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

Nazanin Hosseinkhan[1], Maryam Honardoost[1], Kevin Blighe[2], Tara Moore [2], Mohammad E. Khamseh[1*]

[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 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.

Mutation Type	No of targeted genes	No of patients
Missense	3291	174 (63%)
Small Insertion/Deletions	127	63 (23%)
Silent	4410	171 (62%)
Splice Site	135	67 (24%)
Translation Start Site	14	12 (4%)

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

Signaling pathways	*#Patients with mutations in TSs*	*# Patients with mutations in Oncogenes*
RTK-RAS	3	176
Hippo	13	3
Wnt	6	11
NOTCH	9	3
TP53	8	0
PI3K	0	4
MYC	2	0
TGF-beta	1	0
NRF2	1	0

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

Mutation	*Type*	*Gene Symbol*	*Signaling pathways*
p.Lys740Thr	Missense	CTBP2	NOTCH
p.Asn978Ser	Missense	CTBP2	NOTCH
p.Gly732Glu	Missense	CTBP2	NOTCH
p.Gly732Arg	Missense	CTBP2	NOTCH
p.Ile625Phe	Missense	CTBP2	NOTCH
p.Lys1162Gln	Missense	KDM5A	NOTCH
p.Gln590Pro	Missense	MAML2	NOTCH
p.Phe5028Ser	Missense	HMCN1	HIPPO
p.Asn709Asp	Missense	DEPDC5	PI3K-AKT
p.Leu703His	Missense	KSR2	RTK-RAS
p.Pro1070Ala	Missense	PLXNB1	RTK-RAS
p.Leu1660Val	Missense	PLXNB1	RTK-RAS
p.Val1492Gly	Missense	CHD4	WNT

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:

1. Loomans-Kropp, H.A. and A. Umar, *Cancer prevention and screening: the next step in the era of precision medicine.* NPJ precision oncology, 2019. **3**(1): p. 1-8.
2. Schulte, M., et al., *Cancer evolution, mutations, and clonal selection in relapse neuroblastoma.* Cell and tissue research, 2018. **372**(2): p. 263-268.
3. Shukla, A., et al., *Chromosome arm aneuploidies shape tumour evolution and drug response.* Nature communications, 2020. **11**(1): p. 1-14.
4. Merid, S.K., D. Goranskaya, and A. Alexeyenko, *Distinguishing between driver and passenger mutations in individual cancer genomes by network enrichment analysis.* BMC bioinformatics, 2014. **15**(1): p. 308.
5. Stobbe, M.D., et al., *Recurrent somatic mutations reveal new insights into consequences of mutagenic processes in cancer.* PLoS computational biology, 2019. **15**(11).
6. Brown, A.-L., et al., *Finding driver mutations in cancer: Elucidating the role of background mutational processes.* PLoS computational biology, 2019. **15**(4): p. e1006981.
7. Lusito, E., et al., *Unraveling the role of low-frequency mutated genes in breast cancer.* Bioinformatics, 2019. **35**(1): p. 36-46.
8. Ciriello, G., et al., *Emerging landscape of oncogenic signatures across human cancers.* Nature genetics, 2013. **45**(10): p. 1127-1133.
9. Smith, J.C. and J.M. Sheltzer, *Systematic identification of mutations and copy number alterations associated with cancer patient prognosis.* Elife, 2018. **7**: p. e39217.
10. Jamal-Hanjani, M., et al., *Translational implications of tumor heterogeneity.* Clinical cancer research, 2015. **21**(6): p. 1258-1266.
11. Davis, A., R. Gao, and N. Navin, *Tumor evolution: Linear, branching, neutral or punctuated?* Biochimica et Biophysica Acta (BBA)-Reviews on Cancer, 2017. **1867**(2): p. 151-161.
12. Vitale, M., *Rethinking the role of oncogenes in papillary thyroid cancer initiation.* Frontiers in endocrinology, 2012. **3**: p. 83.
13. Prescott, J.D. and M.A. Zeiger, *The RET oncogene in papillary thyroid carcinoma.* Cancer, 2015. **121**(13): p. 2137-2146.
14. Pierotti, M., et al., *RET/ptc and TRK oncogenes in papillary thyroid carcinoma*, in *Tyrosine Phosphorylation/Dephosphorylation and Downstream Signalling*. 1993, Springer. p. 87-98.
15. Sciavolino, P.J., *NTRK Fusions in Papillary Thyroid Cancer: Expanding Targetable Treatment Options. Target.* Oncol, 2016.

