Large contribution of copy number alterations in early stage of Papillary Thyroid Carcinoma


Link to publication record in Ulster University Research Portal

Published in: Computers in biology and medicine

Publication Status: Published (in print/issue): 31/08/2021

DOI: 10.1016/j.compbiomed.2021.104584

Document Version Author Accepted version

General rights Copyright for the publications made accessible via Ulster University's Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy The Research Portal is Ulster University's institutional repository that provides access to Ulster's research outputs. Every effort has been made to ensure that content in the Research Portal does not infringe any person’s rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact pure-support@ulster.ac.uk.

Download date: 15/09/2023
Large contribution of copy number alterations in early stage of Papillary Thyroid Carcinoma

Nazanin Hosseinkhan¹, Maryam Honardoost¹, Kevin Blighe², Tara Moore ², Mohammad E. Khamseh¹*

¹ Endocrine Research Center, Institute of Endocrinology and Metabolism, Iran University of Medical Sciences, Tehran, Iran, ² Biomedical Sciences Research Institute, Ulster University, Coleraine, Northern Ireland

Correspondence to:

E-mail addresses: khamseh.m@iums.ac.ir

Phone: +98(21)88945247

Fax: +98(21)88945173

ORCID ID: https://orcid.org/0000-0003-4313-8440

Running title: Genomic landscape of Papillary Thyroid Carcinoma initiation
Abstract

Papillary Thyroid Carcinoma (PTC) accounts for approximately 85% of patients with thyroid cancer. Despite its indolent nature, progression to higher stages is expected in a subgroup of patients. Hence, genomic characterization of the early stages of PTC may help to identify this subgroup, leading to better clinical management.

Here, we conducted a comprehensive mutational and somatic copy number alteration (SCNA) investigation on 277 stage one PTC from TCGA.

SCNA analysis revealed amplification and deletion of several cancer related genes. We found amplification of 60 oncogenes (Oncs), from which 15 were recurrently observed. Deletion of 58 tumor suppressors (TSs) was also detected. MAPK, PI3K-Akt, Rap1 and Ras were the signaling pathways with large numbers of amplified Oncs. On the other hand, deleted TSs belonged mostly to cell cycle, PI3K-Akt, mTOR and cellular senescence pathways. This suggests that despite heterogeneity in SCNA events, the final results would be the activation/deactivation of few cancer signaling pathways. Of note, despite large amounts of heterogeneity in stage one PTC, recurrent broad deletion on Chr22 was detected in 21 individuals, leading to deletion of several tumor suppressors.

In parallel, the oncogenic/pathogenic mutations in the RTK-RAS and PI3k-Akt pathways were detected. However, no pathogenic mutation was identified in known tumor suppressor genes. In order to identify a potential subgroup of BRAF (V600E) positive patients, who might progress to higher stages, low frequency mutations accompanying BRAF (V600E) were also identified.

In conclusion, our findings imply that SCNA have a substantial contribution to early stages of PTC. Experimental validation of the observed genomic alterations, could help to stratify patients at the time of diagnosis, and to move toward precision medicine in PTC.

Key words: Papillary thyroid carcinoma, Early stage, Genomic landscape, Mutations, Somatic copy number alterations
1. Introduction

Several decades of cancer research confirms a long incubation time of tumor lesion development.
This provides a great opportunity to detect early precancerous lesions and to intervene during
the initiation and progression of carcinogenesis [1]. During tumor initiation, oncogenes may bear
activating mutations or be subjected to gene amplification, while tumor suppressors commonly
harbor inactivating mutations or acquire gene deletions. In this model, the accumulation of
somatic mutations or copy number alterations can confer a clonal advantage to a single aberrant
cell that subsequently is positively selected for during cancer evolution, resulting in the
generation of a malignant clone [2, 3].

