

Transcriptomic landscape of radiation-induced murine thyroid proliferative lesions

SCHOLARONE™ Manuscripts

- Elena Stauffer: elena.stauffer@helmholtz-muenchen.de
- Dr. Peter Weber: peter.weber@helmholtz-muenchen.de
- Theresa Heider: theresa.heider@helmholtz-muenchen.de
- Dr. Claudia Dalke: dalke@helmholtz-muenchen.de
- Dr. Andreas Blutke: andreas.parzefall@helmholtz-muenchen.de
- Prof. Dr. Axel Walch: axel.walch@helmholtz-muenchen.de
- Dr. Gerald Burgstaller: gerald.burgstaller@helmholtz-muenchen.de
- Dr. Nikko Brix: nikko.brix@med.uni-muenchen.de
- Prof. Dr. Kirsten Lauber: kirsten.lauber@med.uni-muenchen.de
- Prof. Dr. Horst Zitzelsberger: zitzelsberger@helmholtz-muenchen.de
- Dr. Kristian Unger: unger@helmholtz-muenchen.de
- Dr. Martin Selmansberger: martin.selmansberger@helmholtz-muenchen.de
- Helmholtz Zentrum Muenchen
- Ingolstäter Landstrasse 1
- 85764 Oberschleißheim, Germany
-
- brix@med.uni-muenchen.de

per: kirsten.lauber@med.uni-muenchen.de

sberger: zitzelsberger@helmholtz-muenche

ger@helmholtz-muenchen.de

erger: martin.selmansberger@helmholtz-n

Muenchen

se 1

im, Germany

tomics of early-Short title: Transcriptomics of early-state thyroid proliferative lesions
-
- Keywords: thyroid pathogenesis, thyroid carcinogenesis, thyroid neoplasia, histology,
- transcriptomics
-
-
-
-
-

ABSTRACT

istic and transcriptomic analyses of thyroid in the expressive histological examination of 246 the carried out. Proliferative and normal tissue from non-exposed cases were collect owed by RNAseq transcriptomic profiling pr Thyroid carcinoma incidence rates in western societies are among the fastest rising, compared to all malignant tumors over the past two decades. While risk factors such as age and exposure to ionizing radiation are known, early-state carcinogenic processes or pre-lesions are poorly understood or unknown. This study aims at the identification and characterization of early-state radiation-associated neoplastic processes by histologic and transcriptomic analyses of thyroid tissues derived from a mouse model. Comprehensive histological examination of 246 thyroids (164 exposed, 82 non-exposed) was carried out. Proliferative and normal tissues from exposed cases and normal tissue from non-exposed cases were collected by laser-capture microdissection, followed by RNAseq transcriptomic profiling using a low input 3`- library preparation protocol, differential gene expression analysis and functional association by *Gene Set Enrichment Analysis*. Nine exposed samples exhibited proliferative lesions, while none of the non-exposed samples showed histological abnormalities, indicating an association of ionizing radiation exposure with histological abnormalities. Activated immune response signaling and deregulated metabolic processes were observed in irradiated tissue with normal histology compared to normal tissue from non-exposed samples. Proliferative lesions compared to corresponding normal tissues showed enrichment for mainly proliferation- associated gene sets. Consistently, proliferative lesion samples from exposed mice showed elevated proliferation-associated signaling and deregulated metabolic processes compared to normal samples from non-exposed mice. Our findings indicate that a molecular deregulation is already detectable in histologically normal thyroid

tissues and in early proliferative lesions in the frame of multi-step progression from

For Review Only

- irradiated normal tissue to tumorous lesions.
-
- Words: 243
-
-
-
-
-
-
-
-
-
-
-
-
-
-
-
- -
-
-
- -
	-
	-
	-
	-
	-
	-
	-
	-
-
-
-
-
-
-
-

INTRODUCTION

Example II and identify thyroid carcinoma (PTC) subtyp

Fradical surgery and shows low mortality randiation at young age, both internal are

the development of thyroid carcinoma (Ro

in childhood greatly increases the risk Incidence rates for thyroid cancer increased worldwide over the last decades, thus becoming the most common endocrine malignancy (La Vecchia et al., 2015). While environmental and lifestyle factors and the effect of overdiagnosis due to improved screening methods are discussed, their impact on the increased incidence rates remains to be proved (Wiltshire et al., 2016). Despite the high incidence rates, thyroid cancer, especially the papillary thyroid carcinoma (PTC) subtype, is associated with a good prognosis after radical surgery and shows low mortality rates (La Vecchia et al., 2015). Exposure to ionizing radiation at young age, both internal and external, is a well- known risk factor for the development of thyroid carcinoma (Ron et al., 1995). Ionizing radiation exposure in childhood greatly increases the risk for thyroid cancer, particularly PTC. Numerous studies have shown that after the Chernobyl reactor

 accident and the resulting radiation exposure of the population through radioactive fallout (including I-131), the incidence rate of thyroid carcinomas has significantly increased (Efanov et al., 2018, Cahoon et al., 2017, Nikiforov, 2006, Kazakov et al., 1992).

