Gene expression signature discriminates sporadic from post-radiotherapy-induced thyroid tumors

Catherine Ory, Nicolas Ugolin*, Céline Levalois, Ludovic Lacroix[1], Bernard Caillou[1], Jean-Michel Bidart[1], Martin Schlumberger[1], Ibrahima Diallo[2,3,4], Florent de Vathaire[2,3,4], Paul Hofman[5,6,7], José Santini[8], Bernard Malfoy[9,10,11] and Sylvie Chevillard*

CEA, DSV, IRCM, LCE, BP6, Fontenay-aux-Roses F-92265, France
[1]Department of Nuclear Medicine and Endocrine Oncology, Institut Gustave Roussy, Villejuif F-94800, France
[2]Inserm, U1018, Villejuif F-94800, France
[3]Institut Gustave Roussy, Villejuif F-94800, France
[4]Université Paris XI, Villejuif F-94800, France
[5]INSERM ERI-21, Nice F-06002, France
[6]University of Nice-Sophia Antipolis, EA 4319, Nice F-06002, France
[7]Laboratory of Clinical and Experimental Pathology and CHU-CRLCC-UNSA Tumour/Tissue Bank of Nice Area and [8]Department of Otorhinolaryngology, Louis Pasteur Hospital, Nice F-06002, France
[9]Institut Curie, Centre de Recherche, Paris F-75248, France
[10]CNRS, UMR 3244, Paris F-75248, France
[11]Université Paris VI, Paris F-75248, France

(Correspondence should be addressed to C Ory; Email: catherine.ory@cea.fr)

*(C Ory and N Ugolin contributed equally to this work)

Abstract

Both external and internal exposure to ionizing radiation are strong risk factors for the development of thyroid tumors. Until now, the diagnosis of radiation-induced thyroid tumors has been deduced from a network of arguments taken together with the individual history of radiation exposure. Neither the histological features nor the genetic alterations observed in these tumors have been shown to be specific fingerprints of an exposure to radiation. The aim of our work is to define ionizing radiation-related molecular specificities in a series of secondary thyroid tumors developed in the radiation field of patients treated by radiotherapy. To identify molecular markers that could represent a radiation-induction signature, we compared 25K microarray transcriptome profiles of a learning set of 28 thyroid tumors, which comprised 14 follicular thyroid adenomas (FTA) and 14 papillary thyroid carcinomas (PTC), either sporadic or consecutive to external radiotherapy in childhood. We identified a signature composed of 322 genes which discriminates radiation-induced tumors (FTA and PTC) from their sporadic counterparts. The robustness of this signature was further confirmed by blind case-by-case classification of an independent set of 29 tumors (16 FTA and 13 PTC). After the histology code break by the clinicians, 26/29 tumors were well classified regarding tumor etiology, 1 was undetermined, and 2 were misclassified. Our results help shed light on radiation-induced thyroid carcinogenesis, since specific molecular pathways are deregulated in radiation-induced tumors.

Endocrine-Related Cancer (2011) **18** 193–206

Introduction

The link between external radiation during childhood and thyroid cancer has been known since 1950 (Duffy & Fitzgerald 1950); until recently this was the only demonstrated etiological risk factor for thyroid cancers (Ron *et al*. 1995). A higher incidence of thyroid cancer has been reported in epidemiological studies after either internal or external exposure to radiation

DOI: 10.1677/ERC-10-0205
1351–0088/11/018–193 © 2011 Society for Endocrinology *Printed in Great Britain* *Online version via http://www.endocrinology-journals.org*

(Ron *et al*. 1995, Cardis *et al*. 2005). A pooled analysis of seven studies established that the excess relative risk of thyroid cancer in subjects irradiated at a young age was very high – 7.7 per Gray (Gy) – the risk being significant for radiation doses as low as 0.1 Gy and increasing linearly with increasing doses (Ron *et al*. 1995). It has been estimated that 88% of thyroid carcinomas occurring in subjects exposed to radiation doses equal to 1 Gy during childhood are radiation-induced. The risk of developing a thyroid carcinoma is the highest 15–30 years after exposure, but is still present after more than 40 years (Ron *et al*. 1995). If exposure occurred in adulthood, the risk is much lower (Richardson 2009).

In parallel, a worldwide increase in thyroid tumors, mainly PTC, has been observed over the last 30 years (Enewold *et al*. 2009). This has led to debate concerning a potential link with changes in environmental exposure linked to nuclear tests, the nuclear industry, and, in Western Europe, Chernobyl disaster fallout. Some data suggest that this increase is at least partly related to the routine screening of thyroid nodules using neck ultrasound and fine needle biopsy, which permit the detection of small papillary carcinomas that would otherwise have gone undetected (Leenhardt *et al*. 2004, Colonna *et al*. 2007). This high prevalence of such small cancers had already been reported in autopsy studies (Harach *et al*. 1985, Yamamoto *et al*. 1990). However, it is not possible to exclude the possibility that some of these thyroid tumors could have been radiation-induced.

Radiation-induced thyroid tumors have no specific histological characteristics (Rubino *et al*. 2002, Williams *et al*. 2008). They are either follicular thyroid adenomas (FTA) or papillary thyroid carcinomas (PTC). These histological subtypes are also the most frequent sporadic thyroid tumors. For these reasons, it is of major interest to identify specific fingerprints of thyroid cancer developing after thyroid radiation exposure that would indicate, with a high probability, the etiology of any tumor. Molecular differences between sporadic and radiation-induced thyroid tumors were sought using microarray transcriptome analysis. The first study, including sporadic and post-Chernobyl PTC, did not show any specific radiation-induced gene expression signature (Detours *et al*. 2005). However, the authors were able to classify their series of tumors by using a signature that was previously found to discriminate between irradiated- and hydrogen-peroxide-treated lymphocytes (Detours *et al*. 2007). Others studies found radiation-induced signatures in post-Chernobyl PTC (Port *et al*. 2007, Stein *et al*. 2010), but without blind validation of the signature. A recent study compared cell cycle protein expression in sporadic and post-radiotherapy PTC, but none of the tested markers could be associated with the etiology (Achille *et al*. 2009), while combinations of protein markers such as matrix metalloproteinases, cathepsins, and neurotrophic tyrosine kinase receptor 1 allowed discrimination of post-Chernobyl PTC as a function of etiology by immunostaining (Boltze *et al*. 2009). Together, these studies must be considered as preliminary and others should be analyzed to establish a robust characterization of the gene expression differences between sporadic and radiation-induced PTC (rPTC).

The aim of this study is to define the ionizing radiation-related molecular specificities of human thyroid tumors developing in the radiation field of patients treated with radiotherapy. We compared the transcriptome profiles of sporadic and radiation-induced FTA (rFTA) and rPTC ($n=28$) obtained after hybrization on 25K oligonucleotide microarrays and identified a signature of 322 genes that discriminated between radiation-induced tumors (FTA and PTC) and their sporadic counterparts. The blind classification of an independent set of 29 tumors using this signature led to the correct classification of 26/29 tumors regarding tumor etiology, 1 was undetermined, and 2 were misclassified.

Materials and methods

Tumor samples

Fifty-seven frozen tumors were obtained from the Institut Gustave Roussy and the Nice Human Biobank (Cancéropole PACA and CRB INSERM, CHU Nice, France; Tables 1 and 2). Histopathologic diagnosis was performed according to the WHO guidelines. The series comprised 12 rPTCs, 15 rFTAs, 15 sporadic PTCs (sPTCs), and 15 sporadic FTAs (sFTAs). Sporadic tumors were matched by histology, age at diagnosis, and tumor, nodes, metastasis (TNM) classification. Radiation exposure was specifically investigated in each patient by examining their medical records and/or by postal contact. All patients were Caucasian except one who was African black.