16. Cabanillas, M.E., et al., *Acquired Secondary RAS Mutation in BRAFV600E-Mutated Thyroid Cancer Patients Treated with BRAF Inhibitors.* Thyroid, 2020.
17. Melo, M., et al., *TERT, BRAF, and NRAS in primary thyroid cancer and metastatic disease.* The Journal of Clinical Endocrinology & Metabolism, 2017. **102**(6): p. 1898-1907.
18. Zhu, Z., et al., *Prevalence of RET/PTC rearrangements in thyroid papillary carcinomas: effects of the detection methods and genetic heterogeneity.* The Journal of Clinical Endocrinology & Metabolism, 2006. **91**(9): p. 3603-3610.
19. Colaprico, A., et al., *TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data.* Nucleic acids research, 2016. **44**(8): p. e71-e71.
20. Sim, N.-L., et al., *SIFT web server: predicting effects of amino acid substitutions on proteins.* Nucleic acids research, 2012. **40**(W1): p. W452-W457.
21. McLaren, W., et al., *The ensembl variant effect predictor.* Genome biology, 2016. **17**(1): p. 1-14.
22. Goncearenco, A., et al., *Exploring background mutational processes to decipher cancer genetic heterogeneity.* Nucleic acids research, 2017. **45**(W1): p. W514-W522.
23. Morganella, S., S.M. Pagnotta, and M. Ceccarelli, *Gaia: Genomic analysis of important aberrations.* Reference Source.
24. Boursi, B., et al., *Thyroid dysfunction, thyroid hormone replacement and colorectal cancer risk.* JNCI: Journal of the National Cancer Institute, 2015. **107**(6).
25. Kuijpens, J.L., et al., *Hypothyroidism might be related to breast cancer in post-menopausal women.* Thyroid, 2005. **15**(11): p. 1253-1259.
26. Mondul, A.M., et al., *Circulating thyroxine, thyroid-stimulating hormone, and hypothyroid status and the risk of prostate cancer.* PLoS One, 2012. **7**(10): p. e47730.
27. Fighera, T.M., et al., *TSH levels are associated with increased risk of thyroid carcinoma in patients with nodular disease.* Endokrynologia Polska, 2015. **66**(6): p. 480-485.
28. Prete, A., et al., *Update on fundamental mechanisms of thyroid cancer.* Frontiers in endocrinology, 2020. **11**: p. 102.
29. Cisowski, J. and M.O. Bergo, *What makes oncogenes mutually exclusive?* Small GTPases, 2017. **8**(3): p. 187-192.
30. Tang, J., et al., *The clinicopathological significance of Ki67 in papillary thyroid carcinoma: a suitable indicator?* World journal of surgical oncology, 2018. **16**(1): p. 100.
31. Xiong, D.-d., et al., *Ki-67/MKI67 as a Predictive Biomarker for Clinical Outcome in Gastric Cancer Patients: an Updated Meta-analysis and Systematic Review involving 53 Studies and 7078 Patients.* Journal of Cancer, 2019. **10**(22): p. 5339.
32. Yang, C.-k., et al., *Genome-wide association study of MKI67 expression and its clinical implications in HBV-related hepatocellular carcinoma in Southern China.* Cellular Physiology and Biochemistry, 2017. **42**(4): p. 1342-1357.
33. Chai, L., J. Li, and Z. Lv, *An integrated analysis of cancer genes in thyroid cancer.* Oncology reports, 2016. **35**(2): p. 962-970.
34. Erinjeri, N.J., et al., *Whole-exome sequencing identifies two discrete druggable signaling pathways in follicular thyroid cancer.* Journal of the American College of Surgeons, 2018. **226**(6): p. 950-959. e5.
35. Cheng, Q., et al., *A novel integrative risk index of papillary thyroid cancer progression combining genomic alterations and clinical factors.* Oncotarget, 2017. **8**(10): p. 16690.
36. Irvine, A.D., W.I. McLean, and D.Y. Leung, *Filaggrin mutations associated with skin and allergic diseases.* New England Journal of Medicine, 2011. **365**(14): p. 1315-1327.
37. Siołek, M., et al., *CHEK 2 mutations and the risk of papillary thyroid cancer.* International Journal of Cancer, 2015. **137**(3): p. 548-552.

