Thyroid cancer under the scope of emerging technologies

Maxime Tarabichi¹, Pieter Demetter², Ligia Craciun², Carine Maenhaut¹, Vincent Detours¹

¹ Institute of Interdisciplinary Research (IRIBHM), Université Libre de Bruxelles, Brussels, Belgium.

² Department of Pathology, Institut Jules Bordet, Université Libre de Bruxelles, Brussels, Belgium.

corresponding authors: cmaenhau@ulb.ac.be, maxime.tarabichi@ulb.ac.be, vdetours@ulb.ac.be

Keywords (6/6): thyroid cancer; single cell omics; spatial genomics; artificial intelligence; long-read sequencing; immunotherapy
Abstract

Introduction

Thyrocyte-derived tumors

Papillary thyroid carcinoma (PTC)
  Histology and clinical features
  Genetic alterations
  Altered signaling pathways

Follicular thyroid carcinoma (FTC)
  Histology and clinical features
  Genetic alterations
  Altered signaling pathways

Anaplastic thyroid carcinoma (ATC)
  Histology and clinical features
  Genetic alterations
  Altered signaling pathways

Cellular heterogeneity in thyroid cancer

Pending issues and new technologies that could address them

Prevention - understanding the origins of thyroid cancer
  Screening and overdiagnosis
  Disease classification - ambiguous definitions and heterogeneity
  Limited treatment options in advanced cancers
  Treatment monitoring and follow-up

Conclusion

Acknowledgements

Figures and table with legends

References
Abstract

The vast majority of thyroid cancers originate from follicular cells. We outline outstanding issues at each step along the path of cancer patient care, from prevention to post-treatment follow-up and highlight how emerging technologies will help address them in the coming years. Three directions will dominate the coming technological landscape. Genomics will reveal tumoral evolutionary history and shed light on how these cancers arise from the normal epithelium and the genomics alteration driving their progression. Transcriptomics will gain cellular and spatial resolution providing a full account of intra-tumor heterogeneity and opening a window on the microenvironment supporting thyroid tumor growth. Artificial intelligence will set morphological analysis on an objective quantitative ground setting the foundations of a systematic thyroid tumor classification system. It will also integrate into unified representations the molecular and morphological perspectives on thyroid cancer.
Introduction

Tumors originating from thyroid follicular cells are the most frequent endocrine tumors, with an increasing incidence that make this cancer one of the most common in women (Cabanillas, McFadden, and Durante 2016; Deng et al. 2020). Large variations in this incidence have been reported depending on numerous factors such as sex, nationality, or socioeconomic status. Thyroid tumors comprise a spectrum of morphological phenotypes with variable rates of growth, differentiation, biological aggressiveness, and evolving definitions. Distinct types of tumors are associated with different genetic alterations in various signaling pathways. Recently, major advances have deepened the characterization of the morphological, mutational, transcriptional, and genomic profiles of the different types of thyroid tumors, opening new diagnostic, prognostic and therapeutic avenues for these cancers.

In this perspective, we will first describe the main types of thyroid cancers originating from thyrocytes, their clinical presentation, known genetic and pathway alterations, and then discuss new technological advances that will enable researchers to address long-standing questions at all stages of patient care from prevention to post-treatment follow-up.

Our focus on the investigation of clinical samples should not overshadow specific technologies, not covered here, that open new possibilities regarding in vitro and in vivo models.

Thyrocyte-derived tumors

Thyroid tumors occur more frequently in females (2-4 times) than in males and are rare in children. A large proportion of the global increase in thyroid cancer has been attributed to the improvement of the screening techniques such as ultrasonography and other diagnostic methods, as well as to an increase in the detection of small thyroid papillary cancers, less than 1 cm (M. Li, Dal Maso, and Vaccarella 2020). Hence, the contribution of overdiagnosis to the increasing incidence of thyroid cancer is significant. However, more recent studies have suggested that environmental exposures, lifestyle and comorbidities may also contribute to thyroid cancer development. A nationwide Danish study showed that the presence of other benign thyroid conditions were significantly associated with the risk of differentiated thyroid cancer, which could not be attributed to increased screening (Kitahara et al. 2018), but the underlying biological mechanisms remain to be elucidated.
To date, exposure to ionizing radiation has been clearly demonstrated to favor the onset of thyroid cancers (Iglesias et al. 2017), and other factors such as an iodine-deficient diet (Zimmermann and Galetti 2015), being overweight or environmental pollutants such as nitrates, heavy metals and other compounds may also be responsible (Vigneri et al. 2017). Although incidence is rising steadily, mortality has remained low and stable, affecting only about 5% of patients. The main risk associated with these tumors is the persistence and recurrence of the disease which occurs in 10 to 30% of cases of differentiated tumors (Durante et al. 2013).

Thyroid tumors may originate from follicular thyroid cells or parafollicular cells (C cells), which produce the hormone calcitonin. Parafollicular C cells-derived carcinomas are called medullary thyroid carcinoma and account for 2-4% of thyroid cancers. Follicular thyroid cells-derived tumors are more frequent and include benign and malignant forms with distinct histological presentation (Figure 1), potentially amenable to deep-neural-network characterisation, which we will cover here. Autonomously hyperfunctioning and cold follicular adenoma are two types of benign encapsulated tumors, characterized by their respectively high and low capacity to take up iodide and produce thyroid hormones. Malignant carcinomas include follicular thyroid carcinoma (FTC, 10-15%), papillary thyroid carcinoma (PTC, >80 %), poorly differentiated thyroid carcinoma (PDTC, 1-2%) and anaplastic thyroid carcinoma (ATC, 1-2%). The vast majority of thyroid cancers are sporadic. While PTCs and FTCs have a relatively good prognosis and can mostly be treated with surgery and I\textsubscript{131}, ATCs are lethal within months after diagnosis and do not respond to any conventional cancer therapies, i.e. surgery, chemotherapy, I\textsubscript{131}.

### Papillary thyroid carcinoma (PTC)

#### Histology and clinical features

From a histopathological point of view, PTCs are generally characterized by a papillae architecture with a progressive disappearance of the follicular structure; cells are organized around a fibrous structure that is sometimes vascularized. The lesion infiltrates the neighboring normal tissue, invasion of this lesion by lymphocytes is frequent and the presence of a fibrous capsule, partially surrounding the tumor, is rarely observed. The cell nuclei of the thyrocytes are enlarged with a typical ground glass and irregular appearance, i.e. presence of nuclear grooves. Apart from the classic variant which remains the most common, there are many histological variants, the most frequent being microcarcinoma (≤ 1 cm) and the follicular variant composed essentially of follicles with classic papillary nuclear features (J. Liu et al. 2006). Although studies
have associated the follicular variant with a mildly aggressive clinical phenotype (Agrawal et al., 2014), it is still generally considered to have the same prognosis as typical papillary carcinoma (Passler et al. 2003; Zidan et al. 2003). Other variants include the tall cell variant, the most common aggressive variant of PTC, characterized by tall cells (2 - 3 times taller than wide), which usually occurs in older age patients (Asa 2019). PTCs mainly metastasize in the lymph nodes adjacent to the original site of the tumor and can also form distant metastases in lungs and bones. Cervical lymph node metastases are present in nearly 50% of patients, with a frequency increasing with the size and the extrathyroidal extension of the primary tumor. Although PTC is generally indolent and curable, disease recurrence is common and is associated with increased mortality (Ito et al. 2007; Zaydfudim et al. 2008).

