Engineering 28 (2023) 179–192

Contents lists available at ScienceDirect

Engineering

journal homepage: www.elsevier.com/locate/eng

Research
Medical Engineering—Article

Intratumoral Bacteria Dysbiosis Is Associated with Human Papillary Thyroid Cancer and Correlated with Oncogenic Signaling Pathways

Shuang Yu [a,#], Yanqiang Ding [b,#], Xuejie Wang [a,#], Siu Kin Ng [b], Siting Cao [a], Weixin Liu [b], Zhuming Guo [c], Yubin Xie [d], Shubin Hong [a], Lixia Xu [e,f], Xiaoxing Li [e], Jie Li [g], Weiming Lv [g], Sui Peng [d,e], Yanbing Li [a], Joseph J.Y. Sung [e,h], Jun Yu [b,e,*], Haipeng Xiao [a,*]

[a] Department of Endocrinology, The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou 510080, China
[b] Institute of Digestive Disease and Department of Medicine and Therapeutics, State Key Laboratory of Digestive Disease, The Chinese University of Hong Kong, Hong Kong 999077, China
[c] Department of Head and Neck Surgery, Sun Yat-Sen University Cancer Center, Guangzhou 510060, China
[d] Clinical Trials Unit, The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou 510080, China
[e] Institute of Precision Medicine, The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou 510080, China
[f] Department of Oncology, The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou 510080, China
[g] Department of Breast and Thyroid Surgery, The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou 510080, China
[h] Lee Kong Chian School of Medicine, Nanyang Technology University, Singapore 308232, Singapore

ARTICLE INFO

Article history:
Received 5 September 2022
Revised 4 November 2022
Accepted 17 January 2023
Available online 16 March 2023

Keywords:
Papillary thyroid cancer
Benign thyroid nodule
Bacteria
Transcriptome
Hashimoto's thyroiditis

ABSTRACT

Emerging evidence suggests that microbial dysbiosis plays vital roles in many human cancers. However, knowledge of whether the microbial community in thyroid tumor is related to tumorigenesis remains elusive. In this study, we aimed to explore the microbial community in thyroid tissues and its contribution to papillary thyroid cancer (PTC). In parallel, we performed microbial profiling and transcriptome sequencing in the tumor and adjacent normal tissues of a large cohort of 340 PTC and benign thyroid nodule (BTN) patients. Distinct microbial signatures were identified in PTC, BTN, and their adjacent non-tumor tissues. Intra-thyroid tissue bacteria were verified by means of bacteria staining, fluorescence *in situ* hybridization, and immunoelectron microscopy. We found that 17 bacterial taxa were differentially abundant in PTC compared with BTN, which included enrichment in PTC of the pathobionts *Rhodococcus*, *Neisseria*, *Streptococcus*, *Halomonas*, and *Devosia*, and depletion of the beneficial bacteria *Amycolatopsis*. These differentially abundant bacteria could differentiate PTC tumor tissues (PTC-T) from BTN tissues (BTN-T) with an area under the curve (AUC) of 81.66%. Microbial network analysis showed increased correlation strengths among the bacterial taxa in PTC-T in comparison with BTN-T. Immune-function-corresponding bacteria (i.e., *Erwinia*, *Bacillus*, and *Acinetobacter*) were found to be enriched in PTC with Hashimoto's thyroiditis. Moreover, our integrative analysis revealed that the PTC-enriched bacteria had a positive association with key PTC-oncogenic pathway-related genes, including *BRAF*, *KRAS*, *IRAK4*, *CTNNB1*, *PIK3CA*, *MAP3K7*, and *EGFR*. In conclusion, our results suggest that intratumor bacteria dysbiosis is associated with the thyroid tumorigenesis and oncogenic signaling pathways of PTC.

1. Introduction

The incidence of thyroid cancer ranks ninth among all types of cancers and fifth among cancers commonly found in women worldwide. More than 586 000 thyroid cancer cases were newly diagnosed in 2020 [1]. Papillary thyroid cancer (PTC) is the most common form of thyroid cancer, accounting for more than 85% of cases [2]. PTC involves indolent tumors characterized by a good response to standard treatment but showing a perplexing discrepancy between tumor aggressiveness and prognosis [3]. The number of young PTC patients is growing rapidly, but these patients generally have a favorable prognosis—even if there is distant metastasis—whereas a poor prognosis tends to occur in elderly patients. The exact causes of thyroid cancer are still unknown; however, more than 90% of thyroid cancer is sporadic in nature [4]. Risk

* Corresponding authors.
E-mail addresses: junyu@cuhk.edu.hkxiao (J. Yu), xiaohp@mail.sysu.edu.cn (H. Xiao).
# These authors contributed equally to this work.

https://doi.org/10.1016/j.eng.2023.01.007

factors, such as family history, radiation exposure, environment pollutants, or genetic alterations, have been proposed [5]. Among genetic alterations, mutations such as B-Raf proto-oncogene ($BRAF^{V600E}$) and rat sarcoma viral oncogene homolog (*RAS*) involving the effectors of the mitogen-activated protein kinase (MAPK) pathway and the phosphoinositide 3-kinase (PI3K)–protein kinase B (AKT) pathway are important for thyroid cancer initiation and progression [4,6]. However, only a fraction of patients with a genetic predisposition to PTC or who have been exposed to known environmental risk factors actually develop PTC. Therefore, further research on unexplored risk factors is of great significance in the prevention, diagnosis, and treatment of thyroid cancer.

Emerging evidence has revealed that the gut microbiota is increasingly considered to be an important factor associated with tumor development. The effect of the gut microbiota on gastrointestinal cancers has been studied extensively. An investigation of the gut microbial community in PTC patients identified six bacteria genera—including *Bacteroides* and *Roseburia*—that could discriminate between PTC patients and healthy controls with high accuracy [7]. Recent studies have demonstrated the presence of bacteria in several types of solid cancers, including breast, ovary, melanoma, bone, and brain [8,9], suggesting an association between the tissue microbiome and tumor development. There is also increasing interest in microbiome research for thyroid cancer. Gnanasekar et al. [10] characterized the abundance of microbes in PTC using RNA sequencing data obtained from The Cancer Genome Atlas (TCGA). Several recent studies have also reported the presence of bacteria in small cohorts of thyroid cancer tissues subjected to ribosomal RNA (rRNA) sequencing [11,12]. Moreover, Yuan et al. [13] analyzed the PTC intratumor microbiome and found that the bacteria diversity in advanced lesions was greater than that in mild lesions. Nevertheless, further evidence is needed to confirm the presence of bacteria in malignant and benign thyroid tissues, as well as their implication in tumorigenesis for a large sample size.

As most of the microorganisms in tissues cannot be practically cloned and cultured via conventional methods, the link between the tissue microbiome and specific cancer types has been unraveled through 16S rRNA metagenomics. In this study, we used 16S rRNA gene sequencing to perform comprehensive microbial profiling for intratumor microbial communities in 457 tissue biopsies of tumor and adjacent normal tissues from 211 patients with PTC and 63 with benign thyroid nodule (BTN). In parallel, RNA sequencing (RNA-seq) was performed on a total of 356 samples from 148 patients with PTC and 55 patients with BTN. Microbial signatures were confirmed by means of bacteria immunobiological staining and immunoelectron microscopy. We characterized the composition and diversity of the microbiome in malignant and benign thyroid tissues. The relationship between tumor-associated bacteria and Hashimoto's thyroiditis (HT) was also evaluated. Finally, we illustrated the association of microbial heterogeneity with host genetic alteration by performing integrative analyses of intratumoral bacteria profiling and host gene expression profiling.

2. Methods

2.1. Human subjects and sample collection

A total of 340 patients, including patients with PTC (n = 259) and those with BTN (n = 81), were recruited from the Department of Surgery at the First Affiliated Hospital of Sun Yat-Sen University (Guangzhou, China) and Sun Yat-Sen University Cancer Center (Guangzhou, China) between February 2017 and October 2019. The golden reference standard for patients was histopathological diagnosis after surgical resection. Patients diagnosed with other types of thyroid cancers, such as anaplastic thyroid cancer or medullary thyroid cancer, were not included in this study. In addition, patients who failed any of the following criteria were excluded: ① No antibiotics had been administered within the two months prior to sample collection; ② no preoperative chemotherapy or radiotherapy was received; ③ there was no history of malignancy; and ④ there was no history of autoimmune diseases such as rheumatoid arthritis or multiple sclerosis. In total, 211 patients with PTC and 63 patients with BTN were enrolled in this study (Fig. 1(a)). All subjects provided written informed consent before sample collection. This study was approved by the Institutional Research Ethics Committee of Sun Yat-Sen University.

Tissue samples were collected and processed immediately in the operating room to reduce the risk of contamination. Adjacent non-tumor tissue and adjacent benign non-nodule tissue were collected outside the marginal zone, approximately 1 cm away from the nodules. The samples were collected in sterile microtubes, snap frozen in liquid nitrogen, and kept at –80 °C until DNA extraction. In all cases, separate sets of sterile instruments were used to collect the tissues from different positions. After amplification and quality control, a total of 457 tissue biopsies were admitted into our research. These included 170 PTC tumor tissues (PTC-T) and 176 PTC-adjacent non-tumor tissues (PTC-NT) from PTC patients (144 paired tissues), as well as 53 BTN tissues (BTN-T) and 58 BTN-adjacent non-nodule tissue (BTN-NT) from patients with BTN (49 paired tissues). The diagnostic criteria for HT were based on a thyroid autoantibodies test (thyroid peroxidase antibody (TPO-Ab) > 35 IU·mL^{-1} and/or thyroglobulin antibody (Tg-Ab) > 20 IU·mL^{-1}) and/or histopathological features unique to HT [14]. Based on our analysis, 65 of the PTC patients were diagnosed with HT (Figs. 1(b) and (c)).

2.2. DNA extraction and sequencing

Microbial DNA was extracted from the tissue samples using an E.Z.N.A. Stool DNA Kit (Omega Bio-Tek, USA) according to the manufacturer's protocol. In brief, samples were homogenized for 5 min with 200 mg of glass beads and 540 mL of SLX-Mlus buffer using a Geno/Grinder 2010 (SPEX, USA), followed by incubation at 70 °C for 13 min with 60 μL of DS Buffer (Omega Bio-Tek, USA) and 20 μL of Proteinase K Solution (Omega Bio-Tek). The total DNA was eluted in 50 μL of elution buffer. The V3–V4 region of the bacterial 16S rRNA gene was amplified by means of universal primers (341F, 5′-barcode-CCTAYGGGRBGCASCAG-3′ and 806 R, 5′-barcode-GGACTACNNGGGTATCTAAT-3′, where the barcode is an eight-base sequence unique to each sample) using a polymerase chain reaction (PCR) kit (AP221-01; TransGen BioTech, China). PCR reactions were performed in triplicate in a 20 μL mixture containing 4 μL of 5× FastPfu Buffer (TransGen BioTech), 2 μL of 2.5 mmol·L^{-1} deoxy-ribonucleoside triphosphates (dNTPs), 0.8 μL of each primer (5 μmol·L^{-1}), 0.4 μL of FastPfu Polymerase (TransGen BioTech), and 10 ng of template DNA. The procedures were as follows: 95 °C for 5 min; followed by 27 cycles at 95 °C for 30 s, 55 °C for 30 s, and 72 °C for 45 s; and a final extension at 72 °C for 5 min. The PCR products were evaluated using 2% agarose gel electrophoresis.

