ki-67 positive $< 3\%$: 1/8; P value = 0.02, chi-square test) was associated with a higher Ki-67 label index (Figure 4E). These data collectively support that copy number variations contribute to a worse prognosis in PRL-PA patients.

The transcriptional signature related to relapse in PAs was determined by comparing patients who experienced early relapse (< 36 months) and those without any sign of relapse after a long period (> 60 months). Gene expression patterns in the high-CNV group in PRL-PA patients correlated with gene expression patterns that predict early relapse (Figure 4F), suggesting that the genes regulated by aberrant copy number in PRL-PAs are related to relapse.

DISCUSSION

In this study, we profiled transcriptomic alterations in three major subtypes of PAs revealing shared and subtype-specific alterations. The clinical subtypes of PAs were well separated according to their transcriptional signatures in clustering analysis, suggesting that gene expression abnormalities are an essential part of PA etiology. Interestingly, we found that the PA subtypes with abnormal hormonal secretion were more similar to normal pituitary tissue samples and that the number of DEGs between samples of these functional subtypes and normal pituitary tissues samples was less than that of functional subtypes. These findings suggest that the nonfunctional subtype of PA undergoes more widespread gene expression

alterations, as opposed to aberration of specific hormonal pathways.

We utilized WGCNA to generate coexpression networks in PAs and identified subtype-specific modules. The functional subtypes (GH-PAs and PRL-PAs) shared modules enriched in growth hormone pathways; PRL-PAs were also enriched in ribosome genes and the mitochondrial oxidative phosphorylation pathways, suggesting a high demand of energy metabolism in the PRL-PAs. The transcriptional alterations in nonfunctional PAs were enriched in autophagy genes, and induction of autophagy might be beneficial in limiting PA progression. Indeed, autophagic cell death mediates the effects of bromocriptine and temozolomide on PA [25–27]. Nevertheless, the specific roles of autophagy in antagonizing PA progression remain elusive.

Consistent with previous reports, we observed a lack of high-frequency recurrent mutations in PAs [12]. However, we do demonstrate frequent genomic copy number amplifications in PRL-PAs. Moreover, these genomic CNVs in PRL-PAs play a causal role in PA etiology, including prolactin production, proliferative ability, and drug resistance. The impact of copy number amplifications on PA etiology might be caused by its direct influence on abnormal gene expression upregulation, such as genes in ribosome biogenesis, growth signaling and HIF signaling pathways, which facilitate the hypoxic adaptative ability of PAs. The copy number amplification and upregulated expression of BCAT1 may play a central role in the activation of genes downstream of the mTOR pathway through a feedback loop involving prolactin [28, 29]. In addition, PRL-PA-specific upregulation of genes caused by copy number amplification includes chromogranin B (CHGB), which has been linked to tumor progression in PRL-PAs [30], further supporting the role of genomic copy number variation in PA progression.

Altogether, we investigated the genomic and transcriptomic correlation in PAs in the present study and demonstrated for the first time that high CNVs in PRL-PAs play an important role in tumor development and are significantly associated with poor prognosis. We believe the findings have clinical relevance in defining prognostic subgroups as well as implications for developing new targeted treatments for PRL-PA.

MATERIALS AND METHODS

Patients samples

All samples were obtained following transsphenoidal surgery performed at Beijing Tiantan Hospital from May 2013 to May 2017. The fresh tumor samples were

stored in liquid nitrogen. 28 PRL-PAs, 11 GH-PAs and 37 NFPAs from the study population (age range, 20–75 years) were diagnosed according to clinical features and pathological findings. 9 normal pituitary glands were obtained from a donation program. The study protocols were approved by the Internal Review Board of Beijing Tiantan Hospital affiliated to Capital Medical University and conformed to the ethical guidelines of the Declaration of Helsinki (No. KY-2013-02).

High-throughput sequencing

A total amount of 0.6 µg genomic DNA per sample was used as input material for DNA sample preparation. Sequencing libraries were generated using the Agilent SureSelect Human All Exon V6 kit (Agilent Technologies, CA, USA) following the manufacturer's recommendations, and index codes were added to each sample. Clustering of the index-coded samples was performed using a cBot Cluster Generation System with a HiSeq PE Cluster Kit (Illumina, San Diego, CA, USA) according to the manufacturer's instructions. After cluster generation, the DNA libraries were sequenced using the Illumina HiSeq platform, and 150-bp paired-end reads were generated. A total amount of 2 µg RNA per sample was used as input material for RNA sample preparations. Sequencing libraries were generated using NEBNext® UltraTM RNA Library Prep Kit for Illumina® (NEB, Ipswich, USA) following the manufacturer's recommendations, and index codes were added to attribute sequences to each sample. Clustering of the index-coded samples was performed on a cBot Cluster Generation System using TruSeq PE Cluster Kit v3-cBot-HS (Illumina) according to the manufacturer's instructions. After cluster generation, the library preparations were sequenced using the Illumina HiSeq platform, and 125-bp/150-bp paired-end reads were generated.