Due to the lower frequency of background passenger mutations, and also difficulties in predicting
functionality and pathogenicity of genetic variants from sequence data, distinguishing driver
mutations from passengers is a challenging task [4]. The high frequency (or recurring) mutations
across patients with certain types of cancers, is already the most reliable characteristic of driver
mutations [5]. Recent studies, however indicated that the balance between the rate of DNA
replication errors and DNA repair in distinct genomic regions can considerably vary, affecting the
frequency of the occurrence of different mutations. This means that even low frequency
mutations which occur in the regions with low background mutability, can have pathogenic
effects and therefore, taking the background mutability into the account can efficiently assist in
prioritizing mutations [6]. Of note, for several types of cancers, mutations in known cancer driver
genes are rare, with most of these cancers harboring genetic alterations with intermediate (2-
22%) or low (<2%) frequencies in non-driver genes, indicating the possible contribution of low
frequency mutations in promoting cancer initiation[7].

Somatic copy number alterations (SCNAs) can affect larger fractions of genomes than any other
type of somatic variation. A comprehensive investigation of 12 tumor types demonstrated that
the frequency of copy number alterations inversely correlated with mutational events in distinct
tumors and suggested that each cancer type can be considered as mutation- or copy number
alteration-dominant [8]. In another recent pan-cancer study of 16 different tumors, Smith &
Sheltzer investigated the association of mutations and SCNAs with the cancer survival rate, and indicated that prognostic biomarkers are predominantly found among copy number altered genes. They showed, for instance, that amplification of *EGFR*, *PIK3CA* and *BRAF* genes, strongly associated with poor survival in at least 4 different tumor types, while mutations in these oncogenes were largely uninformative [9].

An increasing body of evidence has demonstrated a large degree of genomic heterogeneity among patients with the same tumor type (inter-tumoral), or between tumor cells within a single tumor sample (intra-tumoral). While the former is a major obstacle toward categorizing patients into distinct genomic subtypes, the latter has impacts on response to treatment and also is the main cause of tumor relapse [10, 11].

Papillary Thyroid Carcinoma (PTC) accounts for approximately 85% of all thyroid cancer cases. A body of evidence has partly elucidated the underlying molecular mechanisms of PTC initiation, which include *RET*/PTC and TRK rearrangements, in addition to *BRAF* (V600E) and *RAS* mutations [12-17]. Nevertheless, the incidence and specificity of the suggested tumor markers considerably vary, impeding their clinical applications [18]. In addition, since benign thyroid nodules also show several mutational aberrations, these mutations seem to be insufficient to lead to thyroid carcinoma without accompanying other complementary molecular events [12].

To address the genomic complexities of PTC initiation, we examined the mutational and SCNA landscapes in stage one of PTC. Our major goal was to evaluate the amount of contribution of each of these genomic events in PTC initiation. Due to its indolent nature, small number of stage one PTC tumors progress to higher stages. Thus, we proposed that low frequency mutations accompanied by *BRAF* (V600E) mutations, in a subset of patients, are probably required for the complete pathogenic effect of *BRAF* mutation. We believe that after confirmation the complementary effect of these low frequency mutations for *BRAF* (V600E) mutation, they could be efficiently implemented to identify and stratify stage one PTC patients with potential poor prognosis.
2. Methods

2.1 Mutation analysis

hg19 mutation data for THCA (Thyroid Cancer) was retrieved from TCGA using the TCGAbiolinks package in R [19]; using the GDCquery() function with the following parameters: data.category = ‘simple nucleotide variation’; and file.type = ‘bcgsc.ca_THCA.IlluminaHiSeq_DNASeq.1.somatic.maf’. We applied the ‘Mutect’ pipeline in TCGAbiolinks to get the mutation annotation format (MAF) file. Further investigation of the identified mutations was then carried out using the ‘maftools’ package.

SIFT (Sorting Intolerant From Tolerant, [20]) and VEP (variant effect predictor, [21]) mutation annotation tools were then used to prioritize pathogenic mutations for further analysis.

