Title

The genomic and evolutionary landscapes of anaplastic thyroid carcinoma

Journal

Cell Reports, 43(3)

ISSN

2639-1856

Authors

Zeng, Peter YF
Prokopec, Stephenie D
Lai, Stephen Y
et al.

Publication Date

2024-03-01

DOI

10.1016/j.celrep.2024.113826

Peer reviewed

HHS Public Access

Author manuscript

Cell Rep. Author manuscript; available in PMC 2024 May 08.

Published in final edited form as:

Cell Rep. 2024 March 26; 43(3): 113826. doi:10.1016/j.celrep.2024.113826.

The genomic and evolutionary landscapes of anaplastic thyroid carcinoma

A full list of authors and affiliations appears at the end of the article.

SUMMARY

Anaplastic thyroid carcinoma is arguably the most lethal human malignancy. It often co-occurs with differentiated thyroid cancers, yet the molecular origins of its aggressivity are unknown. We sequenced tumor DNA from 329 regions of thyroid cancer, including 213 from patients with primary anaplastic thyroid carcinomas. We also whole genome sequenced 9 patients using multi-region sequencing of both differentiated and anaplastic thyroid cancer components. Using these data, we demonstrate thatanaplastic thyroid carcinomas have a higher burden of mutations than other thyroid cancers, with distinct mutational signatures and molecular subtypes. Further, different cancer driver genes are mutated in anaplastic and differentiated thyroid carcinomas, even those arising in a single patient. Finally, we unambiguously demonstrate that anaplastic thyroid carcinomas share a genomic origin with co-occurring differentiated carcinomas and emerge from a common malignant field through acquisition of characteristic clonal driver mutations.

In brief

Anaplastic thyroid cancer is one of the most lethal human cancers. Surprisingly, it often evolves alongside a highly non-lethal form of differentiated thyroid cancer. Zeng et al. demonstrate how these two diseases evolve from a common ancestor, leading to differential evolutionary trajectories and dramatic differences in clinical outcomes.

Graphical abstract

*Correspondence: pboutros@mednet.ucla.edu (P.C.B.), anthony.nichols@lhsc.on.ca (A.C.N.).

AUTHOR CONTRIBUTIONS

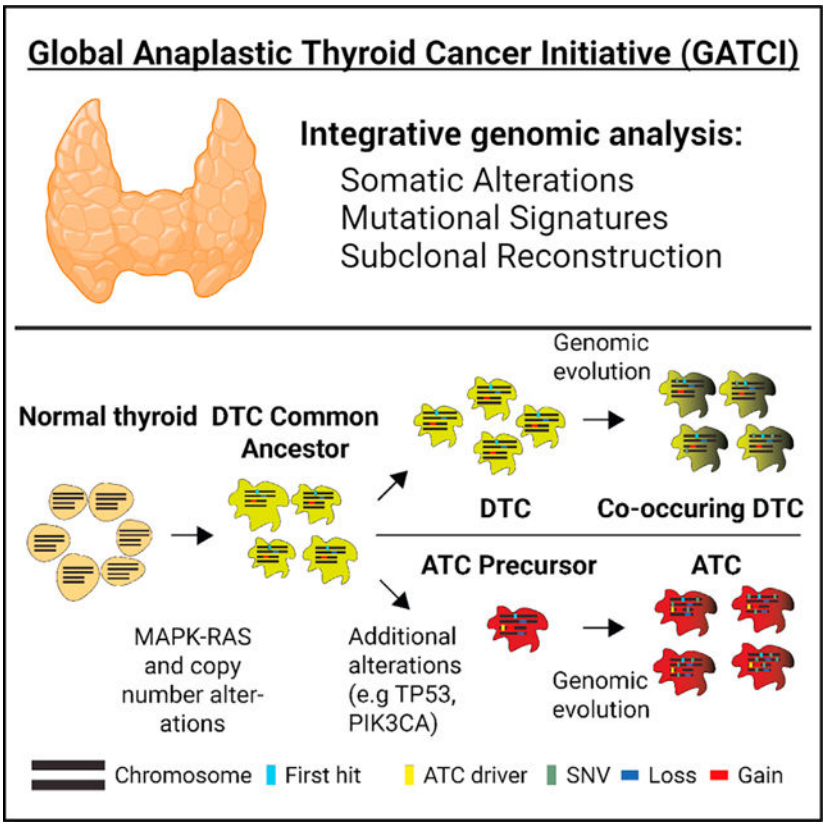
P.Y.F.Z., S.D.P., S.Y.L., N.P., R.X.S., M.A.C.-S.-Y., M.T., T.N.Y., J.L., J.W.B., P.C.B., and A.C.N. analyzed the data. A.C.N., P.C.B., J.W.B., S.D.P., and S.Y.L. conceptualized the study. S.Y.L., N.P., R.C.-B., M.D.W., C.J.H., P.P., M.J.C., A.K.L., I.S., J.W., J.D.W., R.K., T.C., N.A., N.C., D.B., B.N., L.M.R., J.A.B., C.G., K.B., N.G.N., P.W., Y.C.H., C.M.L., A.S., K.P., F.V., A.D., L.X., Y.E.N., R.S., J.A.C., L.A.M., M.D.H., L.D., S. Sidhu, M. Sywak, B.R., K.F., F.G., K.K., S.D.M., A.M., D.A.P., M.I.K., M. Shaikh, K.M.R., B.W., E.W., J.Y., J.S.M., J.W.R., D.W., S. Scherer, T.J.G., W.C.F., A.J.G., and G.C. provided study resources for the study. All authors contributed to data interpretation, writing, and editing of the manuscript. A.C.N., P.C.B., N.A., B.N., D.B., J.A.C., and RCS provided funding for the study. A.C.N. and P.C.B. supervised the study.

DECLARATION OF INTERESTS

All authors declare that they have no competing interests.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2024.113826.



INTRODUCTION

Human malignancies range widely in their lethality. Some tumor types, like differentiated thyroid carcinoma (DTC) and prostatic adenocarcinomas, are typically indolent with long life histories and are managed with minimally invasive therapies or even surveillance protocols.[1,2] Others, like pancreatic adenocarcinomas, progress rapidly after initial diagnosis and are refractory to most therapeutic interventions.[3] This variability between tumor types is mirrored by genetic changes, leading to the development of prognostic tests that predict tumor aggressivity from transcriptomic or proteomic features.[4–6]

At one extreme of this spectrum of tumor lethality lies anaplastic thyroid carcinoma (ATC). ATC is arguably the most lethal human tumor type, with a median survival of ~12 weeks: some patients succumb to their disease within days of diagnosis.[7,8] ATC usually presents dramatically, with rapid onset of airway and esophageal obstruction due to explosively growing neck masses.[7,8] Primary ATCs are frequently inoperable due to encasement of nerves, blood vessels, and the airway. Radio- and chemoresistance are common, and distant metastases are near ubiquitous.[7,9] Thus, ATC is a major clinical dilemma and a model for understanding lethal, treatment-resistant cancer.

The aggressiveness of ATC is particularly fascinating because of its life history: ~80% of ATCs occur in the context of a prior history of thyroid cancer or with a distinct co-occurring region of DTC (co-DTC), most frequently papillary thyroid cancer (PTC).[10–14]

Thus, intriguingly, the most lethal human malignancy frequently arises in the context of one of the most indolent. ATC presents with undifferentiated pathological features, suggesting that ATC may evolve via dedifferentiation from well-differentiated thyroid cancers. There has been little molecular confirmation of this proposed progression from a well to a poorly differentiated tumor. Several studies have sequenced ATCs with limited genome coverage, highlighting frequent *TP53* mutations and a subset of tumors with multiple oncogene driver mutations that potentially drive tumor progression and clinical aggressiveness.[15–23] However, the genomic landscape of ATC remains largely unknown, and the molecular characteristics of its evolutionary relationship to DTC remain elusive.

To fill these gaps in our understanding of ATC pathobiology, we established a 15-site consortium called the Global Anaplastic Thyroid Cancer Initiative (GATCI). We report the mutational landscape of 329 thyroid cancer regions, including 179 primary ATC regions, 1 metastatic ATC, and 34 co-occurring regions of DTC within ATCs. To identify genomic alterations enriched in ATC or its co-DTC region, we sequenced 115 papillary thyroid regions without co-occurring ATC. Together these data reveal distinct genomic subtypes of ATC, with elevated mutational density and signatures. We further show that ATCs harbor a distinct set of driver mutations from DTCs, including multiple recurrent tumor suppressor hits. Finally, we demonstrate that ATC arises in the context of DTCs that share a common evolutionary origin, but acquire additional specific driver mutations, some of which are recurrently altered exclusively in anaplastic carcinoma components.

RESULTS

Cohort summary

We collected 329 thyroid cancer regions from 292 patients, including 213 regions from ATC patients and 115 regions from PTC patients. A detailed breakdown of samples is available in Figure S1 and Tables S1A and S1B. Tumors were subjected to consensus pathology review and manually macrodissected to maximize tumor purity. A total of 259 tumor regions and 14 reference paired normal DNA samples were characterized for copy-number aberrations (CNAs) by SNP microarrays. Single-nucleotide variants (SNVs) were identified by whole-exome (WES) or whole-genome sequencing (WGS), along with RNA sequencing (RNA-seq) when open biopsies provided sufficient material. Deep targeted sequencing was used for validation and nested PCR to detect TERT promoter mutations (Table S1). ATCs were sequenced to a depth of 123× ± 14× and reference samples to 116× ± 13× (median ± standard deviation [SD]) by WES and 33× ± 4× or 34× ± 4×, respectively by WGS. Reads were aligned and somatic variant detection was performed using a validated pipeline.[6,24] Recognizing that ATCs can have low tumor purity, we validated 1,140 candidate mutations in 37 ATCs, including 1 metastasis and 3 DTCs, by deep targeted sequencing on an orthogonal-platform sequencing platform with a median F_1 of 0.914 (Figure S2). After quality control, this yielded 141 ATCs (132 via WES and 9 via WGS), 28 co-occurring DTCs (19 via WES, 9 via WGS), 114 PTCs (all WES), 2 metastases (1 ATC and 1 PTC, both WES), and 13 cell lines characterized with high-quality WES or WGS, and 110 ATCs, 22 co-occurring DTCs, 112 PTCs, 1 metastasis, and 13 cell lines characterized with copy-number arrays (Tables S1A and S1C). In contrast to the strong female preponderance

of PTC in The Cancer Genome Atlas (TCGA) (F:M 2.93:1),[25] our ATC cohort was sex balanced (F:M 1.11:1), consistent with previous reports.[17] Median age at diagnosis was ~70 years for ATC and ~50 years for PTC within our cohort. Most ATC patients presented with locally advanced disease: 87% with T4b disease and 70% with nodal involvement. Distant metastases were common (43%), and median survival was 126 days (18 weeks) from diagnosis. PTCs within the GATCI project were equally distributed by tumor extent (T1–T4), and all tumors were surgically resected, with median follow-up of 10.9 years.

ATCs exhibit moderate SNV mutation burden

While DTCs have fewer somatic mutations of all types than almost any other cancer type,[25] ATCs are thought to be highly mutated, particularly by CNAs.[17,26]

In this large genome-wide ATC cohort (n = 141, both WES and WGS), we identified 3.8 ± 1.2 SNVs/Mb of DNA sequenced (Figures 1A, top, and 1C). These were accompanied by 120 ± 44 CNAs (mean ±99% CI; median = 83.5, n = 119; Figures 1B–1D). The average CNA was 6.21 ± 1.33 Mb (mean ±99% CI; median = 4.95 Mb). Relative to 32 tumor types,[27] ATCs show both more CNAs (Figure 1D) and more SNVs/Mb (Figure 1C) than PTCs, but fewer than most other adult cancer types. Despite their clinical aggressiveness, they are not among the most highly mutated tumor types, nor did they show atypically high intertumor variability relative to other cancer types.

Forty-two genes were affected by recurrent SNVs that are putative drivers in ATC via SeqSig analysis[6] (false discovery rate [FDR] < 0.05, Table S2A). The top five genes with the smallest SeqSig FDR were *TP53*, *NRAS*, *BRAF*, *PIK3CA*, and *USH2A* (Figures 1A and S3A). Non-functional or partially functional[28] non-synonymous SNVs in *TP53* in primary ATC were associated with elevated mRNA abundance (Wilcoxon rank-sum test p = 0.0067, n = 19, Figure S3B; Table S3). The SeqSig recurrent genes also included several other cancer-related genes found in the COSMIC catalog,[29] including *BRCA1*, *BRCA2*, *TET1*, *MET*, *ATR*, *NOTCH1*, *FAT1*, *ATM*, and *SPEN*. Further, recurrent mutations in *USH2A* and *LRP1* have been described previously in ATC.[15] Pathway analysis using GProfiler indicated that the SeqSig genes were enriched for genes related to various cancer types, basement membrane, cell adhesion, DNA repair, and radiation response (Table S2B). Several mutations previously described in ATC were detected at low frequency: *EIF1AX* mutations were detected in five tumors, including three with the p.A113X splice-site mutation reported to be exclusive to thyroid cancers[17] and two missense mutations (G9R). These tended to co-occur with *RAS* mutations, consistent with previous reports (Figure S4A).[17]

We assessed germline variants in known cancer predisposition genes.[30] Nine genes exhibited heterozygous germline coding SNVs in at least two patients (2.4% of ATCs), with a maximum recurrence in four patients (5%) for *RECQL4* (Table S4). This included multiple DNA damage repair genes, for example, alterations in *BRCA2* (n = 3) and *FANCF* (n = 3). No patients with either RECQL4 or BRCA2 germline mutation had additional somatic mutations in RECQL4 and BRAC2 or copy-number loss in the tumor sample. Thus, much larger cohorts will be required to understand the germline predisposition to ATC. We therefore focused on somatic mutations.

Individual ATCs varied significantly in their underlying mutational processes, with activation COSMIC signatures 1, 5, 6, and 13 and a rare novel signature (Figure 2A). The last was characterized by G[T>G]G mutations and was detected in only three tumors (Figure 2B). It may embody unknown sequencing artifacts or an as-yet-undescribed mutational process. Other types of thyroid cancers have been reported to harbor activation of COSMIC signatures 1, 5, and 13,[31,32] and ATC mirrors these. COSMIC 5 describes a diverse range of point mutations consisting of low-frequency pyrimidine transition mutations, and COSMIC signature 13 was attributed to AID/APOBEC activity,[31] but surprisingly, none of these signatures were associated with sex, age, CNA subtype, *TERT* promoter status, or overall survival (ANOVA, FDR > 0.1). Thus, ATC harbored more somatic SNVs than other thyroid cancers, but fewer than other cancer types, with no widely recurrent driver mutations or mutational processes.

Copy-number heterogeneity in ATC

By contrast, ATC harbored a background of multiple recurrent chromosome-scale events, including loss of chromosome 13 (39% of tumors) and gain of chromosome 20q (35% of tumors), which have been reported previously.[17,33,26] Overexpression of a subset of genes on 20q has previously been shown to promote thyroid cancer progression.[34,35] We also identified a novel recurrent gain of chromosome 7 (26% of tumors). Consensus clustering revealed that these large-scale changes reflect five distinct CNA subtypes, denoted A–E (Figures 1B and 2C). Each subtype is defined by characteristic genomic abnormalities, such as wide-scale amplifications with occasional arm-level deletions for subtype A (11% of ATCs) and large-scale deletions with cases of uniparental polysomy, similar to oncocytic carcinoma of the thyroid (OCT; formerly known as Hürthle cell thyroid cancers),[36,37] for subtype B (29%). Subtype C (20%) was dominated by gains of chromosomes 7 and 20 (similar to OCT[36; 37]) and 1q gains (similar to PTC25), while subtype D (15.5%) demonstrated many focal regions of CNA. Subtype E (24.5%) showed a quiet copy-number profile. The lack of copy alterations in subtype E was not due to extensive stromal infiltration, as the variant allele frequency of subtype E tumor that has corresponding WES data and presence of a SeqSig driver gene (n = 10) did not differ from that of subtype C or D (FDR > 0.05), although it was lower than that of subtype A or B (Figure S4M). This was further supported by tumor purity estimated during pathological review. In the 11 ATC tumors of subtype E that had centralized pathology review, the pathologist assessed the tumor purity median to be 50 (minimum 20, maximum 90). Thus, analogous to similar profiles in breast, prostate, and head and neck cancers,[38–40] subtype E showed a quiescent CNA landscape (Figure 2C).