Transcriptomic analysis and identification of differentially expressed genes

For RNA sequencing data, the paired-end clean reads were aligned to the human reference genome (hg19) using Hisat2 (v2.0.5) [31]. HTSeq (v 0.11.2) was used to count the read numbers mapped to each gene [32]. Hierarchical clustering analysis of all transcriptomic samples was performed using the pairwise similarity of each pair of samples determined by the Spearman correlation coefficient. Analysis of differentially expressed genes (DEGs) in each pair of comparisons was performed using the R package DESeq2 [33]. The P value of the differential test was corrected by a multiple hypothesis test, and DEGs were determined by controlling the FDR (false discovery rate) < 0.05.

Genomic analysis

For WGS data, valid sequencing data were mapped to the reference human genome (UCSC hg19) by Burrows-Wheeler Aligner (BWA) software to obtain the original mapping results stored in BAM format [34]. If one or one paired read(s) were mapped to multiple positions, the strategy adopted by BWA is to choose the most likely placement. If two or more likely placements are presented, BWA selects one randomly. SAMtools and Picard (<http://broadinstitute.github.io/picard/>) were used to sort BAM files and perform duplicate marking, local realignment, and base quality recalibration to generate the final BAM file. GATK (v3.4) software [35] was employed for SNP calling. Genome regions with significant amplification or deletion in the samples were evaluated by GISTIC [35], and regions with high frequencies were screened, namely, recurrent CNV regions. The higher the GISTIC score is, the higher is the CNV frequency in this area.

Pathway enrichment analysis

KEGG pathway enrichment for functional analysis of gene lists was performed using the clusterProfiler package under R software (version 3.6.0) [36]. The significance of the enrichment analysis was defined using a hypergeometric test, and the resulting P values were corrected for multiple hypothesis testing with the BH (Benjamini & Hochberg, 1995) method [37]. The final reported enriched terms and pathways were filtered according to adjusted P values < 0.05 or P value < 0.05.

Weighted gene expression network analysis (WGCNA)

The PA coexpression network was constructed using the R package WGCNA (v 1.69) [22]. Biweight midcorrelations between each gene pair were calculated to build an adjacency matrix using the formula: $adjacency = (0.5 * (1 + cor))^power$. A thresholding power of 14 was chosen, and the resulting adjacency matrix was converted to a topological overlap (TO) matrix via the TOM similarity algorithm. The genes were hierarchically clustered based on TO similarity. Modules were assigned by the dynamic tree-cutting algorithm with default parameters. Modules were correlated with each group of samples to identify subtype-specific modules, and the correlation coefficients and P values are indicated in figures.

Immunohistochemistry staining

Tissue from PRL-PAs was fixed in 10% formalin and embedded in paraffin. Three core biopsies (2.0 mm in diameter) were selected from the paraffin-embedded

tissue and transferred to tissue microarrays using a semiautomated system (Aphelys MiniCore, Mitogen, UK). The microarrays were cut into 4- μ m sections and incubated with anti-BCAT1 (rabbit monoclonal, 1:600, ab197941, Abcam), anti-c-Myc (rabbit monoclonal, 1:1000, ab32072, Abcam), anti-Ki-67 (rabbit monoclonal, 1:100, ab16667, Abcam), anti-PRL (rabbit polyclonal, 1:300, ab188229, Abcam), and anti-PIT1 (mouse monoclonal, 1:500, sc393943, Santa-Cruz) primary antibodies. Staining intensity was scored as follows: 0, no staining; 1, weak; 2, moderate; and 3, strong staining. An H-score was calculated based on the percentage of positively stained cells at each intensity level using the following formula: $[1 \times (\% \text{ weakly stained cells}) + 2 \times (\% \text{ moderately stained}) + 3 \times (\% \text{ strongly stained cells})]$.

Statistical analysis

Chi test and Fisher's exact test were employed to determine the significance of categorical variables. Comparisons between two groups were performed using Student's unpaired two-tailed t-test. A P value ≤ 0.05 and/or adjusted P value ≤ 0.05 was considered statistically significant.

Ethics approval and consent to participate

The study protocols were approved by the Internal Review Board of Beijing Tiantan Hospital, which was affiliated to Capital Medical University, and conformed to the ethical guidelines of the Declaration of Helsinki (No. KY2016-035-01).

AUTHOR CONTRIBUTIONS

Chuzhong Li, Yazhuo Zhang and Yiyuan Chen conceived the project and chose the technical route, Yiyuan Chen performed the experiments, analyzed the data, and wrote the manuscript. Hua Gao and Weiyan Xie designed the experiments and the technical route, Songbai Gui and Chunhui Liu helped to revise the manuscript. Jing Guo and Qiuyue Fang performed the IHC. Peng Zhao and Haibo Zhu assisted with the diagnostic assessment. Zhuang Wang and Jichao Wang collected the clinical data and specimens. All authors read and approved the manuscript.

ACKNOWLEDGMENTS

The authors thank the laboratory technicians, data collectors, and medical editors.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

FUNDING

This study was financially supported by the National Natural Science Foundation of China (81771489).

