RESEARCH

Targeted next-generation sequencing of cancer genes in poorly differentiated thyroid cancer

Tiemo S Gerber¹, Arno Schad², Nils Hartmann², Erik Springer², Ulrich Zechner³ and Thomas J Musholt¹

¹Endocrine Surgery Section, Department of General, Visceral and Transplantation Surgery, University Medicine, Mainz, Germany
²Department of Pathology, University Medicine, Mainz, Germany
³Institute of Human Genetics, University Medicine, Mainz, Germany

Correspondence should be addressed to T S Gerber: tiemogerber@gmail.com

Abstract

Poorly differentiated thyroid carcinoma (PDTC) is a rare malignancy with higher mortality than well-differentiated thyroid carcinoma. The histological diagnosis can be difficult as well as the therapy. Improved diagnosis and new targeted therapies require knowledge of DNA sequence changes in cancer-relevant genes. The TruSeq Amplicon Cancer Panel was used to screen cancer genomes from 25 PDTC patients for somatic single-nucleotide variants in 48 genes known to represent mutational hotspots. A total of 4490 variants were found in 23 tissue samples of PDTC. Ninety-eight percent (4392) of these variants did not meet the inclusion criteria, while 98 potentially pathogenic or pathogenic variants remained after filtering. These variants were distributed over 33 genes and were all present in a heterozygous state. Five tissue samples harboured not a single variant. Predominantly, variants in \( P53 \) (43% of tissue samples) were identified, while less frequently, variants in \( APC \), \( ERBB4 \), \( FLT3 \), \( KIT \), \( SMAD4 \) and \( BRAF \) (each in 17% of tissue samples) as well as \( ATM \), \( EGFR \) and \( FBXW7 \) (each in 13% of tissue samples) were observed. This study identified new potential genetic targets for further research in PDTC. Of particular interest are four observed \( ERBB4 \) (alias \( HER4 \)) variants, which have not been connected to this type of thyroid carcinoma so far. In addition, \( APC \) and \( SMAD4 \) mutations have not been reported in this subtype of cancer either. In contrast to other reports, we did not find \( CTNNB1 \) variants.

Key Words

- thyroid
- endocrine cancers
- rare diseases/syndromes

Introduction

Poorly differentiated thyroid carcinoma (PDTC) represents an aggressive variant of thyroid cancer that predominantly arises from the differentiated variants of papillary and follicular thyroid carcinoma (PTC and FTC, respectively) but occasionally from normal follicular cells (1, 2). The incidence varies between 2–3% and 15%, depending on geographical location (3). The Turin consensus, which is based on histological growth pattern, nuclear features, mitosis, necrosis and convoluted nuclei, offered an algorithm to diagnose this entity (4). However, even with this consensus paper published in 2007, PDTC still represents a challenging diagnosis for pathologists and clinicians. Molecular changes in PDTC are heterogeneous. No exclusive mutation has been identified that could potentially facilitate the differentiation process (5, 6, 7). The advanced cancer stage and the impaired or complete lack of radioactive iodine uptake drive the search for effective therapeutic alternatives such as targeted therapies. Although Landa and coworkers (1) as well as Xu and Ghoosein (8) reported an extensive investigation on PDTC, we still need to learn more about the driving molecular alterations.

Using a next-generation sequencing (NGS) approach, we screened 23 PDTC for variants in 48 cancer-relevant
genes. The main goal addressed in this exploratory study was to characterise novel genetic changes as potential targets for further research as well as establishing the NGS technique in our laboratory.

Materials and methods

Patients and tissue acquisition

Patients treated in the Department of General, Visceral and Transplantation Surgery at University Medical Centre, Mainz, Germany between 2007 and 2015 were included in the study. The study was approved by Ethics Committee of the medical association of Rheinland Pfalz. Consent has been obtained from each patient or subject after full explanation of the purpose and nature of all procedures used. Archival tissue from the patients was retrieved and reviewed using a microscope slide from each tumour stained with haematoxylin–eosin, selected material was scratched off and the selected areas for DNA extraction and sample quality control in the tumour cell). An overview of the patient data is presented in Table 1. Routine Sanger sequencing was performed on the BRAF V600E and BRAF wild-type V600 gene locus.