Genetic alterations

PTCs have a lower mutation rate relative to other cancer types (Alexandrov et al. 2013), providing increased signal-to-noise ratio for the identification of the initiating or causative mutations of the disease. Their genetics and molecular phenotypes have been extensively characterized by The Cancer Genome Atlas (TCGA): potential driver mutations could be identified for 96.5% of 496 PTCs (Cancer Genome Atlas Research Network 2014). The majority of PTCs present three genetic alterations which are mutually exclusive: activating point mutations in BRAF and RAS, and RET/PTC chromosomal rearrangements.

(i) BRAF mutations: Activating point mutations in the kinase region of BRAF, a member of the Raf family, mimic phosphorylation of the protein, leading to constitutive kinase activity (Davies et al., 2002). The most common of these mutations is a T1799A missense mutation that results in a valine to glutamic acid substitution at amino acid 600 (V600E). In PTC, BRAF mutations are found with a prevalence of ~60%, which constitutes the most common molecular alteration in this type of tumor (Cancer Genome Atlas Research Network 2014). Although the V600E substitution is the more frequent mutation, other less frequent alterations leading to constitutive BRAF have been described: K601E substitution has been detected in the follicular variant of PTC (Trovisco et al. 2004), and small in-frame insertions or deletions around V600 have been reported (Carta et al. 2006; Hou, Liu, and Xing 2007). Among PTCs, BRAF mutations appear more frequently in the classical and tall cell variants (Cancer Genome Atlas Research Network 2014). They are associated with older patients (Powell et al. 2005) and correlate with distant metastases and tumor aggressivity (Nikiforova et al. 2003; Xing et al. 2005; Y. Wang et al.
2008). They are also associated with low tumour purity, desmoplasia and immune infiltration (Tarabichi et al. 2018).

(ii) RAS mutations: Activating point mutations in codons 12, 13 and 61 of RAS inhibit the GTPase activity of this small G protein, maintaining it in a constitutively active, GTP-bound, state. There are three isoforms of RAS: HRAS, KRAS and NRAS, NRAS and HRAS being predominantly mutated in thyroid tumors, mostly involving codon 12 (G12V) and 61 (Q61R). RAS mutations are not specific of a particular type of thyroid tumors: they are present in 10 to 20% of PTCs and are mainly found in follicular variants (Cancer Genome Atlas Research Network 2014), whereas they are found in 20-40% of benign follicular adenoma and 40-50 % of FTCs (Xing 2013).

(iii) RET/PTC chromosomal rearrangements: in the past, these rearrangements were present in 15-20% of PTCs, essentially in the classical variant (Nikiforov and Nikiforova 2011), but their incidence is gradually decreasing to currently reach 6-7% (Cancer Genome Atlas Research Network 2014; Jung et al. 2014). There are several types of RET rearrangements found in PTCs, formed by the fusion of the intracellular tyrosine kinase domain of the RET receptor with different 5’ heterologous gene fragments which are ubiquitously expressed and possess a dimerization domain. This fusion results in the constitutive activation of the truncated tyrosine kinase portion of RET by autophosphorylation thanks to the dimerization domain of the heterologous gene. RET/PTC1 and RET/PTC3 are the most common combined forms of RET. RET/PTC1 is formed by a paracentric inversion of the long arm of chromosome 10 leading to fusion with a gene named CCDC6 (coiled-coil domain containing 6) (Grieco et al. 1990) RET/PTC3 is formed by fusion with the NCOA4 (nuclear receptor coactivator 4) gene (Santoro et al. 1994). Several other RET/PTC rearrangements have been identified and their development is mainly caused by exposure to ionizing radiation, for instance in PTCs that developed as a consequence of the radioactive fallout of the Chernobyl nuclear accident (Nikiforov 2002). They are also more frequent in children than in adults for both irradiated and non-irradiated PTCs (Romei, Ciampi, and Elisei 2016).

In addition to these main genetic alterations, mutations of the telomerase reverse transcriptase (TERT) promoter have been detected in around 10% of PTCs of all histological types (Cancer Genome Atlas Research Network 2014; Landa et al. 2013; X. Liu et al. 2013; Melo et al. 2014). This enzyme maintains telomere ends and substitutions at nucleotides 1295228 (C228T, C228A) and 1295250 (C250T) of chromosome 5, located in TERT promoter, lead to increased
TERT transcription. These mutations are associated with more aggressive tumors and are concomitant with other mutations such as BRAF^{V600E} or RAS, suggesting a role in the progression rather than in the initiation of the disease (Cancer Genome Atlas Research Network 2014; Melo et al. 2014).

Additional less frequent alterations, never exceeding a few percent of cases, have been described in PTC such as point mutations in EIF1AX, PPM1D, and CHEK2, and gene fusions involving BRAF, NTRK1, NTRK3 or ALK (Cancer Genome Atlas Research Network 2014). Gene fusions are more prevalent in radiation-induced thyroid cancer and are much lower in sporadic cases (Ricarte-Filho et al. 2013).

Altered signaling pathways

PTC progression is essentially driven by constitutive activation of the MAPK pathway, inducing proliferation and dedifferentiation of the thyrocytes (Roger et al. 1988; Zaballos and Santisteban 2017), as revealed by the downregulation of specific thyroid differentiation genes such as NIS, TPO, TG TSHR (Hébrant et al. 2012). This results from mutations or rearrangements in different effectors of this pathway, including BRAF, RAS or RET/PTC as described above, but also from increased expression of local growth factors and their receptors, leading to autocrine loops and autonomous proliferation. Examples include the overexpression of c-met/HGF receptor, sometimes in association with HGF (Di Renzo et al. 1992; Trovato et al. 1998), or overexpression of EGF receptor, TGFα and other genes coding for proteins involved in EGF pathway activation (Aasland et al. 1990; Delys et al. 2007; Lam et al. 2011). Based on their molecular profiles, two distinct molecular subtypes of PTC were identified, named BRAF-like and RAS-like, reflecting their biological differences, dedifferentiation level and aggressiveness. BRAF-like tumors signal preferentially through MAPK while RAS-like PTCs signal through both MAPK and PI3K (Cancer Genome Atlas Research Network 2014). These groups also differ in tumor purity and their microenvironment components (Tarabichi et al. 2018), which confound their molecular profiles at the bulk level.

Another signaling pathway that is activated in PTC is the NFκB pathway, promoting proliferation and inhibiting apoptosis (Cancer Genome Atlas Research Network 2014; Pacifico and Leonardi 2010). In this context, it has been shown that BRAF^{V600E} is able to activate NFκB (Bommarito et al. 2011; Palona et al. 2006).
Follicular thyroid carcinoma (FTC)

Histology and clinical features

Follicular thyroid carcinomas are also well-differentiated thyroid carcinoma. They represent 10 to 15% of thyroid cancers and are of good prognosis. Their incidence has decreased, most likely secondary to dietary iodine supplementation and the elimination of iodine deficiency, and the evolution of criteria for the diagnosis of the follicular variant of papillary cancer (see below). The majority of FTCs resemble benign follicular thyroid adenoma and the distinction between them is based on the presence of vascular and/or capsular invasion. Follicular carcinoma is divided into minimally invasive and invasive variants based on morphologic criteria. Metastases are present in 10-15% of the patients and spread hematogenously, mostly in lung and bones (McHenry and Phitayakorn 2011).

Genetic alterations

Due to their lower incidence, the genomic characterization of FTCs is less advanced than that of PTCs. They have been the subject of fewer high throughput sequencing studies (Jung et al. 2014; Nicolson et al. 2018; Swierniak et al. 2016). The genetic alterations that are most often encountered are present in both follicular adenoma and carcinoma, supporting the notion that a large part of FTCs derive from follicular thyroid adenoma (Dom et al. 2018). The most frequently encountered mutations are RAS activating mutations (predominantly N-RAS), present in 20-40% of follicular adenoma and 40-5% of FTCs, and PAX8-PPARγ rearrangements, identified in 10% of follicular adenoma and 30-40% of FTCs. PAX8-PPARγ rearrangement is a chromosomal translocation resulting in a chimeric protein containing the DNA binding domain of PAX8 and operating as a dominant negative on the transcriptional activity of wild-type PPARγ (Kroll et al. 2000). RAS activating mutations and PAX8-PPARγ rearrangements are generally mutually exclusive.