Sequencing libraries were generated using a TruSeq DNA PCR-Free Sample Preparation Kit (Illumina, Inc., USA) following the manufacturer's recommended protocol, with index codes added. The libraries were quantified on a Qubit@ 2.0 Fluorometer (Thermo Scientific, USA) and using an Bioanalyzer 2100 system (Agilent, USA). Paired-end sequencing (2 × 250 bp) was performed on an NovaSeq platform (Illumnia, Inc.).

2.3. Contamination control setting

To verify that our measurement of intratumor microbial abundance was consistent across samples and was not due to

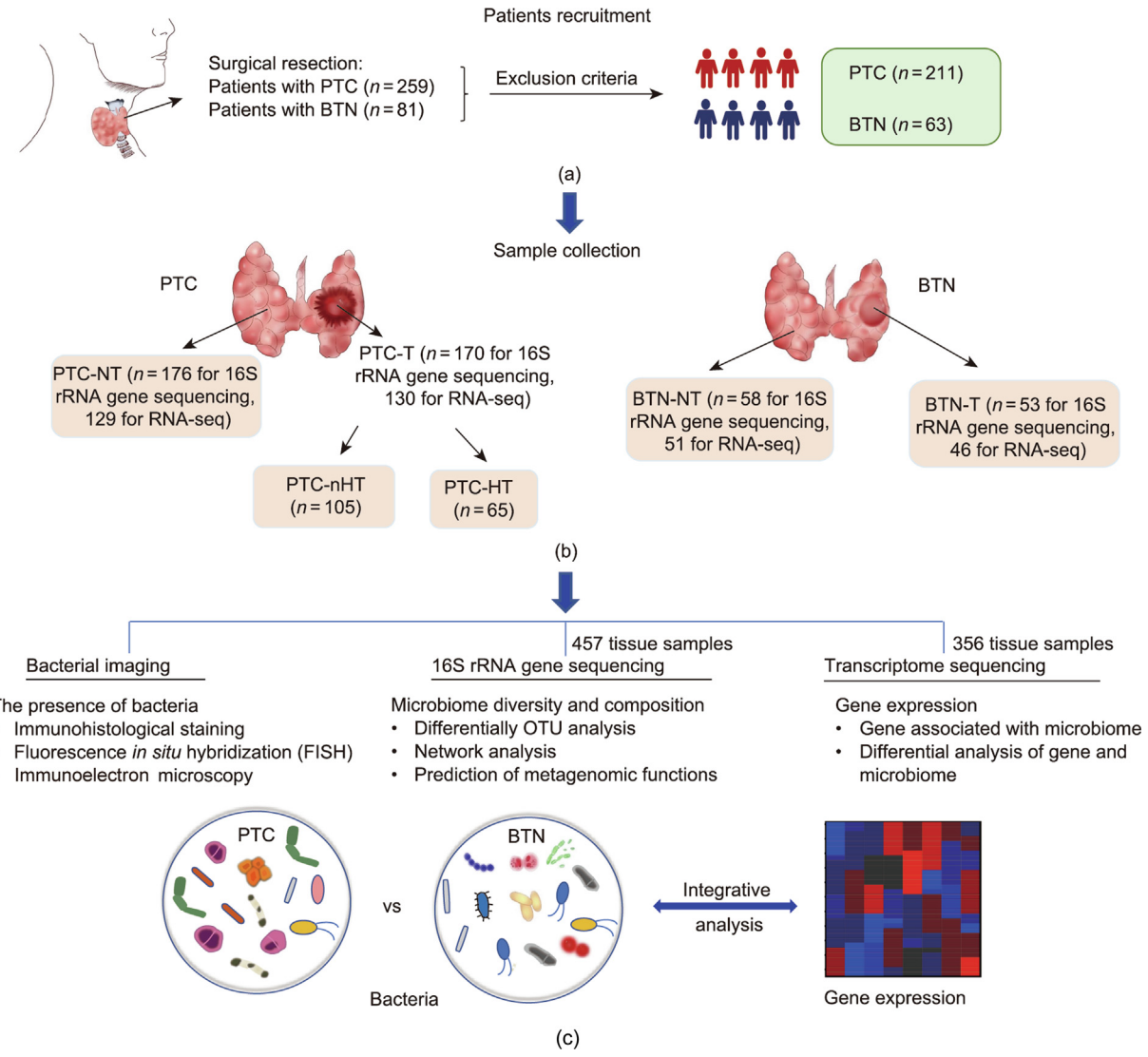

Fig. 1. Sample collection and study design. (a) A total of 340 subjects were enrolled in this study, of which 211 patients with PTC and 63 patients with BTN were included based on the exclusion criteria. (b) Eligible samples (n = 457) included 170 PTC tumor tissues (PTC-T), 176 PTC-adjacent non-tumor tissues (PTC-NT), 53 BTN tissues (BTN-T), and 58 BTN-adjacent non-nodule tissue (BTN-NT) for 16S rRNA gene sequencing. A total of 356 samples included 130 PTC-T, 129 PTC-NT, 46 BTN-T, and 51 BTN-NT for parallel RNA sequencing. A total of 65 PTC patients were diagnosed with HT. (c) Bacterial imaging techniques were applied to detect the presence of bacteria and 16S rRNA gene sequencing to analyze the microbiome characteristics of the tissue samples. PTC-nHT: PTC without HT; PTC-HT: PTC with HT; OTU: operational taxonomic unit.

contamination, environmental controls and negative controls were included in the sequencing, concomitant with the tissue samples. Environmental controls were collected from the operating room, collection tubes, and the equipment used. Three Eppendorf tubes with 1 mL of sterile phosphate-buffered saline (PBS) each were placed open in the operating room during the tissue sample collection. Negative controls were collected from the reagents used in each step of the DNA extraction, library construction, and sequencing process. The environmental controls and negative controls were subjected to 16S rRNA gene sequencing. Based on the sequencing data from the environmental controls and negative controls, the Decontam *R* package was used to remove any contamination ("prevalence" method, $P < 0.05$) [8].

2.4. Sequence curation and annotation

The quality of sequencing reads was evaluated using FastQC v0.11.9 and MultiQC v1.9 with default parameters [15]. The analysis of the 16S rRNA gene sequences was performed using QIIME2 v2019.4.0 [16]. In brief, sequencing reads with a low-quality score were removed via a quality-filter plugin with default parameters prior to read denoising and chimera removal by means of a deblur algorithm with default settings. Taxonomic assignment on the denoised reads was performed using a vsearch taxonomic classifier against the Greengenes reference sequence database preclustered at 99% sequence identity. The resulting operational taxonomic unit (OTU) table contained (13 845 ± 2913) (mean ± standard deviation (SD)) reads per sample (Fig. S1(a) in Appendix A). To compare the microbial diversity of thyroid cancer with those of other cancers (i.e., colorectal cancer (CRC) and hepatocellular carcinoma (HCC)), the specaccum function of the vegan *R* package was used to generate species accumulation curves.

The OTU table was rarefied to 6000 reads prior to the microbial diversity analyses (alpha diversity and beta diversity). Alpha diversity was evaluated by means of the Shannon entropy index with QIIME2 software, and the differences between groups (e.g., PTC-T vs BTN-T) were tested by means of a Wilcoxon rank-sum test. For the ordination analysis, principal coordinate analysis (PCoA) on an unweighted UniFrac distance matrix was performed using

the vegan *R* package and visualized using Origin v2021[†]. The significance of differences among groups was determined by permutational multivariate analysis of variance (PERMANOVA) on the unweighted UniFrac distance with 999 iterations. A redundancy analysis (RDA) was used to evaluate the importance of various clinical traits (i.e., gender, age, tumor size, PTC, BTN, HT, papillary thyroid microcarcinoma (PTMC), and lymph node metastatic carcinoma (LNMC)) to the observed differences in the microbial communities across the samples. After RDA analyses, significant differences in the Bray–Curtis distance between groups (e.g., PTMC and PTC tumor size > 1 cm (nPTMC), Young: age $\leq$ 55 years old and Old: age > 55 years old) was evaluated using Student's two-sample *t*-test.

2.5. Differential abundance bacteria and classification model

Prior to the differential abundance analysis, confounding factors including age, gender, HT, and LNMC were first adjusted using RUVSeq [17]—an algorithm capable of removing unwanted variation. As a continuous factor, age was considered as two groups with a cutoff value of 55. A differential abundance analysis was carried out on the adjusted count data using DESeq2 [18]. The samples were randomly split into training cohorts and validation cohorts with an 8:2 ratio. A logistic regression classifier based on the selected differential abundance bacteria ($P < 0.05$) was implemented with the caret *R* package. Receiver operating characteristic (ROC) curves were computed with the pROC *R* package.

2.6. Microbial association network analysis

Correlation coefficients (r) among bacterial taxa were calculated based on the OTU table separately in the PTC-T, BTN-T, PTC with HT (PTC-HT), and PTC without HT (PTC-nHT) groups with the SparCC algorithm [19], which can limit the number of spurious correlations identified in the compositional data. The *P* value was computed based on 100 bootstraps from the input data. Significant co-occurrence and co-excluding interactions were selected with the igraph *R* package. In detail, for the differential abundance bacteria, interactions with $r > 0.1$ or $r < -0.1$ and $P < 0.05$ were selected; for the overall bacteria, interactions with $r > 0.4$ or $r < -0.4$ and $P < 0.05$ were selected. Significant co-occurrence and co-excluding interactions were visualized using Cytoscape 3.8.2 [20].

2.7. Prediction of functional potential

The functional potential of the microbial communities was inferred based on the abundance of representative sequences using PICRUSt2 [21], which is software for predicting functional abundances based only on marker gene sequences. The MetaCyc Metabolic Pathway Database [22], a curated database of experimentally elucidated metabolic pathways, was used for functional assignment. The functional potential of the microbial communities was determined by the abundance distribution of each MetaCyc metabolic pathway in each sample.