These subtypes were associated with overall genomic instability (Figure 1D), age, TP53 mutation status (Table S5), tumor purity, and ploidy (Figures S4B–S4H). These broad changes were accompanied by many highly recurrent focal driver CNAs (Figure 2E, q < 0.01). Loss of *CDKN2A* was widespread (42% of ATCs),[17,34] as was loss of *BRCA2* (33.6% of ATCs). Other prominent recurrent CNAs included loss of regions on chromosome 8 encoding *ADAM3A*, a 62 kb region on 22q containing *GSTT2* and *GSTTP1*, a region on chromosome 5 involving CDK7, a region on chromosome 17 involving *NF1*, and regions on chromosome 13 involving *FGF9*, *MYCBP2*, *ELF1*, and *FLT3.* A subset of ATC tumors

(18%) also exhibited a 429 kb gain on chromosome 3p containing *VGLL3*, which has been linked with ovarian cancer.[41] Recurrent CNAs in *MYCBP2*, *SPATA13*, and *OMG* were associated with mRNA changes (Table S3; Figures S4I–S4K).

Transcriptomic hallmarks of ATC

Next we assessed transcriptional abundance differences between the ATCs (n = 24) and the cell lines (n = 13) in combination with normal, primary PTC, and metastatic PTC RNA-seq data from TCGA. As expected, we observed a progressive decrease in the thyroid differentiation score (TDS; Figure 2G) from normal thyroid, PTC, and ATC, demonstrating progressive de-dedifferentiation. Further, cell lines possessed a lower TDS compared with primary ATC, in agreement with previous evidence[42] suggesting that thyroid cancer cell lines represent the extreme dedifferentiated forms of thyroid cancers. We also assessed the BRAF-RAS score in our RNA-seq dataset. Although primary PTCs within TCGA demonstrated clear separation between BRAF and RAS-altered tumors (Figure 2H), the RAS-mutant ATC tumors (n = 2) appear more BRAF-like, similar to previous studies.[17]

We further estimated the abundance of stromal and immune cell populations within the samples using MCPCounter, as we have done previously.[43] Compared with the normal thyroid, primary PTCs have a decreased abundance of macrophages/monocytes and T cells and an increased abundance of neutrophils (Figure 2I). On the other hand, the tumor microenvironment of ATC is distinct from that of primary and metastatic PTC, exhibiting dramatically elevated abundance of fibroblasts and macrophages/monocytes and decreased abundance of endothelial cells, myeloid dendritic cells, and neutrophils (Figure 2I).

We next sought to identify fusion genes using RNA-seq data, reasoning that these are common in other thyroid cancers and often attractive therapeutic targets. We identified 84 fusion genes in primary ATC (n = 24) using fusioncatcher,[44] covering 144 total partner genes, and none previously reported in COSMIC (Table S6A). The typical ATC harbored a median of 3 distinct fusion genes, with 25% of samples containing 6 or more fusions and two harboring a maximum of 11 fusions. Only 5 fusion genes occurred in two or more patients, including *LINC01133: SAMHD1*, *MDM4:TRA*, and *KDSR:ANAPC7*. Fusions involving *FN1* and *COL1A1* were associated with higher than average abundance of these genes. For *FN1*, this presented as 4.3× higher transcripts per million (TPM) in tumors with a *FN1* fusion (Wilcoxon p = 0.007), and for *COL1A1*, fusions corresponded to 12× higher TPM (one sample; Tables S6B and S6C). Increased abundance of both genes has previously been associated with tumor cell migration and invasion in multiple tumor types[45,46].

Integrative genomics analysis implicates frequent deletions of tumor suppressors in ATCs lacking an SNV driver mutation

To identify co-occurring mutation processes underlying ATC pathogenesis, we analyzed interassociations of mutation-density, mutation-signature, and driver-gene features (Figure 2F). Multiple associations were detected, including a strong correlation between SNV mutation density (SNVs/Mb sequenced) and COSMIC 1 and 5 signatures (FDR < 10^{-3}). Trinucleotide signatures were not correlated to CNA subtypes (p > 0.05; one-way ANOVA; Figure S4L), suggesting that the processes driving CNA diversification are independent of

those driving SNV patterns. *BRAF*^{V600E} and *RAS* mutations were mutually exclusive, while *BRAF*^{V600E} and *PIK3CA* mutations co-occurred[15,17] (FDR = 0.034; Figure S4A). Promoter mutations upstream of *TERT* were correlated with *TP53* mutations (FDR < 0.01).

We performed integrative analysis in the subset of ATC/co-DTC tumors (n = 104) that were subjected to either both WES and CNA or WGS analysis and that had no identified SNV alterations (in total 21 ATC and 5 co-DTC tumors) in pathways (such as MAPK-RAS, SWI/SNF, tumor suppressors, etc., Figure 3) that were highlighted in previous thyroid cancer studies or SeqSig recurrent genes. The lack of SNV alterations in this collection of genes was not due to differences in tumor purity (Wilcoxon rank-sum test, p = 0.77, Figure S3C) or sequencing coverage (Wilcoxon rank-sum test, p = 0.041, with coverage higher in tumors with no driver SNVs, Figure S3D). Analysis of CNA status revealed that, although the 21 ATC samples with no putative drivers exhibited less overall genomic alteration (Wilcoxon rank-sum test, p = 0.0064, Figure S3E), 10/21 ATC tumors had deletions in one of the tumor suppressors *CDKNA2*, *RB1*, *CDK7*, or *BRCA2* (often multiple alterations, Figure 3). These results suggest that copy-number losses in these tumor suppressors could be the underlying drivers of disease progression and dedifferentiation. Future studies will be needed to comprehensively examine whether alternative mechanisms (such as fusions) could be underlying drivers in the remaining cases.

Clinicogenomics of ATC

To assess the clinical relevance of these somatic mutational features, we first assessed the impact of clinical factors on overall survival (Table S1). Treatment with radiotherapy (FDR = 1.7×10^{-5}) or surgery (FDR = 0.0089) was positively associated with patient outcome, as previously described.[47] Nodal involvement (FDR = 0.009), distal metastasis (FDR = 0.025), leukocytosis (FDR = 0.056), and patient age (FDR = 0.08) were all associated with reduced survival. The use of surgery and patient age strongly stratified ATC survival outcomes (Figure S5A).

We then considered the association of individual driver features with patient outcome, including patient age and treatment as covariates in time-to-event analyses. No genes with point mutations in ≥5% of ATCs were associated with overall survival (FDR > 0.1, Table S7). Although CNA subtypes A–E were not associated with overall survival (p = 0.6, Figure S5B), tumors with fewer CNAs more often presented with distant metastases (Wilcoxon rank-sum test, FDR = 0.034; median percentage genome altered [PGA] for patients with and without distant metastases = 8.6% and 11.6%, respectively). PGA was not associated with overall survival (median dichotomized; HR = 0.93 [95% CI, 0.59–1.47], p = 0.77, adjusted for patient age and surgery), unlike in several other tumor types.[39,48,49] *BRCA2* deletion was surprisingly associated with better survival (HR = 0.48 [95% CI, 0.29–0.80], p = 0.005; Figure S5D), but no other recurrent CNA was (Table S8).

By contrast, CNA subtypes showed distinctive clinical hallmarks: subtype A was enriched for older patients with better overall survival, subtype B was depleted for patients with nodal involvement as well as moderately enriched for patient death, subtype C was enriched for younger patients, subtype D was enriched for male patients with metastatic disease, and subtype E showed no associations with clinical variables (Figure S5E). Despite limited

statistical power, patients with a lower mutation rate (>10 SNVs/Mb) had significantly better survival (HR = 0.51 [95% CI, 0.33–0.77], p = 0.002; Figure S5F). Thus, the mutational features and subtypes of ATC are associated with divergent clinical presentation and outcome.

ATC and DTC share common genomic features

It has been suggested that a subset of ATCs evolve from preexisting DTCs.[11–13,23,50,51] Given the strong association of ATC CNAs with clinical and molecular phenotypes, we first focused on copy-number changes. We compared SNV and CNA profiles between PTCs, co-DTCs, and ATCs using both newly profiled PTC samples from this study and the TCGA dataset.[25] The two recurrent arm-level events of PTC were 1q amplification and 22q deletion,[25] both of which occurred at similar frequencies in ATCs and the co-occurring DTC components of ATCs (Figures 1B and 4C). All recurrent focal CNAs in PTC were also detected in ATC, but several CNA drivers were significantly more frequent in ATC (Figure 4A). Loss of *CDKN2A* was recurrent in both co-occurring DTCs (Figure 4B) and ATCs but was rare in PTCs (~5%). *BRCA2* was also frequently deleted in ATCs (33.6%), uncommon in DTCs (13.6%), and rare in PTCs (4.5%), as was *RB1*, which is on the same chromosome arm. Several other regions showed higher rates of copy-number change in ATCs and co-occurring DTCs, including a broad region on chromosome 20q harboring 328 genes that was preferentially amplified in ATC relative to both cohorts of PTCs (proportion-test, FDR < 0.01; Table S9). This region harbors several cancer driver genes, for example, *ARFGEF2*, *CHD6*, and *GNAS*, all ubiquitously expressed in thyroid tissue,[52] although in our limited RNA-seq data we did not identify an association between CNA status and RNA expression in genes within this region (Table S10).

We next compared the somatic CNA (Figure 4C, n = 21) and SNV profiles (Figure 4D, n = 25) of DTCs co-occurring with ATCs. DTCs co-occurring with ATC showed a larger fraction of their genome affected by CNAs than isolated DTCs (Figure 4C). Similarly, co-occurring DTCs harbored more SNVs than did isolated PTCs or other thyroid cancers arising in individuals without an ATC diagnosis: their overall mutation density was statistically indistinguishable from that of ATC (4.4 ± 2.8 SNVs/Mb). Similarly, to explore somatic driver SNVs, we merged the current study with six previous sequencing studies of ATC, PTC, and OCT[17,21,23,25,36,37] (Table S11). *BRAF* was the only gene more common in differentiated thyroid cancers: it was mutated in 50.9% of PTCs, 50.0% of co-occurring DTCs, and 21.3% of ATCs (Figure 4E). Indeed, in matched co-occurring cases, half of all $BRAF^{V600E}$ variants were detected solely in the DTC component and not in the ATC component. By contrast, multiple other drivers were preferentially mutated in ATCs and co-occurring DTCs, including *ATM*, *ATR*, *BRCA2*, *PIK3CA*, and *TP53* (Figure 4E). As an example, TP53 mutation frequency varied from 0.9% of PTCs to 21.4% of co-occurring DTCs and 36.8% of ATCs (Table S11). We further compared the relative frequency of PTC and ATC tumors exhibiting $BRAF^{V600E}$ and *RAS* mutations, as PTC exists as separate BRAF- or RAS-like subtypes.[25] $BRAF^{V600E}$ mutations exhibited elevated frequency in PTC relative to ATC, suggesting earlier, preferentially clonal, evolutionary timing, whereas no difference was found for *NRAS*.

These patterns of driver mutations were also mirrored for mutational processes. The two most prominent mutational signature activities in PTCs were COSMIC signatures 2 and 13 (AID/APOBEC activity) and 6 (MMR deficiency; Figure S6A). These were uncorrelated to patient sex or age (Student's t test, FDR > 0.1) and were detected in ATC. Of the co-occurring DTCs, about half (12/25) showed mutational signatures identical to their ATC counterparts and the remainder showed either higher activity of a specific signature or a mix of multiple signatures. These data are consistent with a model in which ATCs and DTCs evolved from a common precursor and specific mutations within that common precursor increased the risk of ATC evolving.

ATC and DTC evolved in parallel from a mutagenic field

Although the co-evolution model of ATC and DTC has been proposed previously, based on limited panel sequencing, it has not been confirmed with genome-wide systematic subclonal reconstruction. We therefore utilized the paired ATC and co-occurring DTC multi-region WGS data to infer the clonal hierarchy and mutation timing for each patient using a validated subclonal reconstruction pipeline[53,54] for both SNVs (Figures 4D and S6B; Table S12) and subclonal CNA status (Figure 4F; Table S13). High-coverage (median 567, range: 347–809) WES showed a high degree of correlation in variant allele frequency in coding SNV alterations compared with our WGS data (Figure S6C).

Using the WGS data, we found that, as anticipated, each tumor exhibited a distinctive evolutionary history, but every single case shared two features. First, co-DTC and co-occurring ATC shared a common clonal origin, with a common ancestor splitting to form separate lineages with significant subclonal diversification (Figures 5A–5J). This split of the ATC and DTC lineages typically emerged early in tumor evolution: the common ancestor harbored ~95% of CNAs (Table 1), but only 19.1% ± 7.9% of SNVs. Regions of co-occurring DTC were not restricted to a specific CNA subtype and showed no consistent trends in PGA, total number of CNAs, or other characteristics (Figure 4F), consistent with a lack of clonal interactions.

In most cases the spatially separated ATC component contained no DTC clones, as in ATCWGS37 (Figure 5B) and ATCWGS34 (Figure 5C), where in each case the rapidly growing ATC exhibited only a single subclone at our limits of detection. In several cases, we were able to clearly identify a mutagenic field where a large proportion of cells possessed similar oncogenic alterations, for example, in ATCWGS34 (Figure 5C) and ATCWGS33 (Figure 5D). In these cases, spatially distinct tumors harbored common mutations in both ATC and co-occurring DTC, providing clear evidence for a prior mutagenic field from which both tumor components emerged. ATCWGS33 also provided a remarkable example of the complexity of the subclonal interactions between the two tumor components. The tumor diverged from a common field into two branches (subclones 2 and 3, Figure 5D). Both the differentiated and the anaplastic tumor regions contained clones from both branches of these tumors, but with significant differential evolution: compare subclone 3 (found in DTC at low frequency) with subclone 6 (found in ATC at high frequency).

While several mutations showed biases in evolutionary timing (Figure S6B), *BRAF* SNVs notably occurred in a clonal or early subclonal population. We compared primary with

metastatic ATC for ANPT0021 (Figure S6B): the primary ATC exhibited a clonal or high-prevalence subclonal *BRAF* mutation, while its cervical metastasis had nearly undetectable levels of this variant, suggesting either allelic loss or that the metastasis derived from a low-frequency subclone in the primary tumor. Taken together, these data are consistent with a model in which ATCs and DTCs emerge from a common mutagenic field effect. More studies will be needed to validate these findings in tumors from other ATC patients.

DISCUSSION

The vast majority (>90%) of thyroid cancers are well-differentiated papillary and follicular types with extremely favorable prognoses (>90% survival at 10 years).[55–58] In contrast, ATC is rare and is one of the most lethal human malignancies. It is intriguing that the same organ harbors both the most aggressive and some of the most indolent, manageable tumors. The genomic landscape of PTC has been reported to have a low mutational burden,[25] with either single activating oncogene mutations (*BRAF* or *RAS* family genes) or receptor tyrosine kinase gene fusions.[25] The genomic landscape of OCT is characterized by a small number of chromosome duplications (particularly chromosome 7) and uniparental disomy and few recurrent point mutations.[36,37] In contrast, the genomic landscape of ATC is complex, with significant interpatient heterogeneity; individual tumors carry mutations in multiple oncogenes and tumor suppressors and an overall elevated rate of copy-number changes.

ATC frequently occurs in patients with a previous history of DTC or contains co-existing areas of DTC,[11–13] leading to speculation that some ATC evolves via dedifferentiation from DTC. However, prior evidence has been limited to analysis of a small number of genes, particularly *BRAF*.[50,51] WES and CNA profiling of the DTC and ATC components from 21 tumors, combined with WGS profiling of paired ATC and co-occurring DTC components from 9 tumors, conclusively demonstrated that both components share a common genetic origin. High-coverage validation of a subset of cases showed that many ATC driver mutations are subclonal and exclusive to the ATC component. Our study has provided a detailed reference of somatic SNVs and CNAs, which are more frequently altered in ATC than in PTC. These include known SNVs in *TP53*, *PIK3CA*, and *EIF1AX* and deletions in *CDKN2A*, but also a large number of novel SNVs and CNAs. However, only a subset of these alterations are likely critical drivers of thyroid cancer progression. Functional studies are necessary to determine which mutations lead to the dedifferentiation and increased aggression of well-differentiated thyroid cancer models, as have been carried out by our group and others for *PIK3CA*, *EIFAX*, and SWI/SNF complex genes.[59–61]

Our subclonal reconstruction analysis suggests that ATCs evolve from a DTC subclone (Figure 5K), after accumulation of mostly DNA CNAs. This clone then acquires characteristic additional oncogenic drivers and the majority of its somatic alterations. The original DTC clone and other subclones in the mutagenic field can continue to co-exist within ATC as a clonal admixture, but only the ATC clone harbors metastatic potential. These data suggest a model in which both ATC and DTC arise from a non-malignant mutational field effect, where a set of cells accumulates some SNVs and many CNAs through some combination of environmentally driven and replication-associated mutational processes. Some of these cells then accumulate the characteristic somatic mutations of

DTCs and give rise to curable tumors. A different subset accumulates further specific driver mutations and eventually subclonal driver point mutations in genes like *TP53*. This model suggests that their evolutionary trajectories may lead to the differential aggressiveness of these related tumor types and the expansive subclonal architecture of ATC. ATC would then emerge from precursors within its co-occurring DTCs by significantly increasing genomic diversity and evolve further as it colonizes new metastatic sites. It is unclear if this model also holds for ATCs not arising in the context of DTCs; WGS of non-malignant thyroid and multi-region ATC specimens will be required to address this question.

