



Integrated Whole-Exome and Transcriptome Sequencing of Sporadic Parathyroid Adenoma

Ya Hu^{1†}, Xiang Zhang^{1†}, Ou Wang², Ming Cui¹, Xiaobin Li¹, Mengyi Wang¹, Surong Hua¹ and Quan Liao^{1*}

¹ Department of General Surgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China, ² Laboratory of Endocrinology, Department of Endocrinology, National Health Commission, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China

OPEN ACCESS

Edited by:

Anna Halama,
Weill Cornell Medicine, Qatar

Reviewed by:

Matthias Kroiss,
Julius Maximilian University
of Würzburg, Germany
Sumeet Pal Singh,
Université libre de Bruxelles, Belgium

*Correspondence:

Quan Liao
lqumc@126.com

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Systems Endocrinology,
a section of the journal
Frontiers in Endocrinology

Received: 20 November 2020

Accepted: 15 March 2021

Published: 14 May 2021

Citation:

Hu Y, Zhang X, Wang O, Cui M, Li X,
Wang M, Hua S and Liao Q (2021)
Integrated Whole-Exome and
Transcriptome Sequencing of
Sporadic Parathyroid Adenoma.
Front. Endocrinol. 12:631680.
doi: 10.3389/fendo.2021.631680

Purpose: Hyperparathyroidism is the third most common endocrine disease. Parathyroid adenoma (PA) accounts for approximately 85% of cases of primary hyperparathyroidism, but the molecular mechanism is not fully understood. Herein, we aimed to investigate the genetic and transcriptomic profiles of sporadic PA.

Methods: Whole-exome sequencing (WES) and transcriptome sequencing (RNA-seq) of 41 patients with PA and RNA-seq of 5 normal parathyroid tissues were performed. Gene mutations and characterized expression changes were identified. To elucidate the molecular mechanism underlying PA, unsupervised consensus clustering of RNA-seq data was performed. The correlations between the sequencing data and clinicopathological features of these patients were analyzed.

Results: Previously reported PA driver gene mutations, such as *MEN1* (9/41), *mTOR* (4/41), *ZFX* (3/41), *CASR* (3/41), *EZH2* (2/41) and *FAT1* (2/41), were also identified in our cohort. Furthermore, somatic mutation of *EZH1*, which had not been reported in PA, was found in 4 samples. RNA-seq showed that the expression levels of 84 genes were upregulated and 646 were downregulated in PA samples compared with normal samples. Unsupervised clustering analysis of RNA-seq data clustered these patients into 10 subgroups related to mutation or abnormal expression of a group of potential pathogenic genes.

Conclusion: *MEN1*, *EZH2*, *CASR*, *EZH1*, *ZFX*, *mTOR* and *FAT1* mutations in PA were revealed. According to the RNA-seq data clustering analysis, cyclin D1, β -catenin, VDR, *CASR* and *GCM2* may be important factors contributing to the PA gene expression profile.

Keywords: parathyroid diseases, hyperparathyroidism, high-throughput, nucleotide sequencing, gene expression profiling

INTRODUCTION

Primary hyperparathyroidism (pHPT) is a common endocrine disorder with an incidence of approximately 66 and 25 per 100000 person-years among women and men, respectively (1). It was reported that the prevalence of pHPT was as high as 2.1% in postmenopausal women (2). The parathyroid tumor tissue secretes excessive parathyroid hormone, which results in hypercalcemia and related complications, such as osteoporosis, bone fracture, urolithiasis and renal failure. Parathyroid adenoma (PA) accounts for approximately 85% of pHPT. The genetic mechanisms of PA are not fully known, even though tremendous efforts have been made in previous studies. In PA with a hereditary background, the most well-known driver gene of PA was multiple endocrine neoplasia type 1 (*MEN1*). Other possible driver genes for hereditary pHPT include *RET* in MEN-2, *CDKN1B* in MEN-4, *CDC73* in hyperparathyroidism-jaw tumor syndrome, *GCM2* in familial isolated pHPT, *CASR* in familial hypocalciuric hypercalcemia type 1, *GNA11* in familial hypocalciuric hypercalcemia type 2, and *AP2S1* in familial hypocalciuric hypercalcemia type 3. For sporadic PA, somatic mutation of *MEN1* was identified in approximately 35% of cases (3). Mutation of *CCND1* or overexpression of cyclin D1 was found in 30% of sporadic PA cases. Other genes, such as *CASR*, *EZH2*, *CDKI*, and *CTNNB1*, may also contribute to the tumorigenesis of PA. However, the driver genes could still not be identified in a substantial proportion of sporadic PA.

Several studies have explored the molecular mechanism of PA with next-generation sequencing (NGS) technologies, such as whole-exome sequencing (WES) or transcriptome sequencing (RNA-seq). In 2018, Wei et al. (4) used WES to explore genetic abnormalities in 20 specimens of PAs. In 2019, Chai et al. (5) performed RNA-seq to compare the differentially expressed genes (DEGs) between 10 PA and 5 normal parathyroid (PaN) samples, and 8 hub molecules, including *RPL23*, *RPL26*, *RPN1*, *RPS25*, *SEC11A*, *SEC11C*, *SEC61G* and *SPCS2*, were identified. All these previous studies focused on one platform for tumor genome profiling. However, a multiplatform NGS analysis combined with WES and RNA-seq provides a more comprehensive view of the genetic landscape. RNA-seq is adept for identifying gene expression changes, gene fusions and alternative splicing, while WES is good at detecting copy number variants and gene mutations with low expression levels. With the corresponding RNA-seq data, the functional effects of genetic variants from WES can be interpreted with confidence.

Here, we performed concomitant WES and RNA-seq in 41 PA samples to identify related genetic variants and the consequent transcriptomic changes. The somatic mutations and DEGs were identified, and integrated analysis was employed. Based on RNA-seq data, unsupervised clustering with consensus clustering was applied to classify the patients, and the possible underlying molecular mechanism was explored.

MATERIALS AND METHODS

Patients and Specimens

A total of 41 samples of PA were selected randomly from the tissue bank of parathyroid tumors in the present study. These patients received operations from 2013 to 2019 at Peking Union Medical College Hospital, a tertiary referral university hospital. All the patients had a single PA without a history of multiple endocrine neoplasia, familial hyperparathyroidism syndrome or neck irradiation. The diagnosis of PA was made pathologically according to the WHO 2017 criteria (6). During the operation, the tumor specimens were collected from the central portion of the tumors to avoid contamination from adjacent tissues. Then, the specimens were immediately immersed in RNAlater (Ambion Inc., USA) solution at 4°C and stored at -80°C until use. The peripheral blood samples were collected, and white blood cells (WBCs) were stored at -80°C. Tiny PaN tissues were incidentally obtained from 5 patients with normal parathyroid function during surgery for thyroid diseases. The clinical data, including age, sex, imaging results, biochemical results and pathological evaluation, were collected. Follow-up data was obtained by reviewing outpatient records and telephone interviews. This study was approved by the Institutional Ethics Review Board of Peking Union Medical College Hospital (S-K836). The written informed consent was obtained from all the participants.