2.8. Transcriptome sequencing and analysis

Transcriptome sequencing was performed on paired tumor and adjacent non-tumor tissue samples from 203 patients with PTC ($n = 148$) and BTN ($n = 55$). The adapter sequences of raw sequencing reads were removed using Cutadapt v3.4 [23]. FastQC v0.11.9 was then used to evaluate the quality of the sequencing reads, and low-quality reads were discarded. Filtered reads were mapped to the human reference genome vGRCh37 using the BWA package [24] with default parameters. The expression count value of the gene was calculated using featureCounts [25] with the default parameters. Differential expression analysis was carried out on the count data using DESeq2 [18]. An adjusted $P < 0.05$ and $|\log_2 FC| > 1$ (FC: fold change) were considered to be statistically significant.

2.9. Integrative analysis of microbial profiling and host genomic gene expression

A total of 356 samples with paired transcriptome and 16S rRNA gene sequencing data were selected (130 PTC-T, 129 PTC-NT, 46 BTN-T, and 51 BTN-NT). Bacteria–genes association analyses were conducted based on the bacteria OTU table and genes count table with a Pearson algorithm. Significant positive and negative interactions were selected using the igraph *R* package and then visualized using Cytoscape 3.8.2 [20]. Heatmaps were generated using the ComplexHeatmap *R* package [26]. Genes positively related to bacteria were selected for gene ontology (GO) enrichment analysis using the clusterprofiler *R* package [27].

2.10. Bacteria immunostaining and Gram staining

A total of 30 formalin-fixed paraffin-embedded (FFPE) thyroid tissue samples were subjected to bacteria immunostaining and Gram staining. The FFPE slides were stained for bacterial lipopolysaccharide (LPS) (monoclonal antibody (mAb) WN1 222-5; #HM6011; 1:150 dilution; HycultBiotech, Netherlands) and bacterial lipoteichoic acid (LTA) (mAb 55; #HM2048; 1:100 dilution; HycultBiotech) [28,29]. The slides were subsequently probed with a secondary antibody (#GK500705; Dako System, Denmark). Bacterial Gram staining was performed using a Gram Stain Kit (Modified Brown & Brenn; #BTN160634; BaiAoLaiBo, China). Slides were inspected using an OLYMPUS BX63F #9072693 automated fluorescence microscope at 40× and 100×.

2.11. *Fluorescence* in situ *hybridization*

rRNA fluorescence *in situ* hybridization (FISH) was used to validate the presence of bacterial DNA in the human thyroid tumor tissue. The FFPE samples were deparaffinized and rehydrated. A bacterium FISH pretreatment kit (#D-0015; Focobio, China) was used according to the manufacturer's instruction. A cyanine 3 (Cy3)-labeled universal bacterial probe (EUB338-GCTGCCTCCC GTAGGAGT) was used [30]. Staining was visualized and observed using a ZEISS #LSM880 confocal laser-scanning microscope at 40× and 100×.

2.12. Correlative light and electron microscopy

Fresh tumor tissues were collected under sterile conditions during surgery and stored on ice immediately for further processing. Each tissue was sliced into 100 μm by means of a vibratome (catalog no. VT1200S; Leica, Germany) and immersed in cryoprotectant solution (25% sucrose and 10% glycerol in 0.1 mol·L^{-1} PBS) for 1 h. Liquid nitrogen was used for a quick freeze–thaw of the tissue slices in cryoprotectant solution, which was repeated three times. The tissues were then blocked by 20% goat serum in PBS for 2 h at room temperature. Tissue slices were incubated with LPS overnight at 4 °C. After washing with PBS, tissue slices were incubated in horseradish peroxidase conjugated secondary antibody for 2 h at 37 °C. Post-fixation of the tissues was done using 1% glutaraldehyde in 0.1 mol·L^{-1} phosphate buffer (PB) for 10 min. A 3–3′ diaminobenzidine (DAB) staining kit (#GK500705; Dako System) was used for DAB staining. After DAB staining, each tissue slice

[†] https://www.originlab.com/.

was subjected to osmium tetroxide and gradient dehydration, and was embedded with Epon812 (Ted Pella, USA). Finally, a transmission electron microscope was used to screen the ultrathin tissue sections. This protocol was optimized based on a previous study [31].

2.13. Statistical analysis

The statistical significance of the differences in alpha diversity, Bray–Curtis distance, risk index, and *r* were determined using a Wilcoxon rank-sum test, Student's two-sample *t*-test, or Kolmogorov–Smirnov tests in *R* 4.0.3. Differentially abundant functional compositions of the microbial communities were analyzed using a linear discriminant analysis effective size (Lefse) algorithm [32]. Correlations of the differentially abundant taxa and clinical parameters related to HT diagnosis (TPO-Ab and Tg-Ab) were calculated with the Spearman method, and a smoother line was added with the linear smoothing method in *R* 4.0.3. $P < 0.05$ was considered to be statistically significant. For multiple comparisons, *P* values were adjusted with the false discovery rate (FDR) method.

3. Results

3.1. Clinical characteristics of the study cohorts

To investigate the tumor-associated microbial community, 340 patients including 259 patients with PTC and 81 patients with BTN were enrolled in this study. Among the enrolled subjects, 48 PTC and 18 BTN patients did not meet the selection criteria and were excluded. A final total of 457 tissue biopsies, including 170 PTC-T, 176 PTC-NT, 53 BTN-T, and 58 BTN-NT samples, were included for 16S rRNA gene sequencing. In parallel, RNA-seq was performed on 356 tissue biopsies (130 PTC-T, 129 PTC-NT, 46 BTN-T, and 51 BTN-NT) for a bacteria-genes association analysis (Fig. 1 and Fig. S1(a)). The patients with PTC or BTN were mainly females with a median age of 37.08 and 44.74 years, respectively. Among the PTC patients, 73.7% had a tumor greater than 1 cm in size and 66.8% had lymph node metastasis. The majority of the PTC patients were diagnosed in an early stage (Tumor Node Metastasis Classification (TNM) stages I/II), and only 2.5% were diagnosed in advanced stages (TNM stages III/IV). In addition, 65 patients were diagnosed with HT. The demographic and clinical details of samples are shown in Table S1 in Appendix A.

3.2. Taxonomic profiles of tumor-associated bacteria

We first explored the diversity of the tumor-associated microbial community by profiling microbial 16S rRNA gene sequences. Despite the low microbial biomass in the samples, the sequencing data yielded a high mean quality score (Fig. S1(b) in Appendix A). The overall alpha diversity (Shannon entropy index) was lower in PTC than in BTN ($P = 0.0002$). However, the difference in alpha diversity between the tumor and the corresponding adjacent normal tissues was not significant in PTC and BTN (Figs. 2(a) and (b)). We further evaluated the correlation between clinical parameters (e.g., age, gender, tumor size, lymph node metastasis, extrathyroidal invasion, and TNM stage) and alpha diversity and found that there was a significant difference in alpha diversity between the age groups "35–45" and "45–55" in PTC-T ($P < 0.05$) (Figs. S2(a)–(f) in Appendix A). The PCoA based on the unweighted UniFrac showed that the microbial composition between PTC and BTN was markedly different, based either on tumor samples alone ($P = 0.0003$; PERMANOVA) (Fig. 2(c)), or tumor and adjacent normal tissues combined ($P = 0.00005$; PERMANOVA) (Fig. 2(d)). Consistent with the alpha diversity, we did not observe any significant difference in microbial composition (beta diversity) between tumor and adjacent normal tissues in both PTC and BTN (Figs. S2(g) and (h) in Appendix A). RDA was performed to explore the correlation of clinical factors contributing to microbiota diversity. We found that PTMC ($P = 0.009$) was associated with microbiota diversity in PTC (Fig. 2(e)). Microbiota diversity in BTN was associated with age ($P = 0.04$) (Fig. 2(f)). Moreover, we observed significant differences in microbiota diversity between samples with different tumor sizes (PTMC and nPTMC) in PTC-T and PTC-NT (Fig. 2(g)), as well as between samples in different age groups (Young and Old) in BTN-T and BTN-NT (Fig. 2(h)).

3.3. Validation of microbes in human thyroid tissues

To confirm the presence of bacteria in the thyroid tissue, we first compared our results with a microbiome analysis of CRC and HCC tissues. On average, 36 genera were detected per PTC sample, which was significantly fewer than the number detected in HCC tissue (78 genera per sample) and CRC tissue (230 genera per sample) (Fig. 3(a)). The microbiota diversity of PTC was also significantly lower than those of HCC and CRC (Fig. 3(b)), supporting our assumption that PTC has smaller bacteria populations than the other two tumor types. The presence of bacteria within the tissues of PTC and BTN was verified by Gram staining and immunohistochemistry using bacterial LPS-specific and LTA-specific antibodies (Fig. 3(c)). FISH using a universal bacterial probe and antisense probes as the control further confirmed the presence of bacterial DNA in both the PTC-T and the BTN-T (Fig. 3(d) and Fig. S3 in Appendix A). In addition, through immunoelectron microscopy, we determined that the bacteria were mainly distributed in the cytoplasm of the PTC-T and BTN-T samples (Fig. 3(e)). These findings collectively confirmed the presence of bacteria in both the PTC-T and the BTN-T.

3.4. Tissue microbiome dysbiosis in PTC compared with BTN

We next studied the microbial composition at different taxonomic levels. At the phylum level, Proteobacteria and Actinobacteria were the dominant phyla, collectively representing more than 95% of the microbial communities in PTC and BTN (Fig. 4(a)). At the genus level, *Mycobacterium* (23.28% in PTC-T; 24.92% in BTN-T), *Phyllobacterium* (12.71% in PTC-T; 14.10% in BTN-T), and *Rhodococcus* (12.19% in PTC-T; 6.32% in BTN-T) were the top three most abundant genera, accounting for the major portion of bacteria in PTC and BTN (Fig. 4(b)). After adjusting for potential confounding factors (i.e., age, gender, HT, and LNMC), our differential abundance analysis showed that 17 bacterial taxa from three phyla differed significantly between PTC-T and BTN-T (Fig. 4(c)). Nine bacterial taxa were significantly enriched in PTC-T compared with BTN-T, including *Brevibacterium*, *Neisseria*, *Rhodococcus*, *Staphylococcus*, *Streptococcus*, *Halomonas*, *Corynebacterium*, *Acinetobacter*, and *Devosia*. Of these, *Neisseria*, *Rhodococcus*, *Streptococcus*, *Halomonas*, and *Devosia* have been previously reported to be tumor-promotive pathogens [33–36]. On the other hand, eight bacterial taxa, including tumor-protective *Amycolatopsis* [37], were found to be depleted in PTC-T compared to BTN-T. *Amycolatopsis* was the most significantly depleted bacterial taxa (Fig. 4(c)). By comparing the microbiomes of the tumor tissue and the adjacent normal tissue, we observed that the abundances of 13 bacterial taxa were significantly different. Nine bacterial genera, including the tumor-promotive bacteria *Neisseria*, *Halomonas*, *Streptococcus*, and *Devosia*, were enriched in PTC-T compared with PTC-NT, while the four protective bacteria taxa of *Brevundimonas*, *Lactobacillus*, *Prevotella*, and *Phaeospirillum* were depleted in PTC-T compared with PTC-NT (Fig. 4(d)). Taken together, these results indicate a

