Human Molecular Genetics, 2016, Vol. 25, No. 9 1875–1884

doi: 10.1093/hmg/ddw056
Advance Access Publication Date: 2 March 2016
Association Studies Article

ASSOCIATION STUDIES ARTICLE

Whole exome sequencing identifies lncRNA *GAS8-AS1* and *LPAR4* as novel papillary thyroid carcinoma driver alternations

Wenting Pan[1,†], Liqing Zhou[2,†], Minghua Ge[3,†], Bin Zhang[4], Xinyu Yang[1], Xiangyu Xiong[1], Guobin Fu[6], Jian Zhang[7], Xilin Nie[3], Hongmin Li[5], Xiaohu Tang[1], Jinyu Wei[1], Mingming Shao[5], Jian Zheng[5], Qipeng Yuan[1], Wen Tan[5], Chen Wu[5,*], Ming Yang[1,8,*] and Dongxin Lin[5]

[1]Beijing Laboratory of Biomedical Materials, College of Life Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China, [2]Department of Radiation Oncology, Huaian No. 2 Hospital, Huaian 223002, Jiangsu Province, China, [3]Department of Head and Neck Surgery, Zhejiang Province Cancer Hospital, Hangzhou 310022, Zhejiang Province, China, [4]Department of Head and Neck Surgical Oncology and [5]State Key Laboratory of Molecular Oncology, Cancer Hospital, Chinese Academy of Medical Sciences, Beijing 100021, China, [6]Department of Oncology, Provincial Hospital affiliated to Shandong University, Jinan 250021, Shandong Province, China, [7]Department of Clinical Laboratory, Qilu Hospital of Shandong University, Jinan 250012, Shandong Province, China and [8]Shandong Key Laboratory of Radiation Oncology, Cancer Research Center, Shandong Cancer Hospital and Institute, Jinan 250117, Shandong Province, China

*To whom correspondence should be addressed at: Shandong Key Laboratory of Radiation Oncology, Cancer Research Center, Shandong Cancer Hospital and Institute, Jinan 250117, Shandong Province, China (M.Y.)/State Key Laboratory of Molecular Oncology, Cancer Hospital, Chinese Academy of Medical Sciences, Beijing 100021, China (C.W.). Tel: +86 1064447747 (M.Y.)/+86 1087787395 (C.W.); Fax: +86 1064437610 (M.Y.)/+86 1087787395 (C.W.); Email: aaryoung@yeah.net (M.Y.)/chenwu@cicams.ac.cn (C.W.)

Abstract

Papillary thyroid carcinoma (PTC) is the most common type of thyroid cancer. However, we know little of mutational spectrum in the Chinese population. Thus, here we report the identification of somatic mutations for Chinese PTC using 402 tumor-normal pairs (Discovery: 91 pairs via exome sequencing; validation: 311 pairs via Sanger sequencing). We observed three distinct mutational signatures, evidently different from the two mutational signatures among Caucasian PTCs. Ten significantly mutated genes were identified, most previously uncharacterized. Notably, we found that long non-coding RNA (lncRNA) *GAS8-AS1* is the secondary most frequently altered gene and acts as a novel tumor suppressor in PTC. As a mutation hotspot, the c.713A>G/714T>C dinucleotide substitution was found among 89.1% patients with *GAS8-AS1* mutations and associated with advanced PTC disease ($P = 0.009$). Interestingly, the wild-type lncRNA GAS8-AS1 ($A_{713}T_{714}$) showed consistently higher capability to inhibit cancer cell growth compared to the mutated lncRNA ($G_{713}C_{714}$). Further studies also elucidated the oncogene nature of the G protein-coupled receptor *LPAR4* and its c.872T>G (p.Ile291Ser) mutation in PTC malignant transformation. The *BRAF* c.1799T>A (p.Val600Glu) substitution was present in 59.0% Chinese PTCs, more frequently observed in patients with lymph

[†]W.P., L.Z. and M.G. contributed equally to this work.

Received: September 4, 2015. **Revised:** January 30, 2016. **Accepted:** February 15, 2016

node metastasis ($P = 1.6 \times 10^{-4}$). Together our study defines a exome mutational spectrum of PTC in the Chinese population and highlights lncRNA GAS8-AS1 and LPAR4 as potential diagnostics and therapeutic targets.

Introduction

Thyroid carcinoma is the most common endocrine malignancy with quickly increased incidence over last several decades. In China, the incidence of this cancer is 6.6 per 100 000 individuals according to the Chinese Cancer Registry (1). Papillary thyroid carcinoma (PTC), named for their papillary histological architecture, accounts for 80% of all thyroid carcinomas. Risk factors for PTC development include ionizing radiation, nodular disease of the thyroid and family history (2). Genetic predisposition contributes more for PTC compared to any other cancer (3). By analyzing exome sequence in 402 PTC tumors, the Cancer Genome Atlas (TCGA) project identified *BRAF* and the *RAS* family as dominant driver genetic alterations (4). However, knowledge of the spectrum of somatic mutations in PTC in the Chinese population remains obscure. Since the difference in environmental factors that may cause thyroid carcinoma and the diversity of genetic background in different ethnic populations is expected to play critical roles on PTC etiology, we hypothesized that PTCs in different populations under different environments might present different spectra of mutations.

To address this issue, we performed whole-exome sequencing (WES) of 91 clinically and pathologically characterized PTC tissues and 91 paired peripheral blood samples. Candidate mutations were then validated in additional 311 tumor-normal pairs in the Chinese population. We identified three distinct mutational signatures evidently different from the two mutational signatures in Caucasian PTC patients. Ten significantly mutated genes were discovered, most previously uncharacterized (i.e. *GAS8-AS1* and *LPAR4*) as well as well-known (*BRAF*). Particularly, we found that a long non-coding RNA (lncRNA) *GAS8-AS1* is the secondary most frequently altered gene and acts as a novel tumor suppressor in PTC. Further studies also elucidated the biological relevance of the G protein-coupled receptor *LPAR4* to the PTC malignant phenotype. Our data demonstrated that both coding and non-coding genes may play an essential role in the etiology of this malignant disease.

Results

Mutational signatures of PTC

To gain insight into PTC genetics in the Chinese population, we performed WES of 91 pairs of PTC tissues and peripheral blood samples (Supplementary Material, Table S1), covering a median of 50.3 megabases (Mb) at a median of 126.63× (range: 100.56–314.32×) coverage for PTC samples and 125.58× (range: 100.72–293.81×) coverage for blood samples (Supplementary Material, Table S2), followed by calling of somatic mutations using the MuTect algorithm (5). We identified a total of 1672 somatic mutations across the entire data set, including 1067 missense, 51 stopgain, 4 stoploss, 455 synonymous, and 45 splice site single nucleotide variations (SNVs), as well as 41 frameshift insertions or deletions (indels) and 9 inframe indels (Supplementary Material, Fig. S1). We validated candidate mutations in genes comprising both consensus cancer genes and new recurrently mutated genes in PTC with Sanger sequencing in the original 91 pairs of PTC samples and additional 311 PTC tissues and paired peripheral blood samples (Supplementary Material, Table S1) and found that a true positive rate of 97.2% was achieved. The overall somatic mutation density was 0.38 mutations per Mb (Supplementary Material, Table S4), which is lower compared to other cancers (4,6,7). After analyzing intratumor clonal architecture using the SciClone method (8), we observed monoclonal, biclonal and multiclonal signatures in PTCs (Supplementary Material, Fig. S2).