Recently, single-cell RNA-seq (scRNA-seq) has significantly advanced our understanding of ATC and provided valuable insights into its molecular and cellular heterogeneity. Lu et al.[62] identified three distinct populations of anaplastic cancer cells, labeled inflammatory, mitotic-defective, and mesenchymal ATC cells. Single-cell trajectory analysis in combination with transcriptome-based copy-number estimation revealed that diploid inflammatory ATC cells can acquire aneuploidy and a mesenchymal phenotype that produce a large abundance of cell-adhesion molecules. Similarly, two other studies[63,64] using scRNA-seq and developmental trajectory analysis both identified a group of ATC cells possessing high expression of mesenchymal gene programs that is associated with loss of thyroid markers, highlighting the potential role of this pathway during dedifferentiation. These results align with our large-scale analysis of the genomic landscape in ATC and subclonal reconstruction results, where ATC tumors can possess either very few or many CNAs, sometimes at different time points from the same patient. Further studies combining DNA and RNA-seq at single-cell resolution are needed to precisely define across time points the order of and interplay between genomic and transcriptional alterations that define each step of the progression of thyroid cancer.

On top of genes commonly altered in cancer or previously linked to ATC, such as *BRAF*, *RAS*, *TP53*, *PIK3CA*, *BRCA1*, *BRCA2*, and *USH2A*, our unbiased genomic analysis identified tens of novel alterations enriched in ATC, including recurrent non-synonymous SNVs in the pathways related to MAPK signaling, DNA damage repair, cell division, and cell adhesion. Recurrent alterations in *BRCA1*, *BRCA2*, and *ATM* at both the SNV and the CNA level rationalize the investigations of PARP inhibitors that have transformed treatment of DNA-damage-response-deficient cancers.[65–67] Furthermore, although exact mechanisms are not well understood, frequent alterations in cell-adhesion genes, such as *HMCN1* and *HSPG2*, along with altered transcriptomic programs in scRNA-seq studies, highlight the importance of cell-adhesion and stemness pathways. Specifically, *HMCN1* encodes a large extracellular matrix protein involved in maintaining tissue integrity, promoting cell adhesion, and modulating cell migration.[68–71] HMCN1 abundance is frequently downregulated or lost in various types of cancers, including gastric, breast, colorectal, and ovarian cancer, which is often associated with poor prognosis and survival.[72,73] Disruption of the gene can lead to increased tumor cell invasion and metastasis.[74–76] *HSPG2* encodes the large multidomain proteoglycan perlecan, which is a significant constituent of the extracellular matrix in numerous tissues, contributing to tissue integrity, cell adhesion, and migration.[77] HSPG2 can act as a WNT ligand to trigger downstream WNT signaling, leading to tumor progression and metastasis.[78–81] Further research is needed to examine the alterations in these two genes and their roles in ATC initiation and progression.

Limitations of the study

Development of the 15 institution GATCI was critical to facilitate the characterization of this large number of ATC samples, a pathologic review by world experts, and a platform for future collaborative studies. However, there are limitations to this first work, including the depth of sequencing, which may have resulted in lower mutation rates than reported in other studies,[17,19,21,23,82] heterogeneity in platforms used, incomplete clinical data for all patients, limited sample size of tumors characterized with RNA-seq, and the fact that most samples were collected prior to widespread use of targeted therapy, which may have affected survival outcomes. We foresee that prospectively collected ATC tumors profiled using deep-coverage WGS approaches will more precisely define the coding, non-coding, and copy-number alterations that underlie thyroid cancer progression. We are directly addressing these issues in future projects.

Despite these limitations, our multi-omic platform study defined the driver landscape and disease subtypes and identified a long tail of single-nucleotide alterations and recurrent copy number variations, along with frequent activation of the AID/APOBEC mutational process. Furthermore, we delineated the evolutionary relationship between ATC and differentiated thyroid cancers and provided genomic evidence that these two cancer types share a clonal origin. Taken together, our data not only provide a major resource for the thyroid cancer research community, but also have led to ongoing work evaluating the functional importance of the novel genes and mutational process that we have identified.[60]

STAR★METHODS

RESOURCE AVAILABILITY

Lead contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Dr. Anthony Nichols (anthony.nichols@lhsc.on.ca).

Materials availability—This study did not generate new unique reagents.

Data and code availability

- All ATC WES, WGS and SNP array data is available in EGA under accession EGA: EGAS00001002234. The processed ATC molecular dataset will be deposited such that it can be explored and accessed via CBioPortal. Variants calls has also been deposited as Supplementary Data.
- Original code can be accessed at https://zenodo.org/records/10600153
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Research Ethics Board approval was obtained at all site locations. Tumors from patients previously diagnosed as ATC were reviewed by at least one study pathologist at each site. Samples meeting the American Thyroid Association diagnostic criteria for anaplastic

thyroid cancer[9] were selected for study including both fresh frozen and formalin fixed paraffin embedded (FFPE) samples (Table S1). All WES tumors were from fresh frozen samples or FFPE while matched normal subjected to WES were from FFPE. All tumor samples subjected to WGS were from FFPE. Matched normal were from blood tissue. Well-differentiated and/or poorly differentiated components were macro-dissected when present separately from the undifferentiated ATC components using 1 mm punches. Fresh tumor was prospectively collected at four institutions with matched blood collected for the majority of cases. DNA was extracted using Qiagen kits. Where available, samples were subject to centralized pathology review to estimate tumor purity and confirm diagnoses. PTC samples (n = 115 from 112 patients, Table S1) were selected from the tissue bank at MD Anderson Cancer Center based on age and tumor status (organ confined, regional involvement, or regional and distant metastasis). All cell lines were acquired from (Table S14A) from Japanese Collection of Research Bioresources Cell Bank (JCRB), Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ), University of Uppsala, or the Mayo Clinic and were subjected to regular STR profiling (Table S14B). DNA was isolated and processed[97,98]. All patient samples were collected before 2018, preceding the FDA approval of dabrafenib and trametinib in ATC.

METHOD DETAILS

Detection of copy number aberrations—Affymetrix OncoScan FFPE assays were used to evaluate a total of 157 samples (including 99 ATCs, 24 samples from co-occurring well (20) or poorly (4) differentiated regions, 1 co-occurring cervical ATC metastasis, 20 normal thyroid samples and 13 ATC-derived cell lines) for somatic copy number aberrations (CNAs). Analysis was performed using .OSCHP files generated by OncoScan Console 1.1 using either build 33 of the NetAffx annotation (FFPE samples) or a custom reference consisting of 119 normal blood samples from male patients with prostate cancer, 2 normal blood samples from females with anaplastic thyroid cancer and 10 female HapMap cell line samples (fresh frozen samples and cell lines). BioDiscovery's Nexus Express Software was used to call copy number aberrations using the SNP-FASST2 algorithm with default parameters, with manual re-centering performed as required. Nine samples (six normal tissues, one ATC and two well-differentiated tumor components) were removed due to poor quality and the remaining data evaluated for recurrence. Purity and ploidy of tumor samples was assessed using ASCAT (v2.5) package in R. LRR and BAF values were obtained from the. OSCHP files. Tumor ploidy and aberrant cell fraction (purity) estimates for each sample were obtained using either predicted germline SNPs or by leveraging matched normal arrays where available.

PTC samples (113 tumors), along with an additional 13 ATCs, were processed using Illumina Human Omni2.5–8 or OmniExpress12v1–1 beadchips, [97,98]. Briefly, DNA was denatured, amplified, enzymatically fragmented and hybridized for 16–24 hours at 48°C. Beadchips were imaged using the Illumina iScan system. For improved CNV analysis, B allele frequencies (BAF) were calculated and log2 R ratios (LRR) were extracted after re-clustering the raw data using GenomeStudio cluster algorithms. BAF and LRR were then loaded into Nexus to identify somatic DNA copy number aberrations. UCSC's command-line liftover tool (v359) was used to transfer resulting CNAs to GRCh38 coordinates.

CNAs from both platforms were combined, and UCSC's command-line liftOver tool (v359) was used to convert segment positions to hg38 coordinates (using the hg19-to-hg38 chain file downloaded from UCSC). Gene level copy number aberrations for each patient were identified by overlapping CN segments with Ensembl annotation (Ensembl release 84). Percentage of genome altered (PGA) was calculated for each sample by dividing the number of base-pairs involved in a copy number change by the total length of the genome. GISTIC2.0 (v2.0.22)[84] was used to study the recurrence of gene level CNAs. For each sample, a profile was created that segmented each chromosome into regions with neutral, CN loss, and CN gain events. The average copy number intensity for each segment was obtained from the SNP array analysis. Chromosome-level events were defined as a gain or loss of 25% of the chromosome for each sample.

Subtype discovery was performed on CNA calls for 110 ATCs using consensus clustering to determine the optimal number of clusters and class membership for each sample by stability evidence. Specifically, the ConsensusClusterPlus.custom (v1.8.1) package for R was used to evaluate a wide variety of clustering methods for distributing samples into k clusters (where k is every value from 2 to 10) using 1000 sub-samplings of 80% of the cohort. The method which produced the most stable groups used hierarchical clustering using a Euclidean distance similarity metric with modified Ward's minimum variance method (ward.D) and 5 clusters.

Data visualization was performed in the R statistical environment (≥ v3.2.3). Venn diagrams were created using the VennDiagram package (v1.6.17) and all other data visualizations were generated using the BPG (v5.8.8), lattice (v0.20–35) and latticeExtra (v0.6–28) packages.

DNA whole-exome sequencing (WES)—Tumor samples (and matched normal tissue where available) from 139 patients (including 139 ATCs with 20 co-occurring regions of well or poorly differentiated tissue, 1 cervical ATC metastasis) and 13 ATC-derived cell lines (total = 173 tumors) obtained from eight centres were sequenced at one of six sequencing facilities (Table S1). Raw FASTQ files were provided for a subset of samples while the remainder were provided in BAM format. In the latter case, read extraction was performed using the SamToFastq component of Picard tools (v1.121) to produce FASTQ files. Data were aligned to the GRCh38 human reference genome, including available ALT and decoy alleles using BWA-mem (v0.7.15)[86]. The reference files were downloaded from NCBI (GCA_000001405.15; release date 2013/12/17). Duplicate reads were marked using Picard tools (v1.121). As the exome capture kit regions were provided in hg19 coordinates, USCS's web-based liftover tool (https://genome.ucsc.edu/cgi-bin/hgliftover) was used to convert regions to GRCh38 positions. The Genome Analysis Toolkit (GATK v3.5.0)[87] was used for local realignment and base quality recalibration, with regions limited to those targeted by the applicable exome capture kit (Table S1). As hg38 has not yet been widely adopted, a beta-version of the hg38bundle containing known variants was downloaded from the Broad (ftp://ftp.broadinstitute.org/bundle/hg38/hg38bundle/).

GATK's HaplotypeCaller (v3.5.0) was used to identify germline SNPs and indels using default parameters, however with –output_mode EMIT_VARIANTS_ONLY

–emitRefConfidence GVCF –standard_min_confidence_threshold_for_calling 50. GenotypeGVCFs and SelectVariants tools were used to format variant calls, followed by VariantFiltration with the following criteria: (QD < 10.0 || FS > 60.0 || MQ < 40.0 || DP < 50 || SOR > 4.0 || ReadPosRankSum < −8.0 || MQRankSum < −12.5), (MQ0 >= 4 && ((MQ0 / (1.0 * DP)) > 0.1). Variants were annotated using SnpEff (v4.3)[99] and filtered to keep any position with coverage in the 1000 genome project that had a population allele frequency <1%. These were then filtered to keep variants in clinVar and to remove synonymous variants. Finally, for patients with multiple tumor components, germline variants present in a single component were labelled false positives and removed. Remaining variants were searched against a list of relevant cancer predisposition genes[30] and checked for recurrence.

Coverage was estimated across target regions using BEDTools (v2.18.2) (Figure S2A). For tumor/normal pairs, ContEst (v1.0.24530)[88] was used to estimate cross-sample contamination, using the above generated germline SNPs. As the necessary population allele frequencies were not yet available using GRCh38 coordinates, allele frequencies were obtained from HapMap (hg19) and added to the GRCh38 HapMap file, with variants matched across builds using rsID. Seven samples demonstrated contamination >3% and were removed from downstream analysis (Figure S2B).

For samples with matched normal tissue, somatic SNVs were predicted using SomaticSniper (v1.0.5.0)[89], [6]. A panel of normals (PoN) consisting of the 83 normal thyroid samples was generated using MuTect's (v1.1.7) artifact detection mode; variants detected in two or more normal samples were included in the PoN. For unmatched samples, SNVs were identified using MuTect (v1.1.7)[90] with PoN, with dbSNP (build 150) and COSMIC (v74) filters; COMSIC variants present in five or more samples within the PON were discarded. To combine samples processed at different sequencing facilities, predicted somatic variants were filtered to regions appearing in all exome capture kits used (Table S1). Overlapping regions across kits were identified using bedtools:multiIntersectBed (v2.24.0), resulting in coverage of roughly 26 Mbp. Recurrence analysis was performed using RecSNV (v2.1.5). Briefly, all variants were annotated using Annovar (v2016Feb01) with comprehensive filtering performed to remove known germline variants using previously described datasets[6]; any variant present in COSMIC (v82) was retained, regardless of presence in other datasets. Significance of recurrent non-synonymous mutations was assessed using SeqSig (v3.7.8)[6], using a required mutation threshold of 6 samples across the cohort per gene, individual sample weights of $-\log_{10}(x) + \log_{10}(1-x)$ [where x is the background mutation rate per sample], with p-values calculated using the exact distribution, followed by FDR correction for multiple testing. Genes containing variants in multiple samples were carried forward. Genes harbouring recurrent SNVs were visualized using lollipop plots (Figure S3). For each gene, protein domain information was obtained from NCBI and pfam (2018–02-07), showing the most well conserved domains where feasible.

PTC samples (115 tumors with matched normal and 1 metastatic lymph node) were processed separately, [97,98]. Briefly, DNA was prepared and processed for WES at the Human Genome Sequencing Center (HGSC) in Baylor College of Medicine (BCM). Reads were aligned to GRCh37 by BWA[86], duplicate reads were marked by Picard tools and BAMs realigned/recalibrated by GATK[87]. Somatic SNVs were identified using Atlas-SNP;

filtering was applied to ensure variants had a minimum of 4 high-quality support reads and a minimum VAF of 0.08. For comparison to the above ATC cohort, the resulting SNVs were transferred to GRCh38 coordinates using UCSC's command-line liftOver tool (v359).

Deep targeted IonTorrent sequencing—A total of 1,140 variants were selected for validation based on recurrence within the initial dataset of 30 ATCs. Primer design generated an AmpliSeq-custom panel for 1125 of these variants (995 target regions). A subset of samples was used for targeted validation by IonTorrent sequencing including 30 ATCs (13 with matched normal, 3 with matched co-occurring DTCs and 1 cervical ATC metastatic tumor), 13 ATC-derived cell lines and an additional 7 ATCs without WES data. DNA for each sample was shipped to GeneDX for sequencing. Raw FASTQ files were provided, and reads were aligned to GRCh38, as described above however without marking of duplicate reads. Targeted sequences were converted from hg19 to GRCh38 coordinates using liftOver (UCSC). Read lengths were extracted from FASTQ files and visualized against target region length (Figure S2C). Coverage was estimated across target regions using BEDTools[100] (v2.18.2) (Figure S2D).

