

## Genetic factors for differentiated thyroid cancer in French Polynesia

Zidane, Monia; Haber, Marc; Truong, Thérèse; Rachédi, Frédérique; Ory, Catherine; Chevillard, Sylvie; Blanché, Hélène; Olaso, Robert; Boland, Anne; Conte, Éric; Karimi, Mojgan; Ren, Yan; Xhaard, Constance; Souchard, Vincent; Gardon, Jacques; Taquet, Marc; Bouville, André; Deleuze, Jean-François; Drozdovitch, Vladimir; de Vathaire, Florent

DOI:  
[10.1093/pcmedi/pbad015](https://doi.org/10.1093/pcmedi/pbad015)

License:  
Creative Commons: Attribution (CC BY)

*Document Version*  
Publisher's PDF, also known as Version of record

*Citation for published version (Harvard):*  
Zidane, M, Haber, M, Truong, T, Rachédi, F, Ory, C, Chevillard, S, Blanché, H, Olaso, R, Boland, A, Conte, É, Karimi, M, Ren, Y, Xhaard, C, Souchard, V, Gardon, J, Taquet, M, Bouville, A, Deleuze, J-F, Drozdovitch, V, de Vathaire, F & Cazier, J-B 2023, 'Genetic factors for differentiated thyroid cancer in French Polynesia: new candidate loci', *Precision Clinical Medicine*, vol. 6, no. 2, pbad015. <https://doi.org/10.1093/pcmedi/pbad015>

[Link to publication on Research at Birmingham portal](#)

### General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

### Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact [UBIRA@lists.bham.ac.uk](mailto:UBIRA@lists.bham.ac.uk) providing details and we will remove access to the work immediately and investigate.

# Genetic factors for differentiated thyroid cancer in French Polynesia: new candidate loci

Monia Zidane<sup>1</sup>, Marc Haber<sup>2,3</sup>, Thérèse Truong<sup>4</sup>, Frédérique Rachédi<sup>5</sup>, Catherine Ory<sup>6</sup>, Sylvie Chevillard<sup>6</sup>, Hélène Blanché<sup>7</sup>, Robert Olasso<sup>8</sup>, Anne Boland<sup>8</sup>, Éric Conte<sup>9</sup>, Mojgan Karimi<sup>4</sup>, Yan Ren<sup>1</sup>, Constance Xhaard<sup>10</sup>, Vincent Soucard<sup>1</sup>, Jacques Gardon<sup>11</sup>, Marc Taquet<sup>12</sup>, André Bouville<sup>13</sup>, Jean-François Deleuze<sup>7,8</sup>, Vladimir Drozdovitch<sup>14</sup>, Florent de Vathaire<sup>1,\*</sup> and Jean-Baptiste Cazier<sup>2,3</sup>

<sup>1</sup>University Paris-Saclay, UVSQ, Inserm, Gustave Roussy, CESP, Team "Radiations Epidemiology", Villejuif 94805, France

<sup>2</sup>Centre for Computational Biology, University of Birmingham, Birmingham B152TT, UK

<sup>3</sup>Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham B152TT, UK

<sup>4</sup>University Paris-Saclay, UVSQ, Inserm, Gustave Roussy, CESP, Team "Exposome and Heredity", Villejuif 94805, France

<sup>5</sup>Endocrinology Unit, Territorial Hospital Taaone, F-98713, Papeete, Tahiti 98713, French Polynesia

<sup>6</sup>CEA, Laboratoire de Cancérologie Fondamentale, Institut de Biologie François Jacob, iRCM, SREIT, Laboratoire de Cancérologie Expérimentale (LCE), Université Paris-Saclay, Fontenay aux Roses 92265, France

<sup>7</sup>Fondation Jean Dausset-Centre d'Etude du Polymorphisme Humain, Paris 75010, France

<sup>8</sup>Université Paris-Saclay, CEA, Centre National de Recherche en Génomique Humaine, Evry 91057, France

<sup>9</sup>U.S.R. 2003 (CNRS / UPF), Faa'a, Tahiti 98702, France

<sup>10</sup>University of Lorraine, INSERM CIC 1433, Nancy CHRU, INSERM U1116, Nancy 54500, France