REFERENCES

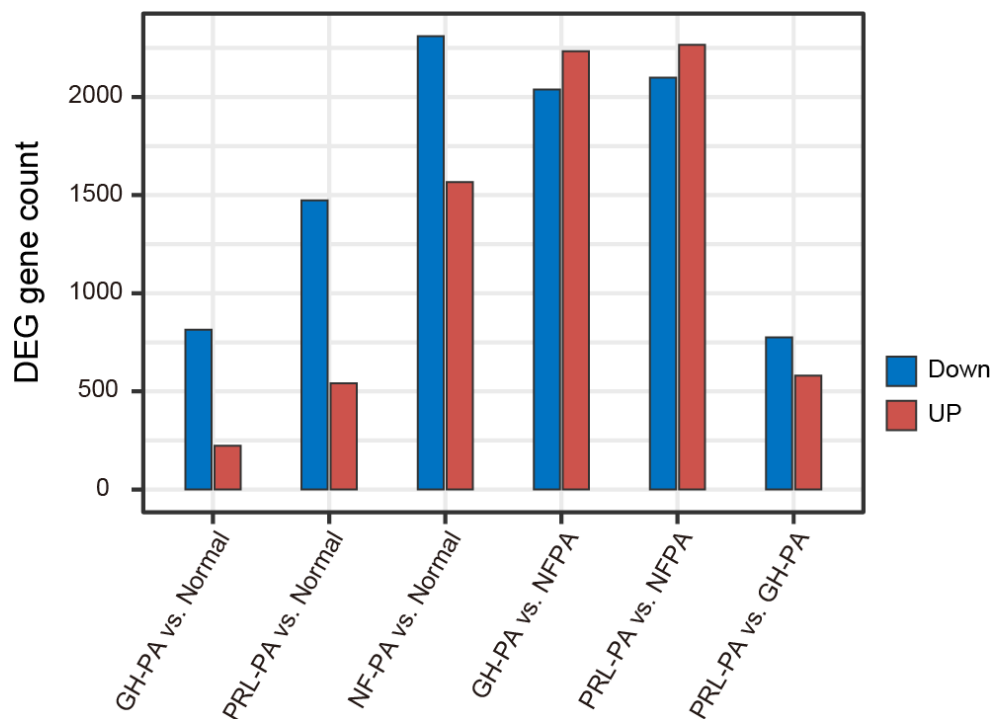
1. Molitch ME. Diagnosis and treatment of pituitary adenomas: a review. *JAMA*. 2017; 317:516–24. <https://doi.org/10.1001/jama.2016.19699> PMID:[28170483](https://pubmed.ncbi.nlm.nih.gov/28170483/)
2. Ezzat S, Asa SL, Couldwell WT, Barr CE, Dodge WE, Vance ML, McCutcheon IE. The prevalence of pituitary adenomas: a systematic review. *Cancer*. 2004; 101:613–19. <https://doi.org/10.1002/cncr.20412> PMID:[15274075](https://pubmed.ncbi.nlm.nih.gov/15274075/)
3. Ntali G, Wass JA. Epidemiology, clinical presentation and diagnosis of non-functioning pituitary adenomas. *Pituitary*. 2018; 21:111–18. <https://doi.org/10.1007/s11102-018-0869-3> PMID:[29368293](https://pubmed.ncbi.nlm.nih.gov/29368293/)
4. Mehta GU, Lonser RR. Management of hormone-secreting pituitary adenomas. *Neuro Oncol*. 2017; 19:762–73. <https://doi.org/10.1093/neuonc/now130> PMID:[27543627](https://pubmed.ncbi.nlm.nih.gov/27543627/)
5. Chanson P, Maiter D. The epidemiology, diagnosis and treatment of prolactinomas: the old and the new. *Best Pract Res Clin Endocrinol Metab*. 2019; 33:101290. <https://doi.org/10.1016/j.beem.2019.101290> PMID:[31326373](https://pubmed.ncbi.nlm.nih.gov/31326373/)
6. Inoshita N, Nishioka H. The 2017 WHO classification of pituitary adenoma: overview and comments. *Brain Tumor Pathol*. 2018; 35:51–56. <https://doi.org/10.1007/s10014-018-0314-3> PMID:[29687298](https://pubmed.ncbi.nlm.nih.gov/29687298/)
7. Moreno CS, Evans CO, Zhan X, Okor M, Desiderio DM, Oyesiku NM. Novel molecular signaling and classification of human clinically nonfunctional pituitary adenomas identified by gene expression profiling and proteomic analyses. *Cancer Res*. 2005; 65:10214–22. <https://doi.org/10.1158/0008-5472.CAN-05-0884> PMID:[16288009](https://pubmed.ncbi.nlm.nih.gov/16288009/)
8. Lazzarini Denchi E, Attwooll C, Pasini D, Helin K. Deregulated E2F activity induces hyperplasia and senescence-like features in the mouse pituitary gland. *Mol Cell Biol*. 2005; 25:2660–72. <https://doi.org/10.1128/MCB.25.7.2660-2672.2005> PMID:[15767672](https://pubmed.ncbi.nlm.nih.gov/15767672/)
9. Seltzer J, Ashton CE, Scotton TC, Pangal D, Carmichael JD, Zada G. Gene and protein expression in pituitary