Tumors from patients who were exposed to external radiation were considered as radiation-induced according to Cahan's criteria (Cahan *et al*. 1948): 1) after radiotherapy, the second neoplasm must arise in the irradiated field and be proved histologically; 2) a latent period of several years must have elapsed between radiation exposure and development of the second neoplasm; and 3) the second tumor must be histologically different from the first tumor. Several patients also received chemotherapy for the treatment of their primary tumors as shown in Tables 1 and 2. Patients hospitalized at the Pasteur Hospital (Department of

Table 1 Clinical data for radiation-induced follicular adenomas and papillary carcinomas

Patient	Histology	Sex	Primary tumor	Age at IR (years)	Age at tumor diagnosis (years)	Thyroid dosimetry (Gy)	Genic alterations	Ethnicity	Tumor size (mm)	Chemotherapy	Detected by
Learning/training set											
RA1	FTA	F	Acne	13	46	20		C	20	−	Screening
RA2	FTA	F	Hodgkin's disease	3	36	42	*HRAS* Q61R	C	22	−	Screening
RA3	FTA	F	Hodgkin's disease	3	36	42		C	13	−	Screening
RA4	FTA	M	Non-Hodgkin lymphoma	8	56	43		C	25	−	Incidental finding
RA5	FTA	M	Nasopharynx carcinoma	9	37	28		C	30	+	Screening
RA6	FTA	F	Non-Hodgkin lymphoma	5	25	43		C	12	−	Incidental finding
RA7	FTA	F	Hodgkin's disease	11	29	21		C	12	+	Screening
RP1	PTC	F	Hodgkin's disease	14	48	43		C	8	−	Incidental finding
RP2	PTC	F	Non-Hodgkin lymphoma	11	22	42		C	11	−	Incidental finding
RP3	PTC	M	Hodgkin's disease	12	30	15		C	15	+	Screening
RP4	PTC	F	Lymphoma	10	40	40		C	10	−	Incidental finding
RP5	PTCFV	M	Neuroblastoma	7	22	12		C	28	+	Screening
RP6	PTC	F	Hodgkin's disease	9	45	40	*BRAF* V600E	C	10	−	Incidental finding
RP7	PTC	F	Acute lymphoblastoid leukemia	6	20	12		C	9	+	Screening
				Mean=8	Mean=35				Mean=16		
Testing set											
XA9	FTA	M	Hodgkin's disease	19	40	40		C	8	+	Screening
XA10	FTA	F	Hodgkin's disease	12	35	8		C	30	+	Screening
XA11	FTA	M	Hodgkin's disease	13	53	Unavailable		Unavailable	Unavailable	Unavailable	Unavailable
XA12	FTA	F	Hodgkin's disease	23	40	43		C	10	+	Screening
XA13	FTA	F	Hodgkin's disease	29	37	41	*KRAS* Q61K	C	10	+	Screening
XA14	FTA	F	Hodgkin's disease	16	60	43		C	13	−	Incidental finding
XA15	FTA	F	Non-Hodgkin lymphoma	19	43	41		C	45	−	Incidental finding
XA16	FTA	F	Uterus	28	60	48		C	30	−	Incidental finding
XP9	PTCFV	M	Hodgkin's disease	23	36	20	RET/PTC3	C	30	+	Screening
XP10	PTC	F	Ovarian teratoma	13	30	0.1	RET/PTC1	AB	3	+	Screening
XP11	PTC	F	Lymphoma	24	59	44		C	12	−	Incidental finding
XP12	PTC	F	Hodgkin's disease	11	61	40	RET/PTC3	C	100	−	Incidental finding
XP13	PTCFV	F	Graves' disease	19	39	Unavailable	*KRAS* Q61R	Unavailable	Unavailable	Unavailable	Unavailable
				Mean=19	Mean=46				Mean=24		

FTA, follicular thyroid adenoma; PTC, papillary thyroid carcinoma; PTCFV, papillary thyroid carcinoma, follicular variant; IR, radiotherapy; C, Caucasian; AB, African black.

Table 2 Clinical data for sporadic tumors

Patient	Histology	Sex	Age at tumor diagnosis (years)	Genic alterations	Ethnicity	Tumor size (mm)	Chemotherapy	Detected by
Learning/training set								
SA1	FTA	F	59		C	26	–	Screening
SA2	FTA	M	63		C	30	–	Screening
SA3	FTA	M	48		C	20	–	Screening
SA4	FTA	F	22		C	40	–	Screening
SA5	FTA	M	44	*HRAS* Q61R	C	33	–	Incidental finding
SA6	FTA	M	24		C	55	–	Screening
SA7	FTA	M	21	*NRAS* Q61R	C	45	–	Incidental finding
SP1	PTCFV	F	54	*BRAF* V600E	C	50	–	Screening
SP2	PTC	F	27	*BRAF* V600E	C	10	–	Screening
SP3	PTC	F	25		C	20	–	Screening
SP4	PTCFV	F	44	RET/PTC3	C	32	–	Screening
SP5	PTC	F	39	*BRAF* V600E	C	18	–	Screening
SP6	PTC	F	34	RET/PTC1	C	13	–	Incidental finding
SP7	PTC	F	23	*BRAF* V600E	C	23	–	Incidental finding
			Mean=37			Mean=29		
Testing set								
XA1	FTA	M	58		C	35	–	Incidental finding
XA2	FTA	F	31		C	20	–	Screening
XA3	FTA	F	29		C	13	–	Screening
XA4	FTA	F	29		C	15	–	Screening
XA5	FTA	F	27		C	30	–	Screening
XA6	FTA	F	59		C	26	–	Screening
XA7	FTA	F	22	*NRAS* Q61K	C	Unavailable	–	Screening
XA8	FTA	F	48		C	38	–	Screening
XP1	PTC	F	17	*BRAF* V600E	C	30	–	Screening
XP2	PTC	F	25		C	25	–	Screening
XP3	PTC	F	39		C	20	–	Screening
XP4	PTC	F	17	RET/PTC1	C	10	–	Screening
XP5	PTC	M	74	*BRAF* V600E	C	25	–	Screening
XP6	PTCFV	F	73	*BRAF* 3 bp Del	C	17	–	Screening
XP7	PTCFV	M	41		C	55	–	Screening
XP8	PTC	F	40	*BRAF* V600E	C	20	–	Screening
			Mean=39			Mean=25		

FTA, follicular thyroid adenoma; PTC, papillary thyroid carcinoma; PTCFV, papillary thyroid carcinoma, follicular variant; C, Caucasian.

Otorhinolaryngology, Nice, France) gave their signed agreement to participate in the study. The protocol was approved by the local ethics committee of the University of Nice ('Comité de Protection des Personnes' and the DRCVI of the CHU of Nice) and by the French Ministry of Research (N°DC-2008-391 and N°AC-2008-83). Written informed consent was obtained from all patients at the Institut Gustave Roussy and the study was performed in accordance with protocols previously approved by the ethics committee of Bicêtre and by the Institutional Review Board of Institut Gustave Roussy.

Calculation of the dose received by the thyroid gland

The approach and software used for retrospective dose evaluations has been previously detailed in Diallo *et al*. (1996, 2009). Technical data concerning the radiotherapy procedure were obtained from radiotherapy records. Direct individual radiotherapy dose calculations were performed using the homemade Dos_EG software package (Diallo *et al*. 1996).

For each patient, a simple mathematical phantom anatomy was generated according to patient dimensions. The shape of external contours and lungs were modeled by means of basic geometric forms. One hundred and eighty-eight sites of interest were specified as points within the phantom. Among these, three were localized at the thyroid. One of these points represents the isthmus and the two others represent the left and right lobes respectively. To calculate the dose to an anatomic site, beams were positioned on the phantom according to details from the patient's record,

as well as information on equipment, common treatment techniques, and guidelines used at the time of the treatment. Radiotherapy parameters included beam size, shape and inclination, location, radiation energy, and delivered treatment dose. The dose calculation algorithm accounted for primary radiation from the treatment machine and scattered radiation from the patient and from beam collimation, leakage radiation, and lung heterogeneity. The local radiation dose at a site was defined as the cumulative absorbed dose resulting from the contribution of all beams involved in the radiotherapy regimen. For all patients who received radiotherapy for a primary cancer, the doses received by the thyroid gland were calculated and are presented in Table 1.