We also searched ‘MutaGene’ [22] web server for the identified mutations, by selecting thyroid cancer as the cancer type. MutaGene provides lists of driver and potential driver mutations based on the local background mutability in different cancers. The authors claim that tumor suppressors with higher background mutability have higher recurrence frequency, while highly recurrent oncogenes are characterized by relatively low background mutability.

2.2- Somatic copy number alteration (SCNA) analysis

PTC copy number variation data was retrieved from the “genome wide snp 6-segmented scna minus germline CNV hg19 (MD5)” file from “FireBrowse.org” (Broad Institute of Harvard & MIT). This is the level 3 Affymetrix SNP 6.0 data of the TCGA and has been pre-processed as follows: the probe level SCNA has been calculated as LRR (log R ratio), that is, the ratio of the signal intensity of tumor samples and paired normal samples. Then, using CBS (circular binary segmentation), LRR values have been segmented at the gene level and the “segment mean values” were produced. Extra processing has then been performed to remove germline CNV, with the final data then being deposited at FireBrowse. Here we used the “GAIA” [23] package in R 3.5.0 to find recurrent SCNAs. We defined absolute 0.3 as the cutoff point for amplification/deletion based on the “segment_mean” values. The “runGAIA()” function was used with a q-value
threshold of 0.15 to select the final recurrent SCNA. Finally, we employed BiomaRt and GenomicRanges packages to annotate the identified aberrations.

3 Results & Discussion

3.1 mutational events of stage one PTC

Table 1 represents the total number of mutation types as well as total number of samples with corresponding mutational events. In total, 237 out of 277 stage one PTC patients (85.6%), showed at least one type of detected mutational alterations (Figure 1.a). Figure 1.b represents top 20 most targeted genes by mutational events. The frequency of somatic mutations in stage one PTC was far less than what is generally indicated for most cancers (i.e., about 0.001 to more than 400 per Mb [4]), ranging from 0.0006 to 0.15 per Mb mutations. There were six patients with more than 100 mutations and 151 patients with less than 20 mutations (Figure 1.b).

Several mutational events associated with the RTK-RAS, NRF2, TP53, MYC, TGF-Beta, PI3K, Hippo, WNT and NOTCH signaling pathways were detected, which target 34 tumor suppressors and 27 oncogenes (Table 2).

Mutational events with “high/moderate” and “damaging/damaging due to stop” impacts that annotated simultaneously by VEP and SIFT tools, were selected for further investigation, respectively. In total 1209, 2 and 4 pathogenic missense, frame shift deletion and translation start site mutations were detected in stage one PTC, respectively (Table S1). Pathway annotations of the genes harboring pathogenic mutations revealed several signaling and cancer associated pathways. Of note, several members of “thyroid cancer”, “thyroid hormone synthesis” and “thyroid hormone signaling” pathways were identified.

Except for BRAF, NRAS and HRAS (all involved in thyroid cancer pathway) mutations which were detected in 128, 23 and 13 stage 1 PTC patients, no other pathogenic mutations were identified in more than 2 samples. However, it is possible that damaging mutations occur in different genes
in a biological pathway in different patients and despite the low frequency of observed mutations, the same phenotypic changes would be resulted. For example, none of the pathogenic mutations in PRKCB, ATP1A3, ADCY2, ADCY9, ATFF2, TSHR, PLCB2, PLCB3, PLCB4, CANX and GNAS genes that are involved in thyrocyte growth, differentiation and thyroid hormone secretion were observed in ≥ 2 patients. However, inspecting their role in “thyroid hormone signaling pathway” showed that the consequence of all of these pathogenic mutations would be the hypothyroidism. Several lines of research have demonstrated the association between subclinical hypothyroidism and the cancer incidence and mortality rates for some malignancies including colorectal, breast, prostate, liver and thyroid cancer [24-27].