- corticotroph adenomas: a systematic review of the literature. *Neurosurg Focus*. 2015; 38:E17.
<https://doi.org/10.3171/2014.10.FOCUS14683>
PMID:25639319
10. Tateno T, Izumiyama H, Doi M, Yoshimoto T, Shichiri M, Inoshita N, Oyama K, Yamada S, Hirata Y. Differential gene expression in ACTH -secreting and non-functioning pituitary tumors. *Eur J Endocrinol*. 2007; 157:717–24.
<https://doi.org/10.1530/EJE-07-0428> PMID:18057378
 11. Li C, Xie W, Rosenblum JS, Zhou J, Guo J, Miao Y, Shen Y, Wang H, Gong L, Li M, Zhao S, Cheng S, Zhu H, et al. Somatic SF3B1 hotspot mutation in prolactinomas. *Nat Commun*. 2020; 11:2506.
<https://doi.org/10.1038/s41467-020-16052-8>
PMID:32427851
 12. Salomon MP, Wang X, Marzese DM, Hsu SC, Nelson N, Zhang X, Matsuba C, Takasumi Y, Ballesteros-Merino C, Fox BA, Barkhoudarian G, Kelly DF, Hoon DS. The epigenomic landscape of pituitary adenomas reveals specific alterations and differentiates among acromegaly, cushing's disease and endocrine-inactive subtypes. *Clin Cancer Res*. 2018; 24:4126–36.
<https://doi.org/10.1158/1078-0432.CCR-17-2206>
PMID:30084836
 13. Reincke M, Sbiera S, Hayakawa A, Theodoropoulou M, Osswald A, Beuschlein F, Meitinger T, Mizuno-Yamasaki E, Kawaguchi K, Saeki Y, Tanaka K, Wieland T, Graf E, et al. Mutations in the deubiquitinase gene USP8 cause cushing's disease. *Nat Genet*. 2015; 47:31–38.
<https://doi.org/10.1038/ng.3166> PMID:25485838
 14. Bi WL, Greenwald NF, Ramkissoon SH, Abedalthagafi M, Coy SM, Ligon KL, Mei Y, MacConaill L, Ducar M, Min L, Santagata S, Kaiser UB, Beroukheim R, et al. Clinical identification of oncogenic drivers and copy-number alterations in pituitary tumors. *Endocrinology*. 2017; 158:2284–91.
<https://doi.org/10.1210/en.2016-1967>
PMID:28486603
 15. Bi WL, Horowitz P, Greenwald NF, Abedalthagafi M, Agarwalla PK, Gibson WJ, Mei Y, Schumacher SE, Ben-David U, Chevalier A, Carter S, Tiao G, Brastianos PK, et al. Landscape of genomic alterations in pituitary adenomas. *Clin Cancer Res*. 2017; 23:1841–51.
<https://doi.org/10.1158/1078-0432.CCR-16-0790>
PMID:27707790
 16. Song ZJ, Reitman ZJ, Ma ZY, Chen JH, Zhang QL, Shou XF, Huang CX, Wang YF, Li SQ, Mao Y, Zhou LF, Lian BF, Yan H, et al. The genome-wide mutational landscape of pituitary adenomas. *Cell Res*. 2016; 26:1255–59.
<https://doi.org/10.1038/cr.2016.114>
PMID:27670697
 17. Gaston-Massuet C, Andoniadou CL, Signore M, Jayakody SA, Charolidi N, Kyeyune R, Vernay B, Jacques TS, Taketo MM, Le Tissier P, Dattani MT, Martinez-Barbera JP. Increased wingless (Wnt) signaling in pituitary progenitor/stem cells gives rise to pituitary tumors in mice and humans. *Proc Natl Acad Sci USA*. 2011; 108:11482–87.
<https://doi.org/10.1073/pnas.1101553108>
PMID:21636786
 18. Recouvreur MV, Camilletti MA, Rifkin DB, Díaz-Torga G. The pituitary TGFβ1 system as a novel target for the treatment of resistant prolactinomas. *J Endocrinol*. 2016; 228:R73–83.
<https://doi.org/10.1530/JOE-15-0451> PMID:26698564
 19. Robbins HL, Hague A. The PI3K/Akt pathway in tumors of endocrine tissues. *Front Endocrinol (Lausanne)*. 2016; 6:188.
<https://doi.org/10.3389/fendo.2015.00188>
PMID:26793165
 20. Xekouki P, Lodge EJ, Matschke J, Santambrogio A, Apps JR, Sharif A, Jacques TS, Aylwin S, Prevot V, Li R, Flitsch J, Bornstein SR, Theodoropoulou M, Andoniadou CL. Non-secreting pituitary tumours characterised by enhanced expression of YAP/TAZ. *Endocr Relat Cancer*. 2019; 26:215–25.
<https://doi.org/10.1530/ERC-18-0330> PMID:30139767
 21. Valiulyte I, Steponaitis G, Skiriute D, Tamasauskas A, Vaitkiene P. Signal transducer and activator of transcription 3 (STAT3) promoter methylation and expression in pituitary adenoma. *BMC Med Genet*. 2017; 18:72.
<https://doi.org/10.1186/s12881-017-0434-3>
PMID:28709401
 22. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008; 9:559.
<https://doi.org/10.1186/1471-2105-9-559>
PMID:19114008
 23. Auriemma RS, De Alcubierre D, Pirchio R, Pivonello R, Colao A. Glucose abnormalities associated to prolactin secreting pituitary adenomas. *Front Endocrinol (Lausanne)*. 2019; 10:327.
<https://doi.org/10.3389/fendo.2019.00327>
PMID:31191454
 24. Maiter D. Management of dopamine agonist-resistant prolactinoma. *Neuroendocrinology*. 2019; 109:42–50.
<https://doi.org/10.1159/000495775>
PMID:30481756
 25. Geng X, Ma L, Li Z, Li J, Li M, Wang Q, Chen Z, Sun Q. Bromocriptine induces autophagy-dependent cell death in pituitary adenomas. *World Neurosurg*. 2017; 100:407–16.