Mutational spectrum showed that the C>T/G>A substitution was the most common pattern of SNVs in the exomic region, comprising 43.3% of all somatic mutations (Fig. 1A and B and Supplementary Material, Table S3). Evaluation of the non-negative matrix factorization decompositions (9) indicated a best estimate of three biologically distinct mutational signatures in Chinese PTCs (Fig. 1C), evidently different from the two mutational signatures in Caucasian PTCs (10). Signature A was characterized by overrepresentation of the C>T/G>A substitution at the NpCpG trinucleotides (N is any base and C is the mutated base) (Fig. 1C), which is consistent with spontaneous deamination and possibly not a carcinogen-induced process (11). However, the underlying molecular mechanisms are still unclear. Signature B was composed predominantly of T>G/A>C mutation at TpTpN or NpTpA trinucleotide (T is the mutated base). Signature C was a mutational pattern with its origin currently unknown.

Significantly mutated genes in PTC

We performed mutation significance analyses on PTC in the Chinese population. There were 23 genes with ≥3 somatic mutations (Fig. 2). Ten significantly mutated genes were determined if mutations were found in that gene at a false discovery rate (FDR) of $q < 0.1$ after correction for multiple testing using the MutSigCV method ($P < 10^{-4}$) (Supplementary Material, Table S4). As expected, *BRAF*, a known oncogenic driver gene and therapeutic target in multiple malignancies (12), was identified as the most frequently mutated gene with mutations present in 58.2% PTC patients (53/91) at the WES discovery stage (Fig. 2 and Supplementary Material, Table S6). Among those 53 cases, 52 only carried the c.1799T>A (p.Val600Glu, V600E) mutation and one case carried both V600E and c.1806T>G (p.Ser602Ser, S602S) mutations (Fig. 2, Supplementary Material, Table S4 and Fig. S3). In the validation study, the *BRAF* V600E mutation was found among 184 PTCs, with a mutation rate of 59.2% (184/311) (Table 1) and the S602S mutation was observed in two cases who also carried the V600E mutation. Therefore, there were 59.0% Chinese patients (237/402) carrying the V600E mutation. As a common molecular change, the *BRAF* V600E mutation is estimated to occur in over 25% of human malignancies, especially in thyroid cancer (Supplementary Material, Fig. S4). In assessing the relationship of the *BRAF* V600E mutation to the clinicopathological characteristics of PTC patients, we found that more PTC patients harboring the *BRAF* V600E mutation had lymph node metastasis compared with patients without the mutation (Discovery, $P = 0.005$; validation, $P = 0.005$; all, $P = 1.6 \times 10^{-4}$) (Table 1). Additionally, there was a significantly higher rate of this mutation identified among females (Discovery, $P = 0.047$; validation, $P = 3.7 \times 10^{-5}$; all, $P = 4.8 \times 10^{-6}$). However, no significantly associations were observed between the mutation and the other clinicopathological characteristics (Table 1).

lncRNA GAS8-AS1 and LPAR4 as novel driver alternations in PTC

Notably, we found that a lncRNA gene *GAS8-AS1* is the most frequently altered gene except *BRAF*. In discovery stage, we observed 8 of 91 PTC cases (8.8%) harboring somatic mutations (Figs 2 and