 Thyroid carcinomas can develop either from follicular cells or c-cells. The resulting malignancies are classified into medullary carcinomas (c-cells) or papillary, follicular and anaplastic carcinomas (follicular cells). Papillary carcinomas are further divided into different subtypes based on the predominant histological architecture (Xing, 2013). The majority of thyroid cancer originates from follicular cells such as PTC and follicular thyroid carcinoma (FTC) (Baloch and LiVolsi, 2018, DeLellis RA, 2004). While

 it is proposed that FTC evolves from follicular adenoma (FA), papillary thyroid carcinoma has no benign precursor lesion (Sponziello et al., 2013).

 Mouse models are frequently used to study thyroid cancer and the underlying carcinogenic processes. However, although mouse and human thyroid cancers show biological similarities, they greatly differ in histological characteristics. In contrast to human thyroid cancer, murine histological cancer entities are not distinguished, e.g. the follicular carcinoma has only a papillary pattern but is classified as follicular carcinoma (Capen, 2001, Boorman G A, 1995).

001, Boorman G A, 1995).

ypes, a series of genetic and epigenetic alternatives, a series of genetic and epigenetic alternatives

16, Vogelstein and Kinzler, 1993). The

the effectors of the MAPK signaling pathway

al role As for other cancer types, a series of genetic and epigenetic alterations is responsible for initiation and progression of thyroid cancer according to a multi-step model (Kondo et al., 2006, Vogelstein and Kinzler, 1993). Those alterations, that consequently stimulate effectors of the MAPK signaling pathway or PI3K/AKT signaling pathway, play a central role in thyroid carcinogenesis (Nikiforova and Nikiforov, 2008). Point mutation of the BRAF and RAS genes and rearrangements of the RET gene are common in PTC, while the rearrangements of PAX8/PPARy are mostly prevalent in FTC and FTA (Nikiforov and Nikiforova, 2011, Xing, 2005). The detection of genetic alterations in follicular adenoma, which can be considered as early lesions, support the concept of a multi-step carcinogenesis model. However, it remains unclear at what point follicle cells develop proliferative potential (Jung et al., 2016). Additionally, a cancer stem cell carcinogenic model and a fetal cell carcinogenesis model are currently discussed as alternatives to the multi-step process (Zane et al., 2016).

 Nowadays, molecular alterations in thyroid cancer are well characterized, while early state events or precursor lesions still remain poorly understood or are unknown (Ye et al., 2017, Xing, 2013). Besides genetic aberrations such as RET- and PAX8-

 $\frac{p}{2}$ rearrangements and BRAF V600E mutations, epigenetic modifications including silencing of RASSF1A via promoter methylation and post-transcriptional regulation via miRNAs and lncRNAs in thyroid carcinogenesis were recently investigated (Brown et al., 2014, Sedaghati and Kebebew, 2019, Colamaio et al., 2015). The aim of this study was the identification and characterization of molecular processes in the early stages of the carcinogenic processes in the thyroid. For this purpose, we used a mouse model and performed comprehensive histological analyses of thyroids from a cohort of mice that received low-dose gamma-irradiation early in life. Subsequently, laser capture microdissection (LCM) enabled transcriptomic profiling of histologically normal and aberrant tissue areas, in order to gain further insight in early events of genetic and molecular deregulation that may drive the early phases of thyroid malignancies. **MATERIAL and METHODS Workflow** The entire workflow of this study is visualized in Figure 1. **Thyroid tissues** Mouse thyroids were obtained from the INSTRA lifetime study and details on mice are presented in Dalke et al. (Dalke et al., 2018). 164 mice were whole body irradiated at the age of 10 weeks with different doses of 0.063 Gy, 0.125 Gy and 0.5 Gy. 82 mice

- exposed to 0 Gy served as control group (sham irradiation). Animals were sacrificed
- at pre-determined time points (6 months, 12 months, 18 months, 24 months) and