Mutation analysis

RET/PTC1 and *RET/PTC3* rearrangements were detected by RT-PCR as described in Smida *et al.* (1999). Mutations of *BRAF* and *HRAS*, *KRAS*, and *NRAS* genes were analyzed by cDNA sequencing (Beckman Coulter Genomics (Cogenics), Grenoble, France) after PCR amplification. Primer sequences are listed in Supplementary Table 1, see section on supplementary data given at the end of this article.

RNA extraction, labeling, and hybridization

The protocol for RNA extraction, RNA amplification, and labeling was as described in Daino *et al.* (2009). Quality of RNA was assessed using the RNA 6000 Nano Lab-On-Chip as developed on the Agilent 2100 Bioanalyzer device (Agilent Technologies, Palo Alto, CA, USA). All specimens included in this study displayed an RNA integrity number over 7.7. Each tumor sample was co-hybridized with a common pool of amplified normal thyroid RNA used as a reference, and all hybridizations were duplicated in dye-swap. Samples were hybridized in 500 µl of 2× SSC (Gibco), 0.1% SDS (Gibco), 5 µg salmon sperm DNA (Invitrogen) on human 25K 50–52mer oligo-microarrays from a national genomic platform (Resogen Program, RNG/MCR, Evry). Spotted slides were prepared as recommended by the manufacturer. The arrays were incubated overnight at 50 °C in an Agilent hybridization oven. After hybridization, the slides were successively washed for 5 min with agitation in 2× SSC plus 0.1% SDS, 1× SSC, 0.2× SSC, and 0.05× SSC, then dried by centrifugation.

Preprocessing of microarrays

The procedure for array scanning, calculation of fluorescence intensities, data normalization, and statistical analysis were performed as described previously (Chevillard *et al.* 2004) with modifications (Bastide *et al.* 2009, Daino *et al.* 2009). All data is deposited in Array Express database (http://www.ebi.ac.uk/arrayexpress/, E-MEXP-2739).

Search for molecular signature of tumor etiology

The thyroid tumors were split into 2 sets: a learning/training set, used to search for the molecular signature, this comprised of 7 rFTAs, 7 sFTAs, 7 rPTCs, and 7 sPTCs, and a testing set, which comprised of the remaining 29 tumors (16 FTAs and 13 PTCs as testing tumors), used for blind classification. At the time of the blind classification the tumor histology, but no radiation exposure history, was known.

The method for classifying the etiology comprised the following four steps:

1) A learning step, based on a classification expectation–maximization algorithm, to select sets of candidate genes whose expression discriminates between the two subgroups (Daino *et al.* 2009);
2) a training step to select from the sets of candidate genes those with the highest potential to classify training tumors correctly (Dudoit *et al.* 2002);
3) the compilation of a unique discriminating set of genes and standardization of their expressions according to gene expression variability in the subgroups; and
4) the blind classification, case-by-case, of testing tumors.

Learning step

Combinatorial matrices of 8 tumors (4 sFTAs or sPTCs out of the 14 sporadic cases versus 4 rFTAs or rPTCs out of the 14 radiation-induced cases) were built from the learning/training set of tumors, the 20 remaining tumors of each combinational matrix being used as training tumors. To avoid bias due to histology, each half matrix (four sporadic or radiation-induced tumors) should include at least one tumor of each histology, either FTA or PTC. In the same way, to avoid bias due to tumor size and specifically to microcarcinomas, we split the tumors of the learning group into three categories: group 1 size $\leq$25th percentile (12 mm); group 2 size $>$25th percentile and $<$75th percentile (30 mm); and group 3 size $\geq$75th percentile. The tumor size distribution in group 2 was similar for sporadic and radiation-induced tumors, whatever the histology, PTC or FTA. Each half matrix should also contain at least one tumor of group 2 size and should

differ from the others by at least 50% in tumor composition. With these rules, we excluded any impact of size or tumor histology in the signature. Finally, 81 8-tumor training matrices were considered.

Training step

Each training matrix was used to classify, by a two-step principal component analysis (PCA) method (see 'Blind classification by a two-step PCA method'), the 20 remaining tumors from the respective combinatorial matrix (training tumors). Then, the rules to select a matrix were applied so that at least one of the training tumors was well classified and none were misclassified, otherwise the matrix was discarded. This was performed for each training matrix. The process continued if at least 90% of the training tumor classifications were validated by at least one of the training matrices. In these conditions, 10% of the tumors may not be validated, but none must be rejected by the retained training matrices. The training step selected 26 matrices that were able to classify at least one training tumor, without false classification.

Compilation of a unique discriminating set of genes

The 26 selected training matrices (e) were compiled in a unique ξ matrix. Among the 974 genes belonging to the ξ matrix, we finally retained, for the final gene signature, in a ξ' matrix, 322 genes as function of a threshold ($P=0.001$) on the frequencies of relevance, which is defined as the frequency at which a given gene and a given tumor are found together in a selected training matrix, weighted by the number of training tumors well classified by this training matrix. Then, the relevance of the 322 genes of the final signature was proved by blind classification of the 29 testing tumors by a two-step PCA method. All the method is detailed in Supplementary Table 2, see section on supplementary data given at the end of this article.

Blind classification by a two-step PCA method

Let us define g and T as ξ' matrix and a given validation tumor respectively or a given training matrix and a corresponding given training tumor respectively. The eigenvectors and eigenvalues were calculated for the considered g matrix. These vectors defined a new space maximizing specifically the asymmetry between the two groups of tumors (sporadic or radiation-induced) of the matrix (classification space). The classification of a T tumor was then realized by its location in the classification space, compared with the location of the two subgroups of tumors of the g matrix.

When taking into account more than three eigenvectors, to assess more precisely the distances between the vectors, we used a decision-making tool based on calculation of the root mean square (Supplementary Table 2, see section on supplementary data given at the end of this article).

Results

To search for a signature of thyroid tumor etiology, we have conducted a 25K microarray transcriptome analysis on a series of sporadic and radiation-induced benign and malignant thyroid tumors. For blind classification of new tumors, we developed a unique strategy based on an expectation-maximisation algorithm that identifies a gene expression signature with the greatest potential to discriminate between the two subgroups of tumors in a learning/training set, and on a two-step PCA analysis, which defines an N dimensional classification space that presents the greatest asymmetry between the two subgroups.

Identification of the discriminating signature between radiation-induced and sporadic tumors

To increase the likelihood that the tumors used to search the signature are radiation-induced and not sporadic, we paid special attention to the choice of patients and tumors included in the learning group. Only patients treated before 15 years of age, which is considered the period of high thyroid radiation sensitivity, were included in the learning/training set of tumors (Table 1). Moreover, to prevent any bias due to multiple comparison failure, patients with sporadic tumors were selected to match, as far as possible, sex, ethnicity, and age at tumor diagnosis of patients with radiation-induced tumors (Tables 1 and 2). Mean age at tumor diagnosis was 38 vs 40 years old for rFTA and sFTA and 32 vs 35 years old for rPTC and sPTC respectively. In the learning set, the female to male ratio was 1.8 (9/5 females/males) in sporadic tumors and 2.5 (10/4 females/males) in radiation-induced tumors (Tables 1 and 2).

A four-step method (learning, training, compilation of a unique discriminating set of genes, and blind classification of new tumors) was applied to a learning/training set of 28 tumors (FTA or PTC) of known etiology, sporadic, or radiation-induced, which included 7 rFTA, 7 sFTA, 7 rPTC, and 7 sPTC. When considering sporadic and radiation-induced tumors as the two subgroups in the search for a signature, whatever the histological subtypes (FTA or PTC),

a final discriminating signature of 322 genes was found to be differentially expressed when comparing the two groups. This set included 137 overexpressed and 185 underexpressed radiation-induced tumors as compared with sporadic tumors (Supplementary Table 3, see section on supplementary data given at the end of this article). The Cartesian coordinate system, defined by the two-step PCA method, determined the testing–classification space, which was found to be organized as three subspaces, corresponding to the sFTA, the sPTC, and all radiation-induced tumors of the learning set (Fig. 1).