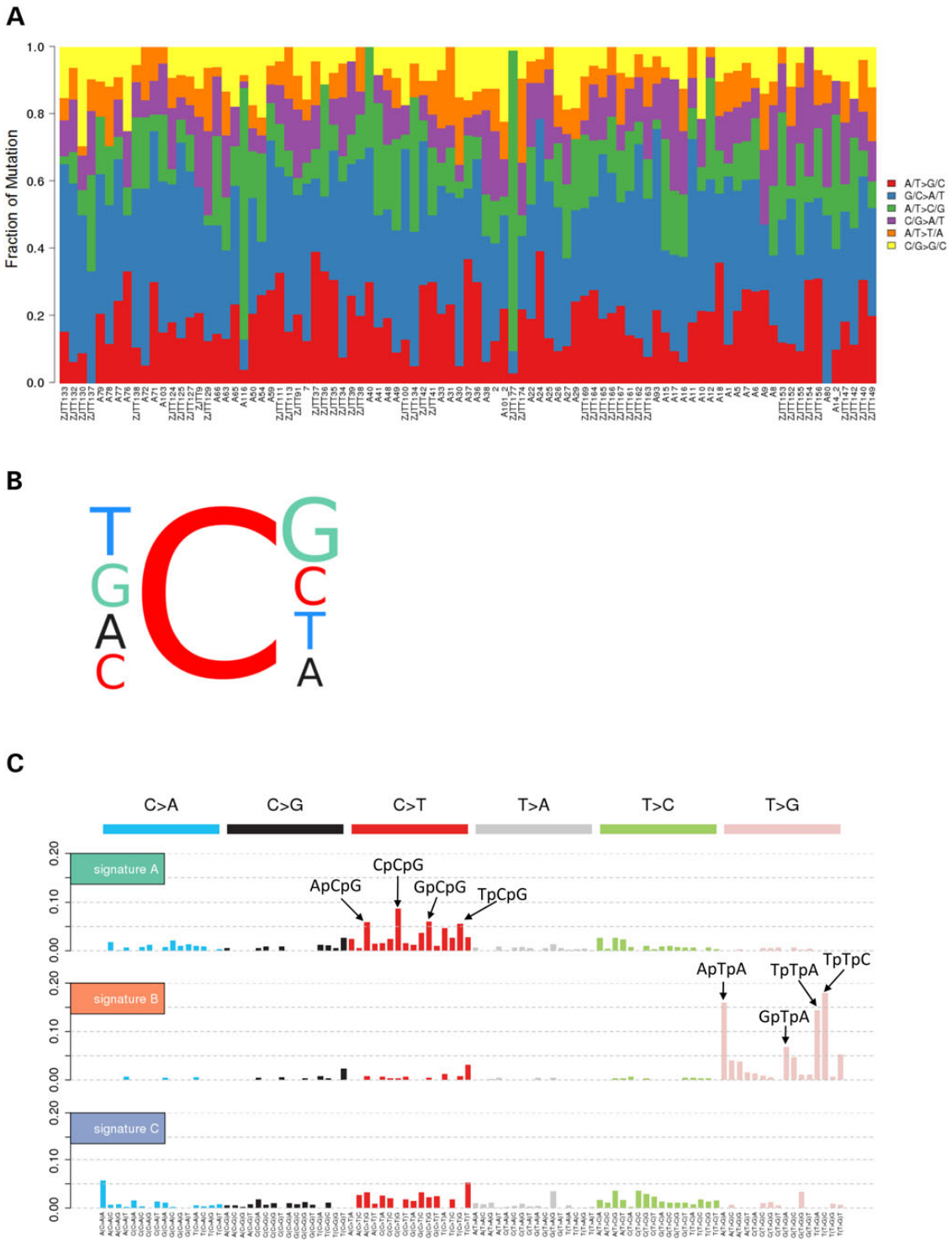

Figure 1. Mutation spectra and three mutational signatures identified in Chinese PTCs across the 91 patients. (**A**) The somatic point mutation spectrum observed in each of the 91 PTC cases. Overall, the C>T/G>A substitution was more commonly observed on average than other mutation types. (**B**) Trinucleotide contexts of mutations occurring at cytosine nucleotides in PTC. Font sizes of the bases at the 5′ and 3′ positions are proportional to their frequencies. (**C**) Ninety-six substitution classifications from WES data derived from the 91 pairs of Chinese PTC samples. Mutation types are displayed in different colors on the horizontal axis. The vertical axis depicts the percentage of mutations attributed to a specific mutation type.

3A and Supplementary Material, Table S4). Interestingly, among these affected patients, seven carried a c.713A>G/714T>C dinucleotide substitution and only one case carried a c.713A>G mononucleotide mutation (Fig. 3A). Among 311 validation patients, there were 29 cases carrying either the c.713A>G/714T>C dinucleotide substitution ($n = 26$) or the c.713A>G mononucleotide mutation ($n = 3$). As shown in Table 2, we found that the dinucleotide substitution was significantly associated with

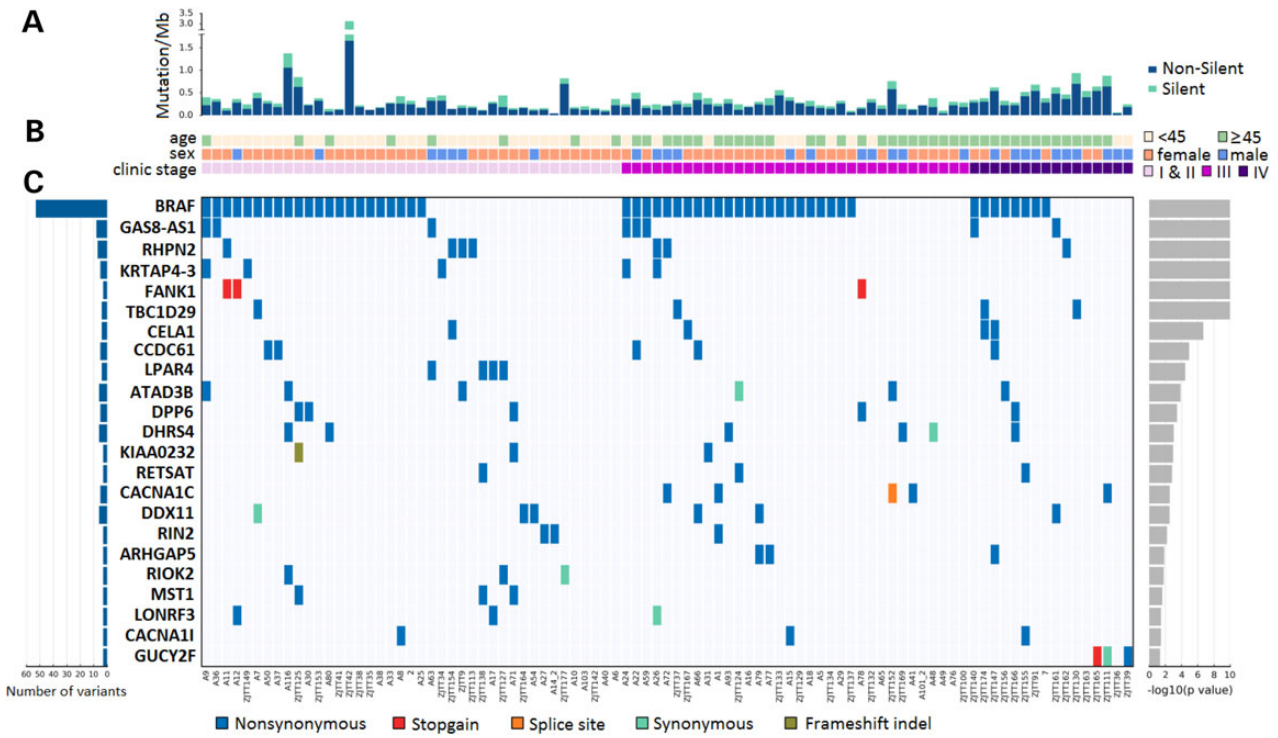

Figure 2. Exome-wide mutational landscape of PTC identified by whole-exome sequencing. (**A**) Number and type of mutations. Non-silent mutations consist of missense, stopgain, stoploss, and splice site mutations, as well as frameshift indels and inframe indels. (**B**) Key clinicopathological characteristics. Stage was determined according to the sixth edition of the Union for International Cancer Control (UICC) staging system for thyroid cancer. (**C**) Middle, mutations in a selection of frequently mutated genes with ≥3 mutations, colored by the type of alteration. Samples are displayed as columns, arranged to emphasize mutual exclusivity among mutations. Left, the total number of somatic alterations that targeted each gene. Right, the logarithmic (log) transformations of *P*-values calculated by MutSigCV (Supplementary Material, Table S6).

Table 1. Associations of *BRAF* V600E mutation status with the clinicopathological characteristics of 402 Chinese PTCs

Variable	Category	Discovery stage ($n = 91$)		Validation stage ($n = 311$)		All ($n = 402$)	
		V600E positive (%)	*P*-value[a]	V600E positive (%)	*P*-value[a]	V600E positive (%)	*P*-value[a]
Age	<45	26 (59.1)	0.874	92 (60.1)	0.733	118 (59.9)	0.706
	≥45	27 (57.4)		92 (58.3)		119 (58.0)	
Sex	Male	12 (42.9)	0.047	43 (42.6)	3.7×10^{-5}	55 (42.6)	4.8×10^{-6}
	Female	41 (65.1)		141 (67.1)		182 (66.7)	
TNM stage	I + II	30 (58.8)	0.911	105 (60.0)	0.714	135 (59.7)	0.653
	III	15 (60.0)		48 (60.8)		63 (60.6)	
	IV	8 (53.3)		31 (54.4)		39 (54.2)	
Lymph node metastasis	Yes	28 (73.0)	0.005	96 (67.6)	0.005	124 (69.3)	1.6×10^{-4}
	No	25 (48.1)		88 (56.1)		113 (50.7)	

PTC, papillary thyroid carcinoma.
[a]Two-sided χ^2 test.

advanced disease (stage IV, 16.7%; stage III, 8.7%; stage I + II, 5.3%; $P = 0.009$). Moreover, we observed a significantly higher rate of this mutation identified among aged cases compared to young cases (13.2 versus 3.0%, $P = 2.2 \times 10^{-4}$).