DNA extraction and sample quality control

Genomic DNA was extracted from the formalin-fixed and paraffin-embedded tumour material. To reduce contamination, microscope slides of the tumour stained with haematoxylin–eosin were used to select areas for macrodissection. The selected material was scratched off from 2 μm thick unstained slides. It was deparaffinised with xylene and ethanol and digested by proteinase K in lysis buffer overnight at 56°C. The samples were purified with MPC Protein Precipitation Solution (Biozym Scientific GmbH, Oldendorf, Germany) and precipitated with isopropanol. The DNA was quantified using a spectrophotometer to measure the absorbance at 260 nm. The DNA quality was tested with the FFPE QC Kit for the TruSeq Amplicon Cancer Panel (Illumina, San Diego, CA, USA) using the manufacturer’s instructions. Two samples failed the quality control criteria, resulting in 23 analysed PDTCs.

Library preparation and MiSeq sequencing

Sequencing libraries were prepared using the TruSeq Amplicon Cancer Panel (Illumina) with the MiSeq Reagent Kit, v2 according to the manufacturer’s instructions. The MiSeq System (Illumina) was used to launch the massively parallel sequencing process to capture the exons of 48 genes with 212 amplicons.

Data analysis

Using the TruSeq Amplicon Workflow on the MiSeq Reporter, the data were demultiplexed and aligned to the human reference genome hg19. The output was generated in the variant caller format. BaseSpace Variant Studio, v2.2.4 (Illumina) was used for post processing. This programme could integrate information about called variants from the Cosmic (v65) and ClinVar (v05.09.2013) scientific databases. The application uses the Variant Effect Predictor (VEP) tool for Annotation and Analysis of coding sequence changes. VEP also harnesses algorithms such as Polymorphism Phenotyping 2 (PolyPhen-2) and Sorting Intolerant From Tolerant (SIFT). To eliminate false positive called variants we used a specific algorithm. First, all variants needed at least a coverage above 20, a sequencing depth of at least 100 and a mutation frequency of 10%, resulting in a threshold of 10 counts for a variant. Second, all variants needed to be nonsynonymous and to have a genotype quality of at least 30. Third, all variants needed to be marked as potentially deleterious in SIFT, as potentially damaging in PolyPhen, as potentially pathogenic or pathogenic in ClinVar or listed in Cosmic. Genes were named in accordance with the guidelines for human gene nomenclature (Supplementary Table 1, see section on supplementary data given at the end of this article; (10)).