Most of the genetic alterations in FTCs involve proteins along the PI3K signaling pathway and include, in addition to RAS activating mutations, essentially PIK3CA copy number gain and activating mutations, and PTEN inactivating mutations or deletions (Hou et al. 2007; Z. Liu et al. 2008). AKT mutations have been described but are very rare. PIK3CA mutations and amplifications occur in 5-15% and 25-30% of FTCs, respectively. PTEN inactivating mutations
and deletions are respectively present in ~ 5 % and 30-40 % of FTCs (Hou et al. 2007; Z. Liu et al. 2008; Y. Wang et al. 2007).

Novel somatic alterations have recently been described (DICER1, EIF1AX, TSHR) but they appear quite heterogeneous and show low prevalence (Nicolson et al. 2018; Swierniak et al. 2016; Duan et al. 2019).

Altered signaling pathways

Altered signaling pathways have been poorly investigated in FTC. Those tumors are mainly characterized by a constitutive activation of the PI3K signaling pathway, resulting from different genetic alterations as described in the previous section (Zaballos and Santisteban 2017). They have many features in common with benign follicular adenomas, with quantitative rather than qualitative gene expression differences, suggesting a biological continuum (Dom et al. 2018).

Variants of follicular thyroid tumors are the Hürthle cell tumors that can be separated into Hürthle cell adenoma and carcinomas. Hürthle cell carcinomas (HCC), also called oxyphilic follicular carcinomas, are oncocytic tumors. They constitute a rare subgroup of thyroid tumors (~4%), more aggressive than non-oncocytic thyroid carcinomas. They show a microfolicular architecture and are characterized by cells with a large eosinophilic and granular cytoplasm, with loss of cell polarity and containing a high number of mitochondria (Kure and Ohashi 2020). These numerous mitochondria result from mutations in mitochondrial DNA, mostly in genes encoding subunits of complex I of the electron transport chain, leading to decreased oxidative phosphorylation (Kumari, Adewale, and Klubo-Gwiezdzinska 2020; Máximo et al. 2002). In addition, HCC show many chromosomal losses with genome-wide loss-of-heterozygosity, resulting in near-haploid cancer genomes (Corver et al. 2018) often rescued by whole-genome doubling, and mutations in nuclear DNA involving several genes from the MAPK and PI3K signaling pathways, such as RAS, PTEN, PIK3CA or TSC1/2 (Ganly et al. 2013; Gopal et al. 2018).

Although Hürthle cell tumors used to be classified as a variant of follicular thyroid tumors, their genetic profiles suggest they should be considered as a separate entity. Accordingly, in the latest WHO classification, they have been identified as a separate type of tumor derived from thyroid follicles.
Anaplastic thyroid carcinoma (ATC)

Histology and clinical features

Anaplastic thyroid carcinomas are very aggressive dedifferentiated tumors, representing less than 5% of all malignant thyroid tumors and with an average median survival < 6 months. No life-saving treatments are so far available. The follicular architecture is completely lost, and large areas of cell necrosis are visible; these tumors are largely vascularized. They are highly invasive and spread metastases locally in lymph nodes and distantly in lungs, bones, or brain. ATC represents the end stage of thyroid tumor progression and may derive from PTC or FTC or arise de novo. Evidence for that include clinicopathological observations: ATC often coexists with PTC or FTC in the same tumor and may arise in patients previously treated for differentiated thyroid cancer (Spires, Schwartz, and Miller 1988). In addition, ATCs share mutations with PTC and FTC (see below). At the bulk level, ATC show distinct mRNA expression profiles, while most genes dysregulated in PTC relative to normal tissues being further dysregulated in ATC (Hébrant et al. 2012). However, these bulk analyses likely reflect both differences in the microenvironment as well as in the cancer cells.

Genetic alterations

The genomic and transcriptional profiles of ATCs have recently been analyzed, revealing that these tumors have a higher tumor mutational burden than other types of thyroid cancer (Kunstman et al. 2015; Landa et al. 2016; Pozdeyev et al. 2018; Yoo et al. 2019). The predominant mutations occur in TP53 and TERTp genes. TP53 point mutations are present in >60-70% of the tumors. Similarly, TERT C228T and C250T mutations, which are present in less than 10% of PTCs, are found in >60-70% of ATCs. These mutations are detected in association with other mutations, for instance with BRAF or RAS, confirming their role in the progression of the disease and their relationship to the aggressive nature of the tumors that present them (Cancer Genome Atlas Research Network 2014; Landa et al. 2016; Melo et al. 2014; Yoo et al. 2019).

In addition, ATCs present genetic alterations found in PTCs and FTCs: \(^{BRAF^{V600E}}\) mutations are present in 45% of ATCs, and mutations in NRAS, HRAS, or KRAS in 24%. These mutations are mutually exclusive. Strikingly, the RET/PTC and PAX8/PPAR\(\gamma\) chromosomal rearrangements have so far not been found in ATCs. Other prevalent alterations include PIK3CA mutations.
(18%) or gene amplification (42%), and PTEN inactivating mutations, present in 15% of ATCs. Less frequent mutations have been observed in NF1/2 (neurofibromin 1/2), EIF1AX, ATM, Rb, AKT1, with a prevalence of <10%. These mutations generally co-occur with others (Landa et al. 2016; Pozdeyev et al. 2018; Yoo et al. 2019). Mutations in β-catenin (CTNNB1) have initially been reported in >60% of ATCs (Garcia-Rostan et al. 1999), but subsequent sequencing results did not replicate these findings: very few tumors carry β-catenin mutations (Kunstman et al. 2015; Landa et al. 2016; Pita et al. 2014). The mutational profile of ATCs thus suggests that a significant portion of them derive from differentiated carcinomas, but also that some initiating mutations of differentiated carcinomas (e.g. RET/PTC3) might lead to different evolutionary routes.

Altered signaling pathways

Whereas the MAPK pathway is involved in PTC development and the PI3K pathway in FTC development, ATCs are characterized by the constitutive activation of both MAPK and PI3K signaling pathways (Zaballos and Santisteban 2017). The progressive accumulation of different alterations activating both pathways drive the progression of differentiated thyroid carcinoma (PTC and FTC) to ATC. Decreased or loss of expression of most of the thyroid differentiation proteins (NIS, TPO, TG, TSHR, etc.) is generally observed. This corresponds to the in vivo biological features of ATCs, which unlike PTC, no longer respond to TSH stimulation, no longer secrete thyroglobulin nor accumulate iodide, and therefore do not respond to radioiodine therapy (Hébrant et al. 2012).

In addition to the activation of the MAPK and PI3K signaling pathways, the WNT-β-catenin signaling pathway is activated in ATC, resulting partly from β-catenin stabilizing mutations but also from the activation of the PI3K pathway, via GSK-3β phosphorylation and inactivation.