LncRNA *GAS8-AS1* is located in the second intron of *GAS8* and transcribes a 994nt ncRNA in the opposite orientation of *GAS8* (Fig. 3A). Prediction with the RNAfold software demonstrated that the c.713A>G/714T>C dinucleotide substitution induces significant change of GAS8-AS1 RNA secondary structure (Fig. 3B). We next sought to determine the involvement of *GAS8-AS1* in PTC etiology. Significant decrease in the expression of lncRNA GAS8-AS1 was observed in PTC tissue specimens compared to normal tissues of patients recruited from different medical centers (all $P < 0.001$) (Fig. 3C). Ectopically expressed either wild-type lncRNA GAS8-AS1 ($A_{713}T_{714}$) or the mutated one ($G_{713}C_{714}$) could significantly suppress cell viability in multiple thyroid cancer cell lines (Fig. 3D and E). However, wild-type lncRNA GAS8-AS1 showed consistently higher capability to inhibit cell growth compared to the mutated lncRNA (Fig. 3D and E). Depletion of endogenous lncRNA GAS8-AS1 expression with different small interfering RNAs (siRNAs) led to markedly increased PTC cell proliferation (Fig. 3F and G).

Lysophosphatidic acid receptor 4 (*LPAR4*) was another significantly mutated gene in Chinese PTC patients ($P = 3.51 \times 10^{-5}$, FDR

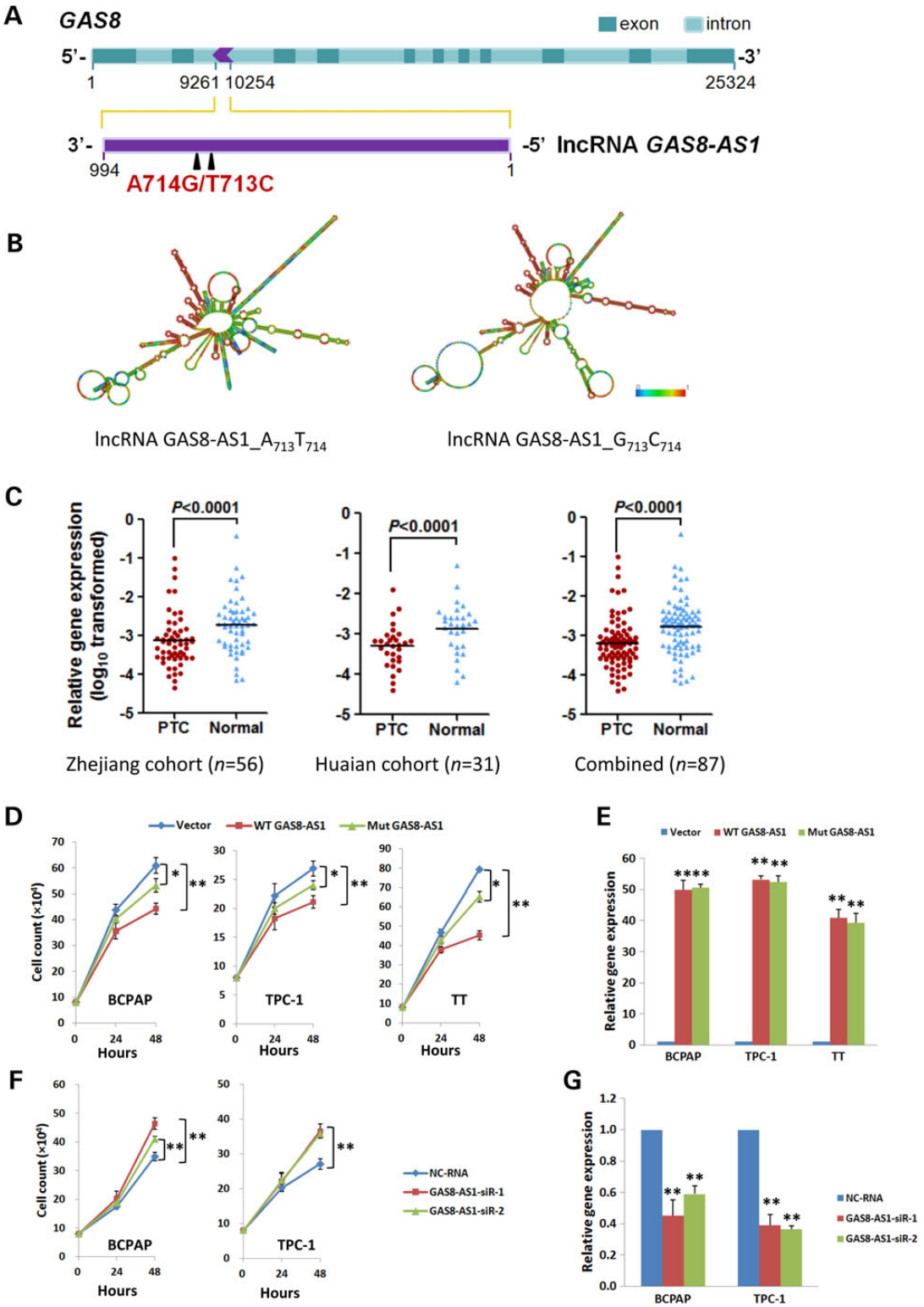

Figure 3. Identification of lncRNA GAS8-AS1 as a novel tumor suppressor gene in PTC. (**A**) Schematics of lncRNA GAS8-AS1, its host gene GAS8 and an $A_{713}T_{714}$>$G_{713}C_{714}$ dinucleotide substitution in lncRNA GAS8-AS1. (**B**) Prediction with the RNAfold software demonstrated that the $A_{713}T_{714}$>$G_{713}C_{714}$ dinucleotide substitution induces significant change of GAS8-AS1 RNA secondary structure. (**C**) lncRNA GAS8-AS1 expression detected in 87 pairs of PTC tissue specimens and normal tissues from different medical centers (56 pairs from Zhejiang Province Cancer Hospital and 31 pairs from Huaian No. 2 Hospital). (**D**) Ectopically expressed either wild-type lncRNA GAS8-AS1 (WT, $A_{713}T_{714}$) or the mutated one (Mut, $G_{713}C_{714}$) could significantly suppress cell viability in PTC cell lines (BCPAP and TPC-1) and thyroid medullary carcinoma TT cell line. (**E**) Ectopic expression levels of WT lncRNA GAS8-AS1 or the Mut GAS8-AS1 are comparable in BCPAP, TPC-1 and TT cell lines. (**F**) Depletion of endogenous lncRNA GAS8-AS1 expression in BCPAP and TPC-1 cells significantly increased cell proliferation. (**G**) Expression levels of lncRNA GAS8-AS1 after RNAi in BCPAP and TPC-1 cells. All experiments were performed in triplicates in three independent transfection experiments and each value represents mean ± SD. *$P < 0.05$; **$P < 0.01$.