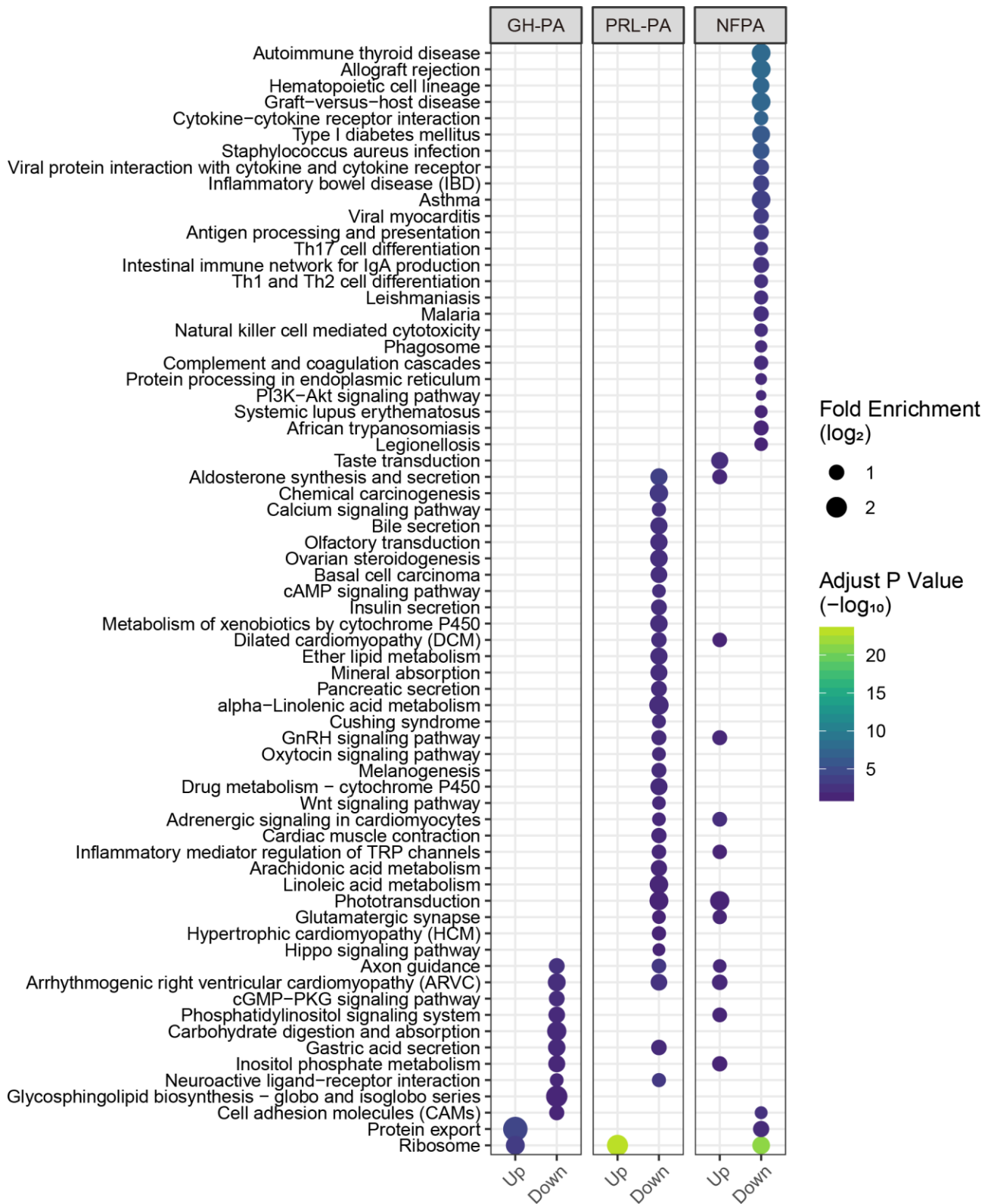
- <https://doi.org/10.1016/j.wneu.2017.01.052>
PMID:[28137551](https://pubmed.ncbi.nlm.nih.gov/28137551/)
26. Kun Z, Yuling Y, Dongchun W, Bingbing X, Xiaoli L, Bin X. HIF-1 α inhibition sensitized pituitary adenoma cells to temozolomide by regulating presenilin 1 expression and autophagy. *Technol Cancer Res Treat*. 2016; 15:NP95–P104.
<https://doi.org/10.1177/1533034615618834>
PMID:[26647409](https://pubmed.ncbi.nlm.nih.gov/26647409/)
27. Leng ZG, Lin SJ, Wu ZR, Guo YH, Cai L, Shang HB, Tang H, Xue YJ, Lou MQ, Zhao W, Le WD, Zhao WG, Zhang X, Wu ZB. Activation of DRD5 (dopamine receptor D5) inhibits tumor growth by autophagic cell death. *Autophagy*. 2017; 13:1404–19.
<https://doi.org/10.1080/15548627.2017.1328347>
PMID:[28613975](https://pubmed.ncbi.nlm.nih.gov/28613975/)
28. Tönjes M, Barbus S, Park YJ, Wang W, Schlotter M, Lindroth AM, Pleier SV, Bai AH, Karra D, Piro RM, Felsberg J, Addington A, Lemke D, et al. BCAT1 promotes cell proliferation through amino acid catabolism in gliomas carrying wild-type IDH1. *Nat Med*. 2013; 19:901–08.
<https://doi.org/10.1038/nm.3217> PMID:[23793099](https://pubmed.ncbi.nlm.nih.gov/23793099/)
29. Mossmann D, Park S, Hall MN. mTOR signalling and cellular metabolism are mutual determinants in cancer. *Nat Rev Cancer*. 2018; 18:744–57.
<https://doi.org/10.1038/s41568-018-0074-8>
PMID:[30425336](https://pubmed.ncbi.nlm.nih.gov/30425336/)
30. Zhang W, Zang Z, Song Y, Yang H, Yin Q. Co-expression network analysis of differentially expressed genes associated with metastasis in prolactin pituitary tumors. *Mol Med Rep*. 2014; 10:113–18.
<https://doi.org/10.3892/mmr.2014.2152>
PMID:[24736764](https://pubmed.ncbi.nlm.nih.gov/24736764/)