Pathways associated with the SWI/SNF chromatin remodeling complex and with histone modifications are deregulated in respectively 36% and 24% of ATCs, caused by mutations in different regulators of these pathways. Similarly, a deficiency in DNA MMR pathways has been described in 12% of ATCs, as a consequence of loss of function mutations in MMR genes (MLH1, MSH2, MSH6) (Landa et al. 2016; Pozdeyev et al. 2018).
Microenvironment and cellular heterogeneity in thyroid cancer

Histopathological analysis of tumors have revealed that they are composed of distinct cell types: in addition to the neoplastic cells, several stromal cell types are present including fibroblasts, myofibroblasts, endothelial cells, adipocytes, lymphocytes, macrophages, and other cells. These cells could either be part of the initial stroma, before the tumor developed, or have been recruited by cancer cells. So far, the contribution of the stromal cells to thyroid tumor growth has been little investigated. Whereas all thyroid tumors contain, apart from cancer cells, endothelial cells making up the blood vessels, the proportion of other cell types varies widely across cancer subtypes, patients and even within patients. Hence, expression of specific markers in bulk tumour tissues should be interpreted with caution, especially for tissues with low tumour purity.

The stroma of FTC has been little investigated, and the few existing data report the presence of immune cells (Yin et al. 2020). The stroma of PTC has been more characterized and is essentially composed of cancer-associated fibroblasts (CAFs) and lymphocytes. PTC cancer cells have been shown to produce several factors, for instance TGF-β1, a paracrine signal for stromal fibroblasts that enhances tumor growth (Zhang et al. 2014). A previous work of our group on BRAFV600E PTCs revealed a major contribution of the stromal fibroblasts to the tumor volume and its expansion, with highly proliferative CAFs (Tarabichi et al. 2018). Thyroid cancer cells and CAFs have also been shown to activate distinct metabolic pathways in PTC (Gill et al. 2016; A. Strickaert et al. 2017). We recently showed that cancer cells have an increased mitochondrial oxidative metabolism which is not observed in stromal cells (Aurélie Strickaert et al. 2019). Hence, our data support the existence of a reverse Warburg effect, characterized by the release into the microenvironment of lactate, produced by aerobic glycolysis by CAFs and available for tumor cells which metabolize it via the TCA cycle (Pavlides et al. 2009; Aurélie Strickaert et al. 2019). In addition to CAFs, PTCs are often infiltrated by immune cells, including lymphocytes, but also macrophages, mast cells, and neutrophils, which positively or negatively influence tumor progression (Ferrari et al. 2019; Xie et al. 2020).
ATCs are characterized by a strong reduction or total absence of lymphoid cells and fibroblasts but display a very dense network of interconnected ramified tumor-associated macrophages (TAM). These ATC-specific TAMs are not found in PTC or FTC and can represent more than 50% of the cells in the tumour mass (Caillou et al. 2011; Hébrant et al. 2012). It has been shown that these are M2 macrophages (Landa et al. 2016), and that higher density correlates with invasion and decreased survival (Ryder et al. 2008). As we have observed for CAFs in PTCs, the proliferation rates of TAMs and cancer cells are similar in ATC (Caillou et al. 2011). TAMs promote tumorigenesis in a paracrine manner by releasing different cytokines (Fang et al. 2014; Landa et al. 2016).

Altogether, stromal-epithelial heterotypic interactions clearly influence cancer cell behavior. However, early omics technologies performed on bulk tissues have shed little light on the mechanisms involved in shaping the cell type composition within the tumour microenvironment, e.g. fibroblasts and lymphocytes in PTC and macrophages in ATCs, and their individual contributions to tumor progression. Emerging technologies allowing to analyze tumors at single-cell resolution and within the spatial context of the tissue are ideally suited to address these issues.

Pending issues and new technologies that could address them

Prevention - understanding the origins of thyroid cancer

As many complex phenotypes, thyroid cancer likely has both genetic and environmental underlying determinants. Genome wide association studies (GWAS) have uncovered single nucleotide polymorphisms (SNPs) correlated with thyroid cancer (Burdett et al. n.d.), some of them are also associated with TSH levels (Zhou et al. 2020). However, most GWAS studies have been restricted to a subset of known variants and even current next-generation sequencing technologies based on short sequencing reads leave recurrent blind spots in the observed variation landscapes along the genome (Tarabichi et al. 2020). Moreover, while each human genome differs from each other at a few million SNP positions, they also differ by many more bases through structural variation, which are overlooked by SNP profiling platforms and most of which are missed through short-read sequencing (Ho, Urban, and Mills 2020).
Long-read third generation sequencing technologies have recently highlighted the complex landscape of structural variation between humans (Ebert et al. 2021). Such variation might also be implicated in defining different risk groups for thyroid cancer but it is so far unknown. While large scale studies using third-generation technologies are yet to become available, it is in theory possible to already “genotype” identified structural variants in short-read databases and link them to the phenotype post-hoc, which might elucidate some of the known SNP associations with thyroid cancer, as these might simply be in linkage disequilibrium with more impactful causative structural variants (Audano et al. 2019).

Besides genetic factors, most of which have low penetrance, iodine deficiency (Zimmermann and Galetti 2015) and radiation (reviewed in (Iglesias et al. 2017)) are the only established factors contributing to thyroid cancer initiation. There are several emerging environmental factors potentially related to thyroid cancers (Fiore et al. 2019) and there might be geographical increase in incidence related to increased exposure to those risk factors (Kim, Gosnell, and Roman 2020). However, these studies are mostly correlative and the biological mechanisms underlying most cancers remain to be discovered. Advances in single-cell and spatial multi-omics technologies, which maximise our histological resolution and allows vertical integration of omics data, i.e. whereby multiple omics layers are measured in the same tissues or cells, will help expand our understanding of the microenvironment, phenotypic and (epi)genetic alterations in cancer formation and to reveal their causes.

A number of DNA repair defects and genotoxic agents have been associated with specific prevalences of trinucleotide sequence contexts and alteration types known as mutational signatures (Alexandrov et al. 2020). Behjati et al. revealed that radiation is associated with an excess of deletions and with other genomic features by comparing the genomes of tumors from radiation-naive patients and from radiation-associated secondary tumors (Behjati et al. 2016). Similarly, a radiation mutational signature was investigated in specimens from the Chernobyl tissue bank and identified radio-induced double-stranded breaks in younger patients leading to fusion genes and PTC formation (Morton et al. 2021). Most thyroid cancers, however, are not caused by radiation. Because H$_2$O$_2$ is genotoxic and produced in massive amounts by follicular cells, it has long been postulated that it may contribute to thyroid cancer initiation (Song et al. 2007). Kucab et al. sequenced the genomes of iPSC exposed to 79 carcinogenic agents (Kucab et al. 2019). Exposure to gamma radiation resulted in more indels, but no remarkable signature was found following H$_2$O$_2$ treatment. While no handle on the genotoxic action of H$_2$O$_2$ emerged
from this first screen, it remains to be seen if this negative result also applies to thyrocytes. In addition, the tri-nucleotide window used to define current signatures could be insufficient to characterize \( \text{H}_2\text{O}_2 \)-inflicted DNA damage at the substitution level.

It is becoming increasingly clear that normal healthy tissues, through ongoing somatic evolution, harbour many of the mutations once attributed to the cancers (Martincorena et al. 2015; Lawson et al. 2020; Suda et al. 2018). For example, non-synonymous substitutions in \( \text{NOTCH1} \), a known oncogene in esophageal cancer, turned out to be more prevalent in healthy esophagus (Martincorena et al. 2018). While cancer-specific substitutions have been repeatedly found across different healthy tissues, structural variants and copy-number aberrations seem to arise later during cancer development and progression. However, there are already exceptions to be found, e.g. in placenta (Coorens et al. 2021), and this remains to be elucidated in the human thyroid, where a normal healthy baseline is yet to be defined. Early work on microdissected normal healthy follicles of two patients suggests low mutational burden and polyclonality (Moore et al. 2020). Benign thyroid lesions show evidence for RAS mutations and RET structural rearrangements (Najafian et al. 2017).