Table 2. Associations of *GAS8-AS1* 713A>G/714T>C dinucleotide mutation status with the clinicopathological characteristics of 402 Chinese PTCs

Variable	Category	No. of cases	All PTC cases ($n = 402$) Mutation positive (%)	P-value[a]
Age	<45	197	6 (3.0)	2.2×10^{-4}
	≥45	205	27 (13.2)	
Sex	Male	129	12 (9.3)	0.583
	Female	273	21 (7.7)	
TNM stage	I + II	226	12 (5.3)	0.009
	III	104	9 (8.7)	
	IV	72	12 (16.7)	
Lymph node metastasis	Yes	179	19 (10.6)	0.115
	No	223	14 (6.3)	

PTC, papillary thyroid carcinoma.
[a]Two-sided χ^2 test.

$q = 7.09 \times 10^{-2}$) (Fig. 2 and Supplementary Material, Table S4), with only one mutation pattern (c.872T>G, p.Ile291Ser, I291S) identified (Fig. 4A). As a G protein-coupled receptor (GPCR) family member, LPAR4 is poorly studied in PTC. Notably, we observed that *LPAR2* (another LPAR family member) and multiple *LPAR4* downstream PI3K-AKT pathway genes (i.e. *PIK3CA*, *PIK3R2*, *AKT2*, *PTEN* and *TP53*) were mutated. Intriguingly, the mutations in this set of genes were mutually exclusive and affected 12.1% PTC patients (11/91) (Fig. 4B). By detecting *LPAR4* mRNA expression in 87 paired PTC and normal tissues, we found higher expression of *LPAR4* in PTC tissues compared to normal tissues (all $P < 0.05$) (Fig. 4C), suggesting an oncogenic nature of *LPAR4* in PTC. In PTC or HEK293 cells, ectopic expression of either wild type LPAR4 (codon291Ile) or mutated LPAR4 (codon291Ser) significantly promoted cell proliferation (Fig. 4D and E). Remarkably, cells expressing mutated LPAR4 showed much higher cell viability compared with those expressing wild type LPAR4 (Fig. 4D and E), indicating that the *LPAR4* I1291S somatic mutation is biologically relevant in PTC. On the contrary, depletion of endogenous *LPAR4* expression significantly suppresses PTC cell proliferation (Fig. 4F and G).

Altered pathways in PTC

We then applied Genome MuSiC path-scan analyses to investigate the core canonical pathways with enrichment of somatic mutations in PTC. The thyroid cancer pathway was identified as the most significantly mutated pathway according to KEGG (FDR $q = 2.8 \times 10^{-9}$, $P = 3.1 \times 10^{-8}$) (Supplementary Material, Table S5). In addition, we found that 69.2% of PTC patients (63/91) harbored protein sequence-altering mutations in the *BRAF*-related signaling pathway (*BRAF*, *EGFR*, *SHC2*, *SHC3*, *NF1*, *NRAS*, *RAF1*, *CACNA1A*, *CACNA1C*, *CACNA1E*, *CACNA1F*, *CACNA1H*, *CACNA1I*, *CACNG6*, *RASGRP2* and *RASGRF2*) (Supplementary Material, Fig. S5). Some of the mutations in the pathway were mutually exclusive, especially mutations in *SHC2*, *SHC3*, *CACNA1C*, *CACNA1F* and *CACNA1H*. There are multiple well-known oncogenes, such as *NRAS*, *EGFR* and *SHC3* in this pathway (13–16). The TCGA PTC study identified multiple codon12 or codon61 mutations in three RAS family genes (*NRAS*, *HRAS* and *KRAS*) in 12.9% patients. *NRAS* c.182A>G (p.Gln61Arg, Q61R) and c.181C>A (p.Gln61Lys, Q61K) mutations were also found in 3.2% Chinese PTCs (13/402), demonstrating its essential role in PTC in different ethnic populations. However, we did not observe any somatic alternation within *HRAS* and *KRAS*. Oncogene *EGFR* overexpression in PTC specimens has been implicated in disease progression (13–15). Moreover, oncoprotein SHC3 could form a protein complex with RET oncoprotein and potentiate PI3K signaling pathway in PTC (16).

Discussion

We describe here the first WES of PTC in the Chinese population with a relative large sample size ($n = 402$), and our data reveal that the spectra of exome mutations of Chinese PTCs are distinct from Caucasian PTCs as well as anaplastic thyroid cancer (17). This supports our hypotheses that PTCs in different populations under different environments might present different mutations. Notably, we firstly identified an lncRNA gene, *GAS8-AS1*, is the most frequently altered gene except *BRAF*, a well-known PTC diver gene, indicating that both coding and non-coding genes are involved in PTC etiology.

That comparable high mutation rate of the *BRAF* V600E mutation in the Chinese population (59.0%) and in the TCGA data (58.5%) (4) (Supplementary Material, Fig. S5) indicates that *BRAF* is the most important driver gene for PTC development. The KEGG path-scan analyses showing thyroid cancer pathway as the most significantly mutated one also confirmed the quality of our WES data. Accumulated evidences demonstrated that lncRNAs may play a crucial role in carcinogenesis (18,19). Unexpectedly, we found that lncRNA *GAS8-AS1* ranks the secondary most mutated gene (9.2%) and acts as a tumor suppressor in PTC (Supplementary Material, Table S4). To the best of our knowledge, we here report for the first time that an lncRNA gene was identified as a recurrently mutated cancer driver gene in PTC. On the contrary, the TCGA study found no mutations in *GAS8-AS1* (5), though the same *GAS8-AS1* mutations were observed in adrenocortical carcinoma supporting the existence of the mutation in different ethnic populations (Supplementary Material, Fig. S6). Additionally, TCGA identified *NRAS* as the secondary most altered gene (8.7%), with the same Q61R mutation pattern observed in 2.0% Chinese PTC patients (8/402) (Supplementary Material, Fig. S3). These data suggest that different environment exposure may lead to various somatic changes and, thus, different etiology of PTC in diverse populations, which needs further investigation.

Six known human *LPARs* are usually divided into two families, including the endothelial differentiation gene (Edg) family (*LPAR1*, *LPAR2* and *LPAR3*) as well as the phylogenetically distant non-Edg family (*LPAR4*, *LPAR5* and *LPAR6*). Several *LPARs* have been implicated in tumorigenesis. For example, activated LPAR1 stimulates cell motility and tumor cell invasion (20). Although *LPAR4* somatic mutations with relative low frequency have been identified in various malignancies (Supplementary Material, Fig. S7), the role of LPAR4 in tumorigenesis is still inconclusive. In colon cancer, suppression of LPAR4 induced cell migration and invasion (21). Similarly, ectopic *LPAR4* expression negatively modulated cell motility in head and neck squamous cell carcinoma (22). Conversely, LPAR4 has been shown to induce fibrosarcoma cell invasion and *in vivo* metastasis formation (23), as well as cell transformation and anchorage-independent growth of Myc-transformed cells (24). In line with these studies, we observed that *LPAR4* acts as a vital oncogene in PTC and, therefore, may serve as a therapeutic target for PTC treatment.