In order to check that this molecular signature of tumor etiology was not specific to some DNA mutations, we searched for *RAS*, *RET/PTC*, and *BRAF* genetic alterations in the learning/training set of tumors (Tables 1 and 2). Mutations at codon 61 of *NRAS*, *HRAS*, or *KRAS* were found in one rFTA and two sFTA. *RET/PTC* rearrangements were identified in two sPTC and *BRAF* mutations were detected in four sPTC and one rPTC (with no overlap with *RET* rearrangements). Thus, it is unlikely that the signature could be specific to any type of mutation.

To understand the molecular specificities of radiation-induced thyroid tumorigenesis, we searched for the biological function and relationship between the 322 genes of the discriminating signature (Supplementary Table 3, see section on supplementary data given at the end of this article). However, to be more exhaustive in the overview of the radiation-induced deregulated pathways, we also included the 651 preselected genes found during the training step to be deregulated (Supplementary Table 4, see section on supplementary data given at the end of this article). While not included in the final signature, these genes were able to classify tumors in several combinations of tumors of the learning/training set without misclassification (see Materials and methods section).

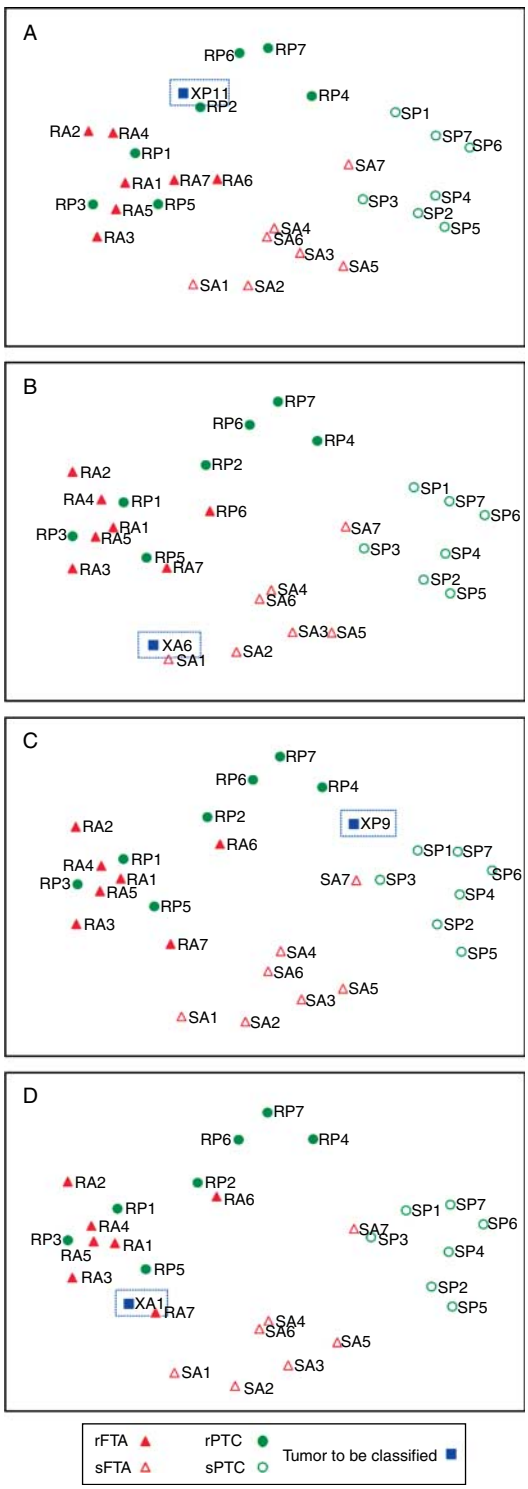

Figure 1 Blind validation of the radiation-induced signature by PCA analysis in the classification space defined by the tumors of the learning/training set. By the two-step PCA method, the tumors of the learning set, being FTA (red triangle) and PTC (green circle), either sporadic (empty symbols) or radiation-induced (full symbols), defined a validation space in which each tumor of the testing set is projected to identify its etiology. The figure represents examples of the relative positioning of four testing tumors (blue square) in this validation space. (A) A well-classified rPTC (XP11), (B) a well-classified sFTA (XA6), (C) the outlier rPTC tumors (XP9), positioned in the validation space between the rPTC and sPTC subgroups.
(D) A misclassified sFTA (XA1). Values of tumors used for hypothesis finding in (A–D) seem to differ slightly. This is an artifact due to data representation in two dimensions.
The validation space is defined in ten dimensions, according to the tumors of the learning/training set, and each tumor of the validation set is projected in this space to be classified.
To visualize the results of tumor classification, the space is restrained to three dimensions and projected in two dimensions. During this reduction, the relative localization of the tumors could appear slightly modified.

Validation of the radiation-induced tumor signature

Using this signature, a two-step PCA analysis was used for blind classification of the 29 tumors of the testing set. Patient and tumor characteristics are given in Tables 1 and 2. For blind validation of the molecular signature, each testing tumor was projected into the classification space, allowing us to propose an etiology depending on the relative positioning of the testing tumors compared with the learning/training tumors (Fig. 1). This signature was robust, since it correctly predicted the etiology of 26 of the 29 tumors. Six of the eight sFTA, the eight rFTA, the eight sPTC, and four of the five rPTC were well classified. One rPTC (XP9) was clusterized between the group of sPTC and the group of radiation-induced tumors, so it was not possible to propose an etiology for this tumor (Fig. 1C), and two sFTA (XA1, Fig. 1A and XA8) were classified among the radiation-induced tumors. Since these tumors developed in patients without known history of radiation exposure, these may be false results but, alternatively, these patients may have high thyroid sensitivity and may have been exposed to ionizing radiation in the past. In any case, the present signature of thyroid tumor etiology has a very good negative predictive value, as all testing tumors diagnosed as sporadic were indeed sporadic tumors and the signature has a rather good positive predictive value, as 12 of the 14 radiation-induced testing tumors were well diagnosed (1 and 0.85 respectively).

It should be mentioned that using a similar approach we also searched for a signature of tumor etiology in each histological subtype. We found two distinct gene expression signatures discriminating sFTA and rFTA, and sPTC and rPTC but, due to the small number of tumors, we did not have enough samples to check the validity of these signatures by blind classification of an independent series of tumors (data not shown).

Discussion

Most methods used to analyze microarray data wish to identify groups of genes that have coherent patterns of expression with large variance across groups of samples. Unfortunately, using these methods, we did not find any signature of etiology. Gene shaving is a useful alternative method, which uses PCA to find the space direction which captures the majority of variance in the whole data set and thus permits the identification of those genes which are able to separate the two groups of tumors, whatever the variance of individual gene expression. Unfortunately, gene shaving is not suitable for a small series of samples, such as our series of tumors. For this reason, we have developed a method based on a similar strategy to that of gene shaving, which is adapted to a small number of samples. It allows us to find the space direction, which maximizes, if they exist, the criteria discriminating the two groups of tumors.