This work is licensed under a Creative Commons Attribution 4.0 International License.
<table>
<thead>
<tr>
<th>No.</th>
<th>Sex</th>
<th>Age</th>
<th>Tumour size (cm)</th>
<th>T category</th>
<th>Node</th>
<th>Distance</th>
<th>Treatment</th>
<th>Local complications*</th>
<th>Vessel invasion</th>
<th>Pathology</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>m</td>
<td>76</td>
<td>8</td>
<td>III</td>
<td>No</td>
<td>Mx</td>
<td>TT + RAI</td>
<td>None</td>
<td>Yes</td>
<td>Lymph</td>
</tr>
<tr>
<td>2</td>
<td>w</td>
<td>78</td>
<td>8.5</td>
<td>III</td>
<td>Yes</td>
<td>Lung, bone</td>
<td>TT + RAI</td>
<td>Dyspnoea, dysphagia</td>
<td>Yes</td>
<td>Blood</td>
</tr>
<tr>
<td>3</td>
<td>w</td>
<td>65</td>
<td>7.7</td>
<td>IVa</td>
<td>Yes</td>
<td>Lung, bone</td>
<td>TT + RA1 + PT</td>
<td>Dyspnoea, dysphagia</td>
<td>Yes</td>
<td>ETE</td>
</tr>
<tr>
<td>4</td>
<td>w</td>
<td>43</td>
<td>2.8</td>
<td>II</td>
<td>No</td>
<td>Mx</td>
<td>TT</td>
<td>None</td>
<td>No</td>
<td>TCF</td>
</tr>
<tr>
<td>5</td>
<td>m</td>
<td>64</td>
<td>6.3</td>
<td>III</td>
<td>Yes</td>
<td>Mx</td>
<td>TT</td>
<td>None</td>
<td>No</td>
<td>Node</td>
</tr>
<tr>
<td>6</td>
<td>w</td>
<td>81</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>7</td>
<td>m</td>
<td>73</td>
<td>9</td>
<td>IVa</td>
<td>Yes</td>
<td>M0</td>
<td>TT + CTx + PT</td>
<td>None</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>8</td>
<td>w</td>
<td>77</td>
<td>7.5</td>
<td>IVa</td>
<td>No</td>
<td>M0</td>
<td>TT + RAI</td>
<td>Dyspnoea</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>9</td>
<td>m</td>
<td>64</td>
<td>9.4</td>
<td>III</td>
<td>No</td>
<td>Bone</td>
<td>TT + CTx</td>
<td>Infiltration trachea</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>10</td>
<td>m</td>
<td>65</td>
<td>8.5</td>
<td>III</td>
<td>No</td>
<td>M0</td>
<td>TT + RAI</td>
<td>Dyspnoea</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>11</td>
<td>w</td>
<td>27</td>
<td>2.2</td>
<td>II</td>
<td>No</td>
<td>M0</td>
<td>TT</td>
<td>None</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>12</td>
<td>m</td>
<td>72</td>
<td>NA</td>
<td>IV</td>
<td>Yes</td>
<td>Mx</td>
<td>TT + CTx + PT + RAI</td>
<td>Infiltration oesophagus</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>13</td>
<td>m</td>
<td>86</td>
<td>6.7</td>
<td>III</td>
<td>Yes</td>
<td>Lung</td>
<td>TT + CTx</td>
<td>None</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>14</td>
<td>m</td>
<td>70</td>
<td>4</td>
<td>IVa</td>
<td>No</td>
<td>Lung</td>
<td>TT + PT + RAI</td>
<td>Infiltration trachea</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>15</td>
<td>w</td>
<td>53</td>
<td>1.2</td>
<td>II</td>
<td>No</td>
<td>M0</td>
<td>TT</td>
<td>None</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>16</td>
<td>w</td>
<td>59</td>
<td>1.5</td>
<td>III</td>
<td>Yes</td>
<td>Lung, bone</td>
<td>TT + PT + RAI</td>
<td>Infiltration oesophagus and trachea</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>17</td>
<td>m</td>
<td>67</td>
<td>NA</td>
<td>IVa</td>
<td>Yes</td>
<td>Lung, bone</td>
<td>TT + PT + CTx + RAI</td>
<td>Infiltration oesophagus and trachea</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>18</td>
<td>w</td>
<td>79</td>
<td>NA</td>
<td>III</td>
<td>Yes</td>
<td>M0</td>
<td>TT + RAI</td>
<td>None</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>19</td>
<td>m</td>
<td>66</td>
<td>2.5</td>
<td>III</td>
<td>Yes</td>
<td>Lung, liver</td>
<td>TT + CTx + PT + RAI</td>
<td>Infiltration larynx</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>20</td>
<td>w</td>
<td>57</td>
<td>0.9</td>
<td>IVa</td>
<td>Yes</td>
<td>Lung</td>
<td>TT + PT</td>
<td>Infiltration oesophagus and trachea</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>21</td>
<td>w</td>
<td>67</td>
<td>4</td>
<td>II</td>
<td>No</td>
<td>M0</td>
<td>TT + RAI</td>
<td>None</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>22</td>
<td>m</td>
<td>68</td>
<td>4.8</td>
<td>III</td>
<td>Yes</td>
<td>Lung, liver</td>
<td>TT + CTx</td>
<td>Dyspnoea</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>23</td>
<td>m</td>
<td>76</td>
<td>11</td>
<td>IVb</td>
<td>Yes</td>
<td>Mx</td>
<td>TT + PT + RAI</td>
<td>Infiltration trachea</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

\*During any time after surgical treatment.

CTx, Chemotherapy; ETE, extrathyroidal extension of tumour; NA, not available; PT, percutaneous radiotherapy; RAI, radioactive iodine treatment; TCF, tumour cell fraction (cell-cell ratio); TT, total thyroidectomy (usually with central node dissection); Wt, wild type.
Eighteen tissue samples of PDTCs showed a total number of 98 variants as presented in Fig. 1. All variants were present in a heterozygous state. Ninety-eight genetic changes were single-nucleotide variants (90 missense variants, two nonsense variants, and 6 splice site variants), while one was a deletion leading to a frameshift. The sequencing depth ranged between 101 and 10258 (median 161). The mutation frequency had a median of 13.2. The coverage varied between 21 and 280 with a median of 102. Eighty-seven variants were identified with a genotype quality of 100. For detailed information on genetic alterations see Supplementary Table 1.

The most commonly altered gene was TP53, affected in 10 of 23 samples (43%). All TP53 variants occurred at different base positions. The FLT3, SMAD4 and BRAF variants were detected in each of four samples. Four variants were observed in the each of the APC, BRAF, ERBB4, FLT3, KIT and SMAD4 gene loci. In three of the four samples with BRAF variants, we found the p.V600E variant and in one, the p.G469V variant. Changes in the ATM, EGFR and FBXW7 genes were identified in each of the three samples (13%).

One-hundred percent concordance in BRAF variant status was observed between NGS and routine conventional Sanger sequencing. Each tumour sample of PDTC diagnosed with the Turin consensus criteria was also classified as PDTC by the MSKCC-criteria that indicated more aggressive PDTC.