In summary, the present study provides a exome mutational spectrum of PTC in the Chinese population. Together with the high frequency of the *BRAF* V600E mutation, the paucity of non-*BRAF* pathway mutations is a prominent feature of Chinese PTCs. We elucidated two novel PTC driver genes (lncRNA *GAS-AS1* as a

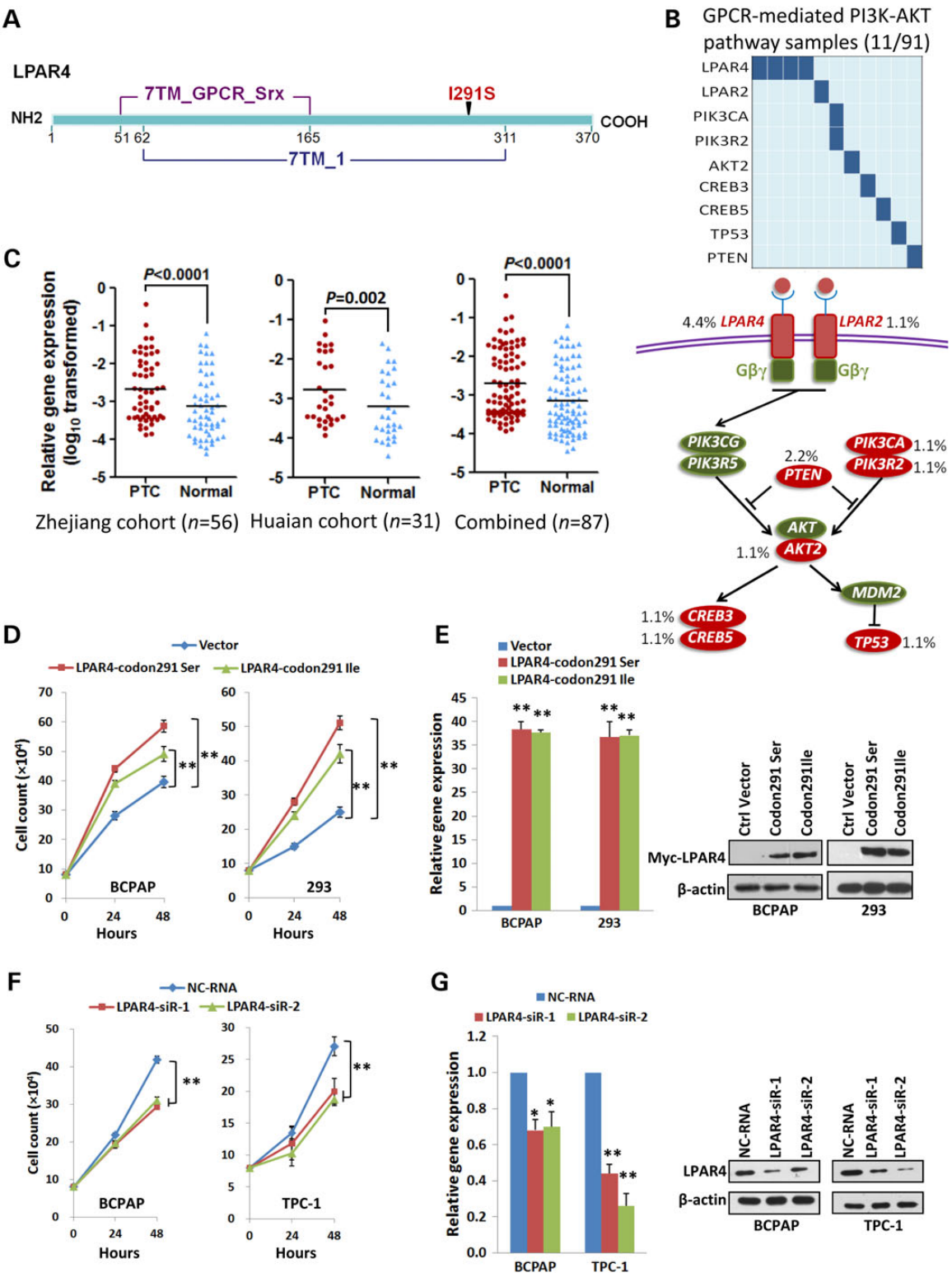

Figure 4. Identification of *LPAR4* as a novel PTC oncogene. (**A**) The schematic of *LPAR4* and an Ile291Ser (I291S) amino acid substitution in *LPAR4*. (**B**) Data matrix showing alterations of genes involved in the GPCR-mediated PI3K-AKT signaling pathway (Upper panel). Simplistic overview of the GPCR-mediated PI3K-AKT signaling pathway in PTC according to KEGG (Lower panel). (**C**) *LPAR4* mRNA expression detected in 87 pairs of PTC tissue specimens and normal tissues from different medical centers (56 pairs from Zhejiang Province Cancer Hospital and 31 pairs from Huaian No. 2 Hospital). (**D**) Ectopically expressed either wild-type *LPAR4* (codon291Ile) or the mutated one (codon291Ser) could significantly stimulate cell viability in BCPAP and HEK293 cells. (**E**) mRNA and protein expression levels of *LPAR4*-codon291Ile or *LPAR4*-codon291Ser were comparable in BCPAP and HEK293 cells. (**F**) Depletion of LPAR4 expression in BCPAP and TPC-1 cells significantly inhibited cell proliferation. (**G**) Expression levels of *LPAR4* mRNA and protein after RNAi in BCPAP and TPC-1 cells. All experiments were performed in triplicates in three independent transfection experiments and each value represents mean ± SD. *$P < 0.05$; **$P < 0.01$.

tumor suppressor and *LPAR4* as an oncogene), which might serve as potential therapeutic targets (25,26). Our findings provide further layers of understanding the tumorigenesis of PTC in the Chinese population and an enhanced road map for development of novel diagnostics and therapeutics for the malignancy.

Materials and Methods

Study subjects

Four hundred and two clinically and pathologically characterized PTC tissues and 402 paired peripheral blood samples for WES discovery and Sanger sequencing validation were collected with written informed consent from individuals with primary PTC who received surgery as their primary treatment at Huaian No. 2 Hospital, Zhejiang Province Cancer Hospital or Cancer Hospital of Chinese Academy of Medical Sciences between June 2011 and March 2014. PTC tissue and matched normal thyroid tissue specimens for gene expression detection were recruited with written informed consent from individuals with primary PTC who received surgery as their primary treatment at Huaian No. 2 Hospital and Zhejiang Province Cancer Hospital between June 2011 and March 2014. Fresh tissue samples were macrodissected within 30 min of surgical resection and stored at liquid nitrogen. Patients' clinicopathological characteristics were extracted from in-patient medical records. All PTC tissues were examined by two certified pathologists independently to confirm histological subtype. Adjacent normal tissue samples were confirmed not to be contaminated by cancer cells. Only DNA samples isolated from PTCs with more than 70% tumor content and their matched blood samples were subjected to WES or validation sequencing. The collection of human samples and the protocols for the investigations were approved by the Institutional Ethics Committees of Zhejiang Province Cancer Hospital, Huaian No. 2 Hospital and Cancer Hospital of Chinese Academy of Medical Sciences.