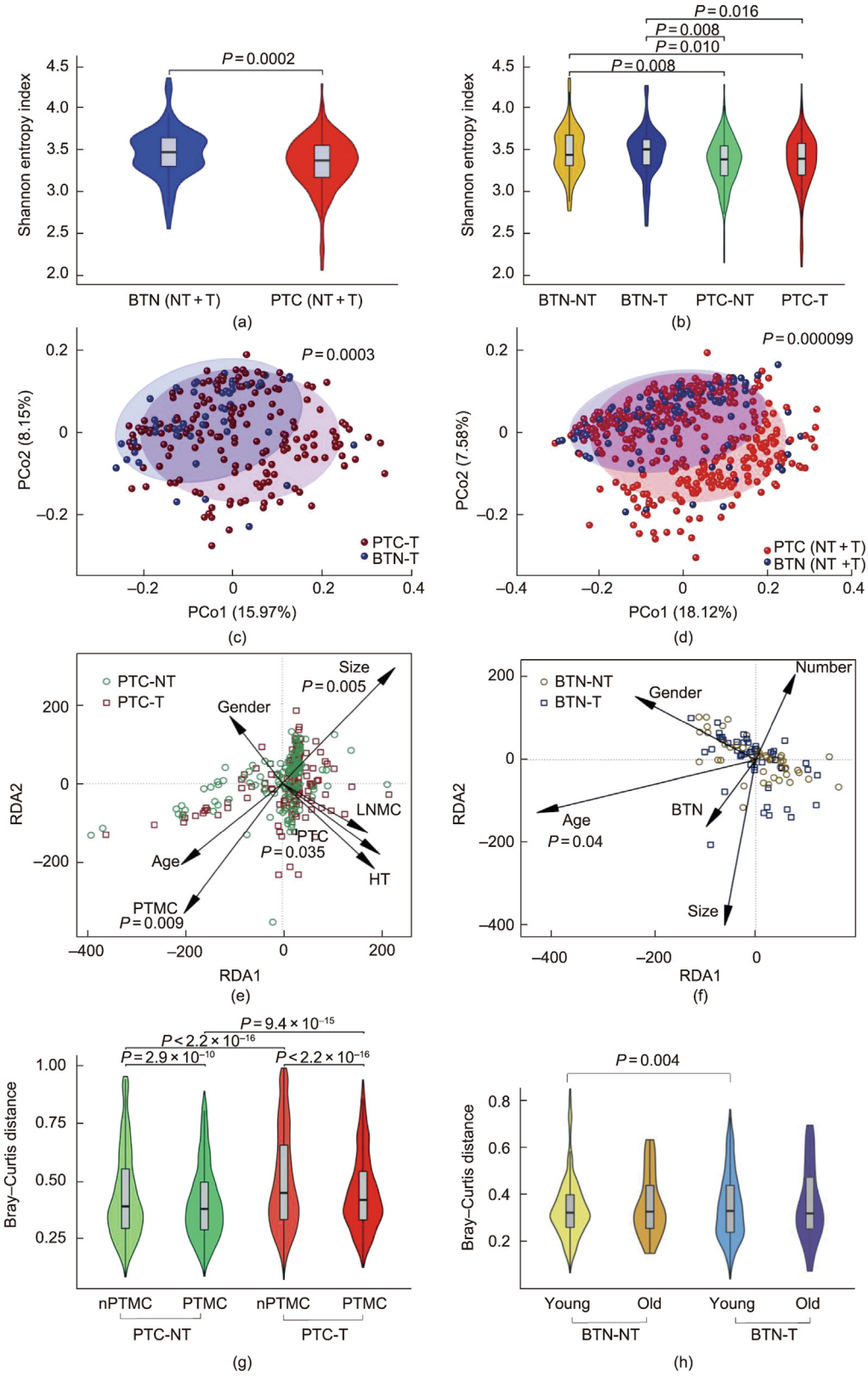

Fig. 2. Overall features of thyroid microbiota communities. Alpha diversity was evaluated by the Shannon entropy index between (a) overall PTC-T and BTN-T; and (b) PTC-T and PTC-NT, and BTN-T and BTN-NT. A beta diversity analysis was done via PCoA on an unweighted UniFrac distance matrix in (c) PTC-T and BTN-T, and (d) overall PTC-T and BTN-T. (e, f) A RDA was used to evaluate the importance of clinical traits (gender, age, tumor size, PTC, BTN, HT, PTMC, and LNMC). An association was found for PTMC and age in PTC and BTN, respectively. (g, h) The influence of PTMC and age on the PTC and BTN microbiomes was evaluated using the Bray–Curtis distance.

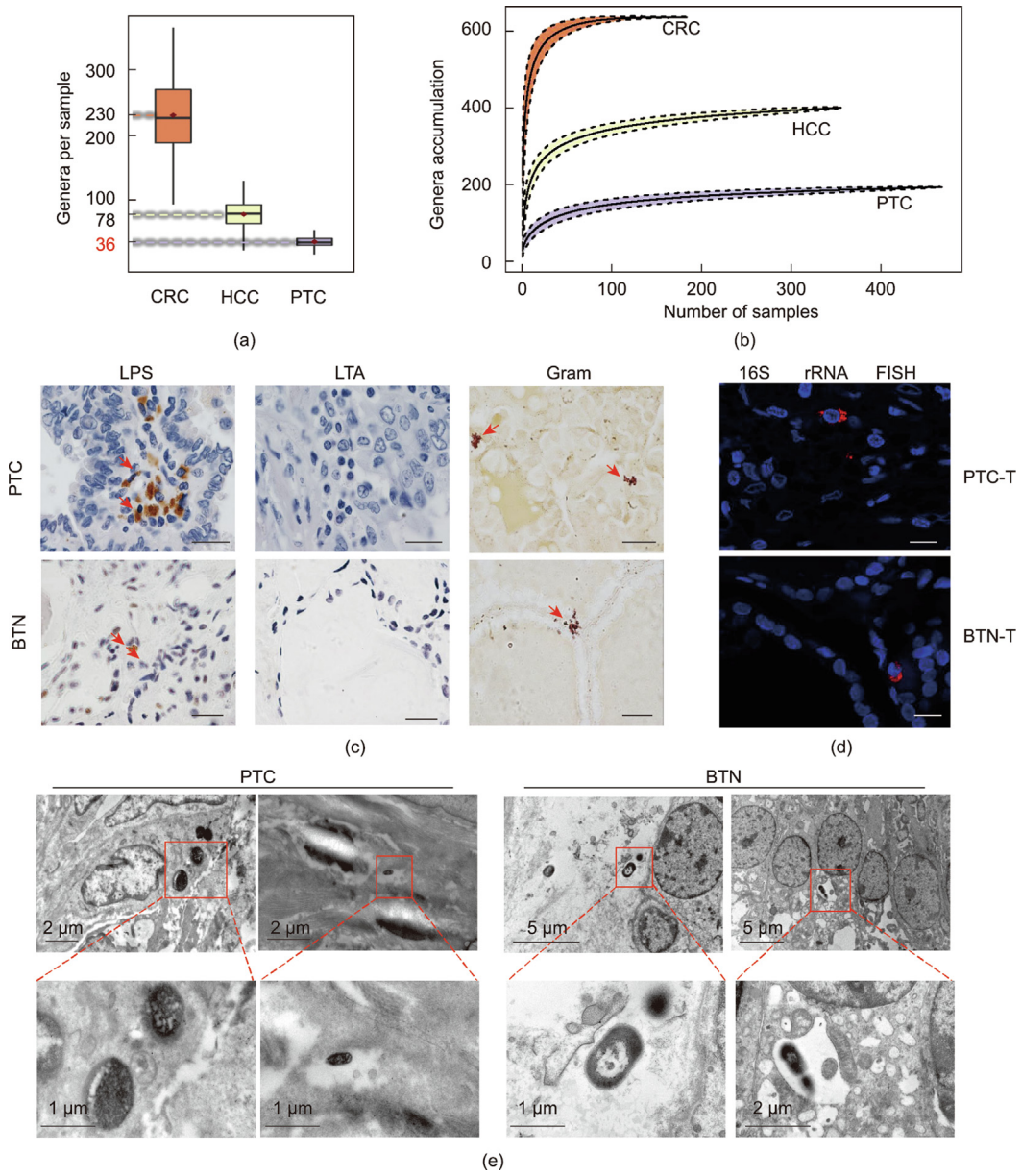

Fig. 3. Bacterial components detected in human thyroid tissue. (a) The number of bacteria genera found per tissue sample for PTC, HCC, and CRC. Dotted lines indicate the average number of bacteria genera in each group. (b) Comparison of the genera accumulation curves for PTC, CRC, and HCC. (c) Bacterial LPS, LTA, and Gram staining were demonstrated in PTC-T and BTN-T. The arrows point to positive staining for PTC-T and BTN-T (scale bar: 20 μm). (d) Bacterial 16S rRNA gene sequences were detected using FISH in PTC-T and BTN-T (red) and cell nuclei stained with 4′,6-diamidino-2phenylindole (blue) (scale bars: 10 μm). (e) Transmission electron microscope images demonstrating intracellular bacteria in PTC-T and BTN-T. The boxed area marks the region selected for higher magnification, which is shown in the second row.

state of dysbiosis, indicated by enriched pathogenic bacteria taxa and depleted protective bacteria taxa, in the tissue microbiome of the patients with PTC.

3.5. Microbial ecological network differs in BTN and PTC

To evaluate the interactions of the microbial taxa within and across different tissue types, we calculated the taxa's pair-wise correlations using SparCC. The distributions of taxonomic pairs with co-occurrence ($r > 0$) or co-excluding ($r < 0$) interaction were significantly different among the tissue types of BTN and PTC ($P < 0.05$) (Fig. S4 in Appendix A).

We then determined the relationships among the differentially abundant bacteria in PTC compared with BTN (nine enriched and eight depleted) in the microbial ecological networks. We observed that both co-occurrence and co-excluding interactions among PTC-enriched and PTC-depleted bacterial taxa were significantly different across the benign and malignant thyroid nodules (Fig. 4(e)). The positive correlations among the enriched bacterial taxa were stronger in PTC-T than in BTN-T ($P < 0.05$). Co-excluding patterns were also observed between the enriched and depleted taxa in PTC-T (Fig. 4(e)).

3.6. Microbial dysbiosis markers differentiate PTC-T from BTN-T

To determine whether the PTC differentially abundant bacteria could be used to differentiate between PTC-T and BTN-T, we performed ROC analyses. The samples were randomly split into training and validation cohorts. The 17 bacteria markers (Fig. 4(c)) distinguished PTC-T from BTN-T with an area under the curve (AUC) of 81.66% (95% confidence interval (95% CI): 74.3%–89.01%) in the training cohort and an AUC of 77.29% (95% CI: 62.81%–91.78%) in the validation cohort (Fig. 4(f)).