### Discussion

Two main questions drive the search for new molecular alterations of PDTC. First, what targeted therapies could add a promising approach to the treatment? Second, which mutational markers could improve the accuracy of diagnosis for advanced thyroid carcinomas? This study aims to contribute to this growing area of research by exploring 48 genes of interest. It is beyond the scope of this explorative study to prove causality or to address the progression from well-differentiated thyroid carcinoma (WDTC) to PDTC.
**BRAF, TP53, RAS**

While tumour-initiating somatic mutations leading to papillary thyroid carcinoma have been researched to a great extent, it remains uncertain whether and to what extent particular genetic alterations participate in the tumour progression to PDTC. The Cancer Genome Atlas Research Network study has identified many oncogenic mutations for PTC. It is generally agreed today that PDTC can originate from WDTC (11).

Landa and coworkers has shown that the changes to the BRAF to RAS relationship are still preserved in PDTC (11). For these reasons, one could assume that certain driver mutations are conserved during the dedifferentiation process. In our study, we found three RAS variants (one each of N-, K- and HRAS) and three BRAF p.V600E variants. However, it is not the same as a BRAF/RAS-score based on mRNA data (11). In accordance with the aforementioned studies, there have been no RAS and BRAF p.V600E variants in the same tissue sample.

Mutations in BRAF, EIF1AX, TP53 and TERT are associated with aggressive behaviour of thyroid malignancies (2, 8, 12). BRAF and TP53 are associated with the dedifferentiation process to anaplastic thyroid carcinomas (ATCs) (13). In our study, there were 10/23 cases of PDTC harbouring a p53 variant. There was no significant difference regarding the clinical course, extrathyroidal extension or vascular invasion between TP53-positive tumours and the other tumours. One possible reason may be the preselection of an already aggressive subtype of cancer. Some previously published works have assessed TP53 mutations in PDTC. They observed lower rates (8, 27 and 38%) of this variant than in our series (1, 14, 15). The variety might be related to the generally low number of patients in our study (23 tissue samples) as well as in compared series (84, 22 and 46 patients). The TERT and EIF1AX gene locations were not included in the applied cancer panel.

In contrast to previous studies, we found less RAS variants (13%) than described in the literature (20–55%) (16, 17, 18, 19, 20). These numbers are slightly lower than the value we expected. It seems plausible that a number of limitations could have influenced the obtained results. On the one hand, we used tissue that was formalin fixed and paraffin embedded for several years, which leads to the necessity of a very conservative filtering process to avoid false-positive findings from artefacts. This implies that there could be true-positive variants that have been lost in the background of artefacts. On the other hand, it is obvious that the overall small number of PDTC cases impairs conclusions based on the comparison of percentages.

**APC, ATM, CTNNB1, EGFR and ERBB4**

PDTC are a genetically heterogeneous entity. In this study, we identified new potential genetic targets for further investigation. Two of these are EGFR (ERBB1) and ERBB4, which have not been previously reported to play a role in thyroid carcinogenesis. ERBB1 is known to be mutated in cancers like gliomas or small-cell lung cancer and associated with the epithelial–mesenchymal transition as well as tumour invasion (21). In colon cancer, the mutation of ERBB4 resulted in a loss of differentiation and an activation of the phosphoinositide 3-kinase (PI3K) signalling pathway (22).

Another valuable target is APC. Recent studies on PTC identified APC as well as ATM as potential driver mutations (11). PDTC may harbour mutations that are present in PTC as they may derive from WDTC (23). Furthermore, APC mutations have been observed in ATC (24). Considering these arguments, it seems possible that this mutation plays a role in PDTC carcinogenesis as well.

Somatic mutations of CTNNB1 were discussed being associated PDTC and ATC (25). However, we detected no variants of this gene in the PDTC examined. The absence of CTNNB1 mutations supports the hypothesis of Rocha and coworkers (26) that loss of E-cadherin expression rather than β-catenin (CTNNB1) gene mutations induces the process of tumour dedifferentiation.

**SMAD4 and KIT**

SMAD4 encodes a mediator protein in the transforming growth factor β (TGF-β) signalling pathway. This can lead to a loss of TGF-β-mediated growth inhibition (27). In our study, five variants affected four patients (17%). The observed mutations were scattered along the sequence of the SMAD4 gene locus, suggesting that there is no mutational hotspot present. This locus is known to be altered in WDTC, oesophageal adenocarcinoma, pancreatic cancer and colorectal cancer (27, 28, 29). A SMAD4 immunohistochemistry expression study has shown a correlation with poor survival in colon cancer (30). An in-depth examination of SMAD4 status demonstrated that abnormalities in this locus occur in benign as well as in malignant thyroid tumours (15/56, 27%), which was interpreted by the authors as indicative of an early event in thyroid tumorigenesis (28). All this leads to the assumption...
that the presence of SMAD4 mutations cannot differentiate between well-differentiated thyroid cancer and PDTC.