Whole-exome sequencing

Human genomic DNA was extracted from frozen tissue samples or blood samples with QIAampDNA Mini kits (Qiagen). After Covaris sonication, genomic DNA was fragmented to 180–280 bp fragments. Adaptors were ligated to both ends of the fragments and adaptor-ligated templates were then purified using AgencourtAMPure SPRI beads. Using the SureSelect Human All Exon V5 kit (Agilent) according to the manufacturer's instructions, whole-exome enrichment was performed. Captured DNA libraries underwent paired-end sequencing on the HiSeq 2000 Genome Analyzer (Illumina), resulting in sequences of 100 bases from each end of the fragments.

Somatic mutation calling

Sequencing reads were aligned to the hg19 reference genome using Burrows-Wheeler Aligner (BWA). The SAM output was converted to BAM output and sorted by Samtools, duplicates were marked using Picard, and indel realignment and quality score recalibration were performed in Genome Analysis Toolkit (GATK). The MuTect algorithm was used to identify somatic SNVs in WES data with default parameters. Somatic indels were identified by Strelka algorithm and somatic CNVs were identified by Control-FREEC.

Sanger validation of mutation calls

A total of 90 somatic SNVs and 20 indels identified by WES were verified by Sanger sequencing in all 402 tumor-normal pairs. In total, 107 loci were successfully amplified. Excluding 3 regions that failed in PCR, 86 of 88 SNVs (97.7%) and 18 of 19 indels (94.7%) were successfully validated. In all, 104 of 107 mutations (97.2%) were validated.

Cell culture

The human HEK293 cell line, human PTC cell lines (BCPAP and TPC-1) and thyroid medullary carcinoma cell line TT were used in the current study. The HEK293 and TT cell lines were obtained from the American Type Culture Collection (ATCC). The thyroid origins of BCPAP and TPC-1 cell lines were confirmed by profiling analyses. BCPAP and TPC-1 cells were cultured in RPMI-1640 medium (Gibco). HEK293 cells were cultured in DMEM medium (Gibco). TT cells were cultured in F-12K medium (Gibco). All culture medium contains 10% fetal bovine serum (Gibco), 100 units/ml penicillin (Invitrogen) and 100 mg/ml streptomycin (Invitrogen).

mRNA quantification

The relative lncRNA or mRNA levels of genes targeted by RNA interference or gene overexpression were examined by quantitative RT-PCR as previously described. In brief, total RNA was extracted from tissue samples or cultured cells using TRIzol Reagent (Invitrogen). Each RNA sample was then treated with RNase-Free DNase to remove genomic DNA (Invitrogen). We utilized SYBR-Green real-time quantitative PCR method to measure lncRNA or mRNA levels in tissues or cell lines (Supplementary Material, Table S6). The expression of individual measurements was calculated relative to expression of *β-actin* using the 2^{-dCt} method as described previously (27–30).

Western blotting

After total cellular proteins were separated with SDS-PAGE gel, proteins were transferred to a polyvinylidene fluoride (PVDF) membrane (Millipore). The PVDF membrane was then incubated with c-Myc primary antibody (Sigma, M4439; 1:1000 dilution) or LPAR4 primary antibody (Santa Cruz, H-60; 1:500 dilution) overnight at 4°C. Target proteins were visualized with enhanced chemiluminescence (ECL) reagents (Millipore).

DNA constructs and mutagenesis

PCR-amplified human *LPAR4* was cloned into pCMV-Myc vector (Clontech) between the *KpnI* and *NotI* restriction sites. The pCMV-Myc-LPAR4 construct containing the I291S mutation was generated using the QuikChange site-directed mutagenesis kit (Stratagene). Restriction analysis and complete DNA sequencing confirmed the orientation and integrity of these constructs.

Cell proliferation assays

Cells were seeded in 6-well plates at a density of 8×10^4 cells per well. For RNAi assays, cells were transfected with *GAS8-AS1* siRNAs (GAS8-AS1-siR-1 and GAS8-AS1-siR-2) or *LPAR4* siRNAs (LPAR4-siR-1 and LPAR4-siR-2) (Genepharma, Shanghai, China) (Supplementary Material, Table S7) combined with Lipofectamine RNAi Max (Invitrogen). For gene overexpression assays, cells were transfected with pcDNA3.1-GAS8-AS1 constructs (wildtype or mutated) or pCMV-Myc-LPAR4 constructs (wildtype or mutated) combined with Lipofectamine 2000 (Invitrogen). Cells were counted at 24 and 48 h after transfection as described previously (27–29). All results of the mean of triplicate assays with standard deviation of the mean were presented.

Statistical analyses

The difference between two groups was calculated using Student's *t*-test. One-way analysis of variance was used to compare differences between three or more groups. A *P*-value of less than 0.05 was used as the criterion of statistical significance. All analyses were performed with SPSS software package (Version 16.0, SPSS Inc.) or GraphPad Prism (Version 5, GraphPad Software, Inc.).

Accession numbers

The accession number for the whole-exome sequences reported in this paper is European Genome-phenome Archive (EGA): EGAS00001001268.

Supplementary Material

Supplementary Material is available at *HMG* online.

Acknowledgements

We thank Drs Ling Lin, Huan Yu and Xi Luo from Novogene Bioinformatics Technology Co., Ltd for their assistance in bioinformatics analyses.

Conflict of Interest statement. None declared.

Funding

This work was supported by the National High-Tech Research and Development Program of China (2015AA020950); National Natural Science Foundation of China (31271382); the Fundamental Research Funds for the Central Universities (YS1407); the open project of State Key Laboratory of Molecular Oncology (SKL-KF-2015-05).