Emerging technologies will undoubtedly help better characterise the normal germline and thyroid baseline: long-read sequencing allows to identify tens of thousands of structural variant candidates for their association with increased thyroid cancer risk; single-cell transcriptomics allows for an almost unbiased cataloguing of cell types (Han et al. 2020; Regev et al. 2017), while single-cell multi-omics approaches provide interpretability of the transcriptional phenotypes and give further clues on their biological underpinnings (Nam, Chaligne, and Landau 2021); spatial (epi-)genomics (Ellis et al. 2021), transcriptomics (Stickels et al. 2021) and proteomics can replace the genotypes and phenotypes in their spatial context (Nam, Chaligne, and Landau 2021). While all of these technologies are available today, they still need to be applied to healthy thyroid tissues to better characterise the different tissue states, and provide a proper baseline for the study of thyroid malignancies.

Screening and overdiagnosis

One of the major public health problems of thyroid tumors is the high rate of thyroid nodules discovered by ultrasonography, ranging from 30 to 50%, of which 5-15% are malignant. Their incidence has been increasing in recent years, and women and the elderly people are the most
concerned (Andrioli, Carzaniga, and Persani 2013). This is concomitant with an over-diagnosis of indolent or very slow-growing cancers that are unlikely to cause symptoms or death (Jegerlehner et al. 2017; Vaccarella et al. 2016). Fine-needle aspiration biopsy (FNAB) is routinely used in the preoperative evaluation of thyroid nodules. However, 15% to 30% of aspirations yield inconclusive cytological findings (Bongiovanni et al. 2012). The indeterminate FNAB results are mainly classified in two different categories: atypia of undetermined significance/follicular lesion of undetermined significance (AUS/FLUS, Bethesda category III) or follicular neoplasm/suspicion for follicular neoplasm (Bethesda category IV). The risk of malignancy is respectively of 10 to 30% and 25 to 40%, inducing a diagnostic uncertainty that often results in a repeat FNAB and/or an unnecessary diagnostic lobectomy/thyroidectomy. This leads to unnecessary costs and risks for the patients. Thyroidectomy may indeed be associated with adverse effects like hypoparathyroidism, the risk of lesion of the vocal cords and one consequence is the need of a hormone supplementation for life. On the other hand, lobectomy might be an insufficient treatment if a cancer is finally diagnosed on histological examination of the nodule, leading the patient to a new surgery.

The characterization of the genome and transcriptome of thyroid cancers, highlighting new markers of malignancy, has led to the development of different molecular approaches to predict malignancy risk for cytologically indeterminate thyroid nodules. A few commercial molecular diagnostic tests have been developed, with variable sensitivities, specificities, and predictive values: Afirma (Veracyte) (Alexander et al. 2012), ThyGeNEXT/ThyraMIR (Interpace Diagnostics) (Labourier et al. 2015; Wylie et al. 2016), and ThyroSeq (CBLPath) (Nikiforov et al. 2015, 2014; Nikiforova et al. 2018). They are based on different technologies: analysis of mRNA or miRNA expressions, detection of mutations, gene fusions or chromosomal copy number alterations, or a combination of these. Each of them presents advantages and limitations, reviewed in Eszlígether et al, 2017 (Eszlinger et al. 2017) and in Nishino and Krane, 2020 (Nishino and Krane 2020). Molecular testing has thus emerged as a new diagnostic and prognostic tool for thyroid nodules.

Machine learning (ML) and artificial intelligence (AI) may help overcome the ambiguities and biases inherent to examination of images by humans at all stages of the diagnostic process, ultrasonography, FNAB, and post-surgery histopathology. Approaches in the field of ultrasonography have been recently reviewed elsewhere (L.-R. Li et al. 2020; Ha and Baek 2021). Several retrospective studies have produced deep convolutional network models performing in par with expert radiologists (L.-R. Li et al. 2020; Ha and Baek 2021; X. Li et al.
Similarly promising results have been reported for FNAB images (Dov et al. 2019), and interesting applications of digital pathology have been proposed to predict molecular phenotypes from histopathological imaging (Fu et al. 2020).

The robustness of these models across variations in the make and settings of image acquisition devices remain to be established. This is known as the domain adaptation problem and is a major topic in AI research, with many promising avenues toward working solutions (Kouw and Loog 2021). A more fundamental limit of AI applications to ultrasonography image analysis, however, is that image quality depends critically on the skills and experience of the radiologist operating the manual scanning device (Park 2021). The same problem also applies to FNAB-derived material — imaged or processed for molecular profiling — and is compounded by the fact that only a small fraction of the possibly heterogeneous tumor mass is being sampled.

The accrued diagnosis of microPTC has led to a subset of these thyroid cancers undergoing active surveillance to avoid unnecessary FNAB (Haugen et al. 2016). This usually involves simple observation. Liquid biopsies represent another emerging avenue in research and in the clinics that could improve screening, monitoring and limit overdiagnosis in thyroid cancer. Non-invasive liquid biopsies, either from the blood or urine, looking at cell-free DNA, circulating tumour cells, or exosomes, and in combination with methylome sequencing, provide a wide range of applications to many cancer types, which make them an ideal early pan-cancer detection candidate (Nassiri et al. 2020; Nuzzo et al. 2020; M. C. Liu et al. 2020). In thyroid cancer, while some have used the supernatant of thyroid FNAB as a liquid biopsy to derive diagnostic molecular profiles (Ye et al. 2019), liquid biopsies from urine or blood hold great non-invasive potential to complement other approaches for the screening and early-detection of thyroid cancer (Khatami et al. 2019; Rappa et al. 2019) and potentially avoid unnecessary FNAB altogether (Rappa et al. 2019).

**Disease classification - ambiguous definitions and heterogeneity**

No less than 15 histological variants are distinguished for PTC alone (Lloyd et al. 2017), each defined by a combination of diverse qualitative or semiquantitative morphological criteria, and in some cases molecular markers. Many of these criteria are not—taken individually—specific to any variant. It is revealing that ground glass nuclei, which are arguably the hallmark of PTCs, are also present in some Hashimoto thyroiditis (LiVolsi 2011). The history of PTC classification over the past 70 years (Xu and Ghossein 2018), in particular for the follicular variant, is
remarkably intricate due to the inherent difficulties of rigorously describing shapes and textures, of setting crisp boundaries between fuzzy categories, and of deciding which features actually matter. For example, the increased diagnosis of the encapsulated follicular variant PTC since 2000 is partly explained by a progressive lowering of the threshold of nuclear alterations required to diagnose a tumor as PTC (Xu and Ghossein 2018). The Cancer Genome Atlas survey brought clarity by positioning PTCs along a transcriptionally-defined RAS-like to BRAF-like axis, establishing that the follicular variant is associated with the RAS-like and the classical papillary variant with the BRAF-like molecular phenotypes, respectively.

It should not be concluded from the above, however, that molecular markers will supersede morphological ones. Most interestingly, deep neural networks can be trained to detect with excellent sensitivity and specificity the presence of the $\text{BRAF}^\text{V600E}$ mutation from histological images alone (Z. Wu et al. 2020; Tsou and Wu 2019; Fu et al. 2020). Interestingly, a deep neural network trained to predict the transcriptional BRAF-RAS score from histological slices, has also shown promise for the diagnosis of Noninvasive follicular thyroid neoplasms with papillary-like nuclear features (NIFTP; (Dolezal et al. 2021)). We contend that molecular phenotypes are not intrinsically superior to morphological phenotypes, but they are currently more effective because some features are actionable in the clinics, and importantly, they can be turned into objective quantitative scores. Yet, convolutional neural networks also have the potential to set morphological analysis onto a quantitative ground.

The above studies used supervised learning, which by design maps images to known categories. But a growing body of work shows that AI systems can learn so-called latent representations that encode the semantic content of images into numerical vectors, without any human supervision (Kopf and Claassen 2021). Unlike human verbal and subjective descriptions, these numerical vectors are adapted to the computation of similarity metrics between images, which could set the foundations of a reproducible and quantitative classification system for thyroid cancer.