38. Borun, P., et al., *Rapid detection method for the four most common CHEK2 mutations based on melting profile analysis.* Molecular diagnosis & therapy, 2015. **19**(6): p. 419-425.
39. Vitte, J., et al., *Timing of Smarcb1 and Nf2 inactivation determines schwannoma versus rhabdoid tumor development.* Nature communications, 2017. **8**(1): p. 1-13.
40. Hosseinkhan, N., et al., *Comprehensive transcriptomic analysis of papillary thyroid cancer: potential biomarkers associated with tumor progression.* Journal of Endocrinological Investigation, 2020: p. 1-13.
41. Wang, L., et al., *Extracellular matrix protein 1 (ECM1) is over-expressed in malignant epithelial tumors.* Cancer letters, 2003. **200**(1): p. 57-67.
42. Gomez-Contreras, P., et al., *Extracellular matrix 1 (ECM1) regulates the actin cytoskeletal architecture of aggressive breast cancer cells in part via S100A4 and Rho-family GTPases.* Clinical & experimental metastasis, 2017. **34**(1): p. 37-49.
43. Calderaro, J., et al., *ESM1 as a marker of macrotrabecular-massive hepatocellular carcinoma.* Clinical Cancer Research, 2019. **25**(19): p. 5859-5865.
44. Kang, Y.H., et al., *ESM-1 regulates cell growth and metastatic process through activation of NF-κB in colorectal cancer.* Cellular signalling, 2012. **24**(10): p. 1940-1949.
45. Qi, M., et al., *DNAJB1 stabilizes MDM2 and contributes to cancer cell proliferation in a p53-dependent manner.* Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms, 2014. **1839**(1): p. 62-69.
46. Vindrieux, D., et al., *PLA2R1 mediates tumor suppression by activating JAK2.* Cancer research, 2013. **73**(20): p. 6334-6345.
47. Li, M., et al., *Uncovering the potential miRNAs and mRNAs in follicular variant of papillary thyroid carcinoma in the Cancer Genome Atlas database.* TRANSLATIONAL CANCER RESEARCH, 2019. **8**(4): p. 1158-1169.
48. Bernard, D. and D. Vindrieux, *PLA2R1: expression and function in cancer.* Biochimica et Biophysica Acta (BBA)-Reviews on Cancer, 2014. **1846**(1): p. 40-44.
49. Xiao, W., et al., *Fibulin-1 is epigenetically down-regulated and related with bladder cancer recurrence.* BMC cancer, 2014. **14**(1): p. 677.
50. Cheng, Y., et al., *Fibulin 1 is downregulated through promoter hypermethylation in gastric cancer.* British journal of cancer, 2008. **99**(12): p. 2083-2087.
51. Watany, M.M., et al., *Serum FBLN1 and STK31 as biomarkers of colorectal cancer and their ability to noninvasively differentiate colorectal cancer from benign polyps.* Clinica Chimica Acta, 2018. **483**: p. 151-155.
52. Wenzl, K., et al., *The nuclear orphan receptor NR4A1 and NR4A3 as tumor suppressors in hematologic neoplasms.* Current drug targets, 2015. **16**(1): p. 38-46.
53. Schlumberger, M.-J., et al., *Nontoxic diffuse goiter, nodular thyroid disorders, and thyroid malignancies*, in *Williams Textbook of Endocrinology*. 2016, Elsevier. p. 449-488.

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.

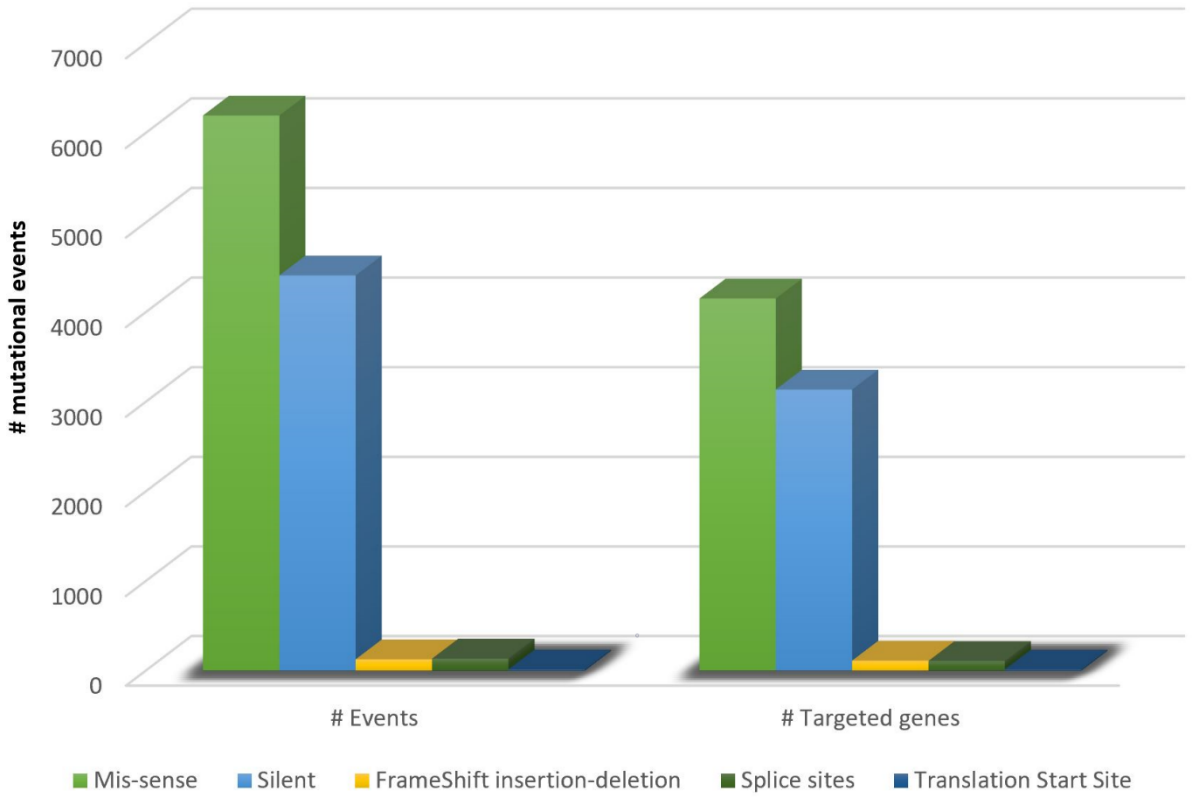
Figure 2: Thirty low frequency mutations accompanying BRAF(V600E) mutation in 24 patients.