We further specified the clinical significance of the identified mutations in tumor suppressors and oncogenes using VEP and SIFT tools. None of the identified mutations in 34 tumor suppressors were of high clinical significance, while Among mutations in the 27 oncogenes, NRAS (rs11554290, rs121913254), HRAS (rs121913233), KRAS (rs121913238, rs121913529), AKT1 (rs121434592), and BRAF (rs121913364, rs113488022: recurrent [V600E]) mutations were “pathogenic/likely pathogenic” in VEP or “damaging” in SIFT. BRAF, HRAS, NRAS and KRAS participate in RTK-RAS pathway and AKT1 and PIK3CA are involved in PI3K pathway. The role of all of these pathogenic mutations in the dysregulation of MAPK and PI3K/AKT pathways in PTC initiation have been well indicated [28].

We further identified mutational mutual exclusivity between BRAF-NRAS, BRAF-HRAS, BRAF-MKI67 and BRAF-FRG1 gene pairs. The observed mutual exclusivity between BRAF and other oncogenes can be explained by the functional redundancy provided through the activation of two oncogenes, particularly for those participating in the same signaling pathway. Moreover, it has been demonstrated that turning on two oncogenes, at the same time, could be harmful for tumor cells, promoting their senescence or death [29]. NRAS and HRAS oncogenes are direct activators of BRAF in the MAPK pathway; thus, oncogenic mutations in NRAS and HRAS seems to be sufficient for MAPK activation, promoting cell proliferation and survival.
Furthermore, recent studies showed the contribution of *MKI67* and *FRG1* in the tumorigenesis of several types of carcinomas including PTC, partly explaining their mutational mutual exclusivity with *BRAF* mutation [30-34].

### 3.1.1 Potential novel driver mutations based on background mutability

As previously stated, the occurrence of mutations in regions with less possibility of bearing mutations (low background mutability), is a way to find if a mutational event in a gene would be harmful and that the targeted gene could be considered as novel driver mutation. Several potential driver mutations were identified based on local background mutability provided by MutaGene web server (Table 3). These potential novel mutations are involved in the NOTCH, HIPPO, PI3K-AKT, RAS-RTK, and MAPK signaling pathways. However, none of the identified potential driver mutations were detected in more than 2 samples and their low frequencies could exclude their contribution in PTC initiation.

### 3.1.2 Low frequency mutations accompanying *BRAF* (V600E) mutation

Several research groups have demonstrated that few *BRAF* (V600E)-positive stage one PTC patients progress to higher stage cancer [35]. Previous studies hypothesized that low frequency mutations accompanying *BRAF* mutation in a subset of *BRAF* positive patients will lead to the complete pathogenic effects of *BRAF* [22]. Here, we identified 3 genes: FLG, KRTAP10-10 and F5, that harbor low frequency mutations, accompanied by *BRAF* (V600E) mutation in ≥ 3 stage 1 PTC patients (Figure 2). Among the identified low frequency mutations, FLG mutations are the most frequent (identified in 11 patients) and has a higher possibility to be the major *BRAF* (V600E) mutation contributor. Filaggrin (FLG) is an important epidermal protein highly expressed in the outer layer of epidermis that establishes the skin barrier. FLG abnormalities are associated with three skin diseases: atopic dermatitis, ichthyosis vulgaris and psoriasis vulgaris [36], and its possible contribution to PTC initiation remains to be understood.