Base counts at target positions were assessed and used to calculate variant allele frequencies (Figure S2E). Validation was performed [101]. Briefly, variant positions were used, along with either the GATK-processed WES BAMs or the IonTorrent BAMs to generate a modified pileup file indicating the base counts at each SNV position. For T/N pairs, metrics including a χ^2 test of the base-count distribution between tumor and normal at each position followed by Bonferroni adjustment of the p-values ($p_{adj} < 0.25$; Figure S2F, left panel), Euclidean distance between variant allele frequencies (>0.15; Figure S2F, right panel), z-score of the sample coverage across target positions (< 2 standard deviations from the mean for tumor and normal separately), and ternary allele proportion for normal samples to quantify potential false positive variants (< 0.05) were used to classify variants as somatic or non-somatic. For tumor only samples, similar metrics were used, however comparing Deep IonTorrent and WES values to classify read-counts as similar or different between the platforms. Accuracy, sensitivity and specificity were assessed for each sample (Figure S2G). Variants identified as false positives using this method were excluded from WES analyses.

DNA whole-genome sequencing (WGS)—Whole-genome sequencing (WGS) was performed on nine ATCs, each with paired co-occurring DTCs and matched normal. Genomic DNA was isolated from FFPE samples and sequenced at The Center for Applied Genomics (Hospital for Sick Children, Toronto, ON). FASTQ files were downloaded and processed using the methods described above. Somatic SNVs were identified using SomaticSniper (v1.0.5.0)[89], as described above and SCNAs were identified using MutationSeq (v4.3.7) and TITAN (v1.20.1)[91], followed by multi-sample subclonal reconstruction using PhyloWGS (v1.5)[92]. Clonal mutations are defined as mutations that are present in all tumour cells (CCF >= 0.8) of a tumour sample or biopsy. A subclones is defined as those that is a descendent of the most recent common ancestor of the tumour sample and with CCF < 0.8 in at least one tumour region.

Validation of TERT promoter mutations by PCR—To determine the mutational status of the TERT promoter, we used a nested PCR strategy with our FFPE patient samples.

Genomic DNA in this region is 78% GC rich and it was necessary to use PCR additives to amplify the region. The primers were designed and tested *in silico* using PRIMER3 software. The first round PCR was performed with primers hTERTF x hTERTR under the following conditions: initial denaturation at 98°C for 30 seconds followed by 36 rounds of 98°C denaturation for 10 s and combined annealing and extension at 72°C for 25 sec. We employed the Q5 high fidelity DNA polymerase (New England Biolabs; NEB) and the supplied 5x PCR buffer and GC enhancer (NEB) to amplify a 474 bp region. The second round of PCR was initiated with 0.2 ul template taken directly from the first round using primers TERT5 F x TERT5 R again with the addition of the GC enhancer and the same conditions increased to 40 cycles. PCR reactions were run on a 2% TBE agarose gel and the expected 193 bp product was excised and gel purified using a Monarch Gel Extraction Kit (NEB) following the product protocol. The purified fragments were Sanger sequenced at the London Regional Genomics Facility with betaine (1M) added to the reaction. Chromatograms were aligned and compared against the TERT sequence.

RNA sequencing—Total RNA was isolated from 13 ATC cell lines and 24 ATC tumor samples using Qiagen AllPrep DNA/RNA kits and shipped to The Center for Applied Genomics (Hospital for Sick Children, Toronto, ON) for sample processing using random primers and sequenced using paired end reads of 75 bp using the manufacturer's protocols. Specifically, sequences were generated using paired-end, strand-specific methods. Data were provided in BAM format; FASTQ files were re-generated using picard (v1.141) CleanSam and samtools (v1.4) fastq functions. A second set of 16 patient tumors were processed [98]. Briefly, frozen tissue was thawed in RNAlater™-ICE Frozen Tissue Transition Solution (ThermoFisher Scientific) to enable easy extraction of high-quality RNA. Total RNA was prepared using TRIzol reagent (Life Technologies) according to the manufacturer's instructions. Poly-A+ Illumina RNA-Seq libraries were prepared and paired-end 75bp sequencing performed using the Illumina HiSeq 2000 and HiSeq 2500 platforms at the HGSC (Baylor College of Medicine, Houston, TX).

For mRNA abundance, DNA sequences and gene annotations (GTF) were downloaded from ensemble.org (release 88) and used to generate reference files needed for RSEM (v1.3.0)[93], using the STAR (v2.3.3a)[94] alignment tool; transcripts per million (TPM) values are available in Table S3. Fusion detection was performed using fusioncatcher (v0.99.7c)[44]. Detected fusions are shown in Table S6; these were compared against the COSMIC fusion database (2019–12-13).

The TCGA PTC cohort RNA-seq read counts processed using STAR and RSEM were retrieved from GDC. Immune population estimation was performed using MCPCounter (v1.1.0)[95], as we previously described[43,102–104]. Genes for the thyroid differentiation score (TDS) and BRAF-RAS score (BRS) were retrieved from the TCGA PTC publication. TDS and BRS were calculated as previously described[25].

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis—Non-negative matrix factorization (NMF) was used to assess the contribution of various trinucleotide signatures within our ATC and PTC cohorts

separately. NMF was performed in R (v3.4.1) using the NMF package (v0.20.6) using counts of abnormal trinucleotides from undifferentiated tumors. Rank estimation metrics, combined with visual inspection of consensus clustering maps, suggested the presence of five trinucleotide signatures for ATC, and two for PTC. Signatures were mapped to previously described trinucleotide signatures from COSMIC (downloaded from COSMIC on 2017-12-08) using the consensus of Euclidean distance, Pearson's or Spearman's correlation, followed by the Hungarian method for linear sum assignment[105] using the clue (v0.3-54) package for R. Spearman's correlation did not identify any associations between these 5 signatures and clinical covariates or CNA clusters.

Gene-wise mutation frequencies from the current dataset were compared to previous studies of thyroid cancer, including a set of 22 ATCs by WES published by Kunstman *et al.*[15], 84 PDTCs and 33 ATCs by targeted-sequencing published by Landa *et al.*[17], 182 PDTCs (including PTC, FTC and OCT) and 134 ATCs by targeted-sequencing published by Pozdeyev *et al.*[21], 27 ATCs and 15 PDTCs from Yoo *et al.*[23], 2 independent studies of 32 and 56 OCTs respectively by WES published by Gopal et al. and Ganly et al. and 496 PTCs from TCGA (MAF downloaded from GDC on 2017-11-02)[106] (Table S9). Where sufficient annotation was available, samples from each study were filtered to remove metastatic or other non-primary thyroid tumors. Variants from the TCGA PTC cohort were filtered to keep only those with predicted functional relevance with an additional coverage threshold applied to the MAF (minimum 10 reads in both tumor and matched normal). SNV data were collapsed to gene-level and proportion tests used to evaluate differences between studies (ATC or co-occurring DTC (GATCI) *vs.* TCGA PTC), followed by FDR adjustment for multiple testing.

Similar comparisons were made for CNAs, again using TCGA PTC data (masked copy number segment data downloaded from GDC on 2018-01-29). TCGA data were converted to ternary CN values and annotated with overlapping genes to produce a gene by sample matrix (primary tumor samples only). Proportion tests were again used to assess differences in population frequencies for each gain/loss separately, with FDR-adjustment (Table S11).

Well or poorly-differentiated co-occurring tumor components were obtained from 30 patients and a single patient provided both an ATC and metastatic tumor obtained from the cervical region. For CNA analysis, data were available for 21 co-occurring DTC tumor components as well as ATC. Hypergeometric tests were used to evaluate the CNA overlap across regions within each patient [P(actual overlap ≥ expected overlap)] (Table S13). Similar analyses were performed for SNV overlap (28 patients with multiple components; Table S12).

Clinical variables were assessed for associations with each other using χ^2 independence tests to assess co-linearity, and with overall survival to identify which should be used as covariates in downstream models. Age, nodal metastases and treatment methods (surgery, radiotherapy and chemotherapy) were all individually associated with survival (FDR < 0.1). Treatment methods showed considerable overlap, as did nodal and distant metastases (FDR < 0.1). Therefore, covariates including age, nodal metastases and surgery were used for survival analyses. Overall survival was evaluated using either CNA or SNV features. For

copy number status, of 961 genes identified by GISTIC, 723 were CN altered in 10% of the cohort with available survival data (n = 83) and were collapsed to 187 genomic regions. Regions were collapsed to binary status where reasonable, such that the CN type with 4 or fewer cases was merged with neutral. Further, CN status of known driver genes, with a minimum recurrence of 10 samples, were also tested (n = 108). Associations with survival were evaluated using either log-rank (for ternary CN status) or Cox proportional hazards regression model with linear adjustments for covariates described above (binary CN status). Binary cases which failed the assumptions for the Cox model were repeated using log-rank tests. FDR-adjustment of the p-values was applied for multiple testing correction (Table S8). For point mutations, 101 patients had available survival data. Feature reduction was performed to remove low frequency variants (≥ 5 patients); this reduced the feature set from 14,543 genes to 514 recurrently altered genes. Associations with survival were using a Cox proportional hazards regression model with linear adjustments using covariates described above, or log-rank test where the Cox model failed its assumptions. FDR-adjustment of the p-values was applied for multiple testing correction (Table S7).

In addition to gene-wise features, clinical and survival associations were also evaluated using PGA or SNVs/Mbp. Associations with clinical variables were performed using Wilcoxon rank sum tests. For survival analyses, PGA and SNVs/Mbp were median dichotomized and a Cox proportional hazards regression model with linear adjustments using covariates described above was performed.

Finally, mutational features were combined and examined for overlaps. Specifically, pairwise comparisons were made between all combinations of CNA features (key GISTIC genes, numbers of CNAs/gains/losses, average CNA length, PGA and CNA subtypes) and SNV features (top recurrently mutated genes as defined using SeqSig, SNVs/Mbp and trinucleotide signatures using either a χ^2 test, Spearman's correlation or non-parametric Wilcoxon or Kruskal-Wallis rank sum test with FDR adjustment of the p-values. Similar comparisons were made between RNA abundance (TPM) and CNA/SNV/fusion status using Wilcoxon rank sum tests.

ADDITIONAL RESOURCES

The processed ATC molecular dataset will be deposited such that it can be explored and accessed via cBioPortal.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Authors

Peter Y.F. Zeng[1,2,3,4,47], Stephenie D. Prokopec[5,45,46,47], Stephen Y. Lai[6,47], Nicole Pinto[1], Michelle A. Chan-Seng-Yue[5], Roderick Clifton-Bligh[7], Michelle D. Williams[8], Christopher J. Howlett[9], Paul Plantinga[9], Matthew J. Cecchini[10], Alfred K. Lam[10], Iram Siddiqui[11], Jianxin Wang[5], Ren X. Sun[5], John D. Watson[5,45,46], Reju Korah[12], Tobias Carling[12], Nishant Agrawal[13], Nicole Cipriani[14], Douglas Ball[15], Barry

Nelkin[16], Lisa M. Rooper[17], Justin A. Bishop[18], Cathie Garnis[19], Ken Berean[19], Norman G. Nicolson[12], Paul Weinberger[20,21], Ying C. Henderson[6], Christopher M. Lalansingh[5], Mao Tian[45,46], Takafumi N. Yamaguchi[5,45,46], Julie Livingstone[5,45,46], Adriana Salcedo[5,22,42], Krupal Patel[1,5], Frederick Vizeacoumar[23], Alessandro Datti[24,25], Liu Xi[26], Yuri E. Nikiforov[27], Robert Smallridge[28], John A. Copland[29], Laura A. Marlow[29], Martin D. Hyrcza[30], Leigh Delbridge[31,32], Stan Sidhu[31,32], Mark Sywak[31,32], Bruce Robinson[32,33], Kevin Fung[1,4], Farhad Ghasemi[1], Keith Kwan[9], S. Danielle MacNeil[1,4], Adrian Mendez[1,4], David A. Palma[2,3,4], Mohammed I. Khan[1], Mushfiq Shaikh[1], Kara M. Ruicci[1], Bret Wehrli[9], Eric Winquist[2,3,4], John Yoo[1,4], Joe S. Mymryk[1,2,4,34], James W. Rocco[35], David Wheeler[26], Steve Scherer[26], Thomas J. Giordano[36], John W. Barrett[1], William C. Faquin[37], Anthony J. Gill[32,38,39], Gary Clayman[40,47], Paul C. Boutros[22,41,42,43,44,45,46,47,*], Anthony C. Nichols[1,2,3,4,47,48,*]

Affiliations

[1]Department of Otolaryngology – Head and Neck Surgery, Western University, London, ON, Canada

[2]London Regional Cancer Program, London, ON, Canada

[3]Lawson Health Research Institute, London, ON, Canada

[4]Department of Oncology, Western University, London, ON, Canada

[5]Ontario Institute for Cancer Research, Toronto, ON, Canada

[6]Department of Head and Neck Surgery, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

[7]Division of Endocrinology, Royal North Shore Hospital, and University of Sydney, Sydney, NSW, Australia

[8]Department of Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

[9]Department of Pathology, Western University, London, ON, Canada

[10]Department of Pathology, School of Medicine, Griffith University, Gold Coast, QLD, Australia

[11]Department of Laboratory Medicine, Hospital for Sick Children, Toronto, ON, Canada

[12]Department of Surgery, Yale University, New Haven, CT, USA

[13]Department of Otolaryngology – Head and Neck Surgery, University of Chicago, Chicago, IL, USA

[14]Department of Pathology, University of Chicago, Chicago, IL, USA

[15]Division of Endocrinology, Department of Medicine, Johns Hopkins University, Baltimore, MD, USA

[16]Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University, Baltimore, MD, USA

[17]Division of Pathology, Johns Hopkins University, Baltimore, MD, USA

[18]Department of Pathology, University of Texas Southwestern, Dallas, TX, USA

[19]BC Cancer Agency, Vancouver, BC, Canada

[20]Department of Otolaryngology – Head and Neck Surgery, Louisiana State University Health Sciences Center, Shreveport, LA, USA

[21]Feist-Weiller Cancer Center, Louisiana State University Health Sciences Center, Shreveport, LA, USA

[22]Department of Medical Biophysics, University of Toronto, Toronto, ON, Canada

[23]Cancer Research Cluster, University of Saskatchewan, Saskatoon, SK, Canada

[24]Network Biology Collaborative Centre, Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, ON, Canada

[25]Department of Agricultural, Food, and Environmental Sciences, University of Perugia, Perugia, Italy

[26]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA

[27]Department of Pathology, University of Pittsburgh Medical Center, Pittsburgh, PA, USA

[28]Division of Endocrinology, Department of Medicine, Mayo Clinic, Jacksonville, FL, USA

[29]Department of Cancer Biology, Mayo Clinic, Jacksonville, FL, USA

[30]Department of Pathology and Molecular Medicine, McMaster University, Hamilton, ON, Canada

[31]Department of Surgery, Royal North Shore Hospital, Sydney, NSW, Australia

[32]University of Sydney, Sydney, NWS, Australia

[33]Department of Endocrinology, Royal North Shore Hospital, Sydney, NSW, Australia

[34]Department of Microbiology and Immunology, Western University, London, ON, Canada

[35]Department of Otolaryngology – Head and Neck Surgery, Ohio State University, Columbus, OH, USA

[36]Department of Pathology, University of Michigan, Ann Arbor, MI, USA

[37]Department of Pathology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

[38]Cancer Diagnosis and Pathology Group, Kolling Institute of Medicine, Royal North Shore Hospital, Sydney, NSW, Australia

[39]NSW Health Pathology, Department of Anatomical Pathology, Royal North Shore Hospital, Sydney, NSW, Australia

[40]The Clayman Thyroid Surgery and Thyroid Cancer Center, The Thyroid Institute, Tampa General Hospital, Tampa, FL, USA

[41]Department of Pharmacology and Toxicology, University of Toronto, Toronto, ON, Canada

[42]Department of Human Genetics, University of California, Los Angeles, Los Angeles, CA, USA

[43]Department of Urology, University of California, Los Angeles, Los Angeles, CA, USA

[44]Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research, University of California, Los Angeles, Los Angeles, CA, USA

[45]Institute for Precision Health, University of California, Los Angeles, Los Angeles, CA, USA

[46]Jonsson Comprehensive Cancer Center, University of California, Los Angeles, Los Angeles, CA, USA

[47]These authors contributed equally

[48]Lead contact

ACKNOWLEDGMENTS

The authors thank all members of the Boutros and Nichols' laboratories for helpful suggestions and technical support. We would like to thank Dr. Arunangshu Sarkar for assistance with data visualization. This work was supported by Canadian Institutes of Health Research grant MOP 377832 to A.C.N. and P.C.B., MOP 145586 to P.C.B., and the Government of Canada through Genome Canada and the Ontario Genomics Institute (OGI-125). N.A., B.N., and D.B. were supported by a Julius Edlow Pilot Research grant. P.C.B. was supported by Canadian Institutes of Health Research and Terry Fox Research Institute New Investigator awards and by the Ontario Institute for Cancer Research through funding provided by the government of Ontario. This work was further supported by National Institutes of Health and Medical Research grant R01 CA136665 (to J.A.C. and R.S.), Florida Department of Health Bankhead-Coley Cancer Research Program grant FL09B202 (J.A.C. and R.S.), Mayo Comprehensive Cancer Center grant P30CA01508343 (R.S.), UCLA Comprehensive Cancer Center grant P30CA016042 (P.C.B.), and the NIH/NCI through support from the ITCR (1U24CA248265-01 (P.C.B.), as well as support from Alfred D. and Audrey M. Petersen (R.S.), the Francis and Miranda Childress Foundation Fund for Cancer Research (J.A.C.), the John A. and Bette B. Klacsmann Fund for Cancer Research at the Mayo Clinic in Florida (J.A.C.), the Bruno V. and Bruce E. Zanoni Endowed Research Fund (J.A.C.), the Betty G. Castigliano Fund in Cancer Research Honoring S. Gordon Castigliano, MD, Cancer Research at the Mayo Clinic Florida (J.A.C.), and the Dewitt C. (Dash) Goff Fund for thyroid cancer research (J.A.C.). This work was also supported through generous donations by (1) the Marty Schaffel Thyroid Cancer Research Fund (MD Anderson Cancer Center, Houston, TX), (2) London Health Sciences Foundation (Robert and Sheila Wilkes, Betty and Harry Ostrander, Mervyn and Joyce Dietz), and (3) the Woodstock Foundation (Laila, Arnold, Mario, Andrea, Tony, and Tomassina Spina; Susan and Bill George; Piero and Maria Manzini; Henry and Rina Deroo; Harry and Shani Loewith; Franco and Carol Castellucci; Tony and Bill Van Haeren; Billy and Jennifer Stevanovich; Tom and Pat Baird; Christine and Jeff Nichols; Cliff and Linda Zaluski; Thomas Vandertuin; Laurie and Paul Green; Kelli and Mike Koopman). A.C.N. was supported by the Wolfe Surgical Research Professorship in the Biology of Head and Neck Cancers Fund. P.Y.F.Z. is supported by a Vanier Canada Graduate Scholarship and a PSI Foundation fellowship.