WES and WES Data Analysis

Genomic DNA was extracted from tumor and WBC samples with a DNeasy Blood & Tissue Kit (QIAGEN, Germany). A total of 0.6 µg of DNA per sample was used to prepare the DNA libraries with the Agilent SureSelect Human All Exon V6 Kit (Agilent Technologies, CA, USA). Then, the DNA libraries of each sample were sequenced with the Illumina HiSeq platform, and 150-bp paired-end reads were produced. Low-quality reads and reads containing adapter contamination were filtered. With Burrows-Wheeler Aligner (BWA) software, valid sequencing reads were mapped to the reference human genome (B37). Samtools, ANNOVAR, 1000 Genomes and other related databases were used to identify and annotate SNPs and insertions and deletions (InDels). The somatic single-nucleotide variant (SNV) was detected by MuTect, and the somatic InDel was detected by Strelka (7, 8). Somatic variants in the segmental duplication or with a frequency > 0.01 in the 1000 Genomes Chinese database were excluded for further analysis. The copy number variant was identified with Control-FREEC (9). The significantly mutated genes (SMGs) were identified by MuSiC software (10).

Transcriptome Sequencing and RNA-seq Data Analysis

RNA was extracted from tissue samples with TRIzol™ Reagent (Invitrogen, USA). A total of 2 µg of RNA per sample was used to generate sequencing libraries with the NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA). RNA-seq was

completed on an Illumina HiSeq platform, and 150-bp paired-end reads were produced. Paired-end clean reads were aligned to the reference genome using HISAT2 v2.0.5. HTSeq was used to count the reads mapped to each gene, and the fragments per kilobase of transcript sequence per million base pairs sequenced (FPKM) value was calculated based on the length of the gene and the read count mapped for each gene. Gene set enrichment analysis (GSEA) was used for functional enrichment analysis of DEGs, including Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses (11). Gene fusion was identified by STAR-Fusion (12). The expression network analysis based on the DEGs was performed using the OmicStudio tools (<https://www.omicstudio.cn/tool>). For each pair of analyzed DEGs, the Pearson correlation coefficient was greater than 0.95 ($p < 0.001$).

Integrative WES and RNA-seq Analysis

To explore the possible molecular mechanism underlying PA, unsupervised clustering of RNA-seq data was performed with the R package ConsensusClusterPlus. Consensus clustering was used to aid in discovering the optimal class. The number of clusters was selected according to the delta area plot and consensus

cumulative distribution function (CDF) plot. Furthermore, the clustering results were explored with the features of the expression profile and genes in each cluster. The flowchart of data analysis was shown in **Figure S1**.

Statistical Analysis

Continuous parameters are shown as the mean \pm standard deviation (SD), and discrete data are reported as numbers or corresponding percentages. Differences in variables between groups were evaluated with the Mann-Whitney U test. Fisher's exact test was used to compare the categorical data. A two-sided p -value < 0.05 was set as the threshold for statistical significance. All statistical analyses were performed with SPSS version 16.0.

RESULTS

Clinical Characteristics of Patients With PA

A total of 41 patients were enrolled in the present study (**Figure 1A**). The average age at diagnosis was 53.9 ± 10.8 (28–72) years. An obvious female predominance was revealed, with a female-to-

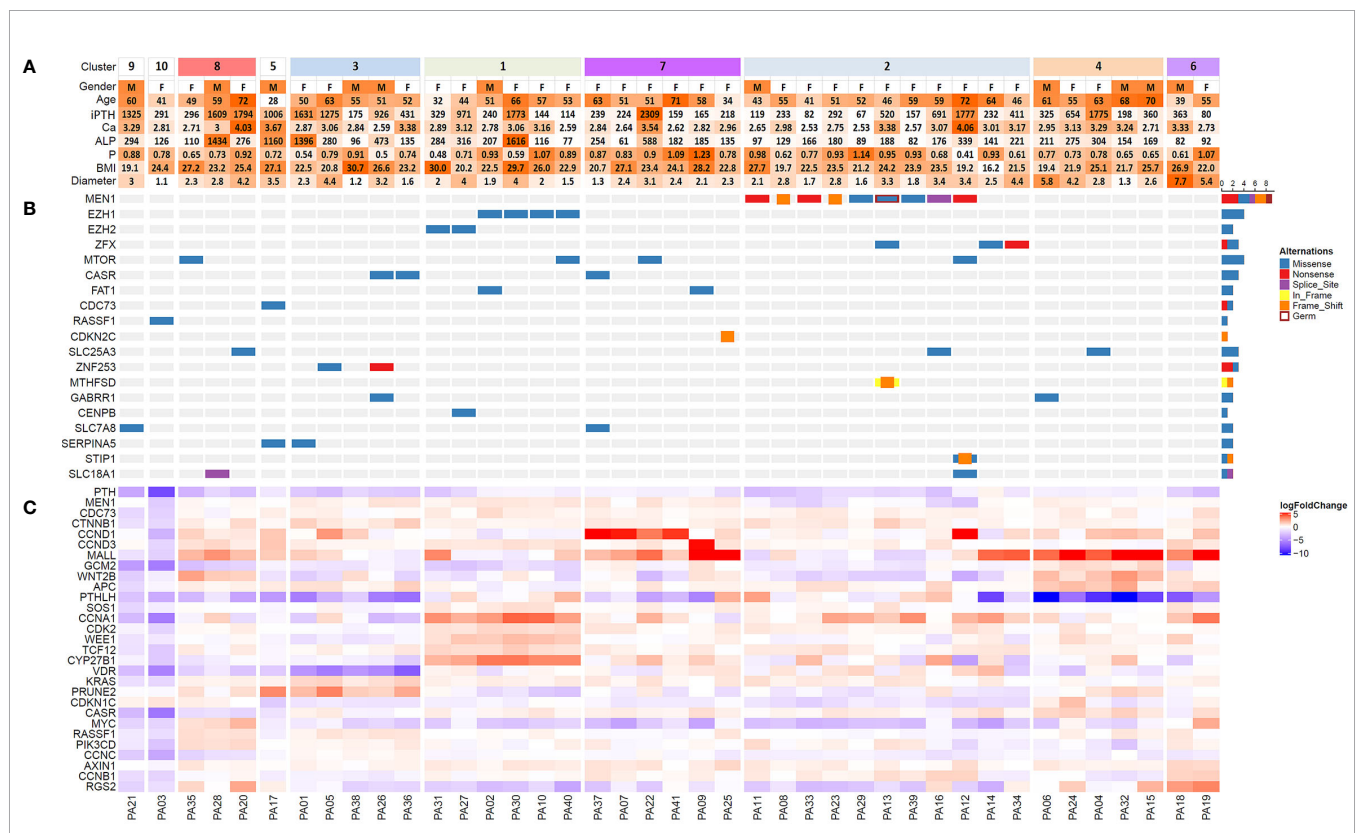


FIGURE 1 | Heatmap of clinical features (A), significantly mutated genes (B) and differentially expressed genes (C) among 41 patients with parathyroid adenoma. Samples were clustered based on RNA-seq data by unsupervised clustering with the ConsensusClusterPlus package. Sample numbers are shown below the corresponding columns. In part (C) the logarithm of fold change in gene expression compared with the expression levels in normal parathyroid tissue is visualized by color intensity. Upregulation and downregulation of gene expression are represented by red and blue, respectively. iPTH, serum intact PTH level (12–67 pg/ml); Ca, serum calcium level (2.13–2.70 mmol/L); P, serum phosphorus level (0.81–1.45 mmol/L); ALP, alkaline phosphatase (35–100 U/L); BMI, body mass index; Diameter, maximum diameter of the tumor (cm).