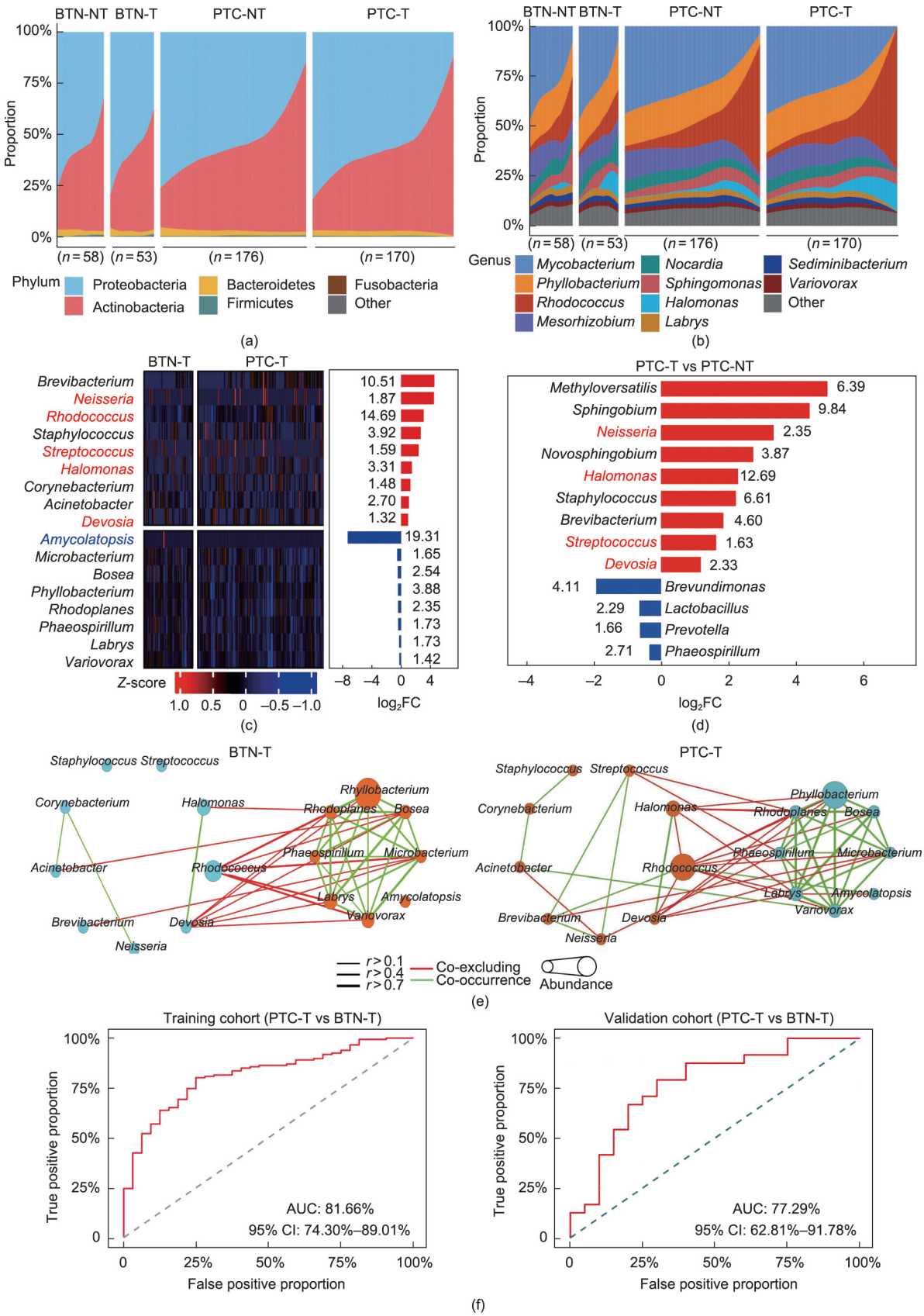

Fig. 4. Microbial compositions and differential abundances between PTC and BTN. Bacterial distribution for the (a) top five bacteria phyla and (b) top ten bacteria genera in PTC and BTN. (c) Heatmap showing the relative fold change of the differentially abundant bacteria genera of PTC-T vs BTN-T. (d) The relative fold change of the differentially abundant bacteria genera of PTC-T vs PTC-NT. The red and blue genera names represent the tumor-promoting and consistent bacteria reported in the previous literature, respectively. (e) Correlation network of PTC-T-enriched and PTC-T-depleted bacteria in PTC-T and BTN-T. *r* were estimated using the SparCC algorithm. Significant ($P < 0.05$) co-occurrence and co-excluding interactions with strengths of at least 0.1 were selected for visualization using Cytoscape 3.8.2. (f) ROC analysis based on differential abundance bacteria for discriminating PTC-T from BTN-T.

3.7. Microbiome functional capacity changes in PTC-T compared with BTN-T

Based on the differential microbial abundance and ecological networks between PTC-T and BTN-T, the functional capabilities of the tissue-associated microbiome were predicted using Phylogenetic Investigation of Communities by Reconstruction of Unobserved States 2 (PICRUSt2) [38] and then subjected to discrimination analysis using the Lefse algorithm. The functional changes in the PTC-T microbiomes included a significantly increased B12 synthesis pathway (adenosylcobalamin salvage from cobinamide), while the depleted pathways included protein/amino acid biosynthesis (aspartate superpathway; superpathway of *L*-methionine biosynthesis; superpathway of *L*-lysine, *L*-threonine, and *L*-methionine biosynthesis), fatty acid biosynthesis (palmitate biosynthesis), and purine nucleotide biosynthesis (5-aminoimidazole ribonucleotide biosynthesis II) (Fig. S5 in Appendix A).

3.8. Microbial profiles of PTC associated with HT

HT is an autoimmune disease and has been shown to have an association with PTC [39]. We assessed the microbial diversity and richness of PTC tumor samples from 65 PTC-HT and 105 PTC-nHT patients. There were no significant differences in microbial diversities (alpha and beta diversity) between PTC-HT and PTC-nHT (Figs. S6(a) and (b) in Appendix A). PTMC was associated with microbiota diversity in PTC-HT and PTC-nHT patients ($P < 0.05$) (Fig. S6(c) in Appendix A). We observed a significant difference in microbiota diversity between samples with different tumor sizes (PTMC and nPTMC) in PTC-nHT patients ($P = 1.9 \times 10^{-14}$) (Fig. S6(d) in Appendix A). After adjusting for potential confounding factors, the abundances of 14 bacterial taxa were found to be significantly different between PTC-HT and PTC-nHT, including eight bacterial taxa being enriched in PTC-HT (*Streptomyces*, *Blastomonas*, *Erwinia*, *Erysipelothrix*, *Sphingobium*, *Bacillus*, *Acinetobacter*, and *Ralstonia*) and six bacterial taxa being depleted (*Candidatus Xiphinematobacter*, *Brevibacterium*, *Staphylococcus*, *Rhodococcus*, *Stenotrophomonas*, and *Sphingomonas*) (Fig. S6(e) in Appendix A). We observed that both the co-occurrence and co-excluding interactions among these PTC-HT-enriched and PTC-HT-depleted taxa were significantly different from those in PTC-nHT (Fig. S6(f) in Appendix A). Moreover, these 14 differential bacteria taxa could be used to distinguish PTC-HT from PTC-nHT with an AUC of 82.33% (95% CI: 75.33%–89.32%) in the training cohort, which was validated in the validation cohort with an AUC of 80.22% (95% CI: 65.23%–95.21%) (Fig. S7(a) in in Appendix A). The functional enrichment analysis revealed that pathways involved in amino acid biosynthesis (e.g., glutamine) and starch metabolism were significantly enriched in PTC-HT (Fig. S7(b) in Appendix A).

The correlation between the differentially abundant taxa in PTC-HT and clinical parameters related to key HT diagnostic factors revealed that *Ralstonia* was positively correlated with anti-Tg-Ab ($r = 0.26$, $P = 0.0026$) and TPO-Ab ($r = 0.23$, $P = 0.0029$), while *Stenotrophomonas* was negatively correlated with Tg-Ab ($r = -0.21$, $P = 0.014$) and TPO-Ab ($r = -0.22$, $P = 0.0048$) (Fig. S8 in Appendix A). Collectively, these findings indicate that there is potential that tissue microbial dysbiosis contributes to PTC-HT.

3.9. Thyroid tumor microbiome is associated with host gene expression in PTC and BTN patients

To understand the potential interplay between microbiome profiling and host differential gene expression, we performed parallel microbiome profiling and transcriptional sequencing from the same set of samples and analyzed their association using Pearson's correlation-based analysis. We observed that the distributions of both positive and negative associations among the bacteria and differentially expressed genes were significantly different between PTC-T and PTC-NT (Fig. 5(a-I)), as well as between BTN-T and BTN-NT (Fig. 5(b-I)). The bacteria found in tumor/nodule tissues had more association with the differentially expressed genes than those in non-tumor/non-nodule tissues. *Succiniclasticum* and *Agrobacterium* were the top two bacteria genera that showed significant associations with differentially expressed genes in PTC-T compared with PTC-NT (Fig. 5(a-II)), while *Bacillus* and *Mycobacterium* were the top two bacteria genera associated with differentially expressed genes in BTN-T compared with BTN-NT (Fig. 5(b-II)). The differential bacteria–host associations between the thyroid tumor tissues and adjacent non-tumor tissues imply a potential contribution of pathogenic bacteria to thyroid tumorigenesis.