 μ (cells/mm²) and was determined in 2-3 se
two groups of tissue samples were det
sive lesions from the irradiated mice, are
gically normal tissue taken from the same at
l thyroid tissues from sex- and age-matche
cula such as adenomas were identified by the presence of demarcated histologic areas with a distinct follicular and/or papillary architecture. To differentiate those neoplastic lesions from carcinomas, the invasion of the surrounding glandular parenchyma or stroma was evaluated. Due to the absence of such invasive areas, there was no indication for carcinomas of the papillary or follicular subtype in any of the neoplastic cases. Furthermore, the presence of immune cells was quantitatively assessed in the evaluated HE slides. Immune cell infiltration was expressed as the number of immune 200 cells per section area (cells/mm²) and was determined in 2-3 sections per case. After histological analysis, two groups of tissue samples were defined: the first group comprised proliferative lesions from the irradiated mice, and the second group contained morphologically normal tissue taken from the same animal, respectively. A 204 third group of normal thyroid tissues from sex- and age-matched non-irradiated mice 205 was added. For molecular investigation 30 μ m thick sections were prepared for LCM. LCM was carried out utilizing a PALM MicroBeam device (Zeiss, Germany), which allows the dissection and collection of targeted tissue formations or groups of cells after microscopic identification, according to the manufacturers protocol and as previously described (Morrogh et al., 2007).

 The histological findings in irradiated and non-irradiated mice were statistically analyzed by Fisher's exact test. Immune cell scores were compared between the three groups by the analysis of variance (ANOVA) and pair-wise compared by post-hoc Tukey test. Statistical significance was considered at p < 0.05.

RNA isolation and RNAseq 3´library preparation

bH, Austria). Library amplification PCR cy
it for Illumina (SKU: 020.96, Lexogen Gm
vere amplified with 17 PCR cycles. Qualit
ated using the Quanti-iT PicoGreen dsD
the Bioanalyzer High Sensitivity DNA Ana
, Inc., USA). 15 The RNA extraction from the laser capture microdissected cells was performed with the RNeasy Micro Kit (Quiagen), according to the manufacturer's protocol. RNA quality was evaluated using a Bioanalyzer 2100 Systems (Agilent Technologies, Inc., USA) with Agilent RNA 6000 Pico Kit (#5067-1513, Agilent Technologies, Inc., USA). RNA integrity was assessed by calculation of the percentage of fragments >200 nucleotides (DV200, SI Table 3). Sequencing libraries were generated using 50 ng of total RNA and the QuantSeq 3 ′ mRNA-Seq Library Prep Kit FWD for Illumina (SKU: 015.96, Lexogen GmbH, Austria). Library amplification PCR cycles were determined using PCR Add-on Kit for Illumina (SKU: 020.96, Lexogen GmbH, Austria) and the individual libraries were amplified with 17 PCR cycles. Quality and quantity of the libraries were evaluated using the Quanti-iT PicoGreen dsDNA Assay Kit (P7589, Invitrogen, USA) and the Bioanalyzer High Sensitivity DNA Analysis Kit (#5067-4626, Agilent Technologies, Inc., USA). 150 bp paired-end sequencing was performed on an Illumina HiSeq4000 platform (Illumina, Inc., USA). Lexogen QuantSeq 3 ′ library preparation was chosen for RNAseq analysis because it has the capability to generate high quality gene expression data from low amounts (down to 10 pg) and/or compromised quality of total RNA. Both properties were expected from the small LCM-specimens used for RNAseq analysis.

RNAseq data analysis, differential expression analysis, Gene Set Enrichment

Analysis (GSEA) and Pathway RespOnsive GENes analysis (PROGENy)

Raw data as well as preprocessed RNAseq data are available from Gene Expression

- Omnibus (GEO accession number: GSE162795). Gene expression quantification was
- carried out by alignment of RNAseq reads to the mouse reference genome (Ensembl

sue (NT vs. NCT, PL vs. NCT), were analyzed

and age. Genes with an absolute log2-fold

1 were considered differentially expressed

and p-value thresholds were used solely to

1

session analysis, but no functional interpr GRCm38 version 94) using the STAR aligner, followed by count quantification using 241 the featureCounts of the Rsubreads R package (Dobin et al., 2013, Liao et al., 2014). Differential gene expression analysis was performed using methods implemented in the *DESeq2* R package (Love et al., 2014). Principal component analysis was carried out on variance stabilized read counts. Differential expression between proliferative lesions and corresponding normal tissue samples (PL vs. NT) was performed with a paired design, while normal tissue vs. normal control tissue and proliferative lesions 247 vs. normal control tissue (NT vs. NCT, PL vs. NCT), were analyzed in group comparisons controlled for sex and age. Genes with an absolute log2-fold change > 0.5 and an 249 adjusted p-value < 0.1 were considered differentially expressed. The log2-fold change (log2-FC) and adjusted p-value thresholds were used solely to visualize the results of the differential expression analysis, but no functional interpretations were based on them. *Gene Set Enrichment Analysis* (GSEA) was conducted in a pre-ranked mode using the log2-FC values obtained from *DESeq2* (log-fold-change-shrinkage) for the ranking of the gene list using the *fgsea* R package (Sergushichev, 2016, Powers et al., 2018). The Hallmark gene sets (n=50), downloaded from Molecular Signature Database (http://software.broadinstitute.org/gsea/msigdb/index.jsp), plus additional thyroid pathogenesis-related or cancer-associated gene sets taken from the KEGG (n=41) and Reactome (n=23) gene set collections were used for enrichment testing (SI Table 4). PROGENy analysis was carried out in order to test for a deregulation in eleven cancer relevant signal transduction pathways using transcriptomics data. In comparison to conventional pathway analysis methods (e.g. GSEA), PROGENy infers pathway activity based on consensus gene signatures obtained from perturbation