male ratio of 30:11. The average serum levels of iPTH and Ca were 633.7 ± 628.1 (67.1-2309.2) pg/ml and 3.01 ± 0.37 (2.53-4.06) mmol/L, respectively.

Gene Mutations and Copy Number Variations With WES

For WES, we achieved a mean sequencing depth of 294× and 140× for the tumor and WBC control, respectively. Thereinto, 91.4% of the exome in tumors and 80.1% of the exome in WBC controls was sequenced with a depth of >50×. According to our filter criteria, a total of 668 nonsynonymous somatic mutations were detected, which included 636 SNVs and 32 InDels (Table S1). The top 7 recurrent mutated genes for PA were *MEN1* (9/41), *EZH1* (4/41), *mTOR* (4/41), *ZFX* (3/41), *CASR* (3/41), *EZH2* (2/41) and *FAT1* (2/41) (Figure 1B, Table 1). In this cohort, 8 somatic mutations and 1 germline mutation of *MEN1* were identified in 9 PA samples. Other potential driver genes reported in COSMIC, such as *RASSF1*, *CDKN2C* and *CDC73*, were found in one sample.

DEGs and Unsupervised Clustering With RNA-seq Data

In 41 PA and 5 PaN samples, a mean of 55.7 million reads were obtained by RNA-seq per sample, and 96.3% reads were successfully aligned to the reference genome. Compared with

PaN samples, 84 genes were upregulated and 646 genes were downregulated (with a fold change of 1.5) in PA samples. GSEA showed that the DEGs were enriched in multiple biological processes, such as ribosome, lysosome, cell cycle and tight junction (Figure 2). Then, the DEGs were selected to build the expression network. The network comprised 113 network nodes and 500 network edges, indicating a complicated regulatory association in PA (Figure S2).

Integrative Analysis of WES and RNA-seq Data

Through consensus clustering, the RNA-seq data of PA samples could optimally be categorized into 10 clusters (Figure S3). Although unsupervised clustering is based merely on a mathematical calculation of gene expression data, prominent molecular features could be identified for most clusters based on existing knowledge (Figure 1C). On one hand, previously reported driver gene mutations could be identified in 3 clusters (Clusters 1, 2 and 5). In Cluster 2, all 9 samples with *MEN1* mutations (8 somatic mutations and 1 germline mutation) and 2 samples with *ZFX* mutations were included, which indicated that *MEN1* and *ZFX* mutations may result in a specific and similar expression profile. The expression of *MEN1* in samples with *MEN1* mutations was decreased significantly ($p < 0.001$).

TABLE 1 | Recurrently mutated genes and potential driver genes in parathyroid adenomas.

Gene	Sample ID	Somatic/Germline	Mutation type	Amino acid change
<i>MEN1</i>	8	Somatic	frameshift deletion	NM_000244: exon2:c.269_270del:p.Y90fs
<i>MEN1</i>	16	Somatic	splicing	NM_000244: exon8:c.928-1G>T
<i>MEN1</i>	23	Somatic	frameshift insertion	NM_000244: exon2:c.309dupG:p.S104fs
<i>MEN1</i>	29	Somatic	missense SNV	NM_000244: exon2:c.C95G:p.P32R
<i>MEN1</i>	33	Somatic	stopgain	NM_000244: exon3:c.C511T;p.Q171X
<i>MEN1</i>	39	Somatic	missense SNV	NM_000244: exon9:c.T1282A;p.W428R
<i>MEN1</i>	11	Somatic	stopgain	NM_000244: exon9:c.C1339T;p.Q447X
<i>MEN1</i>	12	Somatic	stopgain	NM_000244: exon4:c.C787T
<i>MEN1</i>	13	Germline	missense SNV	NM_000244: exon2:c.A1G;p.M1V
<i>EZH1</i>	2	Somatic	missense SNV	NM_001991: exon17:c.A1925T;p.Y642F
<i>EZH1</i>	10	Somatic	missense SNV	NM_001991: exon17:c.A1925T;p.Y642F
<i>EZH1</i>	30	Somatic	missense SNV	NM_001991: exon17:c.A1925T;p.Y642F
<i>EZH1</i>	40	Somatic	missense SNV	NM_001991: exon17:c.A1925T;p.Y642F
<i>MTOR</i>	22	Somatic	missense SNV	NM_004958: exon53:c.A7257T;p.E2419D
<i>MTOR</i>	35	Somatic	missense SNV	NM_004958: exon47:c.C6644T;p.S2215F
<i>MTOR</i>	40	Somatic	missense SNV	NM_004958: exon27:c.T4079C;p.L1360S
<i>MTOR</i>	12	Somatic	missense SNV	NM_004958: exon56:c.A7501T;p.I2501F
<i>ZFX</i>	13	Somatic	missense SNV	NM_001178086: exon4:c.A371G;p.K124R
<i>ZFX</i>	34	Somatic	stopgain	NM_001178086: exon6:c.G658T;p.E220X
<i>ZFX</i>	14	Somatic	missense SNV	NM_001178086: exon6:c.C1603T;p.R535W
<i>CASR</i>	26	Somatic	missense SNV	NM_000388: exon3:c.A307C;p.T103P
<i>CASR</i>	37	Somatic	missense SNV	NM_000388: exon3:c.C413T;p.T138M
<i>CASR</i>	36	Somatic	missense SNV	NM_000388: exon7:c.T1753C;p.C585R
<i>EZH2</i>	31	Somatic	missense SNV	NM_001203249: exon15:c.T1768A;p.Y590N
<i>EZH2</i>	27	Somatic	missense SNV	NM_001203249: exon15:c.A1769T;p.Y590F
<i>FAT1</i>	2	Somatic	missense SNV	NM_005245: exon10:c.G8140C;p.E2714Q
<i>FAT1</i>	9	Somatic	missense SNV	NM_005245: exon10:c.G8671T;p.D2891Y
<i>CDC73</i>	17	Somatic	missense SNV	NM_024529: exon3:c.G268T;p.D90Y
<i>CDC73</i>	17	Somatic	stopgain	NM_024529: exon1:c.G128A;p.W43X
<i>RASSF1</i>	3	Somatic	missense SNV	NM_170713: exon2:c.G157T;p.V53L
<i>CDKN2C</i>	25	Somatic	frameshift deletion	NM_078626: exon2:c.397_413del:p.R133fs