REFERENCES

1. Bruinsma SM, Roobol MJ, Carroll PR, Klotz L, Pickles T, Moore CM, Gnanapragasam VJ, Villers A, Rannikko A, Valdagni R, et al. (2017). Expert consensus document: Semantics in

active surveillance for men with localized prostate cancer - results of a modified Delphi consensus procedure. Nat. Rev. Urol. 14, 312–322. 10.1038/nrurol.2017.26. [PubMed: 28290462]

2. Shaha AR, and Tuttle RM (2018). Editorial: Risk of disease progression during active surveillance of papillary thyroid cancer. Surgery 163, 53–54. 10.1016/j.surg.2017.08.020. [PubMed: 29122326]
3. Garrido-Laguna I, and Hidalgo M. (2015). Pancreatic cancer: from state-of-the-art treatments to promising novel therapies. Nat. Rev. Clin. Oncol. 12, 319–334. 10.1038/nrclinonc.2015.53. [PubMed: 25824606]
4. Gross AM, Orosco RK, Shen JP, Egloff AM, Carter H, Hofree M, Choueiri M, Coffey CS, Lippman SM, Hayes DN, et al. (2014). Multi-tiered genomic analysis of head and neck cancer ties TP53 mutation to 3p loss. Nat. Genet. 46, 939–943. 10.1038/ng.3051. [PubMed: 25086664]
5. Bayani J, Yao CQ, Quintayo MA, Yan F, Haider S, D'Costa A, Brookes CL, van de Velde CJH, Hasenburg A, Kieback DG, et al. (2017). Molecular stratification of early breast cancer identifies drug targets to drive stratified medicine. NPJ Breast Cancer 3, 3. 10.1038/s41523-016-0003-5. [PubMed: 28649643]
6. Fraser M, Sabelnykova VY, Yamaguchi TN, Heisler LE, Livingstone J, Huang V, Shiah YJ, Yousif F, Lin X, Masella AP, et al. (2017). Genomic hallmarks of localized, non-indolent prostate cancer. Nature 541, 359–364. 10.1038/nature20788. [PubMed: 28068672]
7. Smallridge RC, and Copland JA (2010). Anaplastic thyroid carcinoma: pathogenesis and emerging therapies. Clin. Oncol. 22, 486–497. 10.1016/j.clon.2010.03.013.
8. Ranganath R, Shah MA, and Shah AR (2015). Anaplastic thyroid cancer. Curr. Opin. Endocrinol. Diabetes Obes. 22, 387–391. 10.1097/MED.0000000000000189. [PubMed: 26313900]
9. Smallridge RC, Ain KB, Asa SL, Bible KC, Brierley JD, Burman KD, Kebebew E, Lee NY, Nikiforov YE, Rosenthal MS, et al. (2012). American Thyroid Association guidelines for management of patients with anaplastic thyroid cancer. Thyroid 22, 1104–1139. 10.1089/thy.2012.0302. [PubMed: 23130564]
10. Nishiyama RH, Dunn EL, and Thompson NW (1972). Anaplastic spindle-cell and giant-cell tumors of the thyroid gland. Cancer 30, 113–127. 10.1002/1097-0142(197207)30:1<113::aid-cncr2820300118>3.0.co;2-e. [PubMed: 5040735]
11. Aldinger KA, Samaan NA, Ibanez M, and Hill CS Jr. (1978). Anaplastic carcinoma of the thyroid: a review of 84 cases of spindle and giant cell carcinoma of the thyroid. Cancer 41, 2267–2275. 10.1002/1097-0142(197806)41:6<2267::aid-cncr2820410627>3.0.co;2-7. [PubMed: 657091]
12. Spires JR, Schwartz MR, and Miller RH (1988). Anaplastic thyroid carcinoma. Association with differentiated thyroid cancer. Arch. Otolaryngol. Head Neck Surg. 114, 40–44. 10.1001/archotol.1988.01860130044012. [PubMed: 3334817]
13. Venkatesh YS, Ordonez NG, Schultz PN, Hickey RC, Goepfert H, and Samaan NA (1990). Anaplastic carcinoma of the thyroid. A clinicopathologic study of 121 cases. Cancer 66, 321–330. 10.1002/1097-0142(19900715)66:2<321::aid-cncr2820660221>3.0.co;2-a. [PubMed: 1695118]
14. Lam KY, Lo CY, Chan KW, and Wan KY (2000). Insular and anaplastic carcinoma of the thyroid: a 45-year comparative study at a single institution and a review of the significance of p53 and p21. Ann. Surg. 231, 329–338. 10.1097/00000658-200003000-00005. [PubMed: 10714625]
15. Kunstman JW, Juhlin CC, Goh G, Brown TC, Stenman A, Healy JM, Rubinstein JC, 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. 10.1093/hmg/ddu749. [PubMed: 25576899]
16. Sykorova V, Dvorakova S, Vcelak J, Vaclavikova E, Halkova T, Kodetova D, Lastuvka P, Betka J, Vlcek P, Reboun M, et al. (2015). Search for new genetic biomarkers in poorly differentiated and anaplastic thyroid carcinomas using next generation sequencing. Anticancer Res. 35, 2029–2036. [PubMed: 25862857]
17. Landa I, Ibrahimpasic T, Boucai L, Sinha R, Knauf JA, Shah RH, Dogan S, Ricarte-Filho JC, Krishnamoorthy GP, Xu B, et al. (2016). Genomic and transcriptomic hallmarks of poorly differentiated and anaplastic thyroid cancers. J. Clin. Invest. 126, 1052–1066. 10.1172/JCI85271. [PubMed: 26878173]

18. Jeon MJ, Chun SM, Kim D, Kwon H, Jang EK, Kim TY, Kim WB, Shong YK, Jang SJ, Song DE, and Kim WG (2016). Genomic Alterations of Anaplastic Thyroid Carcinoma Detected by Targeted Massive Parallel Sequencing in a BRAF(V600E) Mutation-Prevalent Area. Thyroid 26, 683–690. 10.1089/thy.2015.0506. [PubMed: 26980298]
19. Tiedje V, Ting S, Herold T, Synoracki S, Latteyer S, Moeller LC, Zwanziger D, Stuschke M, Fuehrer D, and Schmid KW (2017). NGS based identification of mutational hotspots for targeted therapy in anaplastic thyroid carcinoma. Oncotarget 8, 42613–42620. 10.18632/oncotarget.17300. [PubMed: 28489587]
20. Bonhomme B, Godbert Y, Perot G, Al Ghuzlan A, Bardet S, Belleannée G, Crinière L, Do Cao C, Fouilloux G, Guyetant S, et al. (2017). Molecular Pathology of Anaplastic Thyroid Carcinomas: A Retrospective Study of 144 Cases. Thyroid 27, 682–692. 10.1089/thy.2016.0254. [PubMed: 28351340]
21. Pozdeyev N, Gay LM, Sokol ES, Hartmaier R, Deaver KE, Davis S, French JD, Borre PV, LaBarbera DV, Tan AC, et al. (2018). Genetic Analysis of 779 Advanced Differentiated and Anaplastic Thyroid Cancers. Clin. Cancer Res. 24, 3059–3068. 10.1158/1078-0432.CCR-18-0373. [PubMed: 29615459]
22. Dong W, Nicolson NG, Choi J, Barbieri AL, Kunstman JW, Abou Azar S, Knight J, Bilguvar K, Mane SM, Lifton RP, et al. (2018). Clonal evolution analysis of paired anaplastic and well-differentiated thyroid carcinomas reveals shared common ancestor. Genes Chromosomes Cancer 57, 645–652. 10.1002/gcc.22678. [PubMed: 30136351]
23. Yoo SK, Song YS, Lee EK, Hwang J, Kim HH, Jung G, Kim YA, Kim SJ, Cho SW, Won JK, et al. (2019). Integrative analysis of genomic and transcriptomic characteristics associated with progression of aggressive thyroid cancer. Nat. Commun. 10, 2764. 10.1038/s41467-019-10680-5. [PubMed: 31235699]
24. Weinreb I, Piscuoglio S, Martelotto LG, Waggott D, Ng CKY, Perez-Ordonez B, Harding NJ, Alfaro J, Chu KC, Viale A, et al. (2014). Hotspot activating PRKD1 somatic mutations in polymorphous low-grade adenocarcinomas of the salivary glands. Nat. Genet. 46, 1166–1169. 10.1038/ng.3096. [PubMed: 25240283]
25. Cancer Genome Atlas Research Network (2014). Integrated genomic characterization of papillary thyroid carcinoma. Cell 159, 676–690. 10.1016/j.cell.2014.09.050. [PubMed: 25417114]
26. Kasaian K, Wiseman SM, Walker BA, Schein JE, Zhao Y, Hirst M, Moore RA, Mungall AJ, Marra MA, and Jones SJM (2015). The genomic and transcriptomic landscape of anaplastic thyroid cancer: implications for therapy. BMC Cancer 15, 984. 10.1186/s12885-015-1955-9. [PubMed: 26680454]
27. ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium (2020). Pan-cancer analysis of whole genomes. Nature 578, 82–93. 10.1038/s41586-020-1969-6. [PubMed: 32025007]
28. Mathe E, Olivier M, Kato S, Ishioka C, Hainaut P, and Tavtigian SV (2006). Computational approaches for predicting the biological effect of p53 missense mutations: a comparison of three sequence analysis based methods. Nucleic Acids Res. 34, 1317–1325. 10.1093/nar/gkj518. [PubMed: 16522644]
29. Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N, Boutselakis H, Cole CG, Creatore C, Dawson E, et al. (2019). COSMIC: the Catalogue Of Somatic Mutations In Cancer. Nucleic Acids Res. 47, D941–D947. 10.1093/nar/gky1015. [PubMed: 30371878]
30. Rahman N. (2014). Realizing the promise of cancer predisposition genes. Nature 505, 302–308. 10.1038/nature12981. [PubMed: 24429628]
31. Alexandrov LB, Nik-Zainal S, Wedge DC, Campbell PJ, and Stratton MR (2013). Deciphering signatures of mutational processes operative in human cancer. Cell Rep. 3, 246–259. 10.1016/j.celrep.2012.12.008. [PubMed: 23318258]
32. Alexandrov LB, Jones PH, Wedge DC, Sale JE, Campbell PJ, Nik-Zainal S, and Stratton MR (2015). Clock-like mutational processes in human somatic cells. Nat. Genet. 47, 1402–1407. 10.1038/ng.3441. [PubMed: 26551669]
33. Wreesmann VB, Ghossein RA, Patel SG, Harris CP, Schnaser EA, Shaha AR, Tuttle RM, Shah JP, Rao PH, and Singh B. (2002). Genome-wide appraisal of thyroid cancer progression. Am. J. Pathol. 161, 1549–1556. 10.1016/S0002-9440(10)64433-1. [PubMed: 12414503]

34. Lee JJ, Au AYM, Foukakis T, Barbaro M, Kiss N, Clifton-Bligh R, Staaf J, Borg A, Delbridge L, Robinson BG, et al. (2008). Array-CGH identifies cyclin D1 and UBCH10 amplicons in anaplastic thyroid carcinoma. Endocr. Relat. Cancer 15, 801–815. 10.1677/ERC-08-0018. [PubMed: 18753363]
35. Pallante P, Berlingieri MT, Troncone G, Kruhoffer M, Orntoft TF, Viglietto G, Caleo A, Migliaccio I, Decaussin-Petrucci M, Santoro M, et al. (2005). UbcH10 overexpression may represent a marker of anaplastic thyroid carcinomas. Br. J. Cancer 93, 464–471. 10.1038/sj.bjc.6602721. [PubMed: 16106252]
36. Gopal RK, Kubler K, Calvo SE, Polak P, Livitz D, Rosebrock D, Sadow PM, Campbell B, Donovan SE, Amin S, et al. (2018). Wide-spread Chromosomal Losses and Mitochondrial DNA Alterations as Genetic Drivers in Hurthle Cell Carcinoma. Cancer Cell 34, 242–255. 10.1016/j.ccell.2018.06.013. [PubMed: 30107175]
37. Ganly I, Makarov V, Deraje S, Dong Y, Reznik E, Seshan V, Nanjangud G, Eng S, Bose P, Kuo F, et al. (2018). Integrated Genomic Analysis of Hurthle Cell Cancer Reveals Oncogenic Drivers, Recurrent Mitochondrial Mutations, and Unique Chromosomal Landscapes. Cancer Cell 34, 256–270. 10.1016/j.ccell.2018.07.002. [PubMed: 30107176]
38. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, Speed D, Lynch AG, Samarajiwa S, Yuan Y, et al. (2012). The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature 486, 346–352. 10.1038/nature10983. [PubMed: 22522925]
39. Lalonde E, Ishkanian AS, Sykes J, Fraser M, Ross-Adams H, Erho N, Dunning MJ, Halim S, Lamb AD, Moon NC, et al. (2014). Tumour genomic and microenvironmental heterogeneity for integrated prediction of 5-year biochemical recurrence of prostate cancer: a retrospective cohort study. Lancet Oncol. 15, 1521–1532. 10.1016/S1470-2045(14)71021-6. [PubMed: 25456371]
40. Cancer Genome Atlas Network (2015). Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature 517, 576–582. 10.1038/nature14129. [PubMed: 25631445]
41. Gambaro K, Quinn MCJ, Wojnarowicz PM, Arcand SL, de Ladurantaye M, Barrès V, Ripeau JS, Killary AM, Davis EC, Lavoie J, et al. (2013). VGLL3 expression is associated with a tumor suppressor phenotype in epithelial ovarian cancer. Mol. Oncol. 7, 513–530. 10.1016/j.molonc.2012.12.006. [PubMed: 23415753]
42. Landa I, Pozdeyev N, Korch C, Marlow LA, Smallridge RC, Copland JA, Henderson YC, Lai SY, Clayman GL, Onoda N, et al. (2019). Comprehensive Genetic Characterization of Human Thyroid Cancer Cell Lines: A Validated Panel for Preclinical Studies. Clin. Cancer Res. 25, 3141–3151. 10.1158/1078-0432.CCR-18-2953. [PubMed: 30737244]
43. Zeng PYF, Cecchini MJ, Barrett JW, Shammas-Toma M, De Cecco L, Serafini MS, Cavalieri S, Licitra L, Hoebers F, Brakenhoff RH, et al. (2022). Immune-based classification of HPV-associated oropharyngeal cancer with implications for biomarker-driven treatment de-intensification. EBioMedicine 86, 104373. 10.1016/j.ebiom.2022.104373.
44. Nicorici D, Satalan M, Edgren H, Kangaspeska S, Murumagi A, Kallioniemi O, Virtanen S, and Kilkku O. (2014). FusionCatcher - a tool for finding somatic fusion genes in paired-end RNA-sequencing data. Preprint at bioRxiv. 10.1101/011650.
45. Xiao J, Yang W, Xu B, Zhu H, Zou J, Su C, Rong J, Wang T, and Chen Z. (2018). Expression of fibronectin in esophageal squamous cell carcinoma and its role in migration. BMC Cancer 18, 976. 10.1186/s12885-018-4850-3 . [PubMed: 30314454]
46. Liu J, Shen JX, Wu HT, Li XL, Wen XF, Du CW, and Zhang GJ (2018). Collagen 1A1 (COL1A1) promotes metastasis of breast cancer and is a potential therapeutic target. Discov. Med. 25, 211–223. [PubMed: 29906404]
47. Maniakas A, Dadu R, Busaidy NL, Wang JR, Ferrarotto R, Lu C, Williams MD, Gunn GB, Hofmann MC, Cote G, et al. (2020). Evaluation of Overall Survival in Patients With Anaplastic Thyroid Carcinoma, 2000–2019. JAMA Oncol. 6, 1397–1404. 10.1001/jamaoncol.2020.3362. [PubMed: 32761153]
48. Hieronymus H, Schultz N, Gopalan A, Carver BS, Chang MT, Xiao Y, Heguy A, Huberman K, Bernstein M, Assel M, et al. (2014). Copy number alteration burden predicts prostate cancer relapse. Proc. Natl. Acad. Sci. USA 111, 11139–11144. 10.1073/pnas.1411446111. [PubMed: 25024180]