- experiments on protein level and constitutes the footprint on gene expression
- (Schubert, 2017).
-

qRT-PCR validation of RNAseq quantification

- 267 In addition to the technical quality control of the RNAseq data using FastQC, their
- validity was investigated by comparing the expression of eight randomly selected
- 269 genes determined by RNAseq with that determined by qRT-PCR in all 27 samples. Four
- of the selected genes (CD36, HSP90AA1, IL1B, SERPINE1) were differentially expressed
- (CD36, HSP90AA1, IL1B, SERPINE1) were d

mparisons carried out and four (ALAS1, LR

htial expression. Total RNA extracted from

malyzed by qRT-PCR as described before (K

values were determined as delta CT-value

nd beta-a 271 in one of the three comparisons carried out and four (ALAS1, LRP1, PLAU, THBS1) that
- 272 did not show differential expression. Total RNA extracted from mouse tissue samples
- was processed and analyzed by qRT-PCR as described before (Krombach et al., 2019).
- Relative expression values were determined as delta CT-values to a reference gene
- matrix of 18S rRNA and beta-actin with efficiency correction via three different cDNA
- dilutions for each sample. Pearson correlation between RNAseq and qRT-PCR
- determined gene expression was used as a measure of concordance.
-

Clustering of samples using a literature-derived gene set

 A literature-derived gene set consisting of 85 thyroid carcinogenesis- and thyroid cancer-associated genes, was utilized for hierarchical clustering of samples using correlation as distance measure (linkage "complete") followed by visualization as heatmap including the cluster dendrogram using the *pheatmap* R package. The considered publications reported the genes to be associated with thyroid carcinogenesis and cancer and in greater detail with proliferation, metabolic deregulation, and immune signaling (Rusinek et al., 2011, Vella and Malaguarnera,

2018, Xing, 2013, Giuliani et al., 2018).

 A subset of these genes was also present in the gene expression data set generated on the normal tissue and PTC tissue from the Ukranian-American cohort (UkrAM) of human individuals who were exposed to ionizing radiation as a consequence of the 291 Chernobyl reactor accident (Abend et al., 2013). Their expression levels were 292 compared to the expression data from the three histological groups generated in this study.

RESULTS

Histological analysis reveals radiation-associated lesions

For Preveals radiation-associated lesions
Spiss revealed nine mice with proliferative
als in total. PLs were exclusively present in
the observed in the 24 months group (p=0.1
No significant difference with regard to the
55 The histological analysis revealed nine mice with proliferative lesions in the thyroid (PL) out of 246 animals in total. PLs were exclusively present in the irradiated group while seven PLs were observed in the 24 months group (p=0.09) and two in the 18 months group (p=1). No significant difference with regard to the sex of animals was detected (male p=0.55; female p=0.18). Fisher´s exact test for the frequency of PLs in the irradiated and non-irradiated group (9/164 exposed, 0/82 non-exposed) indicated a significant association between PL-occurrence and radiation exposure (p=0.03). Three out of nine alterations met the diagnostic criteria of an adenoma according to Capen's Classification (Capen, 2001). The remaining six cases were classified as follicular hyperplasia. Eventually, all nine lesions showed predominantly a papillary pattern, which is characterized by branching infoldings covered by a single-layered epithelium. While follicular hyperplasia is not well demarcated, adenomas often have

ion of the lesions within the thyroid glathe examined tissues, and the observed

it issues are provided in SI Figure 1. S

in each cell densities (infiltration) were identifies
 $\frac{1}{2}$ test between proliferative and no a more complex architecture accompanied by encapsulation (Capen, 2001). Since any signs of invasion of surrounding tissue or blood vessels were absent, none of the neoplastic lesions were classified as carcinoma. Immune cells (lymphocytes, neutrophils, plasma cells, and macrophages) were investigated, both, in irradiated and non-irradiated tissues and cells were counted to quantify the extent of infiltration. The highest density of immune cells was detected in proliferative tissue, followed by normal irradiated tissue and non-irradiated normal tissue. Microscopic images visualizing the location of the lesions within the thyroid glands, exemplary high- resolution images of the examined tissues, and the observed immune cell densities within the different tissues are provided in SI Figure 1. Statistically significant differences in immune cell densities (infiltration) were identified by ANOVA (p<0.01) and post-hoc Tukey's test between proliferative and normal irradiated tissue (PL compared to NT, p=0.01) and between proliferative and non-irradiated normal tissue (PL compared to NCT, p=0.001). No significant difference between irradiated normal tissue and non-irradiated normal tissue was observed (NT compared to NCT, p=0.23, SI Figure 2). Sample information and histological classifications are summarized in SI Table 1 and SI Table 2.