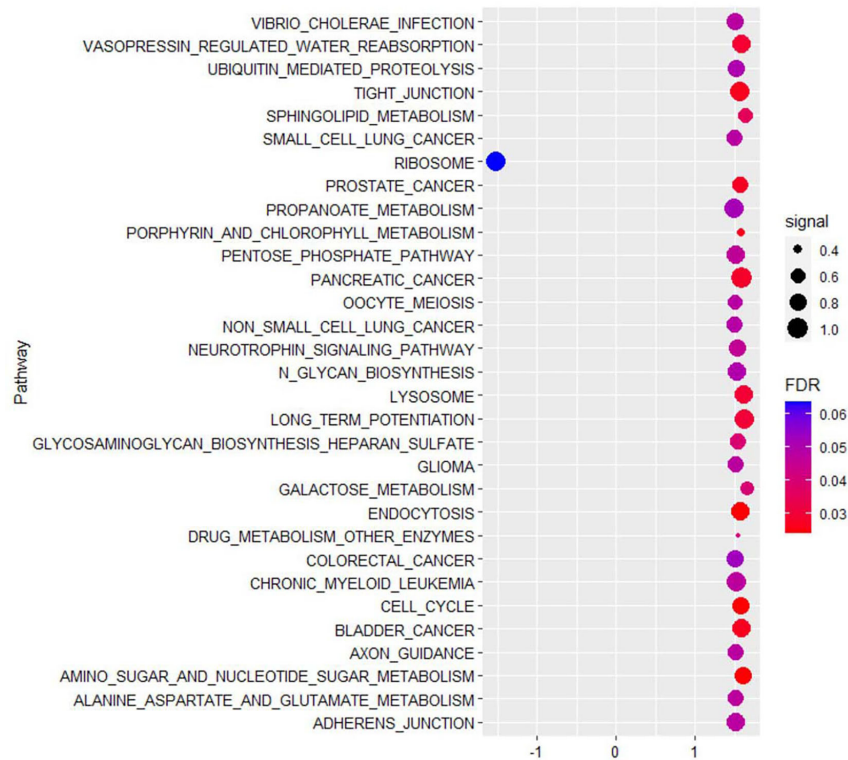


FIGURE 2 | The top 30 KEGG pathways enriched by gene set enrichment analysis (GSEA) showed that differentially expressed genes of parathyroid adenoma were enriched in multiple biological processes, such as ribosome, lysosome, cell cycle and tight junction. The horizontal axis represents the normalized enrichment score (NES), and the vertical axis represents the KEGG pathways. The false discovery rate (FDR) value and enrichment signal are represented by the color and circle diameter, respectively.

In Cluster 1, mutations in *EZH2* and *EZH1* were identified, and the expression levels of *SOS1* ($p=0.002$), *CCNA1* ($p < 0.001$), *CDK2* ($p < 0.001$), *WEE1* ($p < 0.001$), *TCF12* ($p < 0.001$) and *CYP27B1* ($p < 0.001$) were significantly upregulated compared with those in the other subgroups (**Figure 3**). Double somatic mutations of *CDC73*, a stop-gain mutation (NM_024529:exon1:c.G128A:p.W43X) and a missense mutation (NM_024529:exon3:c.G268T:p.D90Y), were identified in PA17 (Cluster 5). Three samples with somatic mutations of *CASR* were distributed in Cluster 3 and Cluster 7.

On the other hand, prominent gene expression abnormalities were found in some clusters as previously reported in PA. The cyclin D1 (*CCND1*) mRNA level in Cluster 7 was much higher than that in any other cluster ($p=0.006$). For *GCM2* and *WNT2B*, all 5 samples with the highest expression levels were clustered into Cluster 4, while no genetic mutation of known driver genes was identified. High expression levels of *AXIN1* and *CCNB1* were found in Cluster 6. High levels of *KRAS* ($p=0.001$) and *PRUNE2* ($p=0.001$), as well as low levels of vitamin D receptor (*VDR*) ($p=0.001$), were prominent in Cluster 3. Low levels of *CASR* and high levels of *MYC* were found in the three samples in Cluster 8. Sample PA21 in Cluster 9 and sample PA03 in Cluster 10 had low levels of *CDC73*, *CASR* and *SOS1*, while no

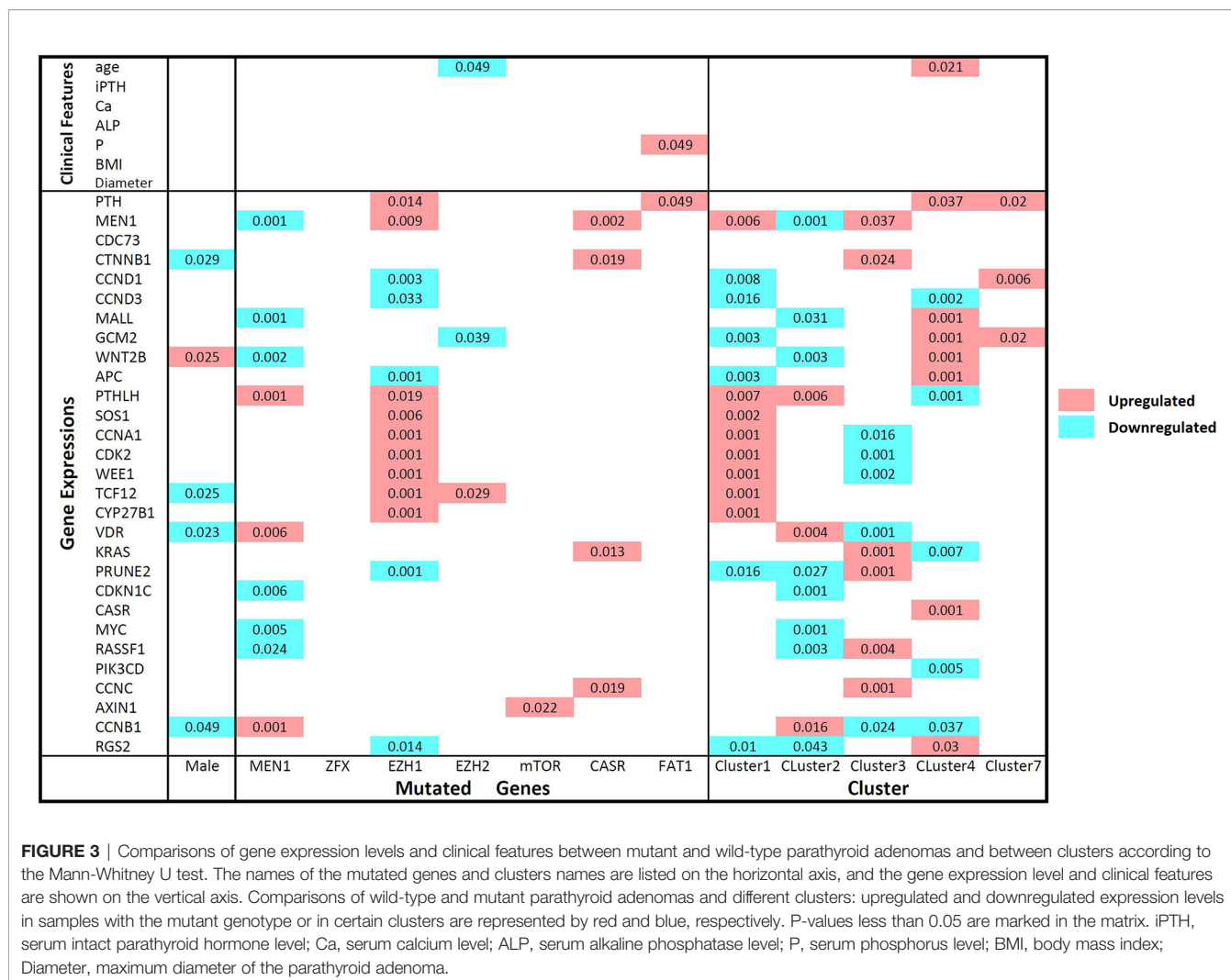
copy number or gene fusion abnormality of these genes was found.

Relationship Between Gene Expression and Clinical Features

The serum level of iPTH was correlated with serum Ca ($p < 0.001$, $rs=0.685$), alkaline phosphatase (ALP) ($p < 0.001$, $rs=0.802$), phosphorus (P) ($p < 0.001$, $rs=-0.525$) and tumor diameter ($p < 0.001$, $rs=0.570$), as well as the expression levels of *VDR* ($p=0.009$, $rs=-0.403$) and *PRUNE2* ($p=0.009$, $rs=0.401$) (**Figure 4**). The expression level of the *PTH* gene was found to be correlated with the expression of multiple genes, such as *CTNNA1* ($p=0.045$, $rs=-0.315$), *MALL* ($p=0.024$, $rs=0.352$), and *GCM2* ($p=0.002$, $rs=0.468$).

DISCUSSION

To our knowledge, this study is the first integrative analysis of WES and RNA-seq in parathyroid neoplasia, providing new insight into the molecular mechanism of PA. Using WES, we revealed a group of mutations in genes such as *MEN1*, *EZH2*, *CDC73*, *ZFX*, *CASR*, *FAT1* and *RASSF1*, almost all of which had



been reported previously. In the RNA-seq analysis, overexpression of *CCND1* and *CTNNB1*, which are recognized as vital factors for PA tumorigenesis, was also found (13, 14). KEGG analysis also found that ribosomal protein dysregulation was related to PA and is a known factor involved in tumor development (15). Previous work focused mainly on the mutation or expression of one or several of these related genes, which may be only one facet of the complex molecular mechanisms of PA tumorigenesis. However, contradictory results were occasionally reported in different publications. Therefore, we performed unsupervised clustering for the RNA-seq data, and 10 clusters were identified. Even though it is not easy to explain the intrinsic molecular mechanism underpinning these clusters from a single study with limited samples, we revealed some intriguing aspects.

In our cohort, 8 somatic mutations and 1 germline mutation of *MEN1* were identified, and the expression levels of *MEN1* in these 9 samples were significantly decreased. All these samples were clustered in one subgroup, indicating that the *MEN1* genotype might be related to the expression profile. *MEN1*, a

tumor suppressor gene, was identified as a genetic driver of multiple endocrine neoplasia type 1 (16). Biallelic inactivation of *MEN1* is also found in approximately 12-35% sporadic PA cases (17, 18). We found that one patient with seemingly sporadic PA had a germline *MEN1* mutation, which was also reported by other authors (19). Three samples (7.3%) harboring somatic *ZFX* mutations were also included in Cluster 2, one of which carried a *MEN1* mutation concurrently. *ZFX*, a transcriptional target of cyclin D, was identified as a candidate driver gene for PA with a mutation rate of 5% (20, 21). In contrast to previous reports, all three mutations of *ZFX* in our cohort were not located at the R767 or R768 position.

Interestingly, 2 patients with the rare activating mutation Y464 of *EZH2* and 4 patients with the somatic mutation Y642F of *EZH1* were classified into one cluster (Cluster 1). Our results are consistent with previous reports showing that *EZH2* mutations are potential genetic drivers of PA. The activating mutation Y464N (previously described as Y641N) in *EZH2* was found previously in 2 of 193 sporadic PA samples (22). Although rare, 2 patients harbored the Y464 mutation in our cohort of 41.

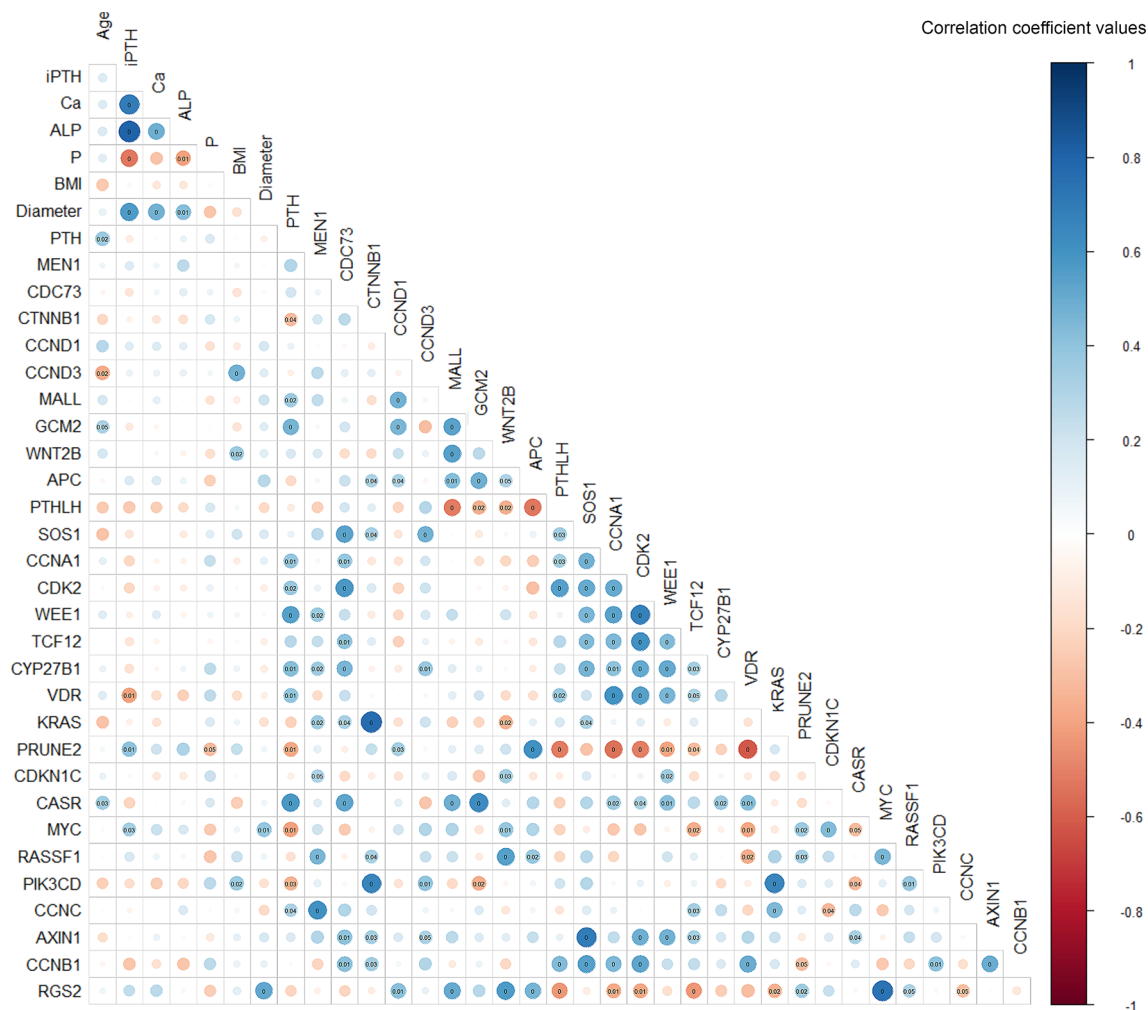


FIGURE 4 | Spearman's correlation matrix between gene expression and clinicopathological features in parathyroid adenoma. The correlation coefficient values are visualized by the size and color intensity of the circles. Positive and negative correlations are represented by blue and red, respectively. P-values below 0.05 are marked in the matrix, with P-values below 0.01 marked as zero. iPTH, serum intact parathyroid hormone level; Ca, serum calcium level; ALP, serum alkaline phosphatase; P, serum phosphorus level; BMI, body mass index; Diameter, maximum diameter of the parathyroid adenoma.

Furthermore, the somatic mutation Y642F of *EZH1* was identified in another 4 patients. Like *EZH2*, *EZH1* is also a component of the noncanonical polycomb repressive complex-2, which is recurrently mutated in multiple hematological malignancies (23). *EZH1* mutations were found in 20% of Hürthle cell adenomas and 10% of Hürthle cell carcinomas, and 53.8% of the mutations were Y642F (24). Though not reported previously, *EZH1* may be a potentially important gene for PA.

A somatic *CDC73* mutation was identified in one sample (PA17). Somatic inactivating mutations of *CDC73*, or loss of staining of the encoded parafibromin, occurred in approximately 60% of parathyroid carcinomas and in a few PAs (25, 26). In addition, significantly low expression of *CDC73* was found in PA03 and PA21, while no mutation or copy number change was found in *CDC73*. Epigenetic changes may be the underlying cause.

In our cohort, overexpression of *CCND1* was found in some samples, and the 4 samples with the highest expression levels were clustered in one group (Cluster 7). *CCND1*, encoding cyclin D1, was first identified as an oncogene in parathyroid tumors, and overexpression of this gene was found in 20-40% of PA samples (17, 27). Overexpression of *CCND1* was caused by rearrangement of the *PTH* 5' regulatory region of the *CCND1* coding region in some PAs. However, this could not be identified by WES in the present study. Neither gene amplification nor mutation of *CCND1* was found in these samples. Furthermore, another cell cycle regulator factor, cyclin D3 (*CCND3*), may also be involved in sporadic PA, as we found that the expression level of cyclin D3 in PA09 was approximately 31 times higher than that in the normal control. This may be caused by the copy gain of *CCND3* identified in this sample. The partners of cyclin D1 in cell cycle regulation, such as *CDKN1B*, *CDKN1C*, *CDKN2C* and

other CDKI genes, were also reported to be involved in sporadic PA (28, 29). As reported, the expression of CDKN1B and CDKN1C was downregulated in some of our samples, while no somatic or germline variant of this gene was found (30). In addition, a frameshift deletion of *CDKN2C* was identified in one sample, and this mutation had been reported previously in a few PA samples (29, 31).

In the present cohort, somatic *CASR* mutations were identified in 3 samples, but the gene expression levels were not different from those in PaN. These 3 samples were clustered into different subgroups. *CASR* encodes a calcium sensing receptor, by which parathyroid cells monitor the extracellular calcium level. Inactivating germline mutations of *CASR* are responsible for familial hypocalciuric hypercalcemia and neonatal severe hyperparathyroidism. However, somatic inactivating *CASR* mutation was rarely found in sporadic PA (32–34). Therefore, *CASR* mutation was believed to be a predisposing factor rather than a genetic driver in PA (17). Another 3 specimens with low *CASR* expression were clustered into an independent subgroup (Cluster 8). Aberrant expression of *CASR* is frequently found in parathyroid tumors (35, 36). Five samples with the lowest *VDR* levels were clustered in Cluster 3, and *VDR* levels were negatively correlated with serum PTH levels, suggesting a vital role in PA. Decreased *VDR* expression was related to the high proliferation of pHPT, but no *VDR* mutation was found in our cohort (35, 37).

We could not identify any somatic or germline variants of *GCM2* in this Chinese cohort. However, a prominent increase in *GCM2* levels was identified in 5 samples in an independent subgroup (Cluster 4). This was consistent with previous findings in which the *GCM2* expression level was significantly upregulated in PA and correlated with a decrease in the response to hypocalcemia (38). *GCM2* encodes a transcription factor that is a critical regulator of the development of the parathyroid gland, and its activating germline variants are responsible for familial isolated hyperparathyroidism (39). In addition, Cluster 4 and 6 revealed no apparent driver mutation but showed aberrant expression of some tumor-related genes, such as *GCM2* and *AXIN1*, indicating that driver gene plays an important, but not indispensable, role in the initiation of PA.

Another interesting mutation might be *SLC25A3*, which was identified in 3 of 41 samples. *SLC25A3* is essential for cytochrome c oxidase of the mitochondrial respiratory chain and *SLC25A3* deletion causes mitochondrial cardiomyopathy (40). Previously, mitochondrial variations may result in oncogenic phenotype of PA (41). However, *SLC25A3*-related studies were not available in PA.

There were several limitations to this study. The main shortcoming was the limited number of participants in this cohort, even though it was quite large considering recent sequencing studies on PA. Additional samples may help validate the gene mutation with low incidence. Second, due to the lack of normal parathyroid specimens for comparison, immunohistochemical staining or Western blot analysis were not suitable to confirm the protein levels of the related gene variants among PA samples. Third, bulk RNA-Seq data are heavily influenced by tissue cellular composition. The gene expression

profile might contain not only the tumor-induced molecular alterations but also the interference from tumor heterogeneity.

CONCLUSIONS

This study revealed almost all previously reported molecular features of PA using integrated WES and RNA-seq analyses. In addition to the previously reported mutant genes, such as *MEN1*, *EZH2*, *CASR* and *CDC73*, somatic mutations of *EZH1* were also identified in the present PA cohort. With clustering based on RNA-seq, abnormal expression of cyclin D1, β -catenin, *VDR*, *CASR* and *GCM2* may be an important factor influencing the gene expression profile of PA.

DATA AVAILABILITY STATEMENT

The data presented in the study are deposited in the Genome Sequence Archive (GSA) in BIG Data Center, Beijing Institute of Genomics (BIG), Chinese Academy of Sciences (CAS), publicly accessible at <https://bigd.big.ac.cn/gsa-human/>, accession number (HRA000665).

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Institutional Ethics Review Board of Peking Union Medical College Hospital. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

YH and XZ designed the study and performed the experiments. YH, XL, and QL performed the surgery. XZ, OW, MC, XL, MW, and SH enrolled the patients and collected the samples. YH, XZ, OW and MC analyzed the data. YH, XZ and QL wrote and modified the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Chinese Academy of Medical Sciences (CAMS) Innovation Fund for Medical Sciences (CIFMS) (2017-I2M-1-001), the Nonprofit Central Research Institute Fund of the Chinese Academy of Medical Sciences (2018PT32014) and the Peking Union Medical College Innovative Team Development Program.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fendo.2021.631680/full#supplementary-material>