49. Vollan HKM, Rueda OM, Chin SF, Curtis C, Turashvili G, Shah S, Lingjærde OC, Yuan Y, Ng CK, Dunning MJ, et al. (2015). A tumor DNA complex aberration index is an independent predictor of survival in breast and ovarian cancer. Mol. Oncol. 9, 115–127. 10.1016/j.molonc.2014.07.019. [PubMed: 25169931]
50. Nikiforova MN, Kimura ET, Gandhi M, Biddinger PW, Knauf JA, Basolo F, Zhu Z, Giannini R, Salvatore G, Fusco A, et al. (2003). BRAF mutations in thyroid tumors are restricted to papillary carcinomas and anaplastic or poorly differentiated carcinomas arising from papillary carcinomas. J. Clin. Endocrinol. Metab. 88, 5399–5404. 10.1210/jc.2003-030838. [PubMed: 14602780]
51. Gauchotte G, Philippe C, Lacomme S, Léotard B, Wissler MP, Allou L, Toussaint B, Klein M, Vignaud JM, and Bressenot A. (2011). BRAF, p53 and SOX2 in anaplastic thyroid carcinoma: evidence for multistep carcinogenesis. Pathology 43, 447–452. 10.1097/PAT.0b013e3283486178. [PubMed: 21716161]
52. Fagerberg L, Hallström BM, Oksvold P, Kampf C, Djureinovic D, Odeberg J, Habuka M, Tahmasebpoor S, Danielsson A, Edlund K, et al. (2014). Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. Mol. Cell. Proteomics 13, 397–406. 10.1074/mcp.M113.035600. [PubMed: 24309898]
53. Espiritu SMG, Liu LY, Rubanova Y, Bhandari V, Holgersen EM, Szyca LM, Fox NS, Chua MLK, Yamaguchi TN, Heisler LE, et al. (2018). The Evolutionary Landscape of Localized Prostate Cancers Drives Clinical Aggression. Cell 173, 1003–1013.e15. 10.1016/j.cell.2018.03.029. [PubMed: 29681457]
54. Salcedo A, Tarabichi M, Espiritu SMG, Deshwar AG, David M, Wilson NM, Dentro S, Wintersinger JA, Liu LY, Ko M, et al. (2020). A community effort to create standards for evaluating tumor subclonal reconstruction. Nat. Biotechnol. 38, 97–107. 10.1038/s41587-019-0364-z. [PubMed: 31919445]
55. Hay ID, Bergstralh EJ, Goellner JR, Ebersold JR, and Grant CS (1993). Predicting outcome in papillary thyroid carcinoma: development of a reliable prognostic scoring system in a cohort of 1779 patients surgically treated at one institution during 1940 through 1989. Surgery 114, 1050–1058. [PubMed: 8256208]
56. D'Avanzo A, Treseler P, Ituarte PHG, Wong M, Streja L, Greenspan FS, Siperstein AE, Duh QY, and Clark OH (2004). Follicular thyroid carcinoma: histology and prognosis. Cancer 100, 1123–1129. 10.1002/cncr.20081. [PubMed: 15022277]
57. LiVolsi VA (2011). Papillary thyroid carcinoma: an update. Mod. Pathol. 24, S1–S9. 10.1038/modpathol.2010.129. [PubMed: 21455196]
58. Carling T, and Udelsman R. (2014). Thyroid cancer. Annu. Rev. Med. 65, 125–137. 10.1146/annurev-med-061512-105739. [PubMed: 24274180]
59. Krishnamoorthy GP, Davidson NR, Leach SD, Zhao Z, Lowe SW, Lee G, Landa I, Nagarajah J, Saqcena M, Singh K, et al. (2019). EIF1AX and RAS Mutations Cooperate to Drive Thyroid Tumorigenesis through ATF4 and c-MYC. Cancer Discov. 9, 264–281. 10.1158/2159-8290.CD-18-0606. [PubMed: 30305285]
60. Pinto N, Ruicci KM, Khan MI, Shaikh MH, Zeng YFP, Yoo J, Fung K, MacNeil SD, Mendez A, Mymryk JS, et al. (2022). Introduction and expression of PIK3CA(E545K) in a papillary thyroid cancer BRAF(V600E) cell line leads to a dedifferentiated aggressive phenotype. J Otolaryngol Head Neck Surg 51, 7. 10.1186/s40463-022-00558-w. [PubMed: 35193694]
61. Saqcena M, Leandro-Garcia LJ, Maag JLV, Tchekmedyian V, Krishnamoorthy GP, Tamarapu PP, Tiedje V, Reuter V, Knauf JA, de Stanchina E, et al. (2021). SWI/SNF Complex Mutations Promote Thyroid Tumor Progression and Insensitivity to Redifferentiation Therapies. Cancer Discov. 11, 1158–1175. 10.1158/2159-8290.CD-20-0735. [PubMed: 33318036]
62. Lu L, Wang JR, Henderson YC, Bai S, Yang J, Hu M, Shiau CK, Pan T, Yan Y, Tran TM, et al. (2023). Anaplastic transformation in thyroid cancer revealed by single cell transcriptomics. J. Clin. Invest. 133. 10.1172/JCI169653.
63. Pu W, Shi X, Yu P, Zhang M, Liu Z, Tan L, Han P, Wang Y, Ji D, Gan H, et al. (2021). Single-cell transcriptomic analysis of the tumor ecosystems underlying initiation and progression of papillary thyroid carcinoma. Nat. Commun. 12, 6058. 10.1038/s41467-021-26343-3. [PubMed: 34663816]

64. Luo H, Xia X, Kim GD, Liu Y, Xue Z, Zhang L, Shu Y, Yang T, Chen Y, Zhang S, et al. (2021). Characterizing dedifferentiation of thyroid cancer by integrated analysis. Sci. Adv. 7, eabf3657. 10.1126/sciadv.abf3657.

65. Hussain M, Mateo J, Fizazi K, Saad F, Shore N, Sandhu S, Chi KN, Sartor O, Agarwal N, Olmos D, et al. (2020). Survival with Olaparib in Metastatic Castration-Resistant Prostate Cancer. N. Engl. J. Med. 383, 2345–2357. 10.1056/NEJMoa2022485. [PubMed: 32955174]

66. Robson M, Im SA, Senkus E, Xu B, Domchek SM, Masuda N, Delaloge S, Li W, Tung N, Armstrong A, et al. (2017). Olaparib for Metastatic Breast Cancer in Patients with a Germline BRCA Mutation. N. Engl. J. Med. 377, 523–533. 10.1056/NEJMoa1706450. [PubMed: 28578601]

67. Tutt ANJ, Garber JE, Kaufman B, Viale G, Fumagalli D, Rastogi P, Gelber RD, de Azambuja E, Fielding A, Balmaña J, et al. (2021). Adjuvant Olaparib for Patients with BRCA1- or BRCA2-Mutated Breast Cancer. N. Engl. J. Med. 384, 2394–2405. 10.1056/NEJMoa2105215. [PubMed: 34081848]

68. Harvey KF, Zhang X, and Thomas DM (2013). The Hippo pathway and human cancer. Nat. Rev. Cancer 13, 246–257. 10.1038/nrc3458. [PubMed: 23467301]

69. Welcker D, Stein C, Feitosa NM, Armistead J, Zhang JL, Lütke S, Kleinridders A, Brüning JC, Eming SA, Sengle G, et al. (2021). Hemicentin-1 is an essential extracellular matrix component of the dermal-epidermal and myotendinous junctions. Sci. Rep. 11, 17926. 10.1038/s41598-021-96824-4. [PubMed: 34504132]

70. Vogel BE, Muriel JM, Dong C, and Xu X. (2006). Hemicentins: what have we learned from worms? Cell Res. 16, 872–878. 10.1038/sj.cr.7310100. [PubMed: 17031392]

71. Vogel BE, and Hedgecock EM (2001). Hemicentin, a conserved extracellular member of the immunoglobulin superfamily, organizes epithelial and other cell attachments into oriented line-shaped junctions. Development 128, 883–894. 10.1242/dev.128.6.883. [PubMed: 11222143]

72. Liu CL, Pan HW, Torng PL, Fan MH, and Mao TL (2019). SRPX and HMCN1 regulate cancer-associated fibroblasts to promote the invasiveness of ovarian carcinoma. Oncol. Rep. 42, 2706–2715. 10.3892/or.2019.7379. [PubMed: 31638245]

73. Lee SH, Je EM, Yoo NJ, and Lee SH (2015). HMCN1, a cell polarity-related gene, is somatically mutated in gastric and colorectal cancers. Pathol. Oncol. Res. 21, 847–848. 10.1007/s12253-014-9809-3. [PubMed: 24912920]

74. Kim T, Issa D, and Onyshchenko M. (2022). Analyzing TCGA Data to Identify Gene Mutations Linked to Hepatocellular Carcinoma in Asians. Gastrointest. Tumors 9, 43–58. 10.1159/000524576. [PubMed: 36590851]

75. Kikutake C, Yoshihara M, Sato T, Saito D, and Suyama M. (2018). Intratumor heterogeneity of HMCN1 mutant alleles associated with poor prognosis in patients with breast cancer. Oncotarget 9, 33337–33347. 10.18632/oncotarget.26071. [PubMed: 30279964]

76. Chen C, Shi C, Huang X, Zheng J, Zhu Z, Li Q, Qiu S, Huang Z, Zhuang Z, Wu R, et al. (2019). Molecular Profiles and Metastasis Markers in Chinese Patients with Gastric Carcinoma. Sci. Rep. 9, 13995. 10.1038/s41598-019-50171-7. [PubMed: 31570735]

77. Gubbiotti MA, Neill T, and Iozzo RV (2017). A current view of perlecan in physiology and pathology: A mosaic of functions. Matrix Biol. 57–58, 285–298. 10.1016/j.matbio.2016.09.003.

78. Mii Y, and Takada S. (2020). Heparan Sulfate Proteoglycan Clustering in Wnt Signaling and Dispersal. Front. Cell Dev. Biol. 8, 631. 10.3389/fcell.2020.00631. [PubMed: 32760727]

79. Kamimura K, Ueno K, Nakagawa J, Hamada R, Saitoe M, and Maeda N. (2013). Perlecan regulates bidirectional Wnt signaling at the Drosophila neuromuscular junction. J. Cell Biol. 200, 219–233. 10.1083/jcb.201207036. [PubMed: 23319599]

80. Grindel BJ, Martinez JR, Pennington CL, Muldoon M, Stave J, Chung LW, and Farach-Carson MC (2014). Matrilysin/matrix metalloproteinase-7(MMP7) cleavage of perlecan/HSPG2 creates a molecular switch to alter prostate cancer cell behavior. Matrix Biol. 36, 64–76. 10.1016/j.matbio.2014.04.005. [PubMed: 24833109]

81. Grindel BJ, Martinez JR, Tellman TV, Harrington DA, Zafar H, Nakhleh L, Chung LW, and Farach-Carson MC (2018). Matrilysin/MMP-7 Cleavage of Perlecan/HSPG2 Complexed with Semaphorin 3A Supports FAK-Mediated Stromal Invasion by Prostate Cancer Cells. Sci. Rep. 8, 7262. 10.1038/s41598-018-25435-3. [PubMed: 29740048]

82. Yoo SK, Song YS, Park YJ, and Seo JS (2020). Recent Improvements in Genomic and Transcriptomic Understanding of Anaplastic and Poorly Differentiated Thyroid Cancers. Endocrinol. Metab. 35, 44–54. 10.3803/EnM.2020.35.1.44.

83. Van Loo P, Nordgard SH, Lingjærde OC, Russnes HG, Rye IH, Sun W, Weigman VJ, Marynen P, Zetterberg A, Naume B, et al. (2010). Allele-specific copy number analysis of tumors. Proc. Natl. Acad. Sci. USA 107, 16910–16915. 10.1073/pnas.1009843107. [PubMed: 20837533]

84. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, and Getz G. (2011). GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol. 12, R41. 10.1186/gb-2011-12-4-r41. [PubMed: 21527027]

85. Wilkerson MD, and Hayes DN (2010). ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics 26, 1572–1573. 10.1093/bioinformatics/btq170. [PubMed: 20427518]

86. Li H, and Durbin R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26, 589–595. 10.1093/bioinformatics/btp698. [PubMed: 20080505]

87. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, and DePristo MA (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 1297–1303. 10.1101/gr.107524.110. [PubMed: 20644199]

88. Cibulskis K, McKenna A, Fennell T, Banks E, DePristo M, and Getz G. (2011). ContEst: estimating cross-contamination of human samples in next-generation sequencing data. Bioinformatics 27, 2601–2602. 10.1093/bioinformatics/btr446. [PubMed: 21803805]

89. Larson DE, Harris CC, Chen K, Koboldt DC, Abbott TE, Dooling DJ, Ley TJ, Mardis ER, Wilson RK, and Ding L. (2012). Somaticallies: identification of somatic point mutations in whole genome sequencing data. Bioinformatics 28, 311–317. 10.1093/bioinformatics/btr665. [PubMed: 22155872]

90. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, Gabriel S, Meyerson M, Lander ES, and Getz G. (2013). Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat. Biotechnol. 31, 213–219. 10.1038/nbt.2514. [PubMed: 23396013]

91. Ha G, Roth A, Khattra J, Ho J, Yap D, Prentice LM, Melnyk N, McPherson A, Bashashati A, Laks E, et al. (2014). TITAN: inference of copy number architectures in clonal cell populations from tumor whole-genome sequence data. Genome Res. 24, 1881–1893. 10.1101/gr.180281.114. [PubMed: 25060187]

92. Deshwar AG, Vembu S, Yung CK, Jang GH, Stein L, and Morris Q. (2015). PhyloWGS: reconstructing subclonal composition and evolution from whole-genome sequencing of tumors. Genome Biol. 16, 35. 10.1186/s13059-015-0602-8. [PubMed: 25786235]

93. Li B, and Dewey CN (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinf. 12, 323. 10.1186/1471-2105-12-323.

94. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. 10.1093/bioinformatics/bts635. [PubMed: 23104886]

95. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, Selves J, Laurent-Puig P, Sautès-Fridman C, Fridman WH, and de Reyniès A. (2016). Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol. 17, 218. 10.1186/s13059-016-1070-5. [PubMed: 27765066]

96. Gaujoux R, and Seoighe C. (2010). A flexible R package for nonnegative matrix factorization. BMC Bioinf. 11, 367. 10.1186/1471-2105-11-367.