### 3.2 SCNA events of stage one PTC

In total, 156 gains and 167 losses were identified in 25 and 96 stage one PTC patients (43.77%), respectively. Figure 3.a represents the distribution of gains and losses in stage one PTC, per
chromosome. We defined alterations < 3Mb as focal, > 3Mb as broad, and those covering >98% of a chromosomal arm as arm-level. Focusing on oncogenes and tumor suppressors, in overall, 60 broad amplifications of proto-oncogenes, of which 15 were recurrently observed in ≥ 3 stage one patients, arm-level gains of q arm of chromosomes 5, 7, 12, 16 and 17, targeting 52 proto-oncogenes (Table S2), arm-level deletion of chromosomes 2, 8, 9, 11 and 13 targeting 41 tumor suppressors, and broad and focal deletion of 56 and 2 tumor suppressors, (Table S3) were identified, respectively. Figure 3. b shows the number of identified arm-level, broad and focal gains-losses in proto-oncogenes and tumor suppressors. MAPK, Rap1, PI3K-Akt, Ras and mTOR signaling pathways were the top 5 pathways with the large number (19, 18, 18, 14 and 10, respectively) of amplified oncogenes (Figure 4.a). The most targeted pathways by loss of tumor suppressors include cell cycle, PI3K-Akt, cellular senescence, mTOR and P53, with respectively 7, 6, 5, 5 and 4 tumor suppressors (Figure 4.b).

Of note, among the identified broad deletions, a recurrent broad deletion on chr22 q arm (in 21 stage one PTC), were identified, encompassing 720 genes, including several members of the Hippo, TGFB, FOXO1, MAPK, RAS, PI3K-AKT, JAK-STAT, P53, and mTOR signaling pathways As well as six tumor suppressors — CHEK2, MN1, NF2, RASL10A, SMARCB1 and SUSD2 (Figure 5). As the activator of P53, the functional product of CHEK2 (Checkpoint kinase 2), regulates cell division. Inactivating mutation of CHEK2 has been reported in a variety of cancers including PTC [37]. In a recent research, Borun and colleagues showed that NF2 (aka Merlin) deletion results into the activation of Ras expression in PTC, and induces cell proliferation [38]. In addition, previous studies showed that inactivation of NF2 and SMARCB1 provoking central nervous system tumors[39].

To identify functional SCNAS, i.e., those that may affect gene expression level, we evaluated the relationship between the gene expression alterations identified in stage one PTC in our previous study [40] with the gains and losses identified in our current study. Results highlighted elevated expression of ECM1 and ESM1 and decreased expression of DNAJB1, PLA2R1, FBLN1, and NR4A3. Except for FBLN1, which we found as recurrently deleted in 21 stage one patients with the broad deletion on chr22, other functional gains or losses were observed in no more than 4 patients. Previous studies have indicated elevated expression of 2 functional genes (ECM1 and ESM1) in
different cancers. The overexpression of extra-cellular matrix 1 (ECM1) results in the migration, invasion and adhesion of tumor cells [41, 42]. Endothelial cell-specific molecule-1 (ESM-1, aka Endocan) also plays a role in tumor growth and angiogenesis through the Akt-dependent activation of NF-κB pathway[43, 44]. Among functional losses, DNAJB1, a member of DNAJ protein family, has anti-apoptotic activity [45] and thus, we expected increase in its activity during stage one PTC. It can be inferred from the presence of DNAJB1 among functional deleted genes in stage one PTC that these kinds of compensatory molecular events, that are not present in aggressive tumors, is responsible for the indolent nature of PTC. PLA2R1 is a positive regulator of DNA damage response, and several investigations have confirmed its tumor suppressor activities in several cancers, including thyroid cancer [46]. Moreover, the epigenetic regulation of PLA2R1 through hyper methylation of its promoter and also by micro RNAs inhibition, support its tumor suppressor activity [47, 48]. Fibulin1 (FBLN1), is a multi-functional extracellular tumor suppressor; with the epigenetic down-regulation in various cancers [49-51]. NR4A3 is a tumor suppressor and direct transcriptional target of P53, that is involved in modulating apoptosis, tumorigenesis and cell cycle [52].

Evaluation of stage one PTC patients by follow-up data could help to evaluate and prioritize the identified mutations and SCNAs according to patients’ outcome (Distant metastasis, locoregional recurrence, or new primary tumor). However, follow-up data was available for very few stage one PTC patients, thus we could not generalize genomic alterations of these patients to entire stage one PTC patients.

4. Conclusion

Figure 6 shows key mutational and copy number alteration features with the potential contribution to the tumor development in PTC initiation.