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.

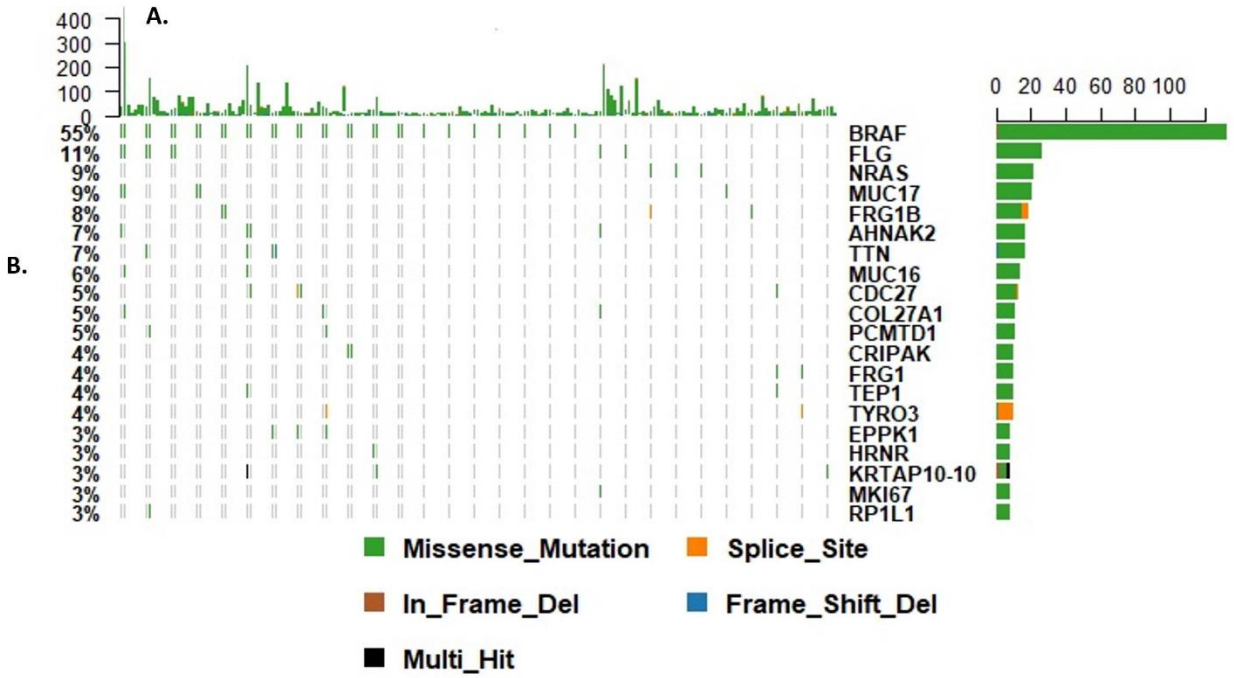


448

449 **Fig.1.a**

450

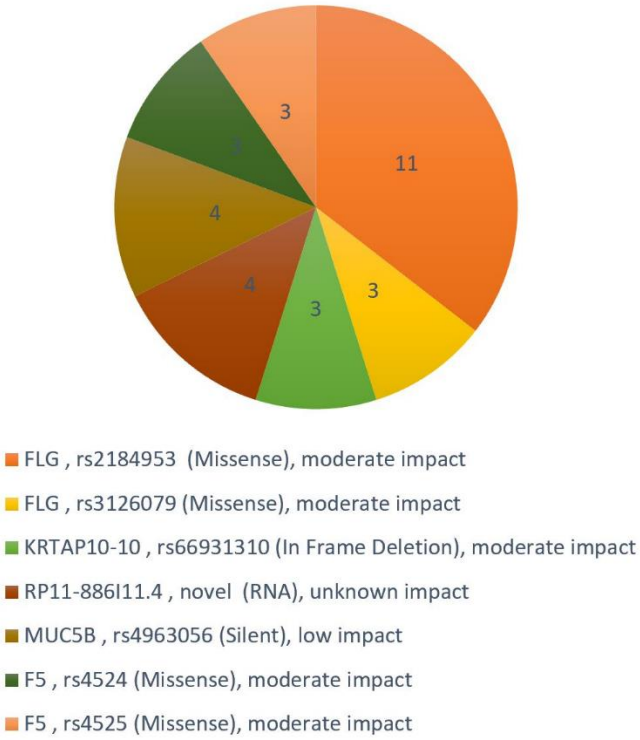
451



452

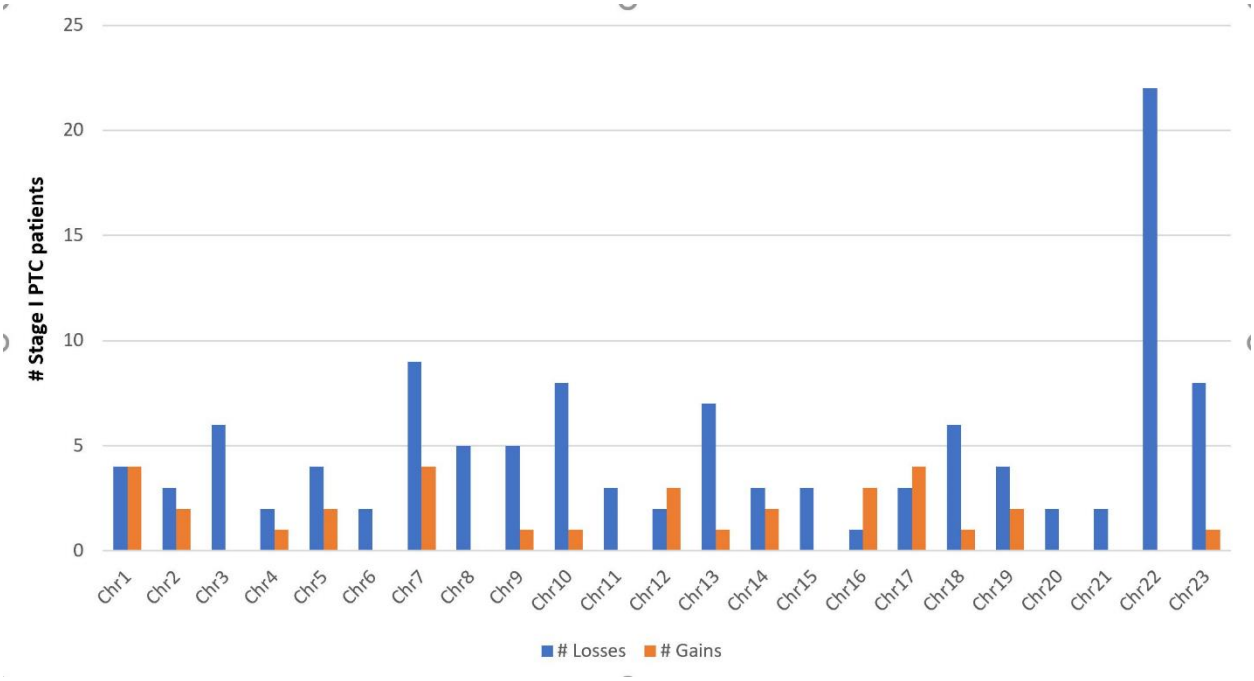
453 **Fig.1.b**

454



455

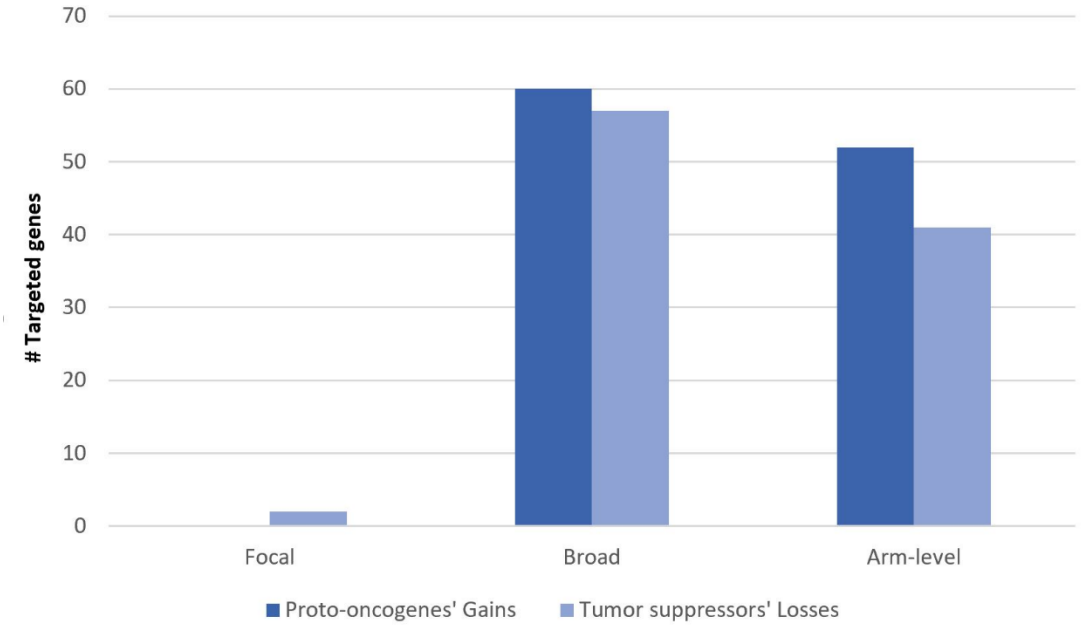
456 **Fig.2**



457