A limit of current thyroid cancer classification systems, including those based on the transcriptome profiling of tissue blocks, is that they poorly reflect intra-tumor heterogeneity. AI systems can operate at any scale within high resolution whole slide images and therefore survey the variation of local morphology across wide tumor areas. More research will be needed to integrate this intra-tumor variability, when present, into relevant tumor-level scores.
AI analysis has been applied to stacks of ~1500 serial pancreatic cancer slices scanned at 20X magnification, resulting in a high resolution 3D rendering of morphological structures (tumor, ducts, islets, etc.) over cubic centimeter volumes, including the exact locations of the ~10⁵ cells present in such volume (Kiemen et al. 2020). Applied to thyroid samples such reconstructions could be viewed as modern versions of the wax models that revealed the thyroid 3D morphology in the late 1920s (Rienhoff 1929), but with cellular resolution, larger volumes, and powered with algorithms that potentially detect any structures visible with H&E staining, calculate distance between these structures, etc. As a glimpse into the 3D PTC morphology, our group has staked slices from two tumors over volumes of about 2 x 2 x 0.2 mm³ from one patient (Tarabichi et al. 2018). Tumor regions were stained with an anti-BRAFV600E antibody and tumor volumes were reconstructed in 3D. In these specimens, the tumor diverged from the textbook model of a tumor with a central core surrounded by invasive fronts. It unfolded as a sparse fractal-like mesh inside the stroma. Every tumor cell was within close distance or in contact with the stroma. Further Ki67 staining demonstrated that the stroma’s proliferative index was comparable to that of the tumor cells. Recent progress in light-sheet microscopy has made possible the 3D imaging of whole human organs, including the thyroid at a resolution of ~5 μm in the three spatial dimensions (Zhao et al. 2020). Thus, high resolution imaging of entire tumors in 3D is within reach. AI is likely to play an important role in the segmentation of morphological structures within these massive images.

Single cell RNA-seq is becoming the new standard for transcriptomics studies opening a window on the cell types and cell states present within the tumor ecosystem. A handful of studies have applied it to PTC (Yan et al. 2021; Peng et al. 2021; Hu et al. 2020), and one addressed ATC among a panel of aggressive tumors affecting other organs (R. Gao et al. 2021). All demonstrated the unbiased identification of cells types and thyroid cancer cell subsets made possible by the technology. Gao and collaborators (R. Gao et al. 2021) presented a novel method to infer DNA copy variation from single cell RNA-seq and showed that it resolved the clonal structure of the cancer cell population.

The tissue digestion required for single cell RNA-seq discards spatial information. Spatially-resolved transcriptomics overcome this limit (Marx 2021). The major distinguishing feature of these technologies, compared to IHC, immunofluorescence, FISH and other probe-based methods is that they are transcriptome-wide and unbiased, thus they have the
potential to uncover both single gene marker and gene expression signatures associated with morphology. This comes at the expense of resolution which is currently supra cellular. Subcellular resolution has been achieved in experimental platforms (J. H. Lee et al. 2015; Vickovic et al. 2019), but the sequencing cost required to achieve a useful per-cell transcriptome coverage over large tissue areas will remain a major limiting factor. Our group has profiled a PTC (Saiselet et al. 2020) with the first generation spatial transcriptomics platform (Ståhl et al. 2016), which offered a resolution of 100 μm over ~1000 locations evenly spaced over an area of 6.2x6.6 mm². The technology is showcased on Fig. 2. RNA-seq profiles are essentially universal readouts for tissue phenotypes as they can be interrogated with any gene or combination of genes. Fig. 2 shows the spatial variation of the expression markers characteristic of cell types (follicular cells, T cells), and cell states (PCNA, IFN type I and TGFβ) across morphologically heterogeneous regions. The meaning of VIM and FN1 expression will be discussed in the next section. In contrast with bulk RNA-seq, which cannot measure absolute RNA levels (Lovén et al. 2012), spatial transcriptomics provides information about the relative RNA production across the tissue (Saiselet et al. 2020).

Combining these descriptive modalities will fuel powerful analyses, either revealing association between the modalities or by producing joint representations. As an example of the first approach, the markers of different cancer cell subpopulations discovered from the single cell RNA-seq profile of a pancreatic cancer were projected onto a spatial transcriptomics slice from the same tumor to reveal their spatial distribution (Moncada et al. 2020). In another publication, a deep neural network was trained from spatial transcriptomics to infer the expression of 102 genes from 150x150 μm² H&E image patches (He et al. 2020). This line of work will ultimately lead to a representation of cell type density, gene expression and genomic alterations as virtual stains overlaid over H&E in the eyepiece of augmented reality microscopes (Chen et al. 2019). On a more fundamental level, it will reveal the transcriptional programs associated with local morphology and help decipher the mechanisms causing morphological changes.

The expression of a hundred of genes could be predicted from local morphology in (He et al. 2020), among >1000 measured in the spatial transcriptomics experiments. Noise and shallow coverage have certainly been a limitation, but it is also possible that some gene expression variations are not associated with significant morphological variations. More generally, the redundancies and complementarities of the different omics and imaging platforms, and the extent to which they produce different tumor classifications have not been formally addressed.
More relevant classification could possibly be discovered by integrating all the perspectives from which tumors can be characterized. Data fusion algorithms (J. Gao et al. 2020) integrate different measurement modalities of a same entity, for example the sound and the image of a video footage. A pioneering application of such tools to spatial transcriptomics demonstrates that super resolution, i.e. the inference of spatial details beyond the resolution or the original transcriptomics measurements, can be achieved from an AI model trained jointly on images and transcriptomes (Bergenstråhle et al. 2020). Comparable approaches could be used to build comprehensive tumor classifications integrating all measurable aspects of their biology.

**Limited treatment options in advanced cancers**

Most thyroid cancers are differentiated, respond well to conventional therapies and, compared to other cancer types, are associated with very good prognosis. However, the more advanced cases, especially the poorly differentiated and undifferentiated thyroid cancers, which represent a small but not neglectable percent of all cases, remain a therapeutic challenge (reviewed in (Lorusso et al. 2021)). Multikinase inhibitors such as lenvatinib and sorafenib are used for the treatment of advanced radioactive iodine-refractory differentiated thyroid cancer. For BRAF\textsuperscript{V600E} mutated ATC, the combination of BRAF and MEK inhibitors (dabrafenib and trametinib), has been approved and showed clinical benefits in the patients (Subbiah et al. 2020). Both drugs were also shown to have potential in the neoadjuvant setting to improve complete surgical resection in otherwise unresectable cases (Wang et al. 2019). These drugs have numerous adverse events, hence a second generation of TKIs has recently been developed which are specific for a particular oncogene: specific inhibitors of RET (selpercatinib, pralsetinib) or NTRK (larotrectinib, entrectinib) fusions have been approved for advanced or metastatic thyroid cancers harboring the corresponding alteration. However, many patients develop resistance to those therapies, and the development of more effective treatments is needed ((Owen et al. 2019)). Clinical trials are ongoing to evaluate the potential use of immunotherapy to treat thyroid cancer, using immune checkpoint inhibitors (nivolumab, pembrolizumab), alone or combined with kinase inhibitors (Lorusso et al. 2021).