Since progression to higher stages (stage three and stage four) are expected for approximately 18% of stage one PTC [53], the prevalence of pathogenic mutations in a large number of stage one PTC patients excludes their exclusive contribution to poor clinical outcome; thus, their oncogenic effects have to be considered alongside other genomic alterations such as accompanying low frequency mutations. Moreover, it has previously been indicated that, copy
number alteration of several oncogenes correlates with poorer outcome, while their mutational changes does not correlate with survival [9]. Thus, we believe that the traditional perspective of considering the mutational alterations as the major and sole contributor to tumor initiation and progression should be revised.

Broad deletion on chr22 and the recurrent amplification of 15 proto-oncogenes, were the only highly recurrent events in stage one PTC.

We identified that in stage one PTC, oncogenic activation occurs through both pathogenic mutations and gene amplification, while tumor suppressor inhibition is exclusively mediated by SCNA. In addition, a large number of gains or losses in oncogenes and tumor suppressors, respectively, highlights the considerable contribution of SCNA, compared with mutational events, in early-stage PTC. The tumorigenic role of the introduced driver mutations (based on background mutability of genes) as well as low frequency mutations accompanying BRAF(V600E), should be evaluated in distinct populations of stage one PTC.

Table 1: Number of targeted genes and patients in each mutation type in stage one PTC.

<table>
<thead>
<tr>
<th>Mutation Type</th>
<th>No of targeted genes</th>
<th>No of patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Missense</td>
<td>3291</td>
<td>174 (63%)</td>
</tr>
<tr>
<td>Small Insertion/Deletions</td>
<td>127</td>
<td>63 (23%)</td>
</tr>
<tr>
<td>Silent</td>
<td>4410</td>
<td>171 (62%)</td>
</tr>
<tr>
<td>Splice Site</td>
<td>135</td>
<td>67 (24%)</td>
</tr>
<tr>
<td>Translation Start Site</td>
<td>14</td>
<td>12 (4%)</td>
</tr>
</tbody>
</table>
Table 2: Number of stage one PTC patients with mutations in tumor suppressors (TSs) and oncogenes in different cancer signaling pathways.

<table>
<thead>
<tr>
<th>Signaling pathways</th>
<th># Patients with mutations in TSs</th>
<th># Patients with mutations in Oncogenes</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTK-RAS</td>
<td>3</td>
<td>176</td>
</tr>
<tr>
<td>Hippo</td>
<td>13</td>
<td>3</td>
</tr>
<tr>
<td>Wnt</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>NOTCH</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>TP53</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>PI3K</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>MYC</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>TGF-beta</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>NRF2</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 3: Potential driver mutations based on background mutability identified by MutaGene webserver.

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Type</th>
<th>Gene Symbol</th>
<th>Signaling pathways</th>
</tr>
</thead>
<tbody>
<tr>
<td>p.Lys740Thr</td>
<td>Missense</td>
<td>CTBP2</td>
<td>NOTCH</td>
</tr>
<tr>
<td>p.Asn978Ser</td>
<td>Missense</td>
<td>CTBP2</td>
<td>NOTCH</td>
</tr>
<tr>
<td>p.Gly732Glu</td>
<td>Missense</td>
<td>CTBP2</td>
<td>NOTCH</td>
</tr>
<tr>
<td>p.Gly732Arg</td>
<td>Missense</td>
<td>CTBP2</td>
<td>NOTCH</td>
</tr>
<tr>
<td>p.Ile625Phe</td>
<td>Missense</td>
<td>CTBP2</td>
<td>NOTCH</td>
</tr>
<tr>
<td>p.Lys1162Gln</td>
<td>Missense</td>
<td>KDM5A</td>
<td>NOTCH</td>
</tr>
<tr>
<td>p.Gln590Pro</td>
<td>Missense</td>
<td>MAML2</td>
<td>NOTCH</td>
</tr>
<tr>
<td>p.Phe5028Ser</td>
<td>Missense</td>
<td>HMCN1</td>
<td>HIPPO</td>
</tr>
<tr>
<td>p.Asn709Asp</td>
<td>Missense</td>
<td>DEPDC5</td>
<td>PI3K-AKT</td>
</tr>
<tr>
<td>p.Leu703His</td>
<td>Missense</td>
<td>KSR2</td>
<td>RTK-RAS</td>
</tr>
<tr>
<td>p.Pro1070Ala</td>
<td>Missense</td>
<td>PLXNB1</td>
<td>RTK-RAS</td>
</tr>
<tr>
<td>p.Leu1660Val</td>
<td>Missense</td>
<td>PLXNB1</td>
<td>RTK-RAS</td>
</tr>
<tr>
<td>p.Val1492Gly</td>
<td>Missense</td>
<td>CHD4</td>
<td>WNT</td>
</tr>
</tbody>
</table>
Funding: No grants received for this study.