<sup>11</sup>Hydrosciences Montpellier, Research Institute for Development, CNRS, University of Montpellier, Montpellier 62307, France

<sup>12</sup>Research Institute for Development, Center IRD on Tahiti, Arue, Tahiti 98713, French Polynesia

<sup>13</sup>National Cancer Institute (retired), Bethesda, MD 20892, USA

<sup>14</sup>Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, DHHS, Bethesda, MD 20892, USA

\*Correspondence: Florent de Vathaire, [florent.devathaire@gustaveroussy.fr](mailto:florent.devathaire@gustaveroussy.fr)

## Abstract

**Background:** Populations of French Polynesia (FP), where France performed atmospheric tests between 1966 and 1974, experience a high incidence of differentiated thyroid cancer (DTC). However, up to now, no sufficiently large study of DTC genetic factors in this population has been performed to reach definitive conclusion. This research aimed to analyze the genetic factors of DTC risk among the native FP populations.

**Methods:** We analyzed more than 300 000 single nucleotide polymorphisms (SNPs) genotyped in 283 DTC cases and 418 matched controls born in FP, most being younger than 15 years old at the time of the first nuclear tests. We analyzed the genetic profile of our cohort to identify population subgroups. We then completed a genome-wide analysis study on the whole population.

**Results:** We identified a specific genetic structure in the FP population reflecting admixture from Asian and European populations. We identified three regions associated with increased DTC risk at 6q24.3, 10p12.2, and 17q21.32. The lead SNPs at these loci showed respective p-values of  $1.66 \times 10^{-7}$ ,  $2.39 \times 10^{-7}$ , and  $7.19 \times 10^{-7}$  and corresponding odds ratios of 2.02, 1.89, and 2.37.

**Conclusion:** Our study results suggest a role of the loci 6q24.3, 10p12.2 and 17q21.32 in DTC risk. However, a whole genome sequencing approach would be better suited to characterize these factors than genotyping with microarray chip designed for the Caucasian population. Moreover, the functional impact of these three new loci needs to be further explored and validated.

**Keywords:** differentiated thyroid cancer, population genetics, genetic susceptibility

## Introduction

French Polynesia (FP) is a French Overseas Territory, only populated by the Māori people from Western Polynesia more than 1000 years ago,<sup>1</sup> up to the first contact with the Europeans, between the 16th and 18th centuries, depending on the archipelago. About 93% of the population born in FP still defined themselves as Māori or of mixed Māori origin in the 1988 census, the last census to record ethnic information.<sup>2</sup>

From 1966 to 1974, France performed 44 atmospheric nuclear tests in the atolls of Mururoa and Fangataufa, in the southeastern part of FP, near the Gambier Islands.<sup>3</sup> Although only about 500 inhabitants permanently lived less than 500 km from nuclear sites, radioactive fallouts of several tests were observed in more distant

islands and atolls, including Tahiti, where more than half of population of FP lived.<sup>4</sup>

Since the beginning of the 2000s, we conducted a population-based case-control study on differentiated thyroid cancer (DTC) risk factors in FP. This study was prompted by the high incidence of DTC recorded in the Polynesian islands, including FP,<sup>5-7</sup> and the known risks of radiation-induced thyroid health effects associated with radioactive fallout from nuclear tests.<sup>8</sup> We have shown that lifestyle-related risk factors such as obesity and a high number of pregnancies are associated with DTC risk in European populations and could explain part of the high incidence of DTC in FP, due to their high prevalence in this territory.<sup>9-11</sup> A family history of thyroid cancer or drinking spring water was also shown to be associated with DTC risk,<sup>12,13</sup> whereas a traditional

Received: April 9, 2023. Accepted: June 9, 2023. Published: 13 June 2023

© The Author(s) 2023. Published by Oxford University Press on behalf of the West China School of Medicine & West China Hospital of Sichuan University. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Polynesian diet largely based on local vegetables and sea products was associated with a lower thyroid cancer risk.<sup>10</sup> Lastly, it was also suggested that there was a small, but significant, role of nuclear tests fallouts in the incidence of DTC in FP.<sup>6,14</sup>