458 **Fig.3.a**

459



460

461 **Fig.3.b**

462

463

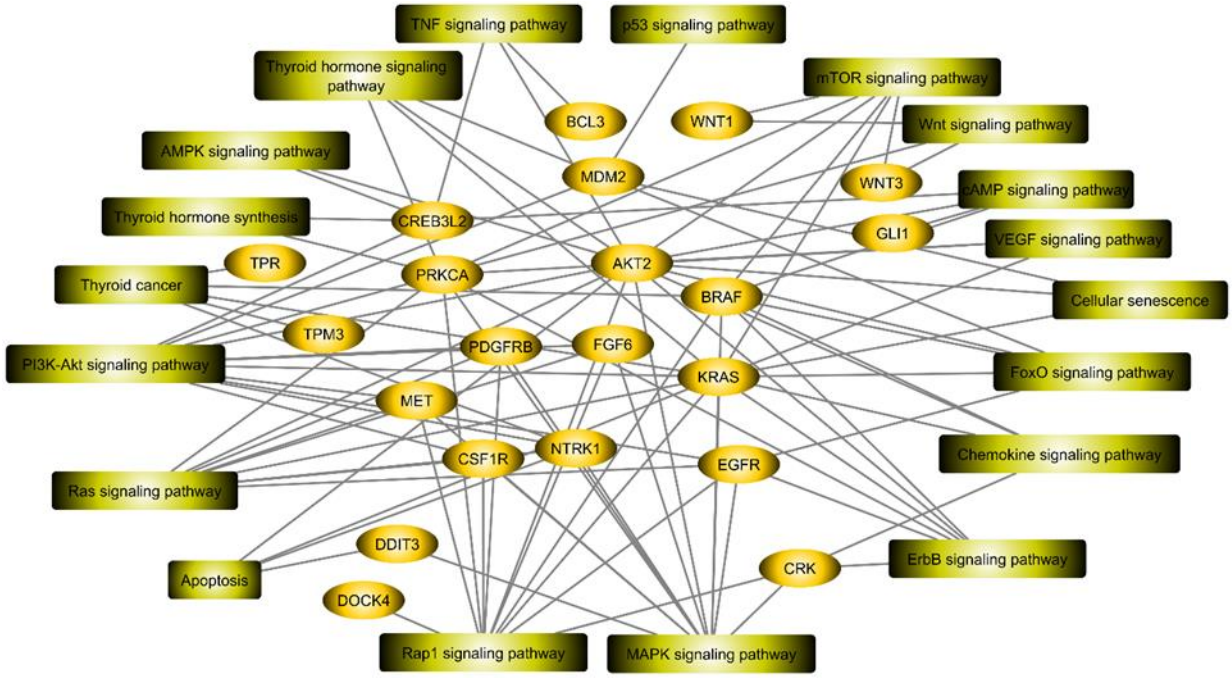

TNF signaling pathway
p53 signaling pathway
Thyroid hormone signaling pathway
mTOR signaling pathway
BCL3
WNT1
Wnt signaling pathway
AMPK signaling pathway
MDM2
WNT3
cAMP signaling pathway
Thyroid hormone synthesis
CREB3L2
GLI1
VEGF signaling pathway
TPR
AKT2
Thyroid cancer
PRKCA
BRAF
Cellular senescence
TPM3
PDGFRB
FGF6
PI3K-Akt signaling pathway
KRAS
FoxO signaling pathway
MET
CSF1R
NTRK1
Chemokine signaling pathway
EGFR
Ras signaling pathway
DDIT3
ErbB signaling pathway
CRK
Apoptosis
DOCK4
Rap1 signaling pathway
MAPK signaling pathway