3.10. Microbial dysbiosis shows significant interaction with oncogene expression in PTC-T

Since 17 bacterial taxa were differentially abundant in PTC-T vs BTN-T (Fig. 4(c)), we next assessed the association of microbiome dysbiosis with host differentially expressed genes. We found that six dysbiosis bacteria taxa were significantly associated with differential gene expression in PTC compared with BTN, including 45 up-regulated genes and 16 down-regulated genes (Fig. 6(a) and Fig. S9 in Appendix A). The enriched bacterial taxa *Halomonas* and *Neisseria* were predominantly positively correlated with up-regulated oncogenes, including dual specificity phosphatase 4 gene (*DUSP4*), dual specificity phosphatase 6 gene (*DUSP6*), four-point-one ezrin-radixin-moesin homology domain-containing protein 5 gene (*FRMD5*), transforming growth factor alpha gene (*TGFA*), and nucleotide-binding oligomerization domain-containing protein 1 gene (*NOD1*), whose function as oncogenes has been well reported [40–44]. The enriched bacterial taxa *Rhodococcus* was positively correlated with the oncogenes myosin binding protein H gene (*MYBPH*) and leucine-rich repeat-containing G-protein-coupled receptor 5 gene (*LGR5*) [45,46], and negatively correlated with the tumor-suppressing gene secretin receptor gene (*SCTR*) [47] (Fig. 6(a)). Among the 17 bacterial taxa, *Rhodococcus*, *Brevibacterium*, *Devosia*, and *Halomonas* had the most significant associations with differentially expressed genes in PTC-T compared with PTC-NT (Fig. 6(b)), while *Rhodococcus*, *Amycolatopsis*, *Neisseria*, and *Streptococcus* had the most significant associations with differentially expressed genes in BTN-T compared with BTN-NT (Fig. 6(c)). We then explored the association of microbiome dysbiosis with gene expression in detail. As shown in Fig. 6(d), *Rhodococcus* exhibited strong positive interactions with the oncogenes *BRAF*, kirsten rat sarcoma viral gene (*KRAS*), interleukin-1 receptor-associated kinase 4 gene (*IRAK4*), catenin beta 1 gene (*CTNNB1*), phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit alpha gene (*PIK3CA*), mitogen-activated protein 3 kinase 7 gene (*MAP3K7*), and epidermal growth factor receptor gene (*EGFR*) in PTC-T; these are key genes in the oncogenic signaling pathways of MAPK, PI3K–AKT and Wnt/β-catenin, which are well-known signaling pathways in PTC. *Brevibacterium*, *Devosia*, and *Halomonas* were also positive associated with the majority of these oncogenic pathway-related genes in PTC-T (Fig. 6(d)). These results suggest that PTC-enriched bacteria may contribute to PTC development by inducing host oncogene expression and the activation of oncogenic signaling pathways.

4. Discussion

In this study, we demonstrated the presence of and characterized the microbes in tumor and adjacent normal tissues from PTC and BTN patients using 16S rRNA gene sequencing. As

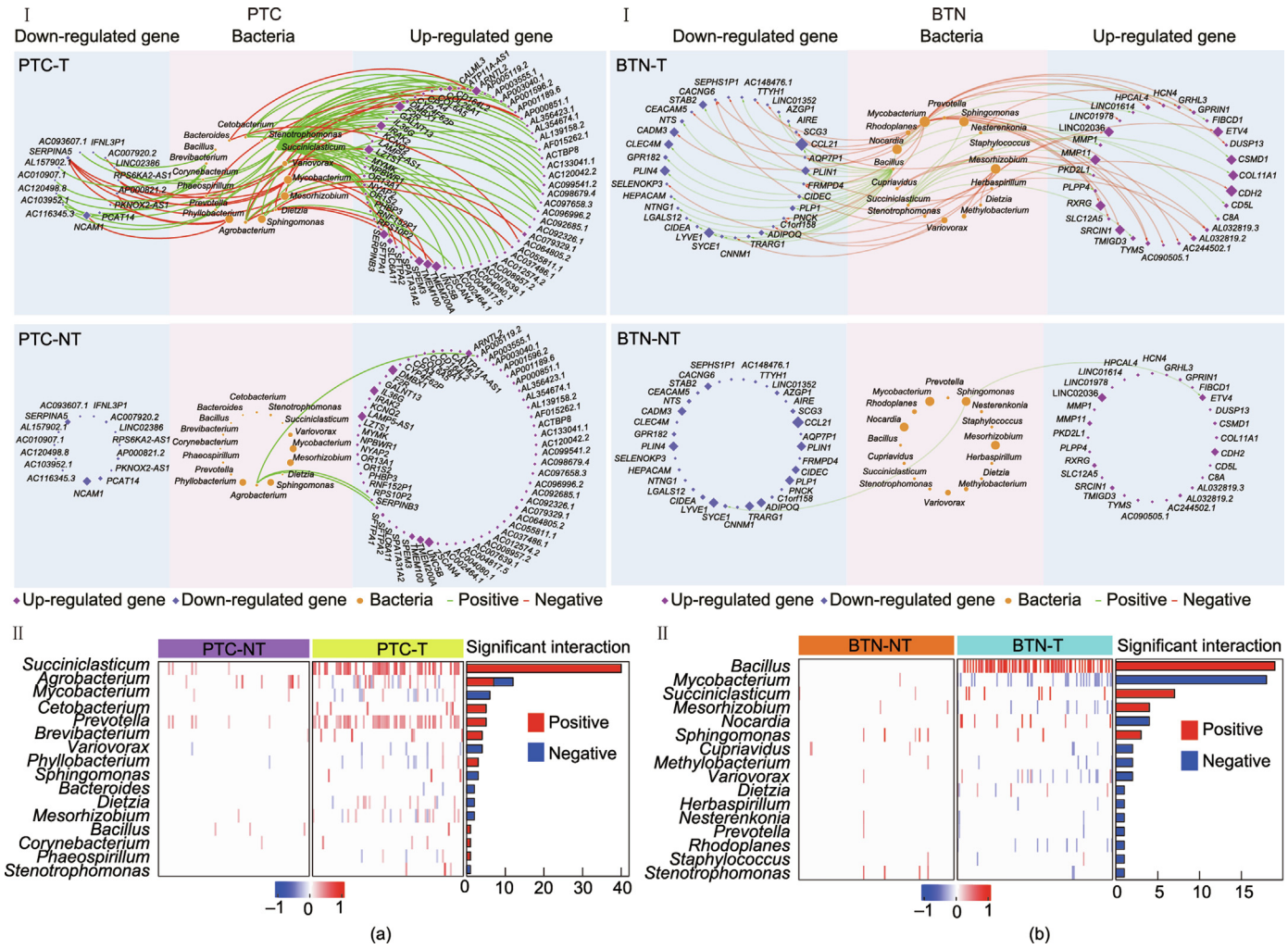

Fig. 5. Correlation of thyroid tumor transcriptome with tissue microbiota. (a) Interaction network of the differentially expressed genes (PTC-T vs PTC-NT) with the microbiota. (b) Interaction network of the differentially expressed genes (BTN-T vs BTN-NT) with the microbiota. Both positive and negative interactions among the differentially expressed genes and associated microbiota were stronger in (a-I) PTC-T than PTC-NT and (b-I) BTN-T than BTN-NT. Heatmaps were appended to the above microbiota-gene interaction network in (a-II) PTC-T and PTC-NT and (b-II) BTN-T and BTN-NT.

expected, the abundance and richness of the PTC-associated microbes were significantly lower and less diverse (36 bacteria genera) than those observed in CRC (230 bacterial genera) and HCC (78 genera) [48]. The intratumoral bacteria were further confirmed by Gram staining, immunobiological staining, FISH, and transmission electronic microscopy. Our study demonstrated the presence of bacteria and, for the first time, confirmed the signature bacteria in thyroid tumor tissues, as identified by multiple approaches.

We observed a significant difference in the microbiota diversities between PTC and BTN, indicating a possible shift in host homeostasis/physiological condition over the course of tumor progression, which subsequently exerts a selection pressure on tumor-associated microbial communities. The association between tumor size and microbiota diversity in PTC further suggests a link between host homeostasis/physiological condition and the composition of tumor-associated microbiota. We noticed no significant difference in the overall composition (beta diversity) of the microbial communities of tumor versus adjacent normal tissues of PTC. This finding aligns with a previous study that reported that the microbial composition of the tumor was similar to that of the para-tumor; in contrast, a difference in bacterial abundance was identified in the analysis of TCGA data [10]. As the thyroid is an encapsulated organ with rich vascular flow, we speculate that bacteria may penetrate into the thyroid via blood circulation, resulting in a relatively homogeneous microbial composition throughout the entire thyroid gland. Moreover, due to the blood circulation, the close distance between the tumor and para-tumor samples we collected might also be a possible reason explaining this result. Taking together the lack of difference in the overall composition of the microbial communities of tumor versus adjacent normal tissues of PTC and the significant difference in microbiota diversity between benign and malignant thyroid nodules suggests the potential relevance of the microbiome composition in mediating thyroid cancer progression.

We found 17 differentially abundant bacterial taxa between PTC-T and BTN-T. The enriched intratumor *Neisseria* and *Streptococcus* in PTC in this study have been previously reported to be enriched in the gut of PTC patients [49]. Since microbes can be detected in the blood of healthy subjects [50], we speculate that gut microbes might enter the blood circulation and colonize the thyroid tissues due to impaired gut barrier function, resulting in an elevated risk of PTC [51]. Previously, a study showed that *Neisseria* could evade the surveillance of the host adaptive immune system, stimulating the infiltration of neutrophils and causing inflammation—a risk factor of thyroid cancer [52]. In keeping with