In order to investigate the potential role of genetic susceptibility in the high thyroid cancer incidence observed in FP, we previously performed a candidate gene-based analysis to confirm in the French Polynesian population the role of 5 variants in 2q35, 9q22.33, 8p12 and 14q13.3 by Maillard *et al.*<sup>11</sup> that were identified in former GWAS in European populations.<sup>13,15</sup> We reported associations of similar magnitude to what was observed in the European population but showed that the at-risk alleles of such variants were not more frequent in the French Polynesian population than in the European Population.<sup>11</sup> Maillard *et al.* performed an analysis on common variants at 9q22.33, 14q13.3, and ATM loci and DTC in the initial case-control study included in the present analysis (about 50% of the population of the present study).<sup>11</sup> The findings of this previous study highlighted an interaction between rs944289 in NKX-1 (14q13.3) and the radiation dose to the thyroid with a higher risk for the genotype “T/T” carriers. In addition, an association was found between increased DTC risk and rs965513 near FOXE1.<sup>11</sup> More detailed results were reported through a multiethnic GWAS and a fine-mapping analysis that included the Melanesian population from New Caledonia and only a part of the Polynesian population included in the present article.<sup>16,17</sup>

Indeed, we extended the previous case-control study conducted in FP by including thyroid cases diagnosed from 2004 to 2016. It resulted in a confirmation of our previous results, although, slightly lowering the estimate of the risk associated with radiation exposure.<sup>18</sup>

In this report, we performed the first GWAS on DTC for individuals exclusively from FP to identify genetic DTC risk factors specific to this high-risk population.

## Materials and method

### Population

Cases were collected from the Cancer Incidence Registry of FP, private endocrinologists and clinics, as well as public hospitals. All patients diagnosed in 1981–2016 with DTC at the age of 55 years or less, born and resident in FP at the time of diagnosis, were eligible. Controls were healthy individuals, selected from the native population using the registry of births that covers the whole territory. Cases and controls were matched on birth date (around 3 months of maximum difference) and sex. Full details about the first study are described elsewhere.<sup>18,19</sup>

All participants answered a detailed questionnaire about education (available as supplementary data), occupation, weight history, personal and familial history of thyroid diseases and cancers, gynecologic and reproductive history, and diet at the time of the interview and during childhood. Pathological and histological thyroid cancer details were also gathered from clinical records for each case with saliva DNA samples (Oragene® kit). The study has been approved by the “Commission Nationale de l’Informatique et des Libertés” (CNIL agreements 996 007 and 900 124) and the Ethical Committee of French Polynesia. Each participant signed an individual written consent.

### Estimation of thyroid radiation dose

The methodology and results of the dose estimation process have already been described in former published studies. Thyroid radiation doses due to (i) intakes of 131I and of short-lived radio-

dine isotopes (132I, 133I, 135I) and 132Te via inhalation and ingestion of foodstuffs and drinking water, (ii) external irradiation from gamma-emitting radionuclides deposited on the ground, and (iii) ingestion of long-lived 137Cs with foodstuffs were reconstructed for each study subject, mainly by using:

- Results of radiation monitoring of the environment and foodstuffs performed by the SMSR (Service Mixte de Sécurité Radiologique) and the SMCB (Service Mixte de Contrôle Biologique), the internal radiation protection services of French Army and CEA after each test, which were declassified in 2013, instead of syntheses sent by France to UNSCEAR, as in the initial study. These reports made it possible to conduct a comprehensive estimation of the ground deposition of 33 radionuclides at the time of arrival of fallout after each test, based on measurements of total ground deposition, measurements of total beta-concentration in air, or measurements of exposure rate at different locations in FP.<sup>20</sup>
- Integrating historical data on population lifestyle related to the period of the tests, which were collected in 2016–2017 using focus-group discussions and key informant interviews monitored in each of the 5 archipelagos of FP, in addition to declarations reported in self-questionnaires.<sup>14</sup>