Post-radiotherapy tumors (R-tumors)

The choice of R-tumors, specifically those used for the learning/testing step, was crucial for the veracity of the results. Due to the high frequency of sporadic thyroid tumors in the general population, we followed very stringent criteria to minimize the risk of including a sporadic tumor (S-tumor) in the group of R-tumors: we strictly followed Cahan's criteria for R-tumor selection. Moreover, these tumors should have developed in children who were exposed to radiotherapy while younger than 15, since this corresponds to the period of highest thyroid radiation sensitivity (Steliarova-Foucher *et al*. 2006). Regarding the risk as a function of the dose received at the thyroid gland, it was shown that, in this age range, the relative risk at 1 Gy was estimated at 8.7 and the attributable risk at 1 Gy was estimated to be 88% (Ron *et al*. 1995). At 1–15 Gy to the thyroid, the relative risk was found to be 3.5 if radiotherapy occurred before 10 years, and 0.9 if it occurred later, the risk then decreases following a symmetrical curve, consistent with a cell killing effect (Sigurdson *et al*. 2005). The relative risk for doses of about 40 Gy is 5 and the attributable risk 80% (Sigurdson *et al*. 2005). It should be estimated that after thyroid radiation exposure during childhood or adolescence, at doses from 20 to 40Gy, the expected frequency of radiation-induced tumors ranged from 80 to 90%. Therefore, 1 of 14 tumors of the R-learning group may be sporadic. Therefore, to evaluate the consequence, if an S-tumor was unfortunately included in the R-group, we have mimicked the situation by searching a signature using a learning set composed of 13 R-tumors plus 1 S-tumor ('false R'-tumor) versus 14 S-tumors (data not shown). Overall, due to the very stringent method applied for gene selection, most often we did not find a signature and when it was found the 'false R'-tumor systematically localized between the two groups of R- and S-tumors. This situation could have consequences if tumors of the validation set were classified only as a function of this 'false R'-tumor. In fact, we could have raised the question regarding RA7 (learning tumor), since XA1 and XA8 (validation tumors) are considered misclassified

only as function of RA7, which was localized at the border of the R-group (Fig. 1D).

Regarding the thyroid radiosensitivity in adulthood, it should be mentioned that four tumors of the validation group, which are classified as R-tumors, developed in patients who received radiotherapy at 30, 35, 36, and 37 years old. These data are in agreement with the fact that although the risk of radiation-induced thyroid cancer decreases as age increases, the excess of relative risk per gray is still positive up to 40 years old (Ron *et al*. 1995). It should be mentioned that in a cohort of patients developing secondary radiation-induced thyroid tumors, 11 of 27 patients received chemotherapy for their primary tumor (Table 1). Importantly, for the validity of the data, it was shown that chemotherapy for the first cancer was not associated with thyroid cancer risk and it did not modify the effect of radiotherapy (Sigurdson *et al*. 2005).

Our study is the first transcriptome analysis of post-radiotherapy radiation-induced thyroid tumors. We found a robust molecular signature of thyroid tumor etiology – demonstrated by the correct blind classification of 26 of 29 tumors. Few studies on transcriptome analysis of post-Chernobyl thyroid tumors have been published. In this study, most of the patients were externally exposed to radiation to treat Hodgkin's disease or non-Hodgkin lymphoma. The estimated mean dose received by the thyroid gland varies between 0.1 and 43 Gy in repeated exposure at high-dose rates (Table 1). In contrast, after the Chernobyl accident, victims' thyroids were mainly chronically contaminated by ^{131}I ingestion. In the exposed population, the cumulative thyroid radiation doses ranged from 0.02 Gy to more than 10 Gy, but most people received doses less than 1 Gy. A low-iodine diet in the exposed population was reported to be an important parameter in the development of these tumors (Williams *et al*. 2008). The relevance of extrapolation of conclusions from data on tumors occurring after exposure to external radiation to post-Chernobyl tumors that occurred after internal ^{131}I contamination is unclear.

An overlap was found between post-radiotherapy deregulated genes, identified in this study, and post-Chernobyl deregulated genes. *EPB41L3*, a tumor suppressor included in the 322-gene signature, and *RERG*, *C13ORF33*, *GZMH*, *MST150*, *RARRES1*, *RIPK4*, and *SFRP1*, found in the enlarged set, were reported in a set of deregulated genes in post-Chernobyl PTC (Port *et al*. 2007). Genes such as *ABI2*, *COL4A5*, *FAT3*, *IGFBP3*, *KRTAP3-2*, and *SPOCK1* and several immunoglobulin chains were deregulated in post-radiotherapy tumors, while genes of the same family and/or function, such as *ABI3*, *COL4A6*, *FAT2*, *IGFBP1*, *KRTAP2-1*, *KRTAP2-4*, *SPOCK2* and other immunoglobulin chains, were also deregulated in the study of Port *et al*. (2007), suggesting common deregulated pathways in post-radiotherapy and post-Chernobyl tumors. Since we also observed genes in common with another previously published post-Chernobyl series of thyroid tumors (Detours *et al*. 2007), (genes from the signature: *C4A*, *CLU*; *DCI*, *DHCR24*, *EGFR*, *EGR3*, *GTF2H2*, *ICAM3*, *NRIP1*, *PLA2R1*, and *RPS19*. From the enlarged set: *EFNA1*, *EIF2AK2*, *FAM38A*, *MED1*, *MGEF8*, *PPL*, *SCARA3*, *SMO*, and *ZFHX4*), several of these genes could be potential markers of radiation-induced thyroid tumors, independently of the histology and the internal–external type of radiation exposure. It should be mentioned that *KLK10*, underexpressed in post-radiotherapy-induced thyroid tumors, was also specifically identified as a downregulated gene in radiation-transformed human mammary epithelial cells (Liu *et al*. 1996).

Molecular specificities of post-radiotherapy-induced tumors

Deregulated genes in radiation-induced tumors are mostly involved in molecular mechanisms such as cellular response to oxidative stress and irradiation, response to hypoxia, regulation of p53 function, immune response, and signal transduction pathways including mitogen-activated protein kinase (MAPK), epidermal growth factor receptor (EGFR), RAC/CDC42, hedgehog, transforming growth factor (TGF)/bone morphogenetic protein (BMP), calcium signaling, and WNT canonical and noncanonical pathways (Fig. 2). The WNT/β-catenin pathway has already been implicated in normal thyroid function and in thyroid carcinogenesis (Garcia-Rostan *et al*. 2001, Castellone *et al*. 2009). Interestingly, genes known to connect the WNT pathway with cellular response to hypoxia (*ARD1A*; Lim *et al*. 2008), to the EGFR, TGF, and NOCHT (*TLE2*) pathways (Buscarlet & Stifani 2007) were also found to be deregulated.

As described by Detours *et al*. (2007), we found that radiation-induced tumors are imprinted by an oxidative stress hallmark since deregulated genes such as *CFL1*, *CLU*, *DHCR24*, *FMOD*, *GSTM1-2*, and *PRNP* are involved in sensitivity to H_2O_2 or ionizing radiation exposure. Interestingly, chronic oxidative stress is known to provoke increasing chromosomal instability (Abdel-Rahman 2008) and we observed that genes regulating the G_2/M phase and/or checkpoints (*PTTG1*, *SPAG5*, *DNAH9*, *C13orf3*, *NUP153*, *CENPE*, *CDCA8*,