qRT-PCR validation of RNAseq quantification

Pearson correlation coefficients ranged from 0.53 to 0.91 for seven out of the eight

genes investigated, only the HSP90AA1 gene exhibited a correlation coefficient of 0.21

- 331 (SI Table 5). These finding validate the 3['] RNAseq data, since only one gene detected
- 332 with primers located near the 5' end indicated low correlation.

 1_i

Transcriptomic analysis points at deregulated pathways in normal tissue and

proliferative lesions

A using the top-500 variant genes in the daidings are provided in SI Table 6 and SI Table
dings are provided in SI Table 6 and SI Table
d EGFR, MAPK and TNFa signal transduction
inent finding in PL compared to NT was
compa Principal component analysis (PCA) of the RNAseq transcriptome data set exhibited grouping of the samples according to their histological classification, separating proliferative lesion samples (PL) very well from the corresponding normal tissue samples (NT) and normal control tissue samples (NCT). The separation of NT and NCT was less pronounced (Figure 1, C). Dimensionality reduction by multi-dimensional- scaling (MDS) and PCA using the top-500 variant genes in the data set are provided in SI Figure 3C/D, PC loadings are provided in SI Table 6 and SI Table 7. PROGENy analysis revealed an activated EGFR, MAPK and TNFa signal transduction in NT compared to NCT. The most prominent finding in PL compared to NT was a downregulated p53 signaling, while PL compared to NCT showed elevated proliferation-associated 346 signaling (MAPK, PI3K, EGFR) and downregulated p53 signaling (Figure 2A, SI Table 8). 347 With respect to log2-FC and false discovery rate (FDR), differential gene expression was least pronounced in NT compared to NCT (128 genes), followed by PL compared 349 to NT (1355 genes), and PL compared to NCT (2379 genes), respectively (Figure 2C, SI Table 9). Common and specific genes for the three comparisons are visualized in SI Figure 4. Gene set enrichment analysis (GSEA) is visualized in Figure 2D. All gene sets with an FDR < 0.1 in at least one out of the three comparisons were included into the plot. Mainly, activated immune signaling and a downregulated metabolic signaling in NT compared to NCT and activated proliferation-associated signaling and downregulated metabolic signaling in PT compared to NT was found (Figure 2D) while 356 PL compared to NCT showed increased proliferation-associated signaling.

in this study, Dalke et al. reported an association of one-time exposure to ionizing

radiation with increased body weight, reduced survival rates and a dose-dependent

association with several other tumor types and eye lens opacity (Dalke et al., 2018).

Mouse models are a valid tool to study the molecular mechanisms underlying thyroid

tumorigenesis since it has been shown that large biological and histological similarities

Therefore, a comparative investigation of a

Instead, we focused on early molecular

Decurred in histologically normal and abe

initiation, promotion and progression in

Isions (e.g. hyperplasia), followed by follice

Iss between human thyroid and mouse thyroid exist (Perlman, 2016). However, the definitions of pre-neoplastic and neoplastic lesions diverge (Baloch and LiVolsi, 2018, DeLellis RA, 2004, Capen, 2001, Boorman G A, 1995). The different histological phenotypes and their histopathological classifications make a comparative diagnosis between the different cancer entities in mice and humans challenging. In this study, tissue samples were examined on the basis of Capen's International Classification of Rodent Tumors (Capen, 2001), which does not distinguish between the different histological entities. Therefore, a comparative investigation of a specific cancer entity was not possible. Instead, we focused on early molecular processes of thyroid tumorigenesis that occurred in histologically normal and aberrant mouse thyroid tissues. While tumor initiation, promotion and progression in murine thyroids starts with proliferative lesions (e.g. hyperplasia), followed by follicular adenoma, which subsequently progress to follicular thyroid carcinoma, in humans it is more complex and is reflected by different thyroid carcinoma subtypes (DeLellis RA, 2004, Capen, 2001, Boorman G A, 1995).

 Comprehensive histological examination of 164 exposed and 82 non-exposed thyroids revealed neoplastic alterations exclusively in exposed mice, which is in line with multiple previous studies reporting ionizing radiation as one of the major risk factors for the development of thyroid carcinoma (Mullenders et al., 2009, Cardis et al., 2005, Williams, 2008, Zablotska et al., 2008, Furukawa et al., 2013, La Vecchia et al., 2015, Ron, 2007). While adenomas were identified within the exposed group, we did not observe any carcinoma in our study. This might be due to the limited life span of a mouse and the fact that the experiment was terminated at 24 months, which might not allow full progression of the carcinogenic process. These observations might

 1°

shown by RNAseq analysis (GSEA and PROGENy) and potentially explain the increased

numbers of immune cells in these tissues (Figure 2, SI Figure 2).