Conflicts of interest/Competing interests: There is no conflict of interest.

Ethics approval: All TCGA data had already been collected from patients considering TCGA Ethics & Policies.

Authors contribution: NH performed the analyses and wrote the manuscript, MH, KB and CM reviewed the manuscript; MK supervised the analyses.

References:


Table legends:

**Table 1:** Number of targeted genes and patients in each mutation type in stage one PTC.

**Table 2:** Number of stage one PTC patients with mutations in tumor suppressors (TSs) and oncogenes in different cancer signaling pathways.

**Table 3:** Potential driver mutations based on background mutability identified by MutaGene webserver.

**Table S1:** The identified pathogenic mutations confirmed by SIFT and VEP tools.

**Table S2:** Proto-oncogenes targeted by arm-level and broad gains.

**Table S3:** Tumor suppressors targeted by arm-level, broad and focal loss.

Figure legends:

**Figure 1:** a. Number of different mutational events (left) and targeted genes (right) in stage one PTC. b. The frequency of mutations across 237 stage one PTC patients: ranging from 2 to 449. B-Top 20 most frequent mutated genes with the number (right) and percent (left) of targeted patients.

**Figure 2:** Thirty low frequency mutations accompanying BRAF(V600E) mutation in 24 patients.

**Figure 3:** a. Distribution of gains and losses across chromosomes in stage one PTC. b. Arm-level vs. broad and focal gains/losses in proto-oncogenes and tumor suppressors.

**Figure 4:** a. Amplified oncogenes and b. tumor suppressors, associated molecular pathways.

**Figure 5:** Signaling pathways targeted by the recurrent broad deletion of chr22 and corresponding genes.

**Figure 6:** Key genomic alteration features in stage one PTC.
Fig. 1.a
Fig. 1.b

Fig. 2

- FLG, rs2184953 (Missense), moderate impact
- FLG, rs3126079 (Missense), moderate impact
- KRTAP10-10, rs66931310 (In Frame Deletion), moderate impact
- RP11-88611.4, novel (RNA), unknown impact
- MUC5B, rs4963056 (Silent), low impact
- F5, rs4524 (Missense), moderate impact
- F5, rs4525 (Missense), moderate impact
Fig. 3.a

Fig. 3.b
Fig. 4.a

Fig. 4.b
Fig. 5

Mutational events
Pathogenic mutations in oncogenes: AKT1, BRAF, HRAS, KRAS, NRAS, PIK3CA oncogenes

Novel driver mutations in HIPPO, NOTCH, PI3K-AKT and RTK-RAS signaling pathways

Low frequency mutations accompanying recurrent BRAF (V600E) mutation in: “FLG, KRTAP10-10, RP11-886II1.4, MUC5B, and F5”

Somatic copy number alteration events
Amplification of 60 proto-oncogenes (15 recurrent)

Non-recurrent deletions of 58 tumor suppressors

Highly recurrent 36.6 Mb broad deletion on Chr22

Highly recurrent functional loss of tumor suppressor FBLN1

Fig. 6