This dosimetric reconstruction led to an estimation of the mean lifetime thyroid dose among the study subjects of around 4.7 mGy.<sup>21</sup> Doses from 131I intake ranged up to 27 mGy, while those from intake of short-lived iodine isotopes (132I, 133I, 135I) and 132Te ranged up to 14 mGy. Thyroid doses from external exposure ranged up to 6 mGy, and those from internal exposure due to 137Cs ingestion did not exceed 1 mGy. Intake of 131I was the main pathway of thyroid exposure accounting for 72% of the total dose.<sup>22</sup>

### Genotyping data

DNA was extracted from saliva samples at the CEPH-Biobank (Centre d’Etude du Polymorphisme Humain, Paris, France) using an automated salt precipitation method on an Autopure (Qiagen). Genotyping was performed with a custom-made Infinium OncoArray-500K BeadChip EPITHYR (Illumina®) at the Centre National de Recherche en Génomique Humaine (CNRGH/CEA, Evry, France), on an Illumina high-throughput genotyping platform automated according to manufacturer’s instructions. The OncoArray-500K is described in detail in Amos *et al.* 2017<sup>23</sup> and contains 499 170 SNPs. An additional 13 759 custom markers were incorporated, based on prior evidence of association with relevant biological pathways for thyroid cancer.<sup>16</sup>

In total, 289 cases and 418 controls with a sufficient amount of DNA, from the first study and the extension, were genotyped. A threshold of 5% was applied for missing call rate per SNP and 10% per individual. Furthermore, three individuals were removed because they were duplicates. Hardy-Weinberg equilibrium (HWE) was also assessed per SNP among controls, using a Fisher exact test with a p-value threshold of 10<sup>-5</sup>. Only SNPs with minor allele frequency (MAF) above 5% were considered in the analyses to ensure sufficient power. The quality control of the genotyping data was performed with Plink2 program (version Alpha 2.3).<sup>24</sup>

As our genotyping array was originally designed based on the human assembly hg19, the SNPs positions were mapped to the reference genome hg38 using the LiftOver tool from UCSC with default parameters<sup>25</sup> (available here: <https://genome.ucsc.edu/cgi-bin/hgLiftOver>). The data was phased with Eagle v2.4.1<sup>26</sup> according to the 1000 Genomes Project phase 3 and the hg38 assembly

as references, with a maximum number of conditioning haplotypes fixed at 10 000. Then imputation was performed with IMPUTE2<sup>27,28</sup> for hits regions on chromosomes 6, 10, and 17 using the 1000 Genomes Project phase 3 (accessed December 2019)<sup>29</sup> and the Human Genome Diversity Project (accessed December 2019)<sup>30</sup> datasets as reference panels. We imputed the region 500kb on both sides of the most significant SNP position for each of the three loci of interest. To avoid strand issues, we excluded palindromic SNPs. After imputation, SNPs with an info score < 0.85 were excluded.

## Statistical methods

### Global Admixture analyses

To study genetic structure in our population we merged our data with published data from worldwide populations previously compiled and available from [https://reichdata.hms.harvard.edu/amh\\_repo/\(v29.1\)](https://reichdata.hms.harvard.edu/amh_repo/(v29.1)). We filtered populations in the dataset keeping: Africans (Yoruba), East-Asians (Chinese, Vietnamese, and Japanese), Central and South Asians (Balochi, Brahui, and Sindhi), North-Eastern Asians (Yakut and Yukagir), other Oceanian populations (Samoan and Papuan), Metropolitan French and ancient South Pacific individuals including three ancient French Polynesians who lived during the eighteenth-century AD.<sup>31</sup> Finally, data from 408 persons from the reference dataset and 56 316 shared SNPs were analyzed.

We first performed a principal component analysis (PCA) using the “smartpca” function from Eigensoft v7.2<sup>32</sup> on the study population alone and found a strong genetic structure (Supplementary Fig. 1). The first 5 principal components explained 90% of the variance in our dataset. We then ran a PCA including worldwide diversity populations and projected the French Polynesians into the defined space. Based on the results from the PCA, we focused our analysis on three ancestries found in French Polynesia today: local Polynesian ancestry represented by ancient French Polynesians, East Asian ancestry represented by modern Chinese Han, and European ancestry represented by modern French. We performed an analysis using the software ADMIXTURE<sup>33</sup> supervised by using only three reference populations, estimating the fraction of these different ancestries in our dataset.