31. Baruzzo G, Hayer KE, Kim EJ, Di Camillo B, FitzGerald GA, Grant GR. Simulation-based comprehensive benchmarking of RNA-seq aligners. *Nat Methods*. 2017; 14:135–39.
<https://doi.org/10.1038/nmeth.4106>
PMID:[27941783](https://pubmed.ncbi.nlm.nih.gov/27941783/)
32. Anders S, Pyl PT, Huber W. HTSeq—a python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015; 31:166–69.
<https://doi.org/10.1093/bioinformatics/btu638>
PMID:[25260700](https://pubmed.ncbi.nlm.nih.gov/25260700/)
33. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014; 15:550.
<https://doi.org/10.1186/s13059-014-0550-8>
PMID:[25516281](https://pubmed.ncbi.nlm.nih.gov/25516281/)
34. Li H, Durbin R. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics*. 2009; 25:1754–60.
<https://doi.org/10.1093/bioinformatics/btp324>
PMID:[19451168](https://pubmed.ncbi.nlm.nih.gov/19451168/)
35. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhi R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol*. 2011; 12:R41.
<https://doi.org/10.1186/gb-2011-12-4-r41>
PMID:[21527027](https://pubmed.ncbi.nlm.nih.gov/21527027/)
36. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012; 16:284–87.
<https://doi.org/10.1089/omi.2011.0118>
PMID:[22455463](https://pubmed.ncbi.nlm.nih.gov/22455463/)
37. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc Series B Stat Methodol*. 1995; 57:289–300.
<https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
38. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, Selves J, Laurent-Puig P, Sautès-Fridman C, Fridman WH, de Reyniès A. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol*. 2016; 17:218.
<https://doi.org/10.1186/s13059-016-1070-5>
PMID:[27765066](https://pubmed.ncbi.nlm.nih.gov/27765066/)