Fig.4.a

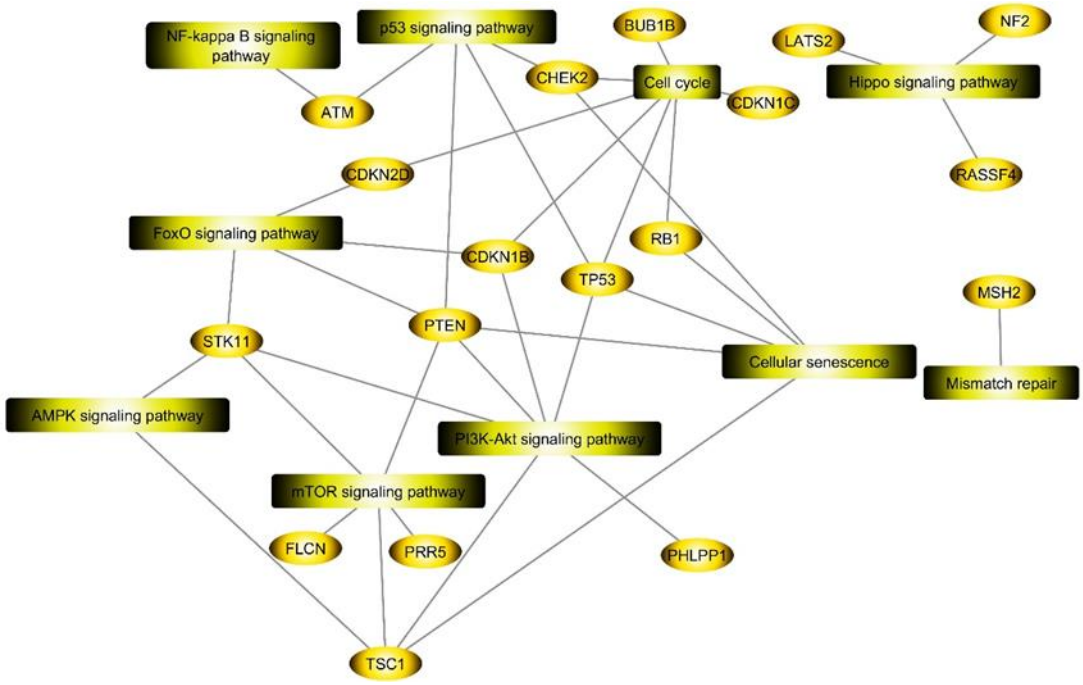

NF-kappa B signaling pathway
p53 signaling pathway
BUB1B
NF2
LATS2
CHEK2
Cell cycle
Hippo signaling pathway
ATM
CDKN1C
CDKN2D
RASSF4
FoxO signaling pathway
RB1
CDKN1B
TP53
MSH2
PTEN
STK11
Cellular senescence
Mismatch repair
AMPK signaling pathway
PI3K-Akt signaling pathway
mTOR signaling pathway
FLCN
PRR5
PHLPP1
TSC1