Experience molecular changes in histologically

Foliferative lesions that underwent a tum

summary use of non-irradiated mice (NCT) with

tire thyroid were used as a reference. Alth

between NT and NCT showed the lowest li In order to analyze the histologically defined tissue regions on molecular level, we established a technically challenging workflow: Firstly, we collected the cells forming the tissue regions of interest (PL, NT, NCT) from the tissue sections by LCM prior to extraction of total RNA from these comparably small numbers of cells, followed by 436 high-quality RNA-sequencing library generation and RNA sequencing. Transcriptomic data were used to delineate molecular changes in histologically normal thyroid tissue (NT) and adjacent proliferative lesions that underwent a tumorigenic process (PL). Normal control tissues of non-irradiated mice (NCT) without any histological aberrations in the entire thyroid were used as a reference. Although, the differential expression analysis between NT and NCT showed the lowest level of transcriptomic 442 deregulation within the three conducted differential expression analysis (log2-FC, 443 adjusted-p-value, Volcano plots, Figure 2B), the deregulation was significantly functionally associated with tumorigenesis by PROGENy and GSEA. Despite the different concepts of functional association of transcriptomic deregulation in PROGENy and GSEA, the derived results integrate well and are biologically plausible in 447 the context of the study.

The PROGENy results fit the concept of a progression model while the level of deregulation accumulates from healthy, histologically normal tissue (NCT) to

histologically normal tissue with a deregulated molecular phenotype, that progresses

- further into tissue with histological abnormalities (PL). As an example (Figure 2A), the
- 452 p53-signaling was weakly upregulated in NT compared to NCT and strongly
- downregulated in PL compared to NT, resulting in less pronounced downregulation in

proliferation and differentiation, contributing to tumor progression (Karin, 2009, Li et

growth by activating cell proliferation and angiogenesis or acts anti-tumorigenically

al., 2013). It is interesting to note that, NF-κB signaling might either promote tumor

- by triggering an immune response and the attraction of immune cells (Pires et al.,
- 477 2018, Pacifico et al., 2004). This could further explain the elevated proliferation-

which are frequently described in thyroid carcinoma would be desirable to further

 support the conclusions drawn from this study. However, this was not feasible due to the limited amounts of tissue derived from the small LCM-specimen.

sts that carcinogenic processes are alread
of this study thus representing earl
xpression levels in the three histologic grou
cancer-associated genes indicate that
l stages of molecular and histologic tran
asing expression While multiple studies discuss the effect of specific genes on a single process of thyroid carcinogenesis, we aimed at an integrated analysis by applying a hierarchical clustering approach on the transcriptomic data. Using 85 thyroid carcinogenesis and thyroid cancer-associated genes derived from the literature (SI Table 10, functional grouping and citations), separation into three defined histological groups was achieved. This suggests that carcinogenic processes are already ongoing within the proliferative lesions of this study thus representing early stages of thyroid tumorigenesis. The expression levels in the three histologic groups of multiple thyroid carcinogenesis- and cancer-associated genes indicate that the analyzed tissues represent sequential stages of molecular and histologic transformation in thyroid tumorigenesis. Increasing expressions from NCT, via NT, to PL for the thyroid tumor- promoting genes TSHR, IDH1, NFKB1, HIF1AN, STAT3, TPO, TRP53, TMEM173 (STING), IGF1, MAPK10 and decreasing expressions for tumor-suppressing genes such as SLC5A8, TIMP3, ITIH5, PCSK2 were observed (Figure 2B, SI Figure 5, SI Table 10). This further underlines that the PL tissues might represent early stages during thyroid tumorigenesis and that histologically normal tissues (NT) already show molecular deregulations pointing to the development of thyroid cancer.

 Additionally, 26 out of the 31 differentially expressed genes in the 85-gene set that were present in the human UkrAm gene expression data set showed similar regulation patterns and significant differential expression between normal tissue and tumor tissue (PTC) samples compared to the expression data generated within this study (SI Figure 5 and 6). This indicates that these genes that are known to be deregulated