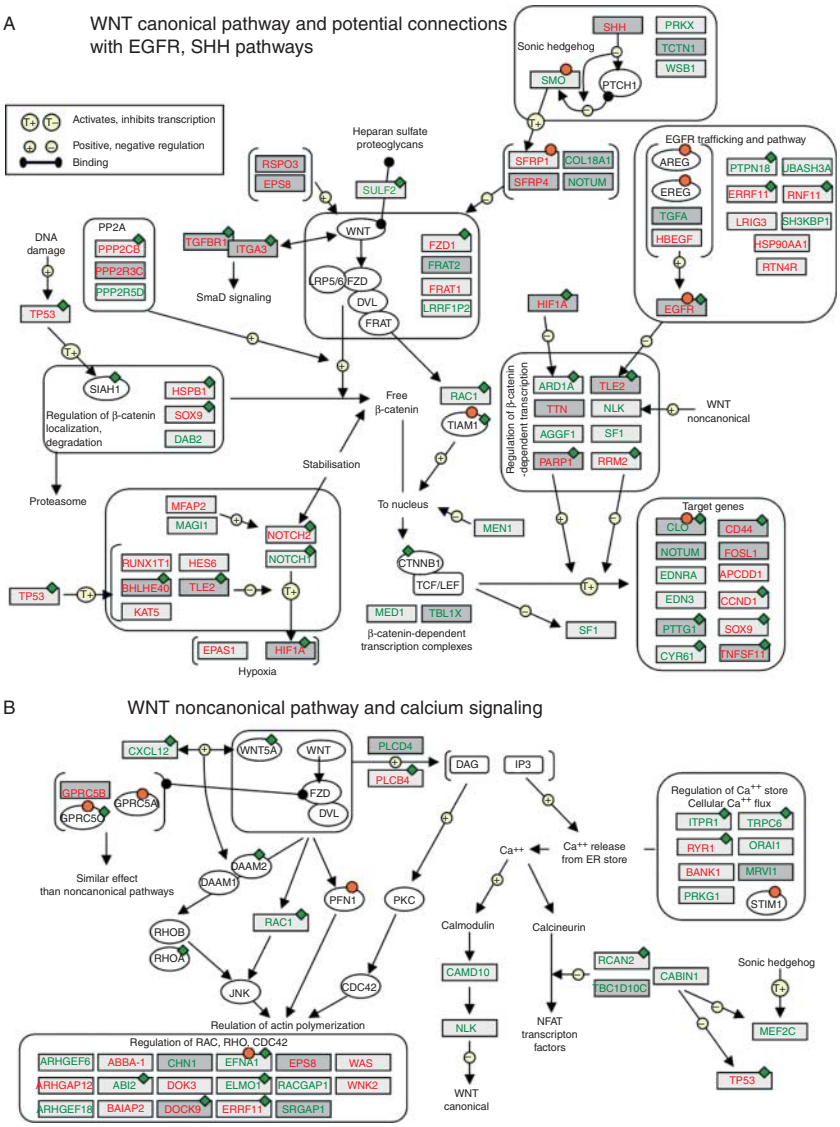

Figure 2 Genes of the WNT canonical and noncanonical pathways deregulated in radiation-induced thyroid tumors. The figures represent a simplified overview of WNT canonical pathway with potential connections with EGFR, SHH, or NOTCH pathways (A) and WNT noncanonical pathway (B). Genes overexpressed (green) or underexpressed (red), either in the discriminating signature (322 genes) or deregulated with less recurrence in post-radiotherapy tumors, are indicated by dark or light rectangles respectively. Orange circles show genes found to be deregulated in post-Chernobyl tumors, while green diamonds indicate deregulated genes in thyroid sporadic tumors (see Discussion section).

STK10, or *NDC80*) were differentially expressed in radiation-induced tumors.

One of the central players in the cellular response to DNA damage and in the maintenance of genome integrity following cell exposure to oxidative stress is the tumor suppressor p53. In thyroid cancer, a very low frequency of *TP53* mutations was observed in sporadic and in post-Chernobyl tumors (Hillebrandt *et al*. 1996, Nikiforov *et al*. 1996). However, p53 protein overexpression was described, in PTC and anaplastic thyroid carcinomas, not associated with gene mutation, suggesting that other mechanisms may regulate p53 in thyroid tumors (Soares *et al*. 1994, Malaguarnera *et al*. 2007). In radiation-induced tumors, a decrease in *TP53* gene expression was observed, but it was not selected in the highly discriminating signature (Fig. 3). Interestingly, several genes involved in the regulation of p53 turnover, p53 phosphorylation, and/or p53 transcriptional activity were differentially expressed depending on the etiology. This included some highly discriminating genes of the signature, such as *MAPKAPK2*, *PARP1*, *IFI16*, *PTTG1*, *RAD32B*, and

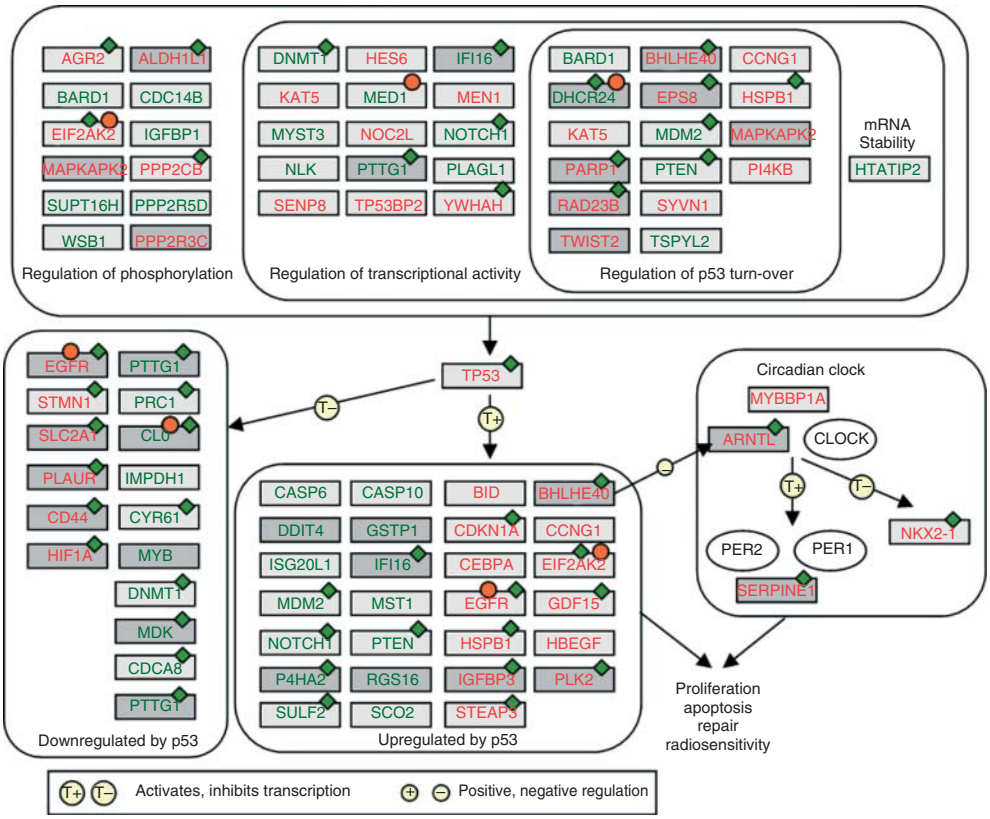

Figure 3 Genes involved in the regulation of p53 turnover and/or function. Genes overexpressed (green) or underexpressed (red), either in the discriminating signature (322 genes) or deregulated with less recurrence in post-radiotherapy tumors, are indicated by dark or light rectangles respectively. Orange circles show genes found to be deregulated in post-Chernobyl tumors, while green diamonds indicate deregulated genes in thyroid sporadic tumors (see Discussion section).

BHLHE40, which participate in cellular response to stress, DNA damage, DNA repair, and genetic stability. *DHCR24*, which protects p53 against *MDM2*-dependent degradation following oxidative stress (Kuehnle *et al*. 2008), has already been found to be deregulated in Chernobyl tumors (Detours *et al*. 2007). The *MDM2* gene is overexpressed in radiation-induced tumors along with genes regulating its function or stabilization, such as *KAT5*, *TWIST2*, *TSPYL2*, *MAPKAPK2*, and *HSPB1*. Recent data have pointed to a link between cancer, DNA damage response, and circadian clock regulation (Sahar & Sassone-Corsi 2009). Three genes involved in the regulation of this pathway were found to be deregulated (*BHLHE40*, *ARNTL/BMAL1*, and *MYBBP1A*), suggesting that the circadian clock, which is physiologically regulated by, and regulates, the endocrine functions of thyroid tissue, may be also be implicated in radiation-induced tumorigenesis.

Finally, another feature of exposure of the thyroid gland to radiation is the increased frequency of autoimmune dysfunctions. In this respect, an immunoglobulin kappa light chain variable region (accession number L12079) was identified in patients with Graves' disease (Chazenbalk *et al*. 1993), as well as genes involved in T lymphocyte function (*POU2AF1*, *CLECL1*, *IL2RG*, *CCL22*, *MADCAM1*, *LY75*, and *TRAT1*) were deregulated and thus could also be related to the immune thyroid status of radiation-induced tumors. In parallel, genes associated with thyroid dysfunction or susceptibility to thyroid disease, such as *IGFBP3*, *DUOX2*, *NKX2 1/TTF1*, *GNA11*, and *SLC26A4/pendrin* (Moreno *et al*. 2002, Moya *et al*. 2006, Kero *et al*. 2007, Kopp *et al*. 2008, Kursunluoglu *et al*. 2009), or with other autoimmune diseases, such as *IL17B*, *IRF4*, *LAIR2*, and *RC3H1*, were also deregulated.