### Association analyses

Logistic mixed effects models implemented in the GENESIS R Bioconductor package were used to test for association between genotype and DTC risk, using a penalized quasi-likelihood approximation to the generalized linear mixed model. To consider the population structure, GENESIS estimates a kinship matrix that accounts for cryptic and known relatedness between subjects.<sup>34</sup> This kinship matrix was derived from a set of LD-pruned genotyped SNPs (at threshold  $R^2 = 0.1$ ) and was included as a random effect in the logistic mixed model. The three first principal components corresponding to our three founding populations were included as fixed-effect covariates. The other covariables included in the model are age (continuous), sex (binary), thyroid cancer family history (binary), goiter history (binary), radiotherapy history (binary), Mururoa site workers (binary), BMI (continuous), height (continuous), educational level (primary school, secondary school), smoking status (binary), among women the number of pregnancies longer than 7 months (0, 1–3, >3), dietary selenium and iodine intake during adulthood and childhood (continuous) and the internal radiation dose delivered to the thyroid gland before the age of 15 (continuous). The same model was also tested after the exclusion of the cases with non-invasive microcarcinoma

(< 10 mm, n = 173), and with an interaction term between the estimated dose delivered to the thyroid gland before the age of 15 and the genotype.

Besides, in an attempt to replicate former results from Maillard et al.,<sup>11</sup> we used an unconditional logistic model adjusted on the same variables as above. Benjamin Hochberg FDR procedure was used for multiple testing correction<sup>35</sup> for the GWAS analysis, to better take into account the non-independence of the SNPs.

### Local ancestry Inference using RFMix

Local ancestry inference, that infers the regional ancestral origin of a chromosomal segment in admixed populations, was performed using RFMix version 2.03,<sup>36</sup> with default parameters. Then, the RFMix local ancestry calls were converted to Plink TPED files, where the encoded dosage is defined as having 0, 1, or 2 copies for each of the ancestry populations at a particular local ancestry segment.

### Admixture mapping

The main aim of this analysis is to test for association between the local ancestry at a genomic region and DTC. These tests were performed using the linear mixed-effect models implemented in the EMMAX software package. These models included local ancestry count as the main effect and a Balding-Nichols kinship matrix as a random effect. This step was run for each of the three identified ancestries: European, South Asian, and East Asian. We adapted the correction for multiple testing from Gao et al.,<sup>37</sup> to be used in the estimation of the number of effective tests in a Bonferroni correction for multiple testing.<sup>38</sup> This number of effective tests is the sum of PC explaining more than 99.5% of genetic variance per chromosome. The significance threshold for our admixture mapping analysis is  $1.13 \times 10^{-4}$  based on 442 effective tests. This approach has already been adopted in a similar study.<sup>39</sup>

### Single-SNP heritability and posthoc power calculation

Based on the summary results from the association analysis we estimated the heritability and the post-hoc power for each tested SNP. The heritability was calculated as the proportion of phenotypic variance explained by a given SNP. The following formula was used to assess single-SNP heritability  $h^2$  from genome-wide association study results, where  $i$  is a given polymorphism:<sup>40</sup>

$$h_i^2 = \frac{2(\beta_i)MAF_i(1 - MAF_i)}{2(\beta_i)MAF_i(1 - MAF_i) + 2N(\beta_i.Se)^2MAF_i(1 - MAF_i)}$$

where  $MAF_i$  is the minor allele frequency for the SNP $_i$ , and allelic causal effect size is  $\beta_i$ ,  $\beta_i.Se$  is the standard error of  $\beta_i$  and  $N$  is the population size. Then, based on  $h^2$  values, we estimated the observed power (post-hoc power) for each tested SNP $_i$ :

$$\text{Observed power}_i = 1 - G(t, \lambda, 1)$$

where  $G$  is the cumulative distribution function of non-central Chi-square distribution, and the non-centrality parameter ( $\lambda$ ) depends on sample size ( $N$ ) and the proportion of phenotypic variance that is explained by the SNP $_i$  and calculated above ( $h^2_i$ ):

$$\lambda = \left( \frac{h^2_i}{1 - h^2_i} \right) N$$

Two different values of significance threshold  $\alpha$  were considered:  $\alpha_1 = 0.05/\text{total number of tested SNPs} = 1.66 \times 10^{-7}$  and  $\alpha_2 = 0.05/\text{number of independent SNPs tested (D' < 0.7 in windows of 1000 SNPs)} = 1.11 \times 10^{-5}$ .