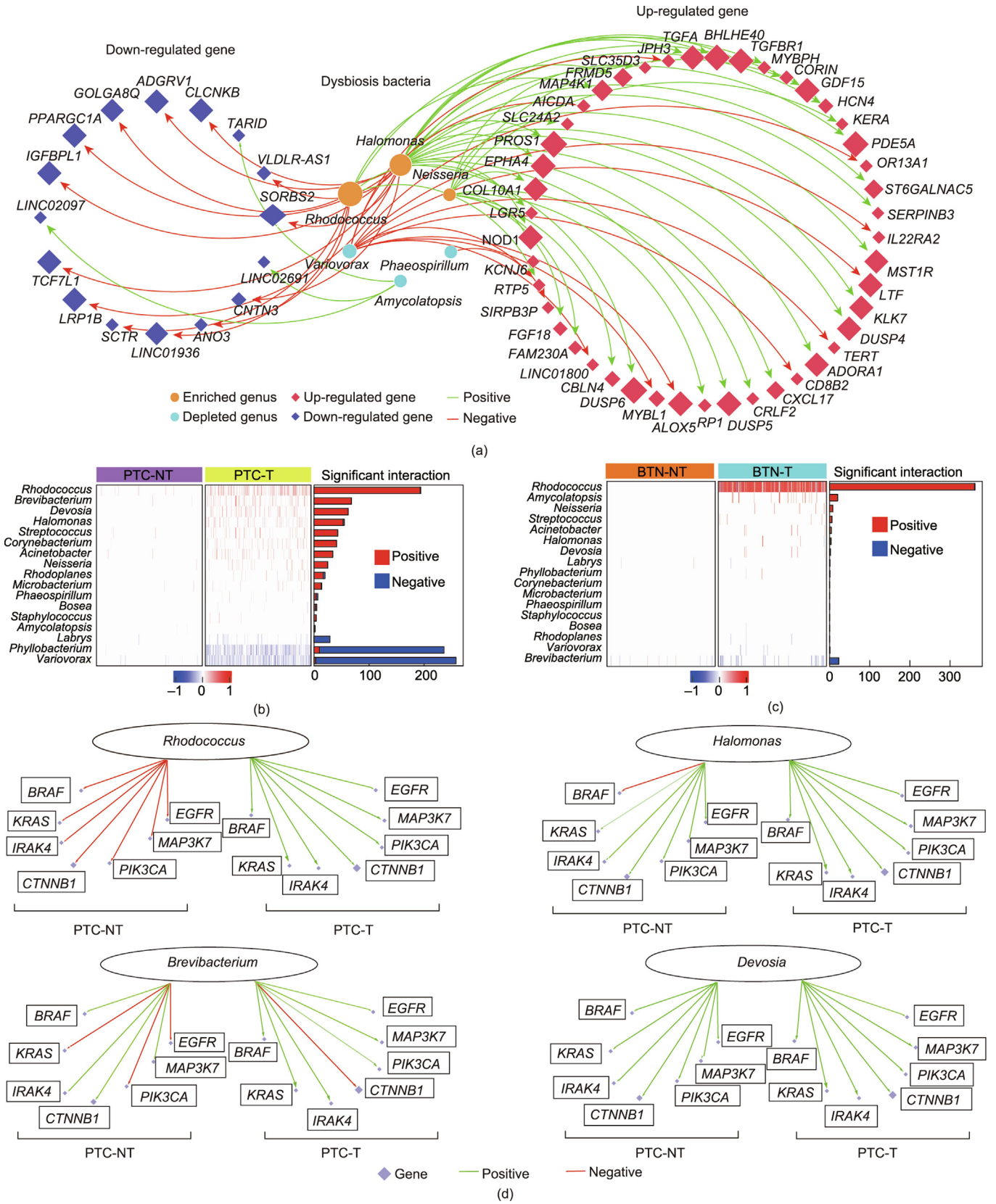

Fig. 6. Microbial dysbiosis interaction with gene expression. (a) Interaction network of the tissue microbiota with different abundances with differentially expressed genes in PTC-T and BTN-T. The PTC-enriched bacterial taxa *Halomonas*, *Neisseria*, and *Rhodococcus* and the PTC-depleted bacterial taxa *Amycolatopsis*, *Phaeospirillum*, and *Variovorax* showed significant associations with differentially expressed genes. (b) Heatmap appended to the microbiota–gene interaction in PTC-T and PTC-NT. (c) Heatmap appended to the microbiota-gene interaction in BTN-T and BTN-NT. (d) Interactions between bacteria involving *Rhodococcus*, *Brevibacterium*, *Devosia*, and *Halomonas* and *BRAF*, *KRAS*, *IRAK4*, *CTNNB1*, *PIK3CA*, *MAP3K7*, and *EGFR* are depicted separately in the PTC-T and PTC-NT groups. Green: positive association; Red: negative association.

our findings, a high abundance of *Rhodococcus* has been reported in prostate cancer [35] and *Halomonas* in lung cancer [34]. *Rhodococcus* is an aerobic and opportunistic pathogen from the phylum Actinobacteria, which has predominantly been detected in immunocompromised patients and thyroid abscesses [53]. *Devosia* is significantly enriched in the gut microbiota of CRC patients [36]. On the other hand, the depleted bacterium *Amycolatopsis* has been reported to have an anticancer effect in breast cancer through the synthesis of thioalbamide, which exhibits an anti-proliferative effect on cancer cells [37]. Collectively, the differential abundances of these microbiomes are associated with PTC, possibly through interaction with the host cells and the activation of oncogenic signaling pathways.

As specific microbial networks can reflect a disease-specific microenvironment, we examined the ecological networks of the associated microbiome in PTC formation. We observed increasing strengths of co-occurrence and co-excluding interactions among PTC-enriched and PTC-depleted bacteria from BTN to PTC. Co-excluding relationships were observed between PTC-enriched and PTC-depleted bacteria. The results of the microbial network analysis highlight the potential roles of these altered bacterial compositions as an important factor, at least, in the progression of thyroid carcinogenesis. The 17 bacterial markers may serve as potential diagnostic tissue markers for differentiating PTC from BTN, as demonstrated by the AUCs of 81.66% and 77.29% in the training and validation cohorts, respectively. Additional comparative analyses of PTC microbiome markers by means of quantitative PCR will be a step forward in the potential use of these bacteria for diagnostic purposes.

HT is an autoimmune disease characterized by a cellular immune response with lymphocytic infiltration of the thyroid gland. Although it is a known risk factor for PTC, the mechanism underlying this relationship remains unclear. We identified a significant association between microbiota diversity and HT status in PTC patients, as well as an ecological network shift between PTC-HT and PTC-nHT. The bacterial dysbiosis could differentiate PTC-HT from PTC-nHT, with AUCs of 82.33% in the discovery and 80.22% in the validation cohorts, indicating an association between intratumor bacteria dysbiosis and the HT status in PTC patients. Higher relative abundances of *Erwinia* and *Acinetobacter* were observed in PTC-HT compared with PTC-nHT. The abundances of these two genera have been reported to be increased in the gut microbiome of patients with autoimmune multiple sclerosis disease [54]. Also, increased abundance of *Bacillus* has been reported to promote the production of T helper 17 (Th17) cells, leading to autoimmune diseases [55]. In particular, we found that the key clinical HT immune parameters Tg-Ab and TPO-Ab have a positive correlation with *Ralstonia* and a negative correlation with *Stenotrophomonas*. *Ralstonia* can induce allergic disease by inhibiting T regulatory cell (Treg) differentiation, as reflected by the positive correlation of the Th2/Treg ratio with *Ralstonia* population [56], while *Stenotrophomonas* cannot promote the up-regulation of the host immune function but may lead to the exhaustion of T lymphocytes [57]. Taken together, these findings suggest that intratumor bacteria—at least in part—play an important role in PTC-HC by modulating host immune function.

To identify significant associations between PTC-associated bacteria and host gene expression, we performed concurrent 16S rRNA gene sequencing and RNA-seq from the same samples. We discovered novel associations between the PTC-enriched bacteria and transcriptome gene expression in PTC compared with BTN (Fig. 5). We characterized the PTC-enriched bacteria *Halomonas*, *Neisseria*, and *Rhodococcus* as potential PTC-promoting bacteria that affect the host transcriptome gene expression of PTC (Fig. 6(a)). *Halomonas* has been shown to be positively correlated with *DUSP6* in the regulation of extracellular signal-regulated kinase 1/2 (ERK1/2) (MAPK3/1)-dependent transcription, as a predictor of invasiveness [40]. *Halomonas* has also been shown to be positively correlated with FRMD5, a downstream effector of telomerase reverse transcriptase gene (*TERT*) that can enhance $BRAF^{V600E}$ mutation in thyroid cancer [42,58]. A strong link between *Neisseria* and *NOD1* expression was observed in PTC. *NOD1* has been linked to inflammatory pathologies and cancer metastasis modulated by the microbiota in subtle ways [43,59]. *Rhodococcus* has been shown to be positively correlated with the oncogene *MYBPH* [45] and negatively correlated with the tumor-suppressing gene *SCTR* [47]. Moreover, when compared to the key PTC cancer pathway-related genes, the PTC-enriched bacteria *Rhodococcus*, *Brevibacterium*, *Devosia*, and *Halomonas* exhibited strong positive interaction with critical pathway-associated genes including *BRAF*, *KRAS*, *IRAK4*, *CTNNB1*, *PIK3CA*, *MAP3K7*, and *EGFR* (Fig. 6(d)), which preferentially activate the MAPK, PI3K–AKT, and Wnt/β-catenin signaling pathways in thyroid tumorigenesis. Collectively, our findings suggest that altered microbiota pathogens—especially the PTC-enriched bacteria—are associated with aberrant tumor-related gene expression and the oncogenic signaling pathways of PTC.

5. Conclusions

In this work, we identified the presence of intratumoral microbial signatures in thyroid tissues. Bacteria dysbiosis and the associated microecological changes were found to be associated with PTC and to critically influence how the thyroid microenvironment is shaped, thereby contributing to thyroid carcinogenesis. A specific type of bacteria dysbiosis was found to be associated with PTC-HT patients, potentially through modulating the host immune function. Moreover, our integrative analysis revealed that the PTC-enriched bacteria had strong positive correlations with key oncogenes of oncogenic signaling pathways in PTC.

Our study has a few limitations. First, additional approaches such as *in vitro* bacteria culture and direct tracking of microbes in mice may be valuable for further identifying microbes in fresh tumor tissues. Second, in-depth exploration by animal experiments would be helpful for interpreting the causal correlation and direct interaction between intratumoral bacteria and PTC. Third, due to the current technical limitations of metagenomic sequencing, it remains highly challenging to achieve a complete microbial genome, especially in tumors with relatively low biomass. In conclusion, our findings provide novel insights into microbiome-relevant thyroid tumorigenesis.

Acknowledgments

We thank Zhen Cheng, Weiman He, and Yingtong Hou for help with the clinical database, Jinlang Wu for technical assistance with the transmission electron microscope, and Bo Lin for access to thyroid cancer samples. We appreciate the help from Dr. Olabisi O. Coker for language editing. This work was supported by the National Natural Science Foundation of China (81772850 and 82273300).

Authors' contribution

Shuang Yu and Xuejie Wang collected the clinical specimens, performed the experiments, and drafted the manuscript; Yanqiang Ding performed the bioinformatic analysis and drafted the manuscript; Siu Kin Ng, Weixin Liu, and Yubin Xie provided bioinformatic support and commented on the study; Siting Cao provided experimental support; Shubin Hong, Sui Peng, Lixia Xu, Xiaoxing Li, and Yanbing Li collected clinical data and commented on the study; Zhuming Guo, Jie Li, and Weiming Lv collected the clinical

specimens; Joseph JY Sung commented on the study; Jun Yu and Haipeng Xiao designed and supervised the study and revised the manuscript.

Compliance with ethics guidelines

Shuang Yu, Yanqiang Ding, Xuejie Wang, Siu Kin Ng, Siting Cao, Weixin Liu, Zhuming Guo, Yubin Xie, Shubin Hong, Lixia Xu, Xiaoxing Li, Jie Li, Weiming Lv, Sui Peng, Yanbing Li, Joseph JY Sung, Jun Yu, and Haipeng Xiao declare that they have no conflict of interest or financial conflicts to disclose.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.eng.2023.01.007.