SUPPLEMENTARY MATERIALS

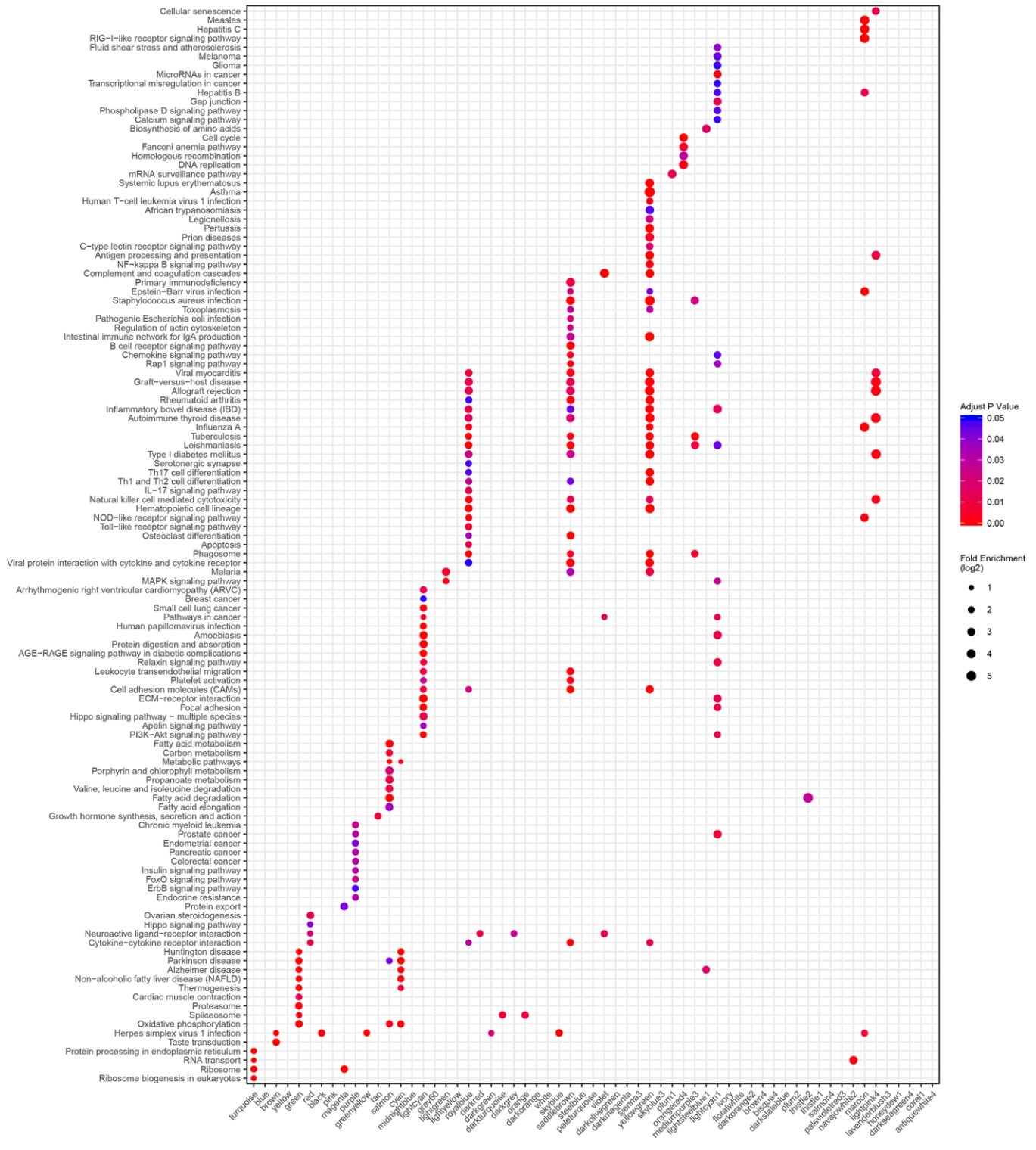
Supplementary Figures



Supplementary Figure 1. Comparison of the number of upregulated DEGs (Red) and downregulated DEGs (Blue) in 6 pairwise DEGs analysis. DESeq2, FDR < 0.05.



Supplementary Figure 2. KEGG pathway enrichment analysis of genes among three PA subtypes vs. Normal. Adjusted P value < 0.05.



Supplementary Figure 3. KEGG pathway enrichment analysis of genes in the modules significantly associated with different PA subtypes. Adjusted P value < 0.05.

Supplementary Table

Supplementary Table 1. Clinical features of 76 PAs.