 2°

535

536 **CONCLUSION**

537

methat is conclusive with respect to the
ates well with findings within the present
mesis and early-stage processes of several research
mesis and early-stage processes of several respectively
proportionally detectable abno In conclusion, we hypothesize that the histologically normal tissues (NT), which contain no macroscopically detectable abnormalities, represent a very early stage in thyroid carcinogenesis and will further progress to thyroid hyperplasia, followed by thyroid adenoma and potentially by thyroid carcinoma. While NT tissues exhibit histologically normal structures, the molecular footprint compared to normal control tissues (NCT) already indicates an initiation of carcinogenic processes. These findings integrate well with the proposed multi-step model and the fetal cell carcinogenesis model as discussed in Zane et al. (Zane et al., 2016). Potential clinical implication might be a refined thyroid cancer diagnosis, which is currently carried out by fine needle biopsy followed by cytologic classification. Further investigation of thyroid tissues with

 $2³$

- normal histology and deregulated molecular levels are necessary in order to develop
- a molecular signature or marker for implementation in routine clinical diagnostics.
-
-
- **Declaration of interest, Funding and Acknowledgements**
-
- Funding:
- The INSTRA project is funded by the Bundesministerium für Bildung und Forschung
- (BMBF) [grant number 02NUK045A], while no influence or suggestions were made
- with respect to the collection, analysis and interpretation of data, the writing of the
- report and the decision to submit the article for publication.
-
- Author Disclosure Statement:
- No conflicts of interests or competing financial interests exist.
-
- Ethical standards:
- s funded by the Bundesministerium für B
er 02NUK045A], while no influence or su
collection, analysis and interpretation of da
on to submit the article for publication.
attement:
sts or competing financial interests exist. All applicable international, national and/or institutional guidelines for the care and
- use of animals were followed. In particular, the study was approved by the
- government of Upper Bavaria (Az. 55.2-1-54-2532-161-12).
-
- Acknowledgements:
- We thank Laura Dajka, Claire Innerlohinger, Aaron Selmeier, Isabella Zagorski and
- Steffen Heuer for their excellent technical support.
-

- 718 XING, M. 2016. Clinical utility of RAS mutations in thyroid cancer: a blurred picture now 719 emerging clearer. *BMC Med,* 14 **,** 12.
- 720 YAPA, S., MULLA, O., GREEN, V., ENGLAND, J. & GREENMAN, J. 2017. The Role of 721 Chemokines in Thyroid Carcinoma. *Thyroid,* 27 **,** 1347-1359.
- 722 YE, L., ZHOU, X., HUANG, F., WANG, W., QI, Y., XU, H., YANG, S., SHEN, L., FEI, X., XIE, J., CAO, 723 M., ZHOU, Y., ZHU, W., WANG, S., NING, G. & WANG, W. 2017. The genetic 724 landscape of benign thyroid nodules revealed by whole exome and transcriptome 725 sequencing. *Nat Commun,* 8 **,** 15533.
- 726 ZABLOTSKA, L. B., BOGDANOVA, T. I., RON, E., EPSTEIN, O. V., ROBBINS, J., LIKHTAREV, I. A., 727 HATCH, M., MARKOV, V. V., BOUVILLE, A. C., OLIJNYK, V. A. et al.. 2008. A cohort 728 study of thyroid cancer and other thyroid diseases after the Chornobyl accident: 729 dose-response analysis of thyroid follicular adenomas detected during first
- 730 screening in Ukraine (1998-2000). *Am J Epidemiol,* 167 **,** 305-12.

734

Friday Sprain 731 ZANE, M., SCAVO, E., CATALANO, V., BONANNO, M., TODARO, M., DE MARIA, R. & STASSI, G. 732 2016. Normal vs cancer thyroid stem cells: the road to transformation. *Oncogene,* 733 **,** 805-15.

erc@bioscientifica.com

Figure 1: Panel A: Workflow of the study. Panel B: Examples of thyroid HE-sections and histological examination (all samples in SI Figure 1). Panel C: Principal component analysis of RNAseq data (vst counts) generated from the three histological groups.

190x170mm (300 x 300 DPI)

Figure 2: Transcriptomic analysis results

Panel 2A: Dot plot of PROGENy results: All pathways included in PROGENy were visualized. Dot size corresponds to signal strength (bigger is stronger), dot color reflects the direction of the deregulation (red: down-regulated, blues: upregulated). Panel 2B: Box plots of gene expression (vst-counts/DESeq2) for a set of eight selected genes in the three histological groups.

Panel 2C: Volcano plot of differential gene expression analysis. Panel 2D: Dot plot of Gene Set Enrichment Analysis (GSEA) results. Dot size corresponds to -log2(adj-p-value) [lager dot = smaller adj-p-value], dot color reflects the Normalized Enrichment Score (NES) [red: downregulated, blue: upregulated]. NCT: histologically normal tissue from non-exposed mice

NT: histologically normal tissue from exposed mice that harbor proliferative lesions

PL: proliferative lesions tissue from exposed mice

200x301mm (300 x 300 DPI)

erc@bioscientifica.com

Manuscript submitted for review to Endocrine the lated Cancer, within the thyroid gland ^{Page 32} of 51 (figure legend on page 3) (scale bar = $200 \mu m$)

follicular hyperplasia

Page BB refigile 1: Page 2/3: Visualizat Manus Grina submitted for review to Foder the Belated Cancer $(scale bar = 200 \mu m)$

30280506

30291015

30260027

30266844

30291035

30280477

ioscientifica.com

SI Figure 1

Page 1: The left column shows HE tissue sections derived from non-exposed mice without histological thyroid aberrations (9 s a m ples/mice). The right column shows HE tissue sections from exposed mice that exhibited histological aberrations in the thyroid and their histological c lassifications (9 s amples/ mice). Page 2/3: The left column shows exemplary tissue areas for the visualization of immune cell densities within the different tissues. The right column shows exposed cases harboring a histological aberration.

The scale bar (black) in each plot equals 200 micrometers. i

Si Figure 2:

Box plots of immune cell densities observed in the different tissues by investigation of multiple tissue sections from each sample are shown. Statistically significant differences (padj < 0.05) were observed between PL and NCT tissues and between PL and NT tissues by ANOVA and post-hoc Tukey test .

NCT: normal control tissue non-exposed mice

NT: normal tissue of exposed mice harboring PLs

PL: proliferative lesions tissue exposed mice

SI Figure 3A:

Hierarchical clustering of 27 tissue samples by their expression levels of 85 thyroid cancer-associated genes derived from the literature. The clustering results in three main clusters and separates the samples by tissue origin. C1 predominantly contains normal control tissues from non-exposed mice without thyroid aberrations. C2 predominantly contains samples af histelogically normal tissue from exposed mice harboring a histological thyroid aberration. C3 exclusively contains samples from the proliferative lesions tissue present in exposed mice.

Page 37 of 51

Manuscript submitted for review to Endocrine-Related Cancer

B) PCA of full expression set: colors indicate the histological group and shapes indicate the cluster based on SI Figure 3A

C) PCA of 500 top variant genes: colors indicate the histological group and shapes indicate the cluster based on SI Figure 3A

Scree Plot: Variance explained in each PC1 - PC10

D) MDS of full expression set: colors indicate the histological group and shapes indicate the cluster based on SI Figure 3A

erc@bioscientifica.com

SI Figure 4: Upset and Venn plot differentially expressed genes Panel A) Upset plot in intersect mode. Panel B) Upset plot in distinct mode Panel C) Venn diagram of differentially expressed genes

NCT: histologically normal tissue from non-exposed mice

NT: histologically normal tissue from exposed mice that harbor proliferative lesions PL: proliferative lesions tissue from exposed mice

 S lfigure 5

Manuscript submitted for review to Endocrine-Related Cancer

erc**@b@up**ientifica.com

Group

Group

Manuscript submitted for review to Endocrine-Related Cancer

SI Figure 5:

TPM (transcripts per million) expression levels of thyroid cancer -associated genes derived from the literature (total 85 genes , SI Figure 3, SI Table 10). The visualized 3 1 genes were differentially expressed in at least one of the three group comparisons carried out (DESeq2). NCT: normal control tissue from mice with thyroid aberrations, NT: normal tissue from mice harboring P L tissue, PL: proliferative lesions tissue

Manuscript submitted for review to Endocrine-Related Cancer

SI Figure 6 (page 1/2):

Gene expression levels (log2, Agilent 44K Array) in a human papillary thyroid carcinoma data set of the UkrAm post-chernobyl cohort. Expression levels of 26 genes are visualized in normal (N) and tumor (T) tissues. The 26 genes are a subset of the thyroid cancer -associated genes derived from the literature (n=85), which showed statistical significance in at least one out of of the three differential gene expression analysis comparisons made between proliferative lesions (PL), normal tissue (NT) and normal control tissue (NCT) of the investigated mice (SI Figure 5) .

A transcriptomic differential expression analysis between normal and tumor tissue was carried out using the *limma* r package.
The results for the 26 genes are visualized in a volcano plot on page3.

erc@bioscientifica.com

Volcano−Plot: differential expression Tumor tissue vs Normal tissue

paired differential expression analysis (limma), human data / UkrAm cohort (Agilent 44K Arrays)

● NS ● P < 0.1 ● P < 0.1 & Log2(FC) > 0.5

 $Total = 26$ variables

SI Table 3: RNA quality control

SI Table 5: qRT-PCR validation

o RNAseq quatifications (27 Samples)

SI Table 8: PROGENY results

Statistical results derived by PROGENy (pathway response gene analysis) (NTC: normal control tissues from nonexposed mice, NT: normal tissues from exposed mice harbouring proliferative lesions, PL: proliverative lesions tissues from exposed mice)

For Review Only

SI Table 10: tyroid pathogenesis/carcinogenesis-associated genes from literature