In most cancers, only a fraction of patients do respond to these harsh and expensive treatments. Some of the key questions for improving patient outcome in advanced cases are to predict and understand different disease progression histories, identify new therapeutic susceptibilities, and predict patient response to treatment.
While the recent efforts for the characterisation of PDTC and ATC mutational landscapes have helped better understand their genetics and identify potential avenues for targeted treatments, most have been limited to single-region bulk sequencing, usually using targeted gene panels or exome sequencing (Landa et al. 2016; Yoo et al. 2019; Jeon et al. 2016; Kunstman et al. 2015; Gerber et al. 2018; Pozdeyev et al. 2018). While targeted deep sequencing is a powerful tool to identify potential targets for advanced thyroid cancers (Ibrahimpasic et al. 2019), whole-genome sequencing efforts will help better characterise the genome-wide driver landscape and the extent of the mutational processes and intra-tumour heterogeneity (Tarabichi et al. 2021; Dentro et al. 2020) in those thyroid cancers. Yoo and colleagues sequenced both the WES and WGS of 13 ATC, and WGS could further identify a TERT enhancer hijacking event and 10,000 APOBEC-related mutations in a hypermutator sample. In the Pan-Cancer Analysis of Whole Genome project, WGS revealed frequent chromoplexy-like catastrophic events in 48 otherwise stable PTC genomes, linking multiple chromosomes and leading to gene fusions between cancer genes and thyroid-specific genes (ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium 2020). WGS is yet to be applied to large cohorts of advanced thyroid cancers, and long-read sequencing technologies are ideally-suited for accurate and exhaustive identification of the complex structural rearrangements seen in thyroid cancers.

Because of their high mortality and limited treatment options, ATC are the subject of many ongoing clinical trials (Ferrari et al. 2020). Already approved MEK-BRAF inhibitors for BRAF-mutant ATC (Subbiah, Baik, and Kirkwood 2020) were also shown to have potential in the neoadjuvant setting to improve complete surgical resection in otherwise unresectable cases (J. R. Wang et al. 2019). Immunotherapies, spearheaded by checkpoint inhibitors, with a curative potential, have seen an increasing number of successes across different cancer types, including thyroid cancer (Ma et al. 2020; French 2020). There is a large body of evidence showing that ATC might be well-suited for these types of treatments (Ma et al. 2020). However, in most cancers only a fraction of patients do respond to these harsh and expensive treatments. Thus predicting patient response or adverse effects (Jing et al. 2020) to those has been an important clinical challenge, for which the power of AI and ML for data fusion can be leveraged, as these are ideally-suited to integrate the complex patterns in single-cell sequencing, multiomics (J. S. Lee and Ruppin 2019; Litchfield et al. 2021), clinical images (M. Wu et al. 2019), such as radiomics (Mu et al. 2020; Butner et al. 2020; Trebeschi et al. 2019) or histopathological (Hildebrand et al. 2021), transcriptomics and genomics data (Litchfield et al. 2021).
Furthermore, ongoing cancer evolution leading to subclonal intra-tumour heterogeneity is a major therapeutic challenge for targeted treatments. Multi-region bulk sequencing combined with single-cell sequencing, among other study designs (Tarabichi et al. 2021), will allow the characterisation of the intra-tumour heterogeneity of advanced thyroid cancers. Similarly, recent advances in molecular archaeology of cancers from sequencing data have leveraged copy-number gains and molecular clocks to reconstruct cancer evolutionary timelines (Gerstung et al. 2020), and could shed light on the evolution of the advanced thyroid cancers. Especially, while the incidence of differentiated thyroid cancer has been increasing, partially due to overdiagnosis and better screening for early-stage cancers, this has not led to a decrease of aggressive subtypes and even rather has been accompanied by a parallel increase of stage IV cancers (Olson et al. 2019). This could reflect an escape of early screening of a late age group in an aging population or through a fast evolution from differentiated to undifferentiated cancers. The latter could be assessed through phylogeny reconstruction in cases of synchronous differentiated and undifferentiated cancers and thanks to molecular timing approaches leveraging molecular clocks (Gerstung et al. 2020). This might shed light on the early clonal targetable drivers, identify the window of opportunity for prevention and the specific early biomarkers for better screening for these advanced thyroid cancers.

Although tumour aneuploidy arises in different contexts (Vasudevan et al. 2021; Ben-David and Amon 2020), it is a common feature of advanced cancers, including advanced thyroid cancers (Gopal et al. 2018; Ganly et al. 2013; Landa et al. 2016; Yoo et al. 2019; Jeon et al. 2016; Kunstman et al. 2015; Gerber et al. 2018), presenting with higher levels of chromosomal instability and genome doubling (Bielski et al. 2018), which fuels the evolution and selection of fitter clones (López et al. 2020; Watkins et al. 2020). Another exciting avenue of treatment exploits aneuploidy as an evolutionary characteristic of advanced cancers, as among other potential targetable pathways, it is sensitive to perturbation of the mitotic-spindle assembly checkpoint (Cohen-Sharir et al. 2021; Chunduri and Storchová 2019).

Beside immunotherapies, most treatments investigated against thyroid cancers are cancer cell-centric, i.e. they target cell autonomous mechanisms of tumor expansion. Yet, as discussed earlier in this review, the microenvironment is quantitatively comparable or larger than the cancer cell population in some ATC and PTC. Thus, tumor expansion implies expansion of the microenvironment in these tumors. Spatial transcriptomics, possibly combined with single cell
RNA-seq, may shed light on the interactions between cell types. To illustrate this possibility, we present in Fig. 3 the preliminary analysis of the spatial transcriptomics profile of a PTC. It shows that although FN1 is generally considered as a mesenchymal marker, in this PTC its expression is positively correlated across space with TG, suggesting a partial EMT. VIM, by contrast, shows a weak but significant negative correlation with TG. A multivariate analysis shows a statistical interaction between VIM and TG in explaining FN1 expression, suggesting a synergy between fibroblasts and epithelial cells in producing FN1. The adjunction of single cell data would settle more accurately which cell type produces FN1, which ligand/receptor pairs could support cell/cell communication, etc.

Treatment monitoring and follow-up

One of the major challenges after treatment of thyroid cancer is the management of persistent and recurrent disease, which are not trivially related to mortality in thyroid cancer. Although early detection of recurrence has significantly improved over the past decades, the screening process still comes at high economical costs (Lamartina et al. 2018). Again, improving risk stratification is needed. At the initial stage, predicting which patient should undergo aggressive follow-up is key (Tuttle and Alzahrani 2019). Molecular profiling of the primary tumor, including methylation (Bisarro Dos Reis et al. 2017) and microRNA (Celano et al. 2017), have shown interesting predictive potential. Then, monitoring and follow-up are required to further assess risk of recurrence and mortality. Liquid biopsies are a promising non-invasive tool for monitoring and follow-up in the clinics (Heitzer et al. 2019; Pantel and Alix-Panabières 2019). In thyroid cancers, liquid biopsies could complement imaging and serum thyroglobulin levels. For example, molecular biomarkers in thyroid cancer have been identified from liquid biopsies such as microRNAs or from circulating epithelial cells (Allin et al. 2018; Lin et al. 2018; Qiu et al. 2018; Nylén et al. 2020).

Conclusion

Thyroid cancers are a basket of multiple sub-entities with evolving definitions and increasing incidence worldwide. While screening efforts catch many of these cancers early in their progression, they also come with increased overdiagnosis. Most cases present excellent prognosis but intriguingly, the incidence of advanced cancers is also increasing, with limited treatment options. Persistent and recurrent disease following treatment is often seen, while
unified management guidelines are lacking and better risk stratification schemes, monitoring and follow-up options are needed.

Table 1 summarizes the research avenues outlined in this review. We have discussed how emerging technologies in genomics and imaging as well as treatments and biopsies, such as spatially-resolved single cell omics, objective AI-based abstraction and integration, immunotherapies and liquid biopsies will help tackle many of the current issues associated with the management of thyroid cancers and reduce their increasing societal burden.