References

[1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021;71(3):209–49.
[2] Cabanillas ME, McFadden DG, Durante C. Thyroid cancer. Lancet 2016;388(10061):2783–95.
[3] Nikiforov YE, Seethala RR, Tallini G, Baloch ZW, Basolo F, Thompson LD, et al. Nomenclature revision for encapsulated follicular variant of papillary thyroid carcinoma: a paradigm shift to reduce overtreatment of indolent tumors. JAMA Oncol 2016;2(8):1023–9.
[4] Xing M. Molecular pathogenesis and mechanisms of thyroid cancer. Nat Rev Cancer 2013;13(3):184–99.
[5] Lamartina L, Grani G, Durante C, Filetti S, Cooper DS. Screening for differentiated thyroid cancer in selected populations. Lancet Diabetes Endocrinol 2020;8(1):81–8.
[6] Caronia LM, Phay JE, Shah MH. Role of BRAF in thyroid oncogenesis. Clin Cancer Res 2011;17(24):7511–7.
[7] Feng J, Zhao F, Sun J, Lin B, Zhao L, Liu Y, et al. Alterations in the gut microbiota and metabolite profiles of thyroid carcinoma patients. Int J Cancer 2019;144(11):2728–45.
[8] Nejman D, Livyatan I, Fuks G, Gavert N, Zwang Y, Geller LT, et al. The human tumor microbiome is composed of tumor type-specific intracellular bacteria. Science 2020;368(6494):973–80.
[9] Link CD. Is there a brain microbiome?. Neurosci Insights 2021;16:26331055211018709.
[10] Gnanasekar A, Castaneda G, Iyangar A, Magesh S, Perez D, Chakladar J, et al. The intratumor microbiome predicts prognosis across gender and subtypes in papillary thyroid carcinoma. Comput Struct Biotechnol J 2021;19:1986–97.
[11] Dai D, Yang Y, Yang Y, Dang T, Xiao J, Wang W, et al. Alterations of thyroid microbiota across different thyroid microhabitats in patients with thyroid carcinoma. J Transl Med 2021;19(1):488.
[12] Liu CJ, Chen SQ, Zhang SY, Wang JL, Tang XD, Yang KX, et al. The comparison of microbial communities in thyroid tissues from thyroid carcinoma patients. J Microbiol 2021;59(11):988–1001.
[13] Yuan L, Yang P, Wei G, Hu X, Chen S, Lu J, et al. Tumor microbiome diversity influences papillary thyroid cancer invasion. Commun Biol 2022;5(1):864.
[14] Caturegli P, de Remigis A, Rose NR. Hashimoto thyroiditis: clinical and diagnostic criteria. Autoimmun Rev 2014;13(4–5):391–7.
[15] Andrews S. FastQC: a quality control tool for high throughput sequence data. Cambridge: Babraham Bioinformatics, Babraham Institute; 2010.
[16] Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol 2019;37(8):852–7.
[17] Risso D, Ngai J, Speed TP, Dudoit S. Normalization of RNA-seq data using factor analysis of control genes or samples. Nat Biotechnol 2014;32(9):896–902.
[18] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15(12):550.
[19] Friedman J, Alm EJ. Inferring correlation networks from genomic survey data. PLOS Comput Biol 2012;8(9):e1002687.
[20] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13(11):2498–504.
[21] Douglas GM, Maffei VJ, Zaneveld JR, Yurgel SN, Brown JR, Taylor CM, et al. PICRUSt2 for prediction of metagenome functions. Nat Biotechnol 2020;38(6):685–8.
[22] Caspi R, Billington R, Keseler IM, Kothari A, Krummenacker M, Midford PE, et al. The MetaCyc database of metabolic pathways and enzymes—a 2019 update. Nucleic Acids Res 2020;48(D1):D445–53.
[23] Kechin A, Boyarskikh U, Kel A, Filipenko M. cutPrimers: a new tool for accurate cutting of primers from reads of targeted next generation sequencing. J Comput Biol 2017;24(11):1138–43.
[24] Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25(14):1754–60.
[25] Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 2014;30(7):923–30.
[26] Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 2016;32(18):2847–9.
[27] Yu G, Wang LG, Han Y, He QY. clusterProfiler: an *R* package for comparing biological themes among gene clusters. OMICS 2012;16(5):284–7.
[28] Kühler W. Bacterial cell wall. In: Ghuysen JM, Hakenbeck R, editors. New comprehensive biochemistry. Amsterdam: Elsevier Science; 1994.
[29] Whitfield C, Trent MS. Biosynthesis and export of bacterial lipopolysaccharides. Annu Rev Biochem 2014;83(1):99–128.
[30] Amann RI, Binder BJ, Olson RJ, Chisholm SW, Devereux R, Stahl DA. Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. Appl Environ Microbiol 1990;56(6):1919–25.
[31] Wu GH, Shi HJ, Che MT, Huang MY, Wei QS, Feng B, et al. Recovery of paralyzed limb motor function in canine with complete spinal cord injury following implantation of MSC-derived neural network tissue. Biomaterials 2018;181:15–34.
[32] Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. Genome Biol 2011;12(6):R60.
[33] Farrell JJ, Zhang L, Zhou H, Chia D, Elashoff D, Akin D, et al. Variations of oral microbiota are associated with pancreatic diseases including pancreatic cancer. Gut 2012;61(4):582–8.
[34] D'Alessandro-Gabazza CN, Méndez-García C, Hataji O, Westergaard S, Watanabe F, Yasuma T, et al. Identification of halophilic microbes in lung fibrotic tissue by oligotyping. Front Microbiol 2018;9:1892.
[35] Alexeyev O, Bergh J, Marklund I, Thellenberg-Karlsson C, Wiklund F, Grönberg H, et al. Association between the presence of bacterial 16S RNA in prostate specimens taken during transurethral resection of prostate and subsequent risk of prostate cancer (Sweden). Cancer Causes Control 2006;17(9):1127–33.
[36] Zhang H, Chang Y, Zheng Q, Zhang R, Hu C, Jia W. Altered intestinal microbiota associated with colorectal cancer. Front Med 2019;13(4):461–70.
[37] Frattaruolo L, Fiorillo M, Brindisi M, Curcio R, Dolce V, Lacret R, et al. Thioalbamide, a thioamidated peptide from Amycolatopsis alba, affects tumor growth and stemness by inducing metabolic dysfunction and oxidative stress. Cells 2019;8(11):1408.
[38] Langille MG, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nat Biotechnol 2013;31(9):814–21.
[39] Resende de Paiva C, Grønhøj C, Feldt-Rasmussen U, von Buchwald C. Association between Hashimoto's thyroiditis and thyroid cancer 64,628 in patients. Front Oncol 2017;7:53.
[40] Lee JU, Huang S, Lee MH, Lee SE, Ryu MJ, Kim SJ, et al. Dual specificity phosphatase 6 as a predictor of invasiveness in papillary thyroid cancer. Eur J Endocrinol 2012;167(1):93–101.
[41] Ma B, Shi R, Yang S, Zhou L, Qu N, Liao T, et al. DUSP4/MKP2 overexpression is associated with BRAF(V600E) mutation and aggressive behavior of papillary thyroid cancer. Onco Targets Ther 2016;9:2255–63.
[42] Gaweł AM, Ratajczak M, Gajda E, Grzanka M, Paziewska A, Cieślicka M, et al. Analysis of the role of FRMD5 in the biology of papillary thyroid carcinoma. Int J Mol Sci 2021;22(13):6726.
[43] Jiang HY, Najmeh S, Martel G, MacFadden-Murphy E, Farias R, Savage P, et al. Activation of the pattern recognition receptor NOD1 augments colon cancer metastasis. Protein Cell 2020;11(3):187–201.
[44] DeglInnocenti D, Alberti C, Castellano G, Greco A, Miranda C, Pierotti MA, et al. Integrated ligand-receptor bioinformatic and *in vitro* functional analysis identifies active TGFA/EGFR signaling loop in papillary thyroid carcinomas. PLoS One 2010;5(9):e12701.
[45] Hosono Y, Yamaguchi T, Mizutani E, Yanagisawa K, Arima C, Tomida S, et al. MYBPH, a transcriptional target of TTF-1, inhibits ROCK1, and reduces cell motility and metastasis. EMBO J 2012;31(2):481–93.
[46] Zhan T, Ambrosi G, Wandmacher AM, Rauscher B, Betge J, Rindtorff N, et al. MEK inhibitors activate Wnt signalling and induce stem cell plasticity in colorectal cancer. Nat Commun 2019;10(1):2197.
[47] Kang S, Kim B, Kang HS, Jeong G, Bae H, Lee H, et al. SCTR regulates cell cycle-related genes toward anti-proliferation in normal breast cells while having pro-proliferation activity in breast cancer cells. Int J Oncol 2015;47(5):1923–31.
[48] Liu W, Zhang X, Xu H, Li S, Lau HC, Chen Q, et al. Microbial community heterogeneity within colorectal neoplasia and its correlation with colorectal carcinogenesis. Gastroenterology 2021;160(7):2395–408.
[49] Zhang J, Zhang F, Zhao C, Xu Q, Liang C, Yang Y, et al. Dysbiosis of the gut microbiome is associated with thyroid cancer and thyroid nodules and correlated with clinical index of thyroid function. Endocrine 2019;64(3):564–74.

[50] Moriyama K, Ando C, Tashiro K, Kuhara S, Okamura S, Nakano S, et al. Polymerase chain reaction detection of bacterial 16S rRNA gene in human blood. Microbiol Immunol 2008;52(7):375–82.
[51] Zhou X, Li J, Guo J, Geng B, Ji W, Zhao Q, et al. Gut-dependent microbial translocation induces inflammation and cardiovascular events after ST-elevation myocardial infarction. Microbiome 2018;6(1):66.
[52] Criss AK, Seifert HS. A bacterial siren song: intimate interactions between Neisseria and neutrophils. Nat Rev Microbiol 2012;10(3):178–90.
[53] Weinstock DM, Brown AE. Rhodococcus equi: an emerging pathogen. Clin Infect Dis 2002;34(10):1379–85.
[54] Cekanaviciute E, Yoo BB, Runia TF, Debelius JW, Singh S, Nelson CA, et al. Gut bacteria from multiple sclerosis patients modulate human T cells and exacerbate symptoms in mouse models. Proc Natl Acad Sci USA 2017;114(40):10713–8.
[55] Sałkowska A, Karaś K, Karwaciak I, Walczak-Drzewiecka A, Krawczyk M, Sobalska-Kwapis M, et al. Identification of novel molecular markers of human Th17 cells. Cells 2020;9(7):1611.
[56] Fu L, Song J, Wang C, Fu S, Wang Y. Bifidobacterium infantis potentially alleviates shrimp tropomyosin-induced allergy by tolerogenic dendritic cell-dependent induction of regulatory T cells and alterations in gut microbiota. Front Immunol 2017;8:1536.
[57] Wang M, Yin S, Qin Q, Peng Y, Hu Z, Zhu X, et al. *Stenotrophomonas maltophilia* inhibits host cellular immunity by activating PD-1/PD-L1 signaling pathway to induce T-cell exhaustion. Mol Immunol 2021;130:37–48.
[58] McKelvey BA, Umbricht CB, Zeiger MA. Telomerase reverse transcriptase (TERT) regulation in thyroid cancer: a review. Front Endocrinol 2020;11:485.
[59] Fernández-García V, González-Ramos S, Martín-Sanz P, García-Del Portillo F, Laparra JM, Boscá L. NOD1 in the interplay between microbiota and gastrointestinal immune adaptations. Pharmacol Res 2021;171:105775.