97. Wang L, Yamaguchi S, Burstein MD, Terashima K, Chang K, Ng HK, Nakamura H, He Z, Doddapaneni H, Lewis L, et al. (2014). Novel somatic and germline mutations in intracranial germ cell tumours. Nature 511, 241–245. 10.1038/nature13296. [PubMed: 24896186]

98. Wang L, Ni X, Covington KR, Yang BY, Shiu J, Zhang X, Xi L, Meng Q, Langridge T, Drummond J, et al. (2015). Genomic profiling of Sezary syndrome identifies alterations of key T cell signaling and differentiation genes. Nat. Genet. 47, 1426–1434. 10.1038/ng.3444. [PubMed: 26551670]

99. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, and Ruden DM (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly 6, 80–92. 10.4161/fly.19695. [PubMed: 22728672]

100. Quinlan AR, and Hall IM (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842. 10.1093/bioinformatics/btq033. [PubMed: 20110278]

101. Boutros PC, Fraser M, Harding NJ, de Borja R, Trudel D, Lalonde E, Meng A, Hennings-Yeomans PH, McPherson A, Sabelnykova VY, et al. (2015). Spatial genomic heterogeneity within localized, multi-focal prostate cancer. Nat. Genet. 47, 736–745. 10.1038/ng.3315. [PubMed: 26005866]

102. Sorgini A, Kim HAJ, Zeng PYF, Shaikh MH, Mundi N, Ghasemi F, Di Gravio E, Khan H, MacNeil D, Khan MI, et al. (2020). Analysis of the TCGA Dataset Reveals that Subsites of Laryngeal Squamous Cell Carcinoma are Molecularly Distinct. Cancers 13, 105. 10.3390/cancers13010105. [PubMed: 33396315]

103. Kim HAJ, Zeng PYF, Sorgini A, Shaikh MH, Khan H, MacNeil D, Khan MI, Mendez A, Yoo J, Fung K, et al. (2022). Tumor molecular differences associated with outcome disparities of Black patients with head and neck cancer. Head Neck 44, 1124–1135. 10.1002/hed.27007. [PubMed: 35187756]

104. Kim HAJ, Zeng PYF, Shaikh MH, Mundi N, Ghasemi F, Di Gravio E, Khan H, MacNeil D, Khan MI, Patel K, et al. (2021). All HPV-negative head and neck cancers are not the same: Analysis of the TCGA dataset reveals that anatomical sites have distinct mutation, transcriptome, hypoxia, and tumor microenvironment profiles. Oral Oncol. 116, 105260. 10.1016/j.oraloncology.2021.105260.

105. Papadimitriou CH, and Steiglitz K. (1982). Combinatorial Optimization: Algorithms and Complexity (Prentice Hall).

106. (2014). Integrated genomic characterization of papillary thyroid carcinoma. Cell 159, 676–690. 10.1016/j.cell.2014.09.050S0092-8674(14)01238-0. [PubMed: 25417114]

Highlights

- Anaplastic thyroid carcinoma (ATC) is an aggressive cancer with poor treatment outcomes
- Large-scale multi-region genomics identify distinct mutational processes
- Subclonal reconstruction suggests ATC shares clonal origin with differentiated thyroid cancers

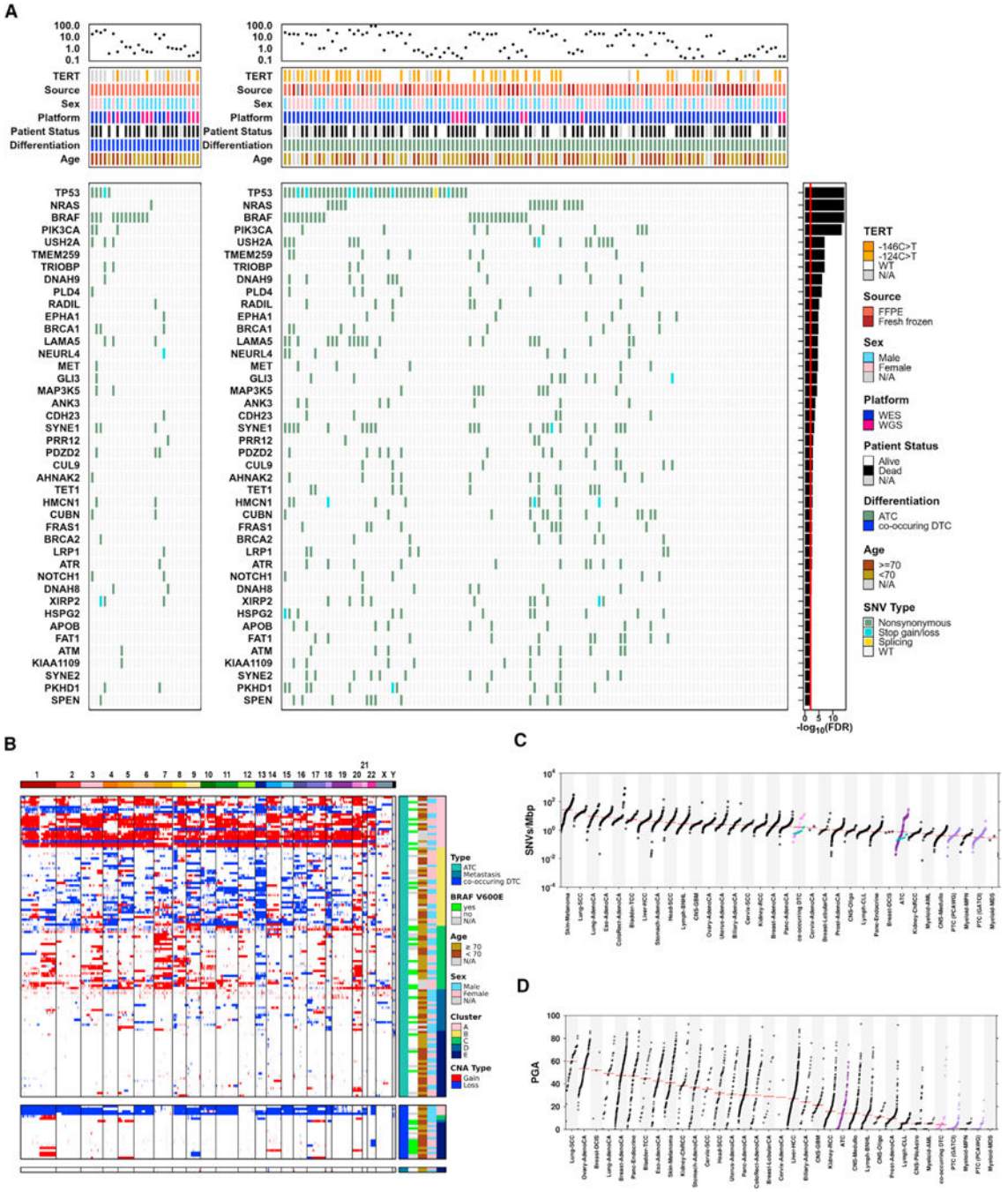


Figure 1. Somatic mutational landscape of anaplastic thyroid cancer

(A) Anaplastic thyroid cancers (ATCs) vary significantly in mutation density (single-nucleotide variants [SNVs] per million base pairs of DNA covered). Genes were selected based on SeqSig analysis p values (FDR < 0.05).

(B) Copy-number aberrations (CNAs) across ATC; consensus clustering was used to identify the optimal method and designation for sample groupings. Each row represents the CNA profile for a single undifferentiated tumor along the genome (chromosome 1 on the left

to chromosome Y on the right), ordered by percentage genome altered (PGA) within each subtype.

(C and D) Metrics of mutation density for ATC, co-occurring DTC, and papillary thyroid cancer (PTC) were compared with 32 additional tumor types available in the PCAWG (Pan-Cancer Analysis of Whole Genomes) dataset: (C) SNVs/Mb and (D) total PGA. Light purple indicates the PCAWG thyroid carcinoma cohort (primarily PTC) and similar GATCI PTC cohort; co-occurring DTC samples are shown in medium purple, while ATCs are in dark purple; for SNVs/Mb, x's indicate samples without a matched normal; these typically have higher than average rates. PGA for the GATCI PTC samples is similar to that of the PCAWG thyroid carcinoma cohort. Blue points for SNVs/Mb indicate results from WGS cohort. Red line indicates the median value. ATC shows higher PGA than either co-occurring DTC or PTC and a lower rate of point mutations than co-occurring DTC.

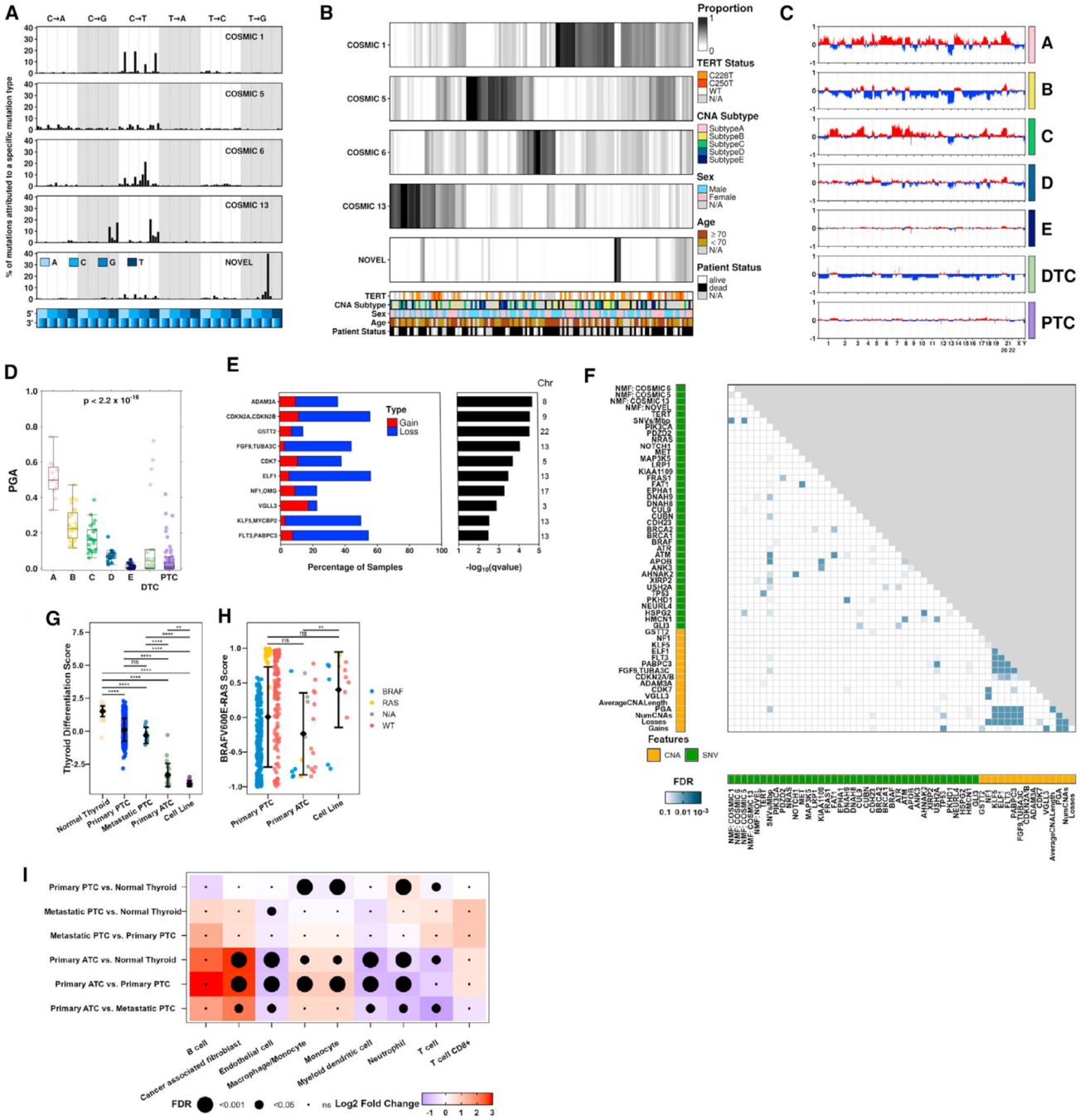


Figure 2. Genomic features of ATC and their associations with clinical features

(A) Non-negative matrix factorization (NMF) identified five trinucleotide signatures within ATC, four of which matched known COSMIC signatures and one that was a novel signature. Within each signature, the percentage of mutations within the cohort presenting each base change is shown (broken down by trinucleotide context).

(B) For each patient, the proportion of SNVs that contribute to each signature. There were no associations between these signatures and the covariates shown (Spearman's correlation $p < 0.1$).

(C and D) Average CNA profiles (C) and distribution of PGA (D) for samples from each ATC subtype, co-occurring DTC, and PTC, showing a significant difference between groups (one-way ANOVA, $p < 0.01$).

(E) GISTIC was used to identify recurrent CNAs within ATC.

(F) Pearson's χ^2, Spearman's correlation, Mann-Whitney U, or Kruskal-Wallis test was used to assess overlap across genomic features in ATC (SeqSig driver genes by SNV or CNAs, SNVs/Mb, PGA, trinucleotide signatures, CNA metrics); shading indicates FDR-adjusted p value.

(G–I) Comparison of thyroid differentiation score (G), BRAF-RAS scores (H), and estimated immune cell subsets (I) between the RNA-seq samples from ATC (n = 24) and cell lines (n = 13) from the current study and normal thyroid (n = 58), primary PTC (n = 502), and metastatic PTC (n = 8) from TCGA. Error bar represents 1 SD. The p values are from FDR-adjusted Mann-Whitney U test. All samples are biological replicates.

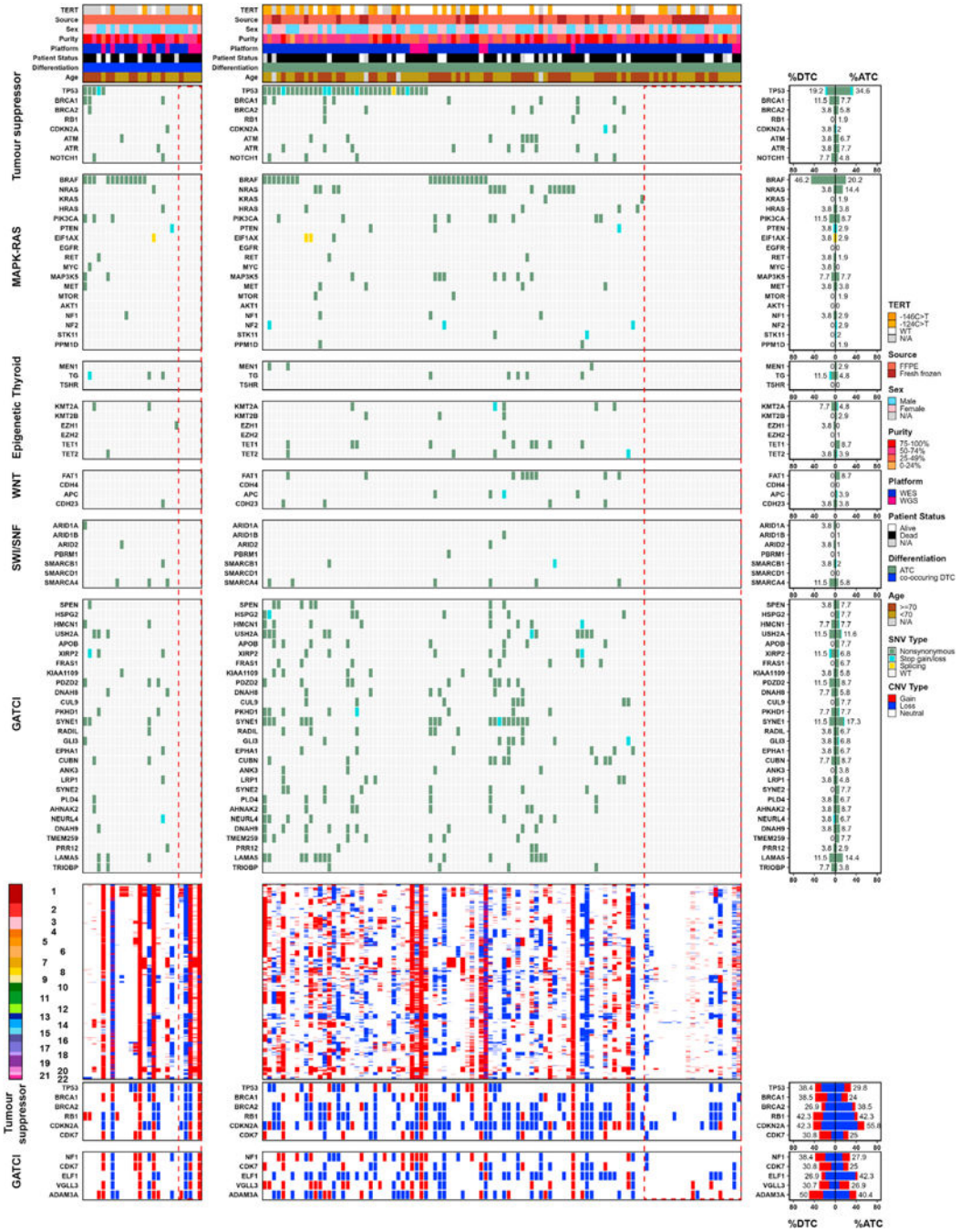


Figure 3. Genomic hallmarks of anaplastic thyroid cancer implicate diverse pathways

Single-nucleotide variants and copy-number alterations of co-occurring differentiated thyroid cancer and anaplastic thyroid cancers reveal recurrent alterations in gene sets such as tumor suppressors, MAPK-RAS, thyroid, epigenetic, WNT, SWI/SNF, and finally other identified GATCI genes (SeqSig FDR < 0.05). Red dashed box indicates samples with no known SNV alteration in the gene sets above.