Acknowledgements

The work performed in the authors’ laboratories was supported by grants from the Belgian Fonds National de la Recherche Scientifique (F.R.S.-FNRS), Welbio, Plan Cancer Belgique (action 29), Télévie, Fondation Rose et Jean Hoguet, Fonds David et Alice van Buuren, Fonds Docteur J.P. Naets, les Amis de l’Institut Bordet, the Belgian Fondation Contre le Cancer.
**Figure 1. Thyroid tumorigenesis.**

**Top.** A normal thyrocyte can evolve into different types of tumors, depending on the signaling pathway that becomes constitutively activated (cAMP, MAPK, PI3K, or both). The size of the pathway label conveys the intensity of activation at the different stages. The main mutations present in the different tumor types are also shown, ordered according to their prevalence.

**Bottom.** Each type of tumor presents distinct morphological features at the tissue and cellular levels, from physiological follicular organisation to different types of disorganisation and anaplasia.
Figure 2. Spatial transcriptomics profile of a PTC.

A. The RNA-seq profiles were determined within 281 spots of 100 μm of diameter arranged on a grid-like array on a PTC slice of approximately 2.5x6 mm². Five types of tissue regions were identified from the H&E image beforehand with the help of a pathologist. Representative spots are shown for each region.

B. Since the entire transcriptome is available at each spot, any gene, or gene signature, expression can be projected onto the image or displayed region-wise as boxplots (lower panels, same color code as panel A). Expression is normalized by total counts and presented on a log-scale. Thyroglobulin (TG, marks thyroid epithelium), vimentin (VIM, fibroblast marker) and T cell gene expression signature accurately delimit the relevant epithelial, fibrotic and immune structures. The quantitative analysis (not shown) reveals a higher expression of VIM at the border of immune foci and dense epithelia. Immune foci are correctly identified beside their small size. Beyond cell type, ST can also detect localized cell states. For example, type I inflammation is present in immune foci but also in sub-parts of the epithelial zone with a gradient pattern. Hence, spatial transcriptomics captures the spatial continuum of expression variation. The spread of the expression values within the boxplots points out the limits of defining discrete areas within this continuum, which is unavoidable with laser microdissection. PCNA expression colocalizes with epithelia. TGFβ-responsive genes are mostly expressed in the cell-dense upper fibrotic area. Importantly, the expression data shown here have been normalized by the total number of sequencing reads collected in each spot. We analyzed in detail the effect of this standard normalisation (Saiselet et al. 2020), and demonstrated that more VIM was actually transcribed in the epithelial areas of this PTC, including in spots where no fibroblasts were present.
Figure 3. Interaction between epithelial cells and fibroblasts in a BRAF\(^{V600E}\) PTC.

A, While fibronectin (FN1), a key structural protein of the extracellular matrix, is believed to be mostly expressed by fibroblasts, we find that its pattern of expression follows that of TG, not VIM (compare to Fig. 2B, TG panel). The same is true of raw read counts (not shown).

B, The five spots where FN1 is most expressed are shown. They contain very few fibroblasts, hardly any is found in the top spot.

C and D, FN1 is correlated positively with both TG and the fraction of epithelial cells, but negatively with VIM and the fraction of fibroblasts. Thus, the epithelium produces a large part of FN1 in the tissue. Are fibroblasts involved indirectly?

E, A multivariate regression of FN1 against TG and VIM shows that VIM is a very significant predictor of FN1 expression when combined with TG. Importantly, the interaction term is also significant. This points to a multiplicative effect of VIM and TG in explaining FN1 expression.

F, Analysis of the TCGA cohort (503 PTCs, including 261 BRAF\(^{V600E}\)) demonstrates that FN1 is overexpressed by two orders of magnitude in BRAF\(^{V600E}\) PTCs compared to either normal thyroid or RAS-mutated PTCs. This massive overexpression is seen in nearly all BRAF\(^{V600E}\) cases, i.e. >50% of the cohort. This suggests that the interaction revealed by this PTC spatial profile could be broadly relevant.
<table>
<thead>
<tr>
<th></th>
<th>Prevention</th>
<th>Screening</th>
<th>Diagnosis</th>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mutational signatures</td>
<td>May point to novel DNA repair defect and/or genotoxic agent causing thyroid cancer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long-read sequencing</td>
<td>Identification of new germline susceptibility markers, base modifications and structural variation</td>
<td>Profiling of base modifications is ideal for early multi-cancer detection through liquid biopsies</td>
<td></td>
<td>Monitoring through liquid biopsies</td>
</tr>
<tr>
<td>Routine targeted sequencing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evolutionary analysis</td>
<td></td>
<td></td>
<td></td>
<td>Identification of key genomics events</td>
</tr>
<tr>
<td>Liquid biopsies</td>
<td></td>
<td></td>
<td></td>
<td>Minimally invasive monitoring of tumor markers</td>
</tr>
<tr>
<td>Single-cell multiomics</td>
<td>Study the normal baseline</td>
<td></td>
<td>Defines cell types and cell states</td>
<td>Define pathways active in cells, ligand/receptor pairs, etc.</td>
</tr>
<tr>
<td>Spatial transcriptomics</td>
<td></td>
<td></td>
<td>Reveal transcriptional heterogeneity, connects morphology and transcriptions</td>
<td>Addresses interactions with microenvironment, locate cell type/states</td>
</tr>
<tr>
<td>Light-sheet microscopy of large specimens</td>
<td></td>
<td></td>
<td>Provide 3D account of morphological heterogeneity</td>
<td>Resolve vasculature and other structures difficult to assess in 2D</td>
</tr>
<tr>
<td>AI, unsupervised</td>
<td></td>
<td></td>
<td>Provides quantitative morphology representations</td>
<td></td>
</tr>
<tr>
<td>AI, supervised</td>
<td>Predict clinical/molecular variables from imaging</td>
<td>Predict clinical and molecular variables from imaging</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AI, data fusion</td>
<td></td>
<td></td>
<td>Unified representation of imaging, transcriptomics, etc.</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Different technologies and some of their potential applications for studying thyroid cancer prevention, screening, diagnosis, and treatment.
References


Carta, Claudio, Sonia Moretti, Lucia Passeri, Flavia Barbi, Nicola Avenia, Antonio Cavaliere, Massimo Monacelli, et al. 2006. “Genotyping of an Italian Papillary Thyroid Carcinoma Cohort Revealed High Prevalence of BRAF Mutations, Absence of RAS Mutations and Allowed the Detection of a New Mutation of BRAF Oncoprotein (BRAF(V599lns)).” Clinical Endocrinology 64 (1): 105–9.


Fu, Yu, Alexander W. Jung, Ramon Viñas Torne, Santiago Gonzalez, Harald Vöhringer, Artem


Peng, Miaoguan, Guohong Wei, Yunjian Zhang, Hai Li, Yingrong Lai, Yan Guo, Yuxin Chen, et al. 2021. “Single-cell Transcriptomic Landscape Reveals the Differences in Cell Differentiation and Immune Microenvironment of Papillary Thyroid Carcinoma between...


Trovisco, Vítor, Inês Vieira de Castro, Paula Soares, Valdemar Máximo, Paula Silva, João Magalhães, Alexander Abrosimov, Xavier Matias Guiu, and Manuel Sobrinho-Simões. 2004. “BRAF Mutations Are Associated with Some Histological Types of Papillary Thyroid...


Xie, Zhenyu, Xin Li, Yuzhen He, Song Wu, Shiyue Wang, Jianjian Sun, Yuchen He, Yu Lun, and Jian Zhang. 2020. “Immune Cell Confrontation in the Papillary Thyroid Carcinoma Microenvironment.” *Frontiers in Endocrinology* 11 (October): 570604.