**Table 1.** Demographic and epidemiologic description.

	Controls (N = 418)	DTC Cases (N = 283)	P-value*
<b>Sex(%)</b>			
Men	58 (13.7%)	51 (18.0%)	0.1
Women	360 (86.3%)	232 (82.0%)	
<b>Smoker (%)</b>			
No	213 (50.8%)	125 (44.2%)	0.1
Yes	205 (49.2%)	158 (55.8%)	
<b>Height (meters)</b>			
Mean (SD)	1.66 (0.0762)	1.66 (0.0771)	0.6
Median [Min, Max]	1.66 [1.44, 2.01]	1.66 [1.40, 1.90]	
<b>Personal history of radiotherapy (%)</b>			
No	407 (97.4%)	265 (93.6%)	0.02
Yes	11 (2.6%)	18 (6.4%)	
<b>Diploma</b>			
No	138 (32.6%)	104 (36.7%)	0.3
Yes	281 (67.4%)	179 (63.3%)	
<b>Age at diagnosis in years (%)</b>			
1–19	19 (4.3%)	3 (1.1%)	<0.001
20–29	65 (15.6%)	20 (7.1%)	
30–39	103 (24.7%)	76 (26.9%)	
40–49	145 (34.8%)	37 (13.1%)	
50–59	86 (20.6%)	147 (51.9%)	
<b>BMI (kg/m<sup>2</sup>)</b>			
Mean (SD)	28.3 (6.90)	31.4 (7.38)	<0.001
Median [Min, Max]	27.1 [16.0, 56.5]	30.7 [17.1, 66.1]	
<b>Dietary iodine intake during childhood (µg/day)</b>			
Mean (SD)	125 (86.7)	150 (130)	0.005
Median [Min, Max]	107 [0, 841]	116 [0, 934]	
<b>Dietary selenium intake during childhood (µg/day)</b>			
Mean (SD)	109 (74.8)	138 (151)	0.003
Median [Min, Max]	92.8 [0, 432]	99.1 [0, 1080]	
<b>Radiation dose at thyroid gland before the age of 15 (mGy)</b>			
Mean (SD)	1.48 (2.54)	1.39 (2.69)	0.7
Median [Min, Max]	0 [0, 13.0]	0 [0, 22.0]	

\*Univariate logistic regression

## Results

### Demographic and epidemiologic characteristics

At the end of quality control of the genotyping data, the present study included 283 cases and 418 controls and 300 908 SNPs were available for the analysis (study flowchart in Supplementary Fig. 1). Main characteristics of all cases and controls are described in Table 1: more than 80% of cases were women. The mean age of the thyroid cancer diagnosis was 44 years, and 56% of the cases were smokers. Anthropometric factors (BMI and body height) and dietary intake of selenium and iodine during childhood were found to significantly differ between cases and controls. The distributions of the thyroid doses for the cases and for the controls are presented in the Supplementary Fig. 2.

### Population structure

PCA of the French Polynesian population revealed a significant genetic structure in the population (Supplementary Fig. 3). When the French Polynesian population was projected on a PCA defined by worldwide populations we identified two ancestry clines showing a group of French Polynesians that were drawn towards

East Asians compared to the position of ancient Polynesians while a second group was drawn towards Europeans (Fig. 1A & B). Our admixture test confirmed East Asian-related and European-related admixture in French Polynesia (Fig. 2, Supplementary Fig. 4). PCA results including African and other Asian ethnicities are not shown.

### Association and admixture mapping results

The GWAS allowed us to identify 3 loci as highlighted in the Manhattan plot presented in Fig. 3A. The genomic inflation factor  $\lambda$  was = 1.03, confirming a lack of systematic genetic bias between our cases and controls (Fig. 3B). The lowest p-value was observed for rs1358902 ( $