In conclusion, we found a highly specific post-radiotherapy gene signature to diagnose thyroid tumor etiology independently of histological subtype. Several genes of this signature were also deregulated in post-Chernobyl PTC, although the type of exposure, range of dose, and other parameters such as iodine consumption are different. It could be interesting, using our methodology, to search for a signature in post-Chernobyl

thyroid tumors and indeed to assess the relevance of our signature in classifying post-Chernobyl tumors compared with sporadic tumors.

Supplementary data

This is linked to the online version of the paper at http://dx.doi.org/10.1677/ERC-10-0205.

Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding

This work was supported by Electricité de France (CEA DRR and RB 2005–10) and the European Union 6th Framework GenRiskT.

References

Abdel-Rahman WM 2008 Genomic instability and carcinogenesis: an update. *Current Genomics* **9** 535–541. (doi:10.2174/138920208786847926)

Achille M, Boukheris H, Caillou B, Talbot M, de Vathaire F, Sabatier L, Desmaze C, Schlumberger M & Soria JC 2009 Expression of cell cycle biomarkers and telomere length in papillary thyroid carcinoma: a comparative study between radiation-associated and spontaneous cancers. *American Journal of Clinical Oncology* **32** 1–8. (doi:10.1097/COC.0b013e3181783336)

Bastide K, Guilly MN, Bernaudin JF, Joubert C, Lectard B, Levalois C, Malfoy B & Chevillard S 2009 Molecular analysis of the Ink4a/Rb1-Arf/Tp53 pathways in radon-induced rat lung tumors. *Lung Cancer* **63** 348–353. (doi:10.1016/j.lungcan.2008.06.007)

Boltze C, Riecke A, Ruf CG, Port M, Nizze H, Kugler C, Miethke C, Wiest N & Abend M 2009 Sporadic and radiation-associated papillary thyroid cancers can be distinguished using routine immunohistochemistry. *Oncology Reports* **22** 459–467. (doi:10.3892/or_00000457)

Buscarlet M & Stifani S 2007 The 'Marx' of Groucho on development and disease. *Trends in Cell Biology* **17** 353–361. (doi:10.1016/j.tcb.2007.07.002)

Cahan WG, Woodward HQ, Higinbotham NL, Stewart FW & Coley BL 1948 Sarcoma arising in irradiated bone; report of 11 cases. *Cancer* **1** 3–29. (doi:10.1002/1097-0142(194805)1:1<3::AID-CNCR2820010103>3.0.CO;2-7)

Cardis E, Kesminiene A, Ivanov V, Malakhova I, Shibata Y, Khrouch V, Drozdovitch V, Maceika E, Zvonova I, Vlassov O *et al.* 2005 Risk of thyroid cancer after exposure to ^{131}I in childhood. *Journal of the National Cancer Institute* **97** 724–732. (doi:10.1093/jnci/dji129)

Castellone MD, DeFalco V, Rao DM, Bellelli R, Muthu M, Basolo F, Fusco A, Gutkind JS & Santoro M 2009 The β-catenin axis integrates multiple signals downstream from RET/papillary thyroid carcinoma leading to cell proliferation. *Cancer Research* **69** 1867–1876. (doi:10.1158/0008-5472.CAN-08-1982)

Chazenbalk GD, Portolano S, Russo D, Hutchison JS, Rapoport B & McLachlan S 1993 Human organ-specific autoimmune disease. Molecular cloning and expression of an autoantibody gene repertoire for a major autoantigen reveals an antigenic immunodominant region and restricted immunoglobulin gene usage in the target organ. *Journal of Clinical Investigation* **92** 62–74. (doi:10.1172/JCI116600)

Chevillard S, Ugolin N, Vielh P, Ory K, Levalois C, Elliott D, Clayman GL & El-Naggar AK 2004 Gene expression profiling of differentiated thyroid neoplasms: diagnostic and clinical implications. *Clinical Cancer Research* **10** 6586–6597. (doi:10.1158/1078-0432.CCR-04-0053)

Colonna M, Guizard AV, Schvartz C, Velten M, Raverdy N, Molinie F, Delafosse P, Franc B & Grosclaude P 2007 A time trend analysis of papillary and follicular cancers as a function of tumour size: a study of data from six cancer registries in France (1983–2000). *European Journal of Cancer* **43** 891–900. (doi:10.1016/j.ejca.2006.11.024)

Daino K, Ugolin N, Altmeyer-Morel S, Guilly MN & Chevillard S 2009 Gene expression profiling of α-radiation-induced rat osteosarcomas: identification of dysregulated genes involved in radiation-induced tumorigenesis of bone. *International Journal of Cancer* **125** 612–620. (doi:10.1002/ijc.24392)

Detours V, Wattel S, Venet D, Hutsebaut N, Bogdanova T, Tronko MD, Dumont JE, Franc B, Thomas G & Maenhaut C 2005 Absence of a specific radiation signature in post-Chernobyl thyroid cancers. *British Journal of Cancer* **92** 1545–1552. (doi:10.1038/sj.bjc.6602521)

Detours V, Delys L, Libert F, Weiss Solis D, Bogdanova T, Dumont JE, Franc B, Thomas G & Maenhaut C 2007 Genome-wide gene expression profiling suggests distinct radiation susceptibilities in sporadic and post-Chernobyl papillary thyroid cancers. *British Journal of Cancer* **97** 818–825. (doi:10.1038/sj.bjc.6603938)

Diallo I, Lamon A, Shamsaldin A, Grimaud E, de Vathaire F & Chavaudra J 1996 Estimation of the radiation dose delivered to any point outside the target volume per patient treated with external beam radiotherapy. *Radiotherapy and Oncology* **38** 269–271. (doi:10.1016/0167-8140(96)01713-6)

Diallo I, Haddy N, Adjadj E, Samand A, Quiniou E, Chavaudra J, Alziar I, Perret N, Guerin S, Lefkopoulos D *et al.* 2009 Frequency distribution of second solid cancer locations in relation to the irradiated volume among 115 patients treated for childhood cancer. *International Journal of Radiation Oncology, Biology, Physics* **74** 876–883. (doi:10.1016/j.ijrobp.2009.01.040)

Dudoit S, Fridlyand J & Speed TP 2002 Comparison of discrimination methods for the classification of tumors using gene expression data. *Journal of the American Statistical Association* **97** 77–87. (doi:10.1198/016214502753479248)

Duffy BJ Jr & Fitzgerald PJ 1950 Thyroid cancer in childhood and adolescence; a report on 28 cases. *Cancer* **3** 1018–1032. (doi:10.1002/1097-0142(1950)3:6<1018::AID-CNCR2820030611>3.0.CO;2-H)

Enewold L, Zhu K, Ron E, Marrogi AJ, Stojadinovic A, Peoples GE & Devesa SS 2009 Rising thyroid cancer incidence in the United States by demographic and tumor characteristics, 1980–2005. *Cancer Epidemiology, Biomarkers and Prevention* **18** 784–791. (doi:10.1158/1055-9965.EPI-08-0960)

Garcia-Rostan G, Camp RL, Herrero A, Carcangiu ML, Rimm DL & Tallini G 2001 β-Catenin dysregulation in thyroid neoplasms: down-regulation, aberrant nuclear expression, and CTNNB1 exon 3 mutations are markers for aggressive tumor phenotypes and poor prognosis. *American Journal of Pathology* **158** 987–996.