Sample	Gender	Age (Y)	Clinical Type	Volume ¹	Invasive/non-Invasive	PFS Time ²	Relapse ²	Sequencing strategy
P906	F	25	PRL-PAs	Large	Invasive	81	N	WGS+RNA-seq
P918	M	49	PRL-PAs	Large	non-Invasive	80	N	WGS+RNA-seq
P1133	F	61	PRL-PAs	Large	non-Invasive	66	N	WGS+RNA-seq
P1154	F	46	PRL-PAs	Giant	Invasive	65	N	WGS+RNA-seq
P1037	M	43	PRL-PAs	Giant	Invasive	37	Y	WGS+RNA-seq
P1408	M	45	PRL-PAs	Giant	non-Invasive	52	N	WGS
P1100	M	37	PRL-PAs	Large	non-Invasive	41	Y	WGS
P1694	F	44	PRL-PAs	Large	Invasive	39	N	WGS+RNA-seq
P1711	F	30	PRL-PAs	Large	non-Invasive	-	-	WGS+RNA-seq
P1712	M	31	PRL-PAs	Giant	Invasive	8	Y	WGS+RNA-seq
P1736	M	55	PRL-PAs	Giant	Invasive	17	Y	WGS+RNA-seq
P1809	F	20	PRL-PAs	Large	Invasive	34	N	WGS+RNA-seq
P25502	M	65	PRL-PAs	Large	Invasive	-	-	WGS+RNA-seq
P28607	F	28	PRL-PAs	Large	Invasive	-	-	WGS+RNA-seq
P2890	F	15	PRL-PAs	Large	non-Invasive	-	-	WGS+RNA-seq
P1824	M	22	PRL-PAs	Giant	Invasive	33	N	WGS+RNA-seq
P25505	M	27	PRL-PAs	Large	Invasive	-	-	WGS
P1821	F	22	PRL-PAs	Large	non-Invasive	-	-	WGS
P1070	F	50	PRL-PAs	Large	non-Invasive	71	N	WGS+RNA-seq
P961	M	42	PRL-PAs	Giant	Invasive	53	Y	WGS+RNA-seq
P1483	M	20	PRL-PAs	Giant	Invasive	49	N	WGS+RNA-seq
P1587	F	26	PRL-PAs	Large	Invasive	44	N	WGS+RNA-seq
P_N6_30	F	48	PRL-PAs	Large	non-Invasive	52	N	WGS
P1199	F	41	PRL-PAs	Large	Invasive	4	Y	WGS+RNA-seq
P1144	M	48	PRL-PAs	Large	Invasive	66	N	WGS
P1825	F	38	PRL-PAs	Giant	Invasive	33	N	WGS+RNA-seq
P640829	M	43	PRL-PAs	Large	Invasive	-	-	WGS+RNA-seq
P29115	M	37	PRL-PAs	Large	Invasive	-	-	WGS
P1087	M	46	NFPAs	Giant	Invasive	25	Y	WGS
P1169	F	57	NFPAs	Giant	Invasive	65	N	WGS
P1271	M	56	NFPAs	Large	non-Invasive	59	N	WGS+RNA-seq
P1301	F	55	NFPAs	Large	non-Invasive	14	Y	WGS
P1315	F	50	NFPAs	Large	non-Invasive	57	N	WGS
P1339	M	43	NFPAs	Large	Invasive	55	N	WGS+RNA-seq
P1356	M	47	NFPAs	Large	Invasive	36	Y	WGS+RNA-seq
P1391	M	51	NFPAs	Large	non-Invasive	52	N	WGS
P1409	M	60	NFPAs	Large	Invasive	52	N	WGS
P1423	F	66	NFPAs	Large	non-Invasive	51	N	WGS+RNA-seq
P1448	F	49	NFPAs	Giant	Invasive	44	Y	WGS+RNA-seq
P1454	F	71	NFPAs	Giant	non-Invasive	50	N	WGS+RNA-seq
P1467	F	54	NFPAs	Giant	Invasive	24	Y	WGS+RNA-seq
P1487	M	65	NFPAs	Giant	Invasive	49	N	WGS+RNA-seq
P1574	M	65	NFPAs	Giant	Invasive	3	Y	WGS+RNA-seq
P1582	M	66	NFPAs	Giant	Invasive	45	N	WGS+RNA-seq
P1594	M	67	NFPAs	Large	non-Invasive	44	N	WGS
P1605	F	44	NFPAs	Large	non-Invasive	44	N	WGS+RNA-seq
P1613	M	51	NFPAs	Large	Invasive	43	N	WGS
P1643	M	43	NFPAs	Large	non-Invasive	36	Y	WGS+RNA-seq
P_N6_15	M	49	NFPAs	Giant	Invasive	55	N	WGS+RNA-seq
P_N6_16	M	34	NFPAs	Giant	Invasive	55	N	WGS+RNA-seq
P_N6_20	M	61	NFPAs	Giant	Invasive	54	N	WGS+RNA-seq
P_N6_21	F	51	NFPAs	Large	Invasive	54	N	WGS+RNA-seq
P_N6_29	M	47	NFPAs	Large	non-Invasive	52	N	WGS+RNA-seq

P_N6_32	M	64	NFPAs	Giant	Invasive	52	N	WGS+RNA-seq
P_N6_34	M	55	NFPAs	Giant	Invasive	52	N	WGS+RNA-seq
P_N6_4	F	32	NFPAs	Large	Invasive	57	N	WGS+RNA-seq
P_N6_48	M	64	NFPAs	Giant	Invasive	51	N	WGS+RNA-seq
P_N6_5	F	35	NFPAs	Giant	Invasive	57	N	WGS+RNA-seq
P_N6_50	F	41	NFPAs	Large	Invasive	51	N	WGS
P_N6_56	M	45	NFPAs	Large	Invasive	51	N	WGS
P_N6_61	F	49	NFPAs	Giant	Invasive	49	N	WGS+RNA-seq
P1183	F	57	NFPAs	Giant	Invasive	66	N	WGS
P1068	M	48	NFPAs	Giant	Invasive	19	Y	WGS
P1182	M	67	NFPAs	Large	Invasive	66	N	WGS
P1195	M	35	NFPAs	Large	non-Invasive	65	N	WGS
P1284	F	21	GH-PAs	Giant	Invasive	12	Y	WGS+RNA-seq
P1298	F	69	GH-PAs	Large	non-Invasive	57	N	WGS+RNA-seq
P1332	M	57	GH-PAs	Giant	Invasive	14	N	WGS+RNA-seq
P1352	M	31	GH-PAs	Giant	Invasive	3	Y	WGS+RNA-seq
P23	F	51	GH-PAs	Giant	Invasive	4	Y	WGS+RNA-seq
P46	F	44	GH-PAs	Large	non-Invasive	51	N	WGS+RNA-seq
P1603	M	45	GH-PAs	Large	non-Invasive	44	N	WGS+RNA-seq
P1563	M	32	GH-PAs	Large	Invasive	12	Y	WGS
P1547	F	33	GH-PAs	Large	non-Invasive	46	N	WGS+RNA-seq
P1520	M	21	GH-PAs	Giant	Invasive	17	Y	WGS+RNA-seq
P1660	F	51	GH-PAs	Large	non-Invasive	41	N	WGS+RNA-seq

¹ Tumor classification by volume, micro: <1 cm, large: 1-4 cm, giant: >4 cm.

² Patients were lost to follow-up: -

Supplementary Material

Supplementary Material 1. Sense mutations of CDS in 11 pairs GH-PAs.

Please browse Full Text version to see the data of Supplementary Material 1.