Fig.4.b

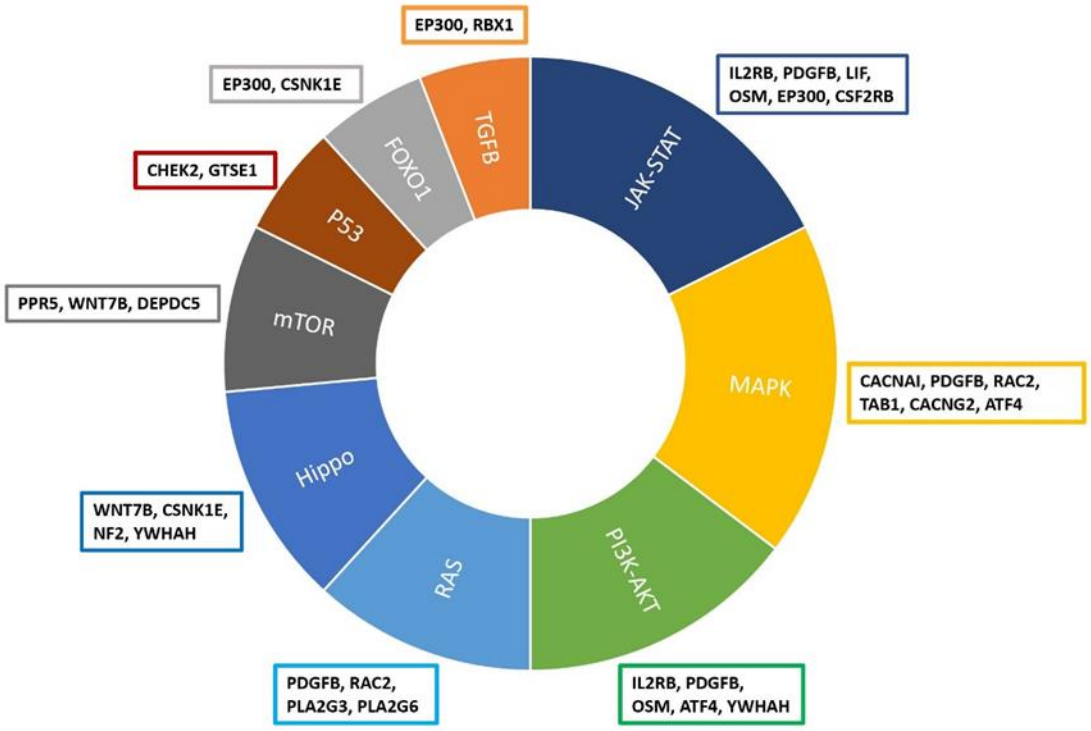

EP300, RBX1
IL2RB, PDGFB, LIF, OSM, EP300, CSF2RB
EP300, CSNK1E
TGFB
FOXO1
JAK-STAT
CHEK2, GTSE1
P53
PPR5, WNT7B, DEPDC5
mTOR
MAPK
CACNAI, PDGFB, RAC2, TAB1, CACNG2, ATF4
Hippo
WNT7B, CSNK1E, NF2, YWHAH
RAS
PI3K-AKT
PDGFB, RAC2, PLA2G3, PLA2G6
IL2RB, PDGFB, OSM, ATF4, YWHAH

Fig.5

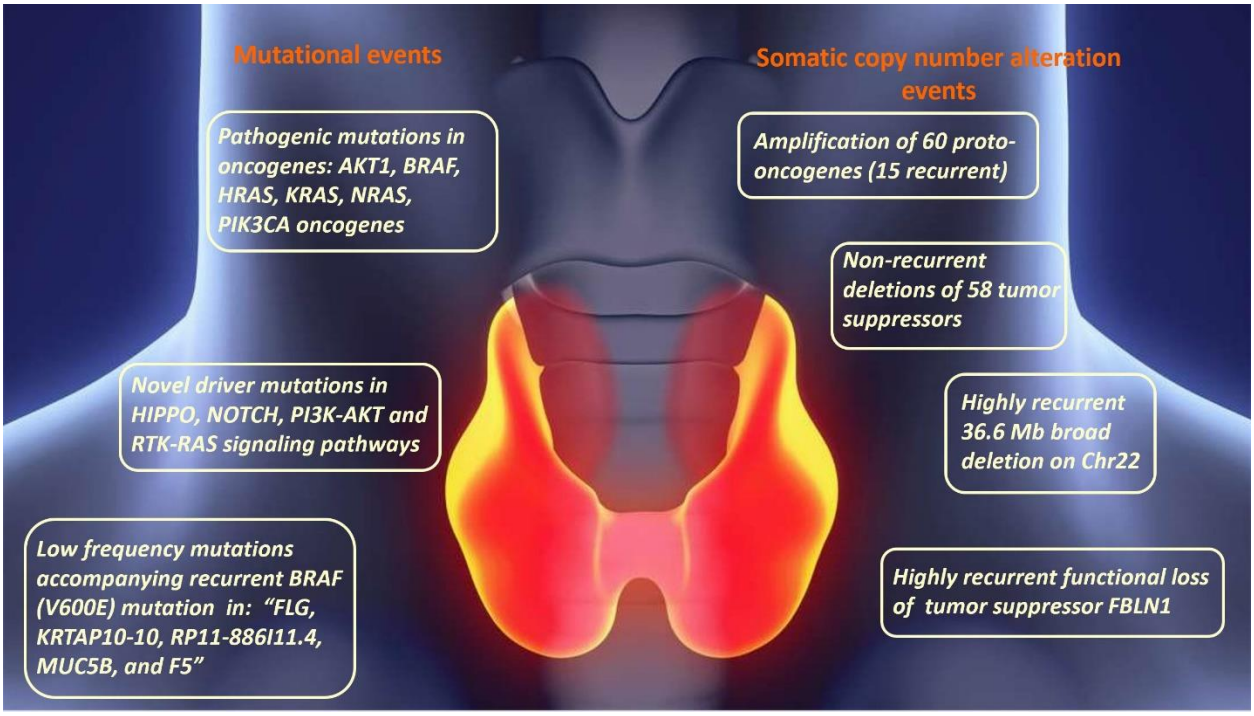

Mutational events
Somatic copy number alteration events
Pathogenic mutations in oncogenes: AKT1, BRAF, HRAS, KRAS, NRAS, PIK3CA oncogenes
Amplification of 60 proto-oncogenes (15 recurrent)
Non-recurrent deletions of 58 tumor suppressors
Novel driver mutations in HIPPO, NOTCH, PI3K-AKT and RTK-RAS signaling pathways
Highly recurrent 36.6 Mb broad deletion on Chr22
Low frequency mutations accompanying recurrent BRAF (V600E) mutation in: "FLG, KRTAP10-10, RP11-886I11.4, MUC5B, and F5"
Highly recurrent functional loss of tumor suppressor FBLN1

Fig.6

475

476

477