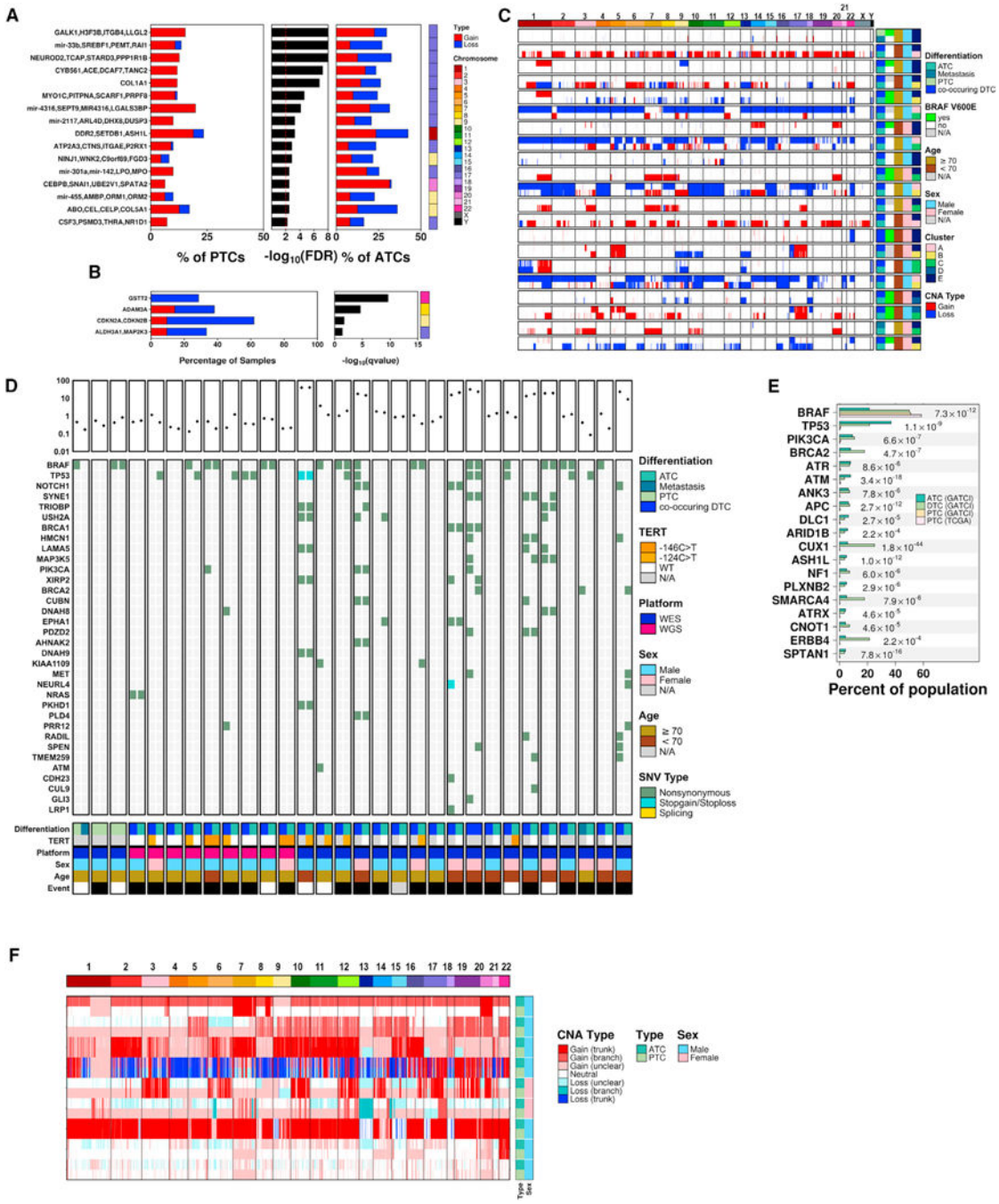


Figure 4. ATC and co-occurring DTC share clonal origins, but with distinct driver events

(A) GISTIC was used to identify recurrent CNAs within PTC; this identified multiple amplifications of chromosome 17 as being significant and recurrent compared with ATC.

(B) GISTIC was used to identify recurrent CNAs within co-occurring DTC. Copy-number losses of CDKN2A were more frequent in ATC than co-occurring well/poorly differentiated samples (DTC).

(C) Twenty-one patients had CNA profiles generated through copy-number arrays for multiple tumor regions. Each plot shows the genomic CNA profiles for a single patient,

with different tumor regions as rows. Covariates on the right indicate ATC subtype (where available), tumor type (ATC, co-occurring DTC, or metastasis), patient sex, patient age at diagnosis, and *BRAF* V600E status.

(D) Thirty patients had WES or WGS in multiple regions. Each subplot shows the mutational profile of the SeqSig genes (FDR < 0.05) in the paired samples.

(E) Gene-wise mutation frequencies were contrasted between datasets (GATCI-ATC, GATCI-DTC, GATCI-PTC, and TCGA-PTC) to identify candidate drivers of tumor progression. Genes with a statistically significant difference (proportion-test) between ATC and TCGA-PTC were selected and further filtered to show only known driver genes (defined using COSMIC or Candidate Cancer Gene Database). FDR-adjusted p values from proportion-tests across the three tumor types (PTC [combined], co-occurring DTC, ATC) are shown. Results for pairwise comparisons are available in Table S11.

(F) Landscape of CNAs determined by WGS for patients with both ATC and paired co-occurring DTC, along with matched normal tissue. CNAs are color-coded as to their origin (trunk, branch, or unclear). All samples are biological replicates.

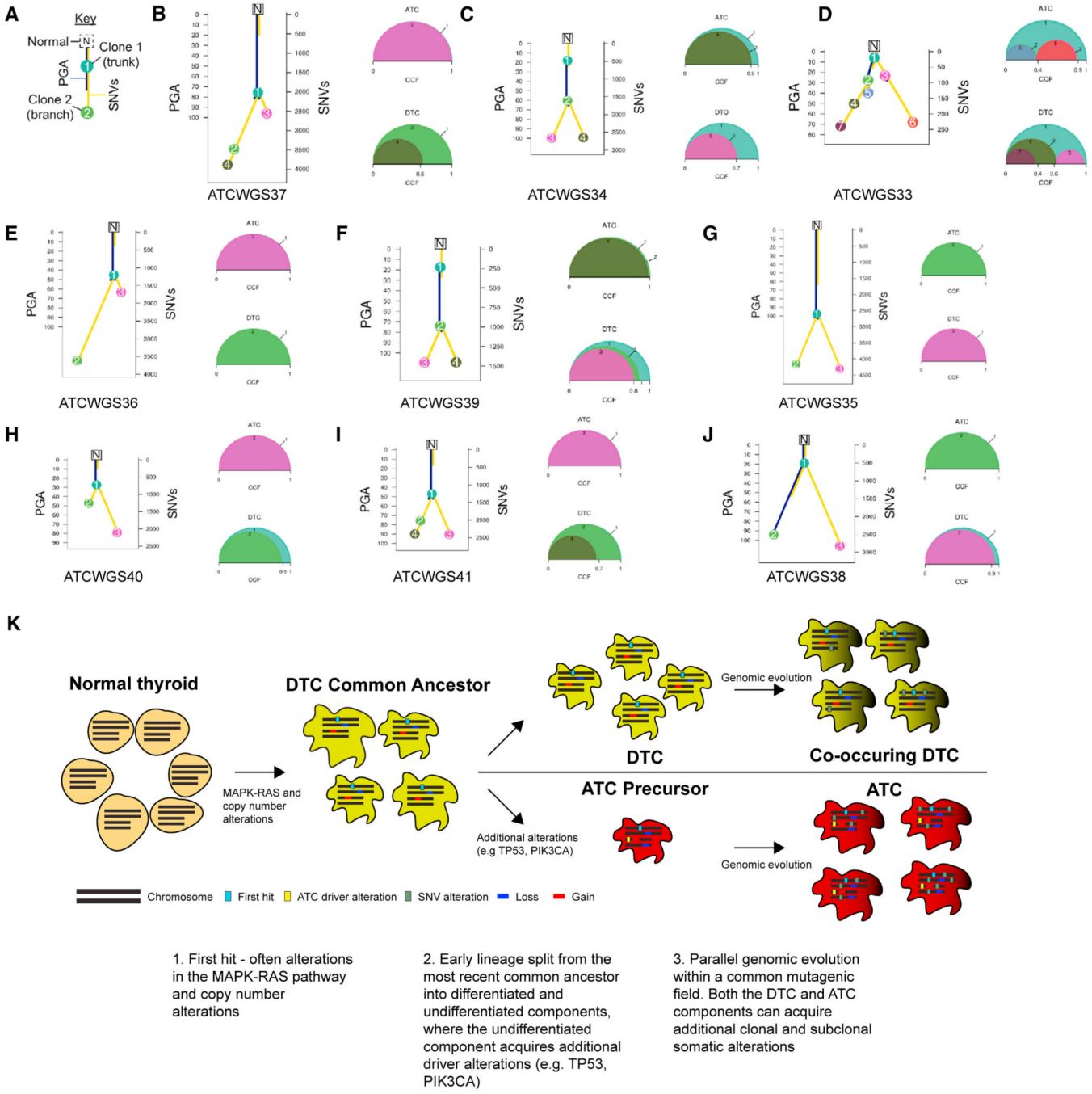


Figure 5. ATC and DTC evolve in a mutagenic field

(A) Key for interpreting plots.

(B–J) Subclonal reconstruction of the nine patients profiled using WGS. Left: changes in percentage genome altered (PGA) are shown by blue lines and correspond to the left axis, while accumulation of SNVs is shown by gold lines and corresponds to the right axis. Right: representation of the total cancer cell fraction (CCF) for each distinct subclone for either the ATC (top) or the co-occurring DTC (bottom) component. Interestingly, ATCWGS33 (D) showed outgrowth of two distinct co-mixed lineages from a mutagenic field.

(K) Proposed model of thyroid cancer progression.

Table 1.

Summary of phylogenetic reconstruction

	SNVs				PGA				CNAs			
	Trunk	**C**	**ATC**	**DTC**	**Trunk**	**C**	**ATC**	**DTC**	**Trunk**	**C**	**ATC**	**DTC**
ATCWGS33	3	51	148	104	0.00	12.35	0.06	0.00	0	241	4	0
ATCWGS34	350	26	857	871	0.00	32.40	0.00	0.00	0	437	1	0
ATCWGS35	1,686	0	1,420	1,575	92.64	0.00	0.05	0.32	604	0	14	37
ATCWGS36	364	0	262	2,347	40.10	0.00	0.20	1.45	731	0	49	4
ATCWGS37	517	0	316	1,486	70.75	0.00	0.60	1.27	1,207	0	6	3
ATCWGS38	302	0	852	2,245	13.87	0.00	71.38	0.00	150	0	360	0
ATCWGS39	166	58	376	380	0.00	46.30	0.02	0.00	0	1,194	14	1
ATCWGS40	250	0	1,237	309	22.13	0.00	0.00	0.00	238	0	0	0
ATCWGS41	437	0	979	703	42.04	0.00	0.00	1.25	917	0	3	123

Multisample subclonal reconstruction was performed on nine samples with paired regions of ATC and co-occurring DTC tumor. SNVs, single-nucleotide variants; PGA, percentage genome altered; CNAs, copy-number aberrations; C, common; ATC, anaplastic thyroid cancer; DTC, differentiated thyroid cancer.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
GATCI Cohort	This study	N/A
Chemicals, peptides, and recombinant proteins		
Betaine solution	Sigma	B0300-1VL
Deparaffinization Solution	Qiagen	19093
Q5 DNA Polymerase	NEB	M0491L
Critical commercial assays		
QIAamp DNA FFPE Tissue Kit	Qiagen	56404
RNeasy FFPE Kit	Qiagen	73504
AllPrep DNA/RNA Mini Kit	Qiagen	80204
QIAshredder	Qiagen	79656
Monarch DNA Gel Extraction Kit	NEB	T1020L
Deposited data		
GATCI Genomic Data	This study	EGA: EGAS00001002234
TCGA-THCA		https://portal.gdc.cancer.gov/projects/TCGA-THCA
ATC WES	Kunstman et al.[15]	Kunstman et al.[15]
PDTC and ATC WES	Landa et al.[17]	Landa et al.[17]
ATC and PDTC WES	Yoo et al.[23]	Yoo et al.[23]
PDTC and ATC targeted sequencing	Pozdeyev et al.[21]	Pozdeyev et al.[21]
Experimental models: Cell lines		
ASH3	JCRB	N/A
BHT101	DSMZ	N/A
C643	University of Uppsala	N/A
KMH2	JCRB	N/A
SW1736	University of Uppsala	N/A
THJ-11T	Mayo Clinic	N/A
THJ-16T	Mayo Clinic	N/A
THJ-21T	Mayo Clinic	N/A
THJ-29T	Mayo Clinic	N/A
UHTh7	University of Uppsala	N/A
UHTh74 clone7	University of Uppsala	N/A
CAL62	DSMZ	N/A
8505C	JCRB	N/A
Oligonucleotides		

REAGENT or RESOURCE	SOURCE	IDENTIFIER
hTERT F	IDT	5'-ACGAACGTGGCCAGCGGCAG-3'
hTERT R	IDT	5'-CTGGCGTCCCTGCACCCTGG-3'
TERT5 F	IDT	5'-CACCCGTCCTGCCCCTTCACCTT-3'
TERT5 R	IDT	5'-GGCTTCCCACGTGCGCAGCAGGA-3'
Software and algorithms		
ASCAT (v2.5)	Van Loo et al.[83]	https://github.com/VanLoo-lab/ascat
liftover tool (v359)		https://hgdownload.soe.ucsc.edu/admin/exe/linux.x86_64/
GISTIC2.0 (v2.0.22)	Mermel et al.[84]	https://github.com/broadinstitute/gistic2
ConsensusClusterPlus (v1.8.1)	Wilkerson and Hayes[85]	https://bioconductor.org/packages/release/bioc/html/ConsensusClusterPlus.html
VennDiagram package (v1.6.17)		https://cran.r-project.org/web/packages/VennDiagram/index.html
BPG (v5.8.8)		https://cran.r-project.org/web/packages/BoutrosLab.plotting.general/index.html
lattice (v0.20-35)		https://cran.r-project.org/web/packages/lattice/index.html
latticeExtra (v0.6-28)		https://cran.r-project.org/web/packages/latticeExtra/index.html
Picard tools (v1.121)		https://github.com/broadinstitute/picard
BWA-mem (v0.7.15)	Li and Durbin[86]	https://github.com/lh3/bwa
GATK (v3.5.0)	McKenna et.al.[87]	https://github.com/broadinstitute/gatk
HaplotypeCaller (v3.5.0)	McKenna et.al.[87]	https://github.com/broadinstitute/gatk
BEDTools (v2.18.2)		https://github.com/arq5x/bedtools2
ContEst (v1.0.24530)	Cibulskis[88]	https://github.com/broadinstitute/gatk
SomaticSniper (v1.0.5.0)	Larson et al.[89]	https://github.com/genome/somatic-sniper
MuTect (v1.1.7)	Cibulskis[90]	https://github.com/broadinstitute/mutect
TITAN (v1.20.1)	Ha et al.[91]	https://github.com/gavinha/TitanCNA
PhyloWGS (v1.5)	Deshwar[92]	https://github.com/morrislab/phylowgs
RSEM (v1.3.0)	Li and Dewey[93]	https://deweylab.github.io/RSEM/
STAR (v2.3.3a)	Dobin[94]	https://github.com/alexdobin/STAR
fusioncatcher (v0.99.7c)	Nicorici[44]	https://github.com/ndaniel/fusioncatcher
MCPCounter (v1.1.0)	Becht et al.[95]	https://github.com/ebecht/MCPcounter
NMF package (v0.20.6)	Gaujoux and Seoighe[96]	https://cran.r-project.org/web/packages/NMF/index.html
R (≥ v3.2.3)	The R Project	https://www.r-project.org/