KIT encodes for a tyrosine kinase receptor and mediates growth regulation. Mutations in this gene are associated with haematologic diseases and gastrointestinal stromal tumours (31). KIT expression is reduced in PTC compared to benign nodules (32). Its assessment can improve the diagnosis in fine-needle aspiration (33, 34). There are, however, little data available for KIT mutational status in PDTC, although it is known that alterations in KIT transcripts increase with tumour progression from WDTC to PDTC (35, 36). In a study with 118 ATCs, none showed a KIT mutation (37). In our study, we detected four KIT variants. This discrepancy may indicate a slowing effect on cancer growth that is lost during tumour dedifferentiation to anaplastic thyroid cancer, which may be lost in favour of a more deleterious mutation in the downstream signalling pathways.

Coexistent follicular or papillary tumours

Surprisingly, we found many tumours with coexistent follicular or papillary carcinoma components (Figs 1 and 2). Eight PDTCs showed follicular coexistent components while five showed proportions with papillary carcinomas. Given the higher incidence of PTC than FTC, this could indicate a more frequent progression of FTC to PDTC. This is in accordance with the findings of Soares and coworkers (38) who emphasised that PDTC are more closely related to FTC than to PTC. The authors established their assumption on the absence of BRAF variants in 19 PDTC. In contrast to their results, we found three BRAF V600E variants. Two of the three p.V600E BRAF variants were in PDTC with PTC proportions. The discrepancy is most likely related to the fact that Soares and coworkers excluded PDTC with PTC components in their series.

Figure 2
Macrodissection of a tumour with two different entities. Complete slide of the specimen before (A) and after (B) dissection of the PDTC. (C) A solid growing PDTC is demonstrated on the upper left next to a coexisting follicular thyroid tumour. Magnified clipping of a as indicated. (A, B and C: hematoxylin-eosin stain; A and B: 5x magnification; C: 200x magnification).
Methodological considerations

Five PDTCs did not show any genetic alteration at all, which implies that yet unknown driving genetic alterations exist or that the applied methods failed to detect them. According to Bozic and coworkers (39), the actual selective growth advantage of a cancer by typical mutations is much smaller than commonly assumed. The majority of genes leading to a carcinoma (or supporting its dedifferentiation consecutively) have a low mutation frequency (40). Thus, it is possible that either these PDTCs had genetic variants not covered by the NGS panel or they had a mutation frequency less than 10% and were not distinguishable from artefacts.

The results of the present study demonstrate that there are many noteworthy genes that could be responsible for the dedifferentiation of thyroid carcinoma, but a number of restrictions of our study should be mentioned. One should note that we did not analyse blood or normal tissue to discriminate between true mutations or SNP. In a detailed examination of germline mutations in tumour tissue 3% of all samples showed germline mutations (41). We can assume similar results. Given the exploratory nature of this study, the variants presented here can only serve as an indicator for more targeted studies that should consider pathogenic germline alterations.

Furthermore, it is important to emphasise that the data have not been completely validated by Sanger sequencing. Acknowledging the limitations of validation of NGS-derived variants using Sanger sequencing, we can surmise that the validation rate is precise enough for an exploratory study (42). Admittedly, contrary to the aforementioned study, we used paraffin-embedded formalin-fixed tissue. This pre-treated tissue could potentially be the source of sequencing artefacts. Following recommendations to reduce this effect, we macrodissected tumour-rich tissue, measured the quantity of DNA and excluded insufficient tissue samples (43). Given the nature of the TruSeq Amplicon Cancer Panel, we only sequenced exonic DNA with a limitation on certain genes. Due to this limitation, it was not possible to detect rearrangements of oncogenes like RET and NTRK1 or mutations of TERT. The latter are suspected to be connected to the disease progression of WDTC and to distinguish PDTCs from ATCs (1). However, rearrangements are rarely described in PDTC possibly because WDTC expressing RET/PTC hybrid oncogenes will usually not dedifferentiate further (44).

Even with the recent progress in the knowledge of molecular changes in PDTC, we think that our study reveals some novel variants, possibly influencing future research.


are more frequently detected in undifferentiated than in insular and insular-like poorly differentiated carcinomas. *Virchows Archiv* 2004 444 572–576. ([https://doi.org/10.1007/s00428-004-1018-0](https://doi.org/10.1007/s00428-004-1018-0))