Harach HR, Franssila KO & Wasenius VM 1985 Occult papillary carcinoma of the thyroid A "normal" finding in Finland. A systematic autopsy study. *Cancer* **56** 531–538. (doi:10.1002/1097-0142(19850801)56:3<531::AID-CNCR2820560321>3.0.CO;2-3)

Hillebrandt S, Streffer C, Reiners C & Demidchik E 1996 Mutations in the p53 tumour suppressor gene in thyroid tumours of children from areas contaminated by the Chernobyl accident. *International Journal of Radiation Biology* **69** 39–45. (doi:10.1080/095530096146165)

Kero J, Ahmed K, Wettschureck N, Tunaru S, Wintermantel T, Greiner E, Schutz G & Offermanns S 2007 Thyrocyte-specific G_q/G_{11} deficiency impairs thyroid function and prevents goiter development. *Journal of Clinical Investigation* **117** 2399–2407. (doi:10.1172/JCI30380)

Kopp P, Pesce L & Solis SJ 2008 Pendred syndrome and iodide transport in the thyroid. *Trends in Endocrinology and Metabolism* **19** 260–268. (doi:10.1016/j.tem.2008.07.001)

Kuehnle K, Crameri A, Kalin RE, Luciani P, Benvenuti S, Peri A, Ratti F, Rodolfo M, Kulic L, Heppner FL *et al.* 2008 Prosurvival effect of DHCR24/Seladin-1 in acute and chronic responses to oxidative stress. *Molecular and Cellular Biology* **28** 539–550. (doi:10.1128/MCB.00584-07)

Kursunluoglu R, Turgut S, Akin F, Bastemir M, Kaptanoglu B, Genc O & Turgut G 2009 Insulin-like growth factor-I gene and insulin-like growth factor binding protein-3 polymorphism in patients with thyroid dysfunction. *Archives of Medical Research* **40** 42–47. (doi:10.1016/j.arcmed.2008.10.009)

Leenhardt L, Bernier MO, Boin-Pineau MH, Conte Devolx B, Marechaud R, Niccoli-Sire P, Nocaudie M, Orgiazzi J, Schlumberger M, Wemeau JL *et al.* 2004 Advances in diagnostic practices affect thyroid cancer incidence in France. *European Journal of Endocrinology* **150** 133–139. (doi:10.1530/eje.0.1500133)

Lim JH, Chun YS & Park JW 2008 Hypoxia-inducible factor-1α obstructs a Wnt signaling pathway by inhibiting the hARD1-mediated activation of β-catenin. *Cancer Research* **68** 5177–5184. (doi:10.1158/0008-5472.CAN-07-6234)

Liu XL, Wazer DE, Watanabe K & Band V 1996 Identification of a novel serine protease-like gene, the expression of which is down-regulated during breast cancer progression. *Cancer Research* **56** 3371–3379.

Malaguarnera R, Vella V, Vigneri R & Frasca F 2007 p53 family proteins in thyroid cancer. *Endocrine-Related Cancer* **14** 43–60. (doi:10.1677/erc.1.01223)

Moreno JC, Bikker H, Kempers MJ, van Trotsenburg AS, Baas F, de Vijlder JJ, Vulsma T & Ris-Stalpers C 2002 Inactivating mutations in the gene for thyroid oxidase 2 (THOX2) and congenital hypothyroidism. *New England Journal of Medicine* **347** 95–102. (doi:10.1056/NEJMoa012752)

Moya CM, Perez de Nanclares G, Castano L, Potau N, Bilbao JR, Carrascosa A, Bargada M, Coya R, Martul P, Vicens-Calvet E *et al.* 2006 Functional study of a novel single deletion in the TITF1/NKX2.1 homeobox gene that produces congenital hypothyroidism and benign chorea but not pulmonary distress. *Journal of Clinical Endocrinology and Metabolism* **91** 1832–1841. (doi:10.1210/jc.2005-1497)

Nikiforov YE, Nikiforova MN, Gnepp DR & Fagin JA 1996 Prevalence of mutations of ras and p53 in benign and malignant thyroid tumors from children exposed to radiation after the Chernobyl nuclear accident. *Oncogene* **13** 687–693.

Port M, Boltze C, Wang Y, Roper B, Meineke V & Abend M 2007 A radiation-induced gene signature distinguishes post-chernobyl from sporadic papillary thyroid cancers. *Radiation Research* **168** 639–649. (doi:10.1667/RR0968.1)

Richardson DB 2009 Exposure to ionizing radiation in adulthood and thyroid cancer incidence. *Epidemiology* **20** 181–187. (doi:10.1097/EDE.0b013e318196ac1c)

Ron E, Lubin JH, Shore RE, Mabuchi K, Modan B, Pottern LM, Schneider AB, Tucker MA & Boice JD Jr 1995 Thyroid cancer after exposure to external radiation: a pooled analysis of seven studies. *Radiation Research* **141** 259–277. (doi:10.2307/3579003)

Rubino C, Cailleux AF, Abbas M, Diallo I, Shamsaldin A, Caillou B, De Vathaire F & Schlumberger M 2002 Characteristics of follicular cell-derived thyroid carcinomas occurring after external radiation exposure: results of a case control study nested in a cohort. *Thyroid* **12** 299–304. (doi:10.1089/10507250252949423)

Sahar S & Sassone-Corsi P 2009 Metabolism and cancer: the circadian clock connection. *Nature Reviews. Cancer* **9** 886–896. (doi:10.1038/nrc2747)

Sigurdson AJ, Ronckers CM, Mertens AC, Stovall M, Smith SA, Liu Y, Berkow RL, Hammond S, Neglia JP, Meadows AT *et al.* 2005 Primary thyroid cancer after a first tumour in childhood (the Childhood Cancer Survivor Study): a nested case–control study. *Lancet* **365** 2014–2023. (doi:10.1016/S0140-6736(05)66695-0)

Smida J, Salassidis K, Hieber L, Zitzelsberger H, Kellerer AM, Demidchik EP, Negele T, Spelsberg F, Lengfelder E, Werner M *et al.* 1999 Distinct frequency of ret rearrangements in papillary thyroid carcinomas of children and adults from Belarus. *International Journal of Cancer* **80** 32–38. (doi:10.1002/(SICI)1097-0215(19990105)80:1<32::AID-IJC7>3.0.CO;2-L)

Soares P, Cameselle-Teijeiro J & Sobrinho-Simoes M 1994 Immunohistochemical detection of p53 in differentiated, poorly differentiated and undifferentiated carcinomas of the thyroid. *Histopathology* **24** 205–210. (doi:10.1111/j.1365-2559.1994.tb00511.x)

Stein L, Rothschild J, Luce J, Cowell JK, Thomas G, Bogdanova TI, Tronko MD & Hawthorn L 2010 Copy number and gene expression alterations in radiation-induced papillary thyroid carcinoma from chernobyl pediatric patients. *Thyroid* **20** 475–487. (doi:10.1089/thy.2009.0008)

Steliarova-Foucher E, Stiller CA, Pukkala E, Lacour B, Plesko I & Parkin DM 2006 Thyroid cancer incidence and survival among European children and adolescents (1978–1997): report from the Automated Childhood Cancer Information System project. *European Journal of Cancer* **42** 2150–2169. (doi:10.1016/j.ejca.2006.06.001)

Williams ED, Abrosimov A, Bogdanova T, Demidchik EP, Ito M, LiVolsi V, Lushnikov E, Rosai J, Tronko MD, Tsyb AF *et al.* 2008 Morphologic characteristics of Chernobyl-related childhood papillary thyroid carcinomas are independent of radiation exposure but vary with iodine intake. *Thyroid* **18** 847–852. (doi:10.1089/thy.2008.0039)

Yamamoto Y, Maeda T, Izumi K & Otsuka H 1990 Occult papillary carcinoma of the thyroid. A study of 408 autopsy cases. *Cancer* **65** 1173–1179. (doi:10.1002/1097-0142(19900301)65:5<1173::AID-CNCR2820650524>3.0.CO;2-2)

Received in final form 4 November 2010
Accepted 9 December 2010
Made available online as an Accepted Preprint
9 December 2010