REFERENCES

1. Yeh MW, Ituarte PHG, Zhou HC, Nishimoto S, Liu I-LA, Harari A, et al. Incidence and Prevalence of Primary Hyperparathyroidism in a Racially Mixed Population. *J Clin Endocrinol Metab* (2013) 98:1122–9. doi: 10.1210/jc.2012-4022
2. Lundgren E, Rastad J, Thurffjell E, Åkerström G, Ljunghall S. Population-Based Screening for Primary Hyperparathyroidism with Serum Calcium and Parathyroid Hormone Values in Menopausal Women. *Surgery* (1997) 121:287–94. doi: 10.1016/s0039-6060(97)90357-3
3. Newey PJ, Nesbit MA, Rimmer AJ, Attar M, Head RT, Christie PT, et al. Whole-Exome Sequencing Studies of Nonhereditary (Sporadic) Parathyroid Adenomas. *J Clin Endocrinol Metab* (2012) 97:E1995–2005. doi: 10.1210/jc.2012-2303
4. Wei Z, Sun B, Wang ZP, He JW, Fu WZ, Fan YB, et al. Whole-Exome Sequencing Identifies Novel Recurrent Somatic Mutations in Sporadic Parathyroid Adenomas. *Endocrinology* (2018) 159:3061–8. doi: 10.1210/en.2018-00246
5. Chai YJ, Chae H, Kim K, Lee H, Choi S, Lee KE, et al. Comparative Gene Expression Profiles in Parathyroid Adenoma and Normal Parathyroid Tissue. *J Clin Med* (2019) 8:297. doi: 10.3390/jcm8030297
6. Lloyd RV, Osamura RY, Klöppel G, Rosai J. WHO/IARC Classification of Tumours. In: *WHO Classification of Tumours of Endocrine Organs. Lyon: IARC, 4th Edition*, Volume 10. France: International Agency for Research on Cancer (IARC) (2017).
7. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, et al. Sensitive Detection of Somatic Point Mutations in Impure and Heterogeneous Cancer Samples. *Nat Biotechnol* (2013) 31:213–9. doi: 10.1038/nbt.2514
8. Saunders CT, Wong WSW, Swamy S, Becq J, Murray LJ, Cheetham RK. Strelka: Accurate Somatic Small-Variant Calling from Sequenced Tumor-Normal Sample Pairs. *Bioinformatics* (2012) 28:1811–7. doi: 10.1093/bioinformatics/bts271
9. Boeva V, Popova T, Bleakley K, Chiche P, Cappo J, Schliepacher G, et al. Control-FREEC: a Tool for Assessing Copy Number and Allelic Content Using Next-Generation Sequencing Data. *Bioinform (Oxf Engl)* (2012) 28:423–5. doi: 10.1093/bioinformatics/btr670
10. Dees ND, Zhang Q, Kandoth C, Wendl MC, Schierding W, Koboldt DC, et al. MuSiC: Identifying Mutational Significance in Cancer Genomes. *Genome Res* (2012) 22:1589–98. doi: 10.1101/gr.134635.111
11. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene Set Enrichment Analysis: A Knowledge-Based Approach for Interpreting Genome-Wide Expression Profiles. *Proc Natl Acad Sci USA* (2005) 102:15545–50. doi: 10.1073/pnas.0506580102
12. Haas BJ, Dobin A, Li B, Stransky N, Pochet N, Regev A. Accuracy Assessment of Fusion Transcript Detection via Read-Mapping and de novo Fusion Transcript Assembly-Based Methods. *Genome Biol* (2019) 20:213. doi: 10.1186/s13059-019-1842-9
13. Arnold A. Major Molecular Genetic Drivers in Sporadic Primary Hyperparathyroidism. *Trans Am Clin Climatol Assoc* (2016) 127:235–44.
14. Alberto F. Genetics of Parathyroids Disorders: Overview. *Best Pract Res Clin Endocrinol Metab* (2018) 32:781–90. doi: 10.1016/j.beem.2018.09.011
15. Goudarzi KM, Lindström MS. Role of Ribosomal Protein Mutations in Tumor Development (Review). *Int J Oncol* (2016) 48:1313–24. doi: 10.3892/ijo.2016.3387
16. Chandrasekharappa SC. Positional Cloning of the Gene for Multiple Endocrine Neoplasia-type 1. *Science* (1997) 276:404–7. doi: 10.1126/science.276.5311.404
17. Costa-Guda J, Arnold A. Genetic and Epigenetic Changes in Sporadic Endocrine Tumors: Parathyroid Tumors. *Mol Cell Endocrinol* (2014) 386:46–54. doi: 10.1016/j.mce.2013.09.005
18. Farnebo F. Alterations of the MEN1 Gene in Sporadic Parathyroid Tumors. *J Clin Endocrinol Metab* (1998) 83:2627–30. doi: 10.1210/jc.83.8.2627
19. Starker LF, Åkerström T, Long WD, Delgado-Verdugo A, Donovan P, Udelsman R, et al. Frequent Germ-Line Mutations of the MEN1, CASR, and HRPT2/CDC73 Genes in Young Patients with Clinically Non-Familial Primary Hyperparathyroidism. *Horm Cancer* (2012) 3:44–51. doi: 10.1007/s12672-011-0100-8
20. Soong CP, Arnold A. Recurrent ZFX Mutations in Human Sporadic Parathyroid Adenomas. *Oncoscience* (2014) 1:360–6. doi: 10.18632/oncoscience.116
21. Casimiro MC, Crosariol M, Loro E, Ertel A, Yu Z, Dampier W, et al. ChIP Sequencing of Cyclin D1 Reveals a Transcriptional Role in Chromosomal Instability in Mice. *J Clin Invest* (2012) 122:833–43. doi: 10.1172/JCI60256
22. Cromer MK, Starker LF, Choi M, Udelsman R, Nelson-Williams C, Lifton RP, et al. Identification of Somatic Mutations in Parathyroid Tumors Using Whole-Exome Sequencing. *J Clin Endocrinol Metab* (2012) 97:E1774–81. doi: 10.1210/jc.2012-1743
23. Nakagawa M, Kitabayashi I. Oncogenic Roles of Enhancer of Zeste Homolog 1/2 in Hematological Malignancies. *Cancer Sci* (2018) 109:2342–8. doi: 10.1111/cas.13655
24. Jung CK, Kim Y, Jeon S, Jo K, Lee S, Bae JS. Clinical Utility of EZH1 Mutations in the Diagnosis of Follicular-Patterned Thyroid Tumors. *Hum Pathol* (2018) 81:9–17. doi: 10.1016/j.humpath.2018.04.018
25. Hu Y, Liao Q, Cao S, Gao X, Zhao Y. Diagnostic Performance of Parafibromin Immunohistochemical Staining for Sporadic Parathyroid Carcinoma: A Meta-Analysis. *Endocrine* (2016) 54:612–9. doi: 10.1007/s12020-016-0997-3
26. Cetani F, Banti C, Pardi E, Borsari S, Viacava P, Miccoli P, et al. CDC73 Mutational Status and Loss of Parafibromin in the Outcome of Parathyroid Cancer. *Endocr Connect* (2013) 2:186–95. doi: 10.1530/EC-13-0046
27. Motokura T, Bloom T, Kim HG, Jüppner H, Ruderman JV, Kronenberg HM, et al. A Novel Cyclin Encoded by a bcl1-linked Candidate Oncogene. *Nature* (1991) 350:512–5. doi: 10.1038/350512a0
28. Arya AK, Bhadada SK, Singh P, Sachdeva N, Saikia UN, Dahiya D, et al. Promoter Hypermethylation Inactivates CDKN2A, CDKN2B and RASSF1A Genes in Sporadic Parathyroid Adenomas. *Sci Rep* (2017) 7:3123. doi: 10.1038/s41598-017-03143-8
29. Costa-Guda J, Soong CP, Parekh VI, Agarwal SK, Arnold A. Germline and Somatic Mutations in Cyclin-Dependent Kinase Inhibitor Genes CDKN1A, CDKN2B, and CDKN2C in Sporadic Parathyroid Adenomas. *Horm Cancer* (2013) 4:301–7. doi: 10.1007/s12672-013-0147-9
30. Buchwald PC, Akerstrom G, Westin G. Reduced p18INK4c, p21CIP1/WAF1 and p27KIP1 mRNA Levels in Tumours of Primary and Secondary Hyperparathyroidism. *Clin Endocrinol* (2004) 60:389–93. doi: 10.1111/j.1365-2265.2004.01995.x
31. Gluck T, Yuan Z, Libutti SK, Marx SJ. Mutations in CDKN2C (p18) and CDKN2D (p19) May Cause Sporadic Parathyroid Adenoma. *Endocr Relat Cancer* (2013) 20:L27–9. doi: 10.1530/erc-13-0445
32. Hosokawa Y. Mutational Analysis of the Extracellular Ca(2+)-Sensing Receptor Gene in Human Parathyroid Tumors. *J Clin Endocrinol Metab* (1995) 80:3107–10. doi: 10.1210/jc.80.11.3107
33. Cetani F, Pinchera A, Pardi E, Cianferotti L, Vignali E, Picone A, et al. No Evidence for Mutations in the Calcium-Sensing Receptor Gene in Sporadic Parathyroid Adenomas. *J Bone Miner Res* (1999) 14:878–82. doi: 10.1359/jbmr.1999.14.6.878
34. Guarnieri V, Canaff L, Yun FHJ, Scillitani A, Battista C, Muscarella LA, et al. Calcium-Sensing Receptor (CASR) Mutations in Hypercalcemic States: Studies from a Single Endocrine Clinic Over Three Years. *J Clin Endocrinol Metab* (2010) 95:1819–29. doi: 10.1210/jc.2008-2430
35. Carling T, Rastad J, Szabó E, Westin G, Åkerström G. Reduced Parathyroid Vitamin D Receptor Messenger Ribonucleic Acid Levels in Primary and Secondary Hyperparathyroidism. *J Clin Endocrinol Metab* (2000) 85:2000–3. doi: 10.1210/jc.85.5.2000
36. Aycicek GS, Aydogan BI, Sahin M, Cansiz Ersoz C, Sak SD, Baskal N. Clinical Impact of p27Kip1 and CaSR Expression on Primary Hyperparathyroidism. *Endocr Pathol* (2018) 29:250–8. doi: 10.1007/s12022-018-9524-9
37. Yano S, Sugimoto T, Tsukamoto T, Chihara K, Kobayashi A, Kitazawa S, et al. Decrease in Vitamin D Receptor and Calcium-Sensing Receptor in Highly Proliferative Parathyroid Adenomas. *Eur J Endocrinol* (2003) 148:403–11. doi: 10.1530/eje.0.1480403
38. Kebebew E, Peng M, Wong MG, Ginzinger D, Duh QY, Clark OH. GCMB Gene, a Master Regulator of Parathyroid Gland Development, Expression, and Regulation in Hyperparathyroidism. *Surgery* (2004) 136:1261–6. doi: 10.1016/j.surg.2004.06.056

39. Guan B, Welch JM, Sapp JC, Ling H, Li Y, Johnston JJ, et al. GCM2-Activating Mutations in Familial Isolated Hyperparathyroidism. *Am J Hum Genet* (2016) 99:1034–44. doi: 10.1016/j.ajhg.2016.08.018
40. Mayr JA, Zimmermann FA, Horvath R, Schneider HC, Schoser B, Holinski-Feder E, et al. Deficiency of the Mitochondrial Phosphate Carrier Presenting as Myopathy and Cardiomyopathy in a Family with Three Affected Children. *Neuromuscul Disord* (2011) 21:803–8. doi: 10.1016/j.nmd.2011.06.005
41. Costa-Guda J, Tokura T, Roth SI, Arnold A. Mitochondrial DNA Mutations in Oxyphilic and Chief Cell Parathyroid Adenomas. *BMC Endocr Disord* (2007) 7:8. doi: 10.1186/1472-6823-7-8

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Hu, Zhang, Wang, Cui, Li, Wang, Hua and Liao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.