References

1. Chen, W., Zheng, R., Zeng, H., Zhang, S. and He, J. (2015) Annual report on status of cancer in China, 2011. *Chin. J. Cancer Res.*, **27**, 2–12.
2. Schneider, A.B. and Sarne, D.H. (2005) Long-term risks for thyroid cancer and other neoplasms after exposure to radiation. *Nat. Clin. Pract. Endocrinol. Metab.*, **1**, 82–91.
3. Goldgar, D.E., Easton, D.F., Cannon-Albright, L.A. and Skolnick, M.H. (1994) Systematic population-based assessment of cancer risk in first-degree relatives of cancer probands. *J. Natl. Cancer Inst.*, **86**, 1600–1608.
4. Cancer Genome Atlas Research Network. (2014) Integrated genomic characterization of papillary thyroid carcinoma. *Cell*, **159**, 676–690.
5. Cibulskis, K., Lawrence, M.S., Carter, S.L., Sivachenko, A., Jaffe, D., Sougnez, C., Gabriel, S., Meyerson, M., Lander, E.S. and Getz, G. (2013) Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.*, **31**, 213–219.
6. Lawrence, M.S., Stojanov, P., Polak, P., Kryukov, G.V., Cibulskis, K., Sivachenko, A., Carter, S.L., Stewart, C., Mermel, C.H., Roberts, S.A. *et al.* (2013) Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature*, **499**, 214–218.
7. Lawrence, M.S., Stojanov, P., Mermel, C.H., Robinson, J.T., Garraway, L.A., Golub, T.R., Meyerson, M., Gabriel, S.B., Lander, E.S. and Getz, G. (2014) Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature*, **505**, 495–501.
8. Miller, C.A., White, B.S., Dees, N.D., Griffith, M., Welch, J.S., Griffith, O.L., Vij, R., Tomasson, M.H., Graubert, T.A., Walter, M.J. *et al.* (2014) SciClone: inferring clonal architecture and tracking the spatial and temporal patterns of tumor evolution. *PLoS Comput. Biol.*, **10**, e1003665.
9. Lee, D.D. and Seung, H.S. (1999) Learning the parts of objects by non-negative matrix factorization. *Nature*, **401**, 788–791.
10. Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Aparicio, S.A., Behjati, S., Biankin, A.V., Bignell, G.R., Bolli, N., Borg, A., Børresen-Dale, A.L. *et al.* (2013) Signatures of mutational processes in human cancer. *Nature*, **500**, 415–421.
11. Waters, T.R. and Swann, P.F. (2000) Thymine-DNA glycosylase and G to A transition mutations at CpG sites. *Mutat. Res.*, **462**, 137–147.
12. Holderfield, M., Deuker, M.M., McCormick, F. and McMahon, M. (2014) Targeting RAF kinases for cancer therapy: BRAF-mutated melanoma and beyond. *Nat. Rev. Cancer*, **14**, 455–467.
13. Tang, C., Yang, L., Wang, N., Li, L., Xu, M., Chen, G.G. and Liu, Z.M. (2014) High expression of GPER1, EGFR and CXCR1 is associated with lymph node metastasis in papillary thyroid carcinoma. *Int. J. Clin. Exp. Pathol.*, **7**, 3213–3223.
14. Sharma, G.K., Dhillon, V.K., Masood, R. and Maceri, D.R. (2015) Overexpression of EphB4, EphrinB2, and epidermal growth factor receptor in papillary thyroid carcinoma: a pilot study. *Head Neck*, **37**, 964–969.
15. Lam, A.K., Lau, K.K., Gopalan, V., Luk, J. and Lo, C.Y. (2011) Quantitative analysis of the expression of TGF-alpha and EGFR in papillary thyroid carcinoma: clinicopathological relevance. *Pathology*, **43**, 40–47.
16. De Falco, V., Guarino, V., Malorni, L., Cirafici, A.M., Troglio, F., Erreni, M., Pelicci, G., Santoro, M. and Melillo, R.M. (2005) RAI(ShcC/N-Shc)-dependent recruitment of GAB 1 to RET oncoproteins potentiates PI 3-K signalling in thyroid tumors. *Oncogene*, **24**, 6303–6313.
17. Kunstman, J.W., Juhlin, C.C., Goh, G., Brown, T.C., Stenman, A., Healy, J.M., Rubinstein, J.C., Choi, M., Kiss, N., Nelson-Williams, C. *et al.* (2015) Characterization of the mutational landscape of anaplastic thyroid cancer via whole-exome sequencing. *Hum. Mol. Genet.*, **24**, 2318–2329.
18. Prensner, J.R. and Chinnaiyan, A.M. (2011) The emergence of lncRNAs in cancer biology. *Cancer Discov.*, **1**, 391–407.
19. Batista, P.J. and Chang, H.Y. (2013) Long noncoding RNAs: cellular address codes in development and disease. *Cell*, **152**, 1298–1307.
20. Hama, K., Aoki, J., Fukaya, M., Kishi, Y., Sakai, T., Suzuki, R., Ohta, H., Yamori, T., Watanabe, M., Chun, J. *et al.* (2004) Lysophosphatidic acid and autotaxin stimulate cell motility of neoplastic and non-neoplastic cells through LPA1. *J. Biol. Chem.*, **279**, 17634–17639.
21. Lee, Z., Cheng, C.T., Zhang, H., Subler, M.A., Wu, J., Mukherjee, A., Windle, J.J., Chen, C.K. and Fang, X. (2008) Role of LPA4/p2y9/GPR23 in negative regulation of cell motility. *Mol. Biol. Cell*, **19**, 5435–5445.
22. Matayoshi, S., Chiba, S., Lin, Y., Arakaki, K., Matsumoto, H., Nakanishi, T., Suzuki, M. and Kato, S. (2013) Lysophosphatidic acid receptor 4 signaling potentially modulates malignant behavior in human head and neck squamous cell carcinoma cells. *Int. J. Oncol.*, **42**, 1560–1568.
23. Harper, K., Arsenault, D., Boulay-Jean, S., Lauzier, A., Lucien, F. and Dubois, C.M. (2010) Autotaxin promotes cancer invasion via the lysophosphatidic acid receptor 4: participation of the cyclic AMP/EPAC/Rac1 signaling pathway in invadopodia formation. *Cancer Res.*, **70**, 4634–4643.
24. Taghavi, P., Verhoeven, E., Jacobs, J.J., Lambooij, J.P., Stortelers, C., Tanger, E., Moolenaar, W.H. and van Lohuizen, M.

(2008) In vitro genetic screen identifies a cooperative role for LPA signaling and c-Myc in cell transformation. *Oncogene*, **27**, 6806–6816.
25. Lappano, R. and Maggiolini, M. (2011) G protein-coupled receptors: novel targets for drug discovery in cancer. *Nat. Rev. Drug Discov.*, **10**, 47–60.
26. Wahlestedt, C. (2013) Targeting long non-coding RNA to therapeutically upregulate gene expression. *Nat. Rev. Drug Discov.*, **12**, 433–446.
27. Wang, X., Li, M., Wang, Z., Han, S., Tang, X., Ge, Y., Zhou, L., Zhou, C., Yuan, Q. and Yang, M. (2015) Silencing of long noncoding RNA MALAT1 by miR-101 and miR-217 inhibits proliferation, migration, and invasion of esophageal squamous cell carcinoma cells. *J. Biol. Chem.*, **290**, 3925–3935.
28. Zhang, X., Zhou, L., Fu, G., Sun, F., Shi, J., Wei, J., Lu, C., Zhou, C., Yuan, Q. and Yang, M. (2014) The identification of an ESCC susceptibility SNP rs920778 that regulates the expression of lncRNA HOTAIR via a novel intronic enhancer. *Carcinogenesis*, **35**, 2062–2067.
29. Zhang, X., Wei, J., Zhou, L., Zhou, C., Shi, J., Yuan, Q., Yang, M. and Lin, D. (2013) A functional BRCA1 coding sequence genetic variant contributes to risk of esophageal squamous cell carcinoma. *Carcinogenesis*, **34**, 2309–2313.
30. Liu, L., Zhou, C., Zhou, L., Peng, L., Li, D., Zhang, X., Zhou, M., Kuang, P., Yuan, Q., Song, X. *et al.* (2012) Functional FEN1 genetic variants contribute to risk of hepatocellular carcinoma, esophageal cancer, gastric cancer and colorectal cancer. *Carcinogenesis*, **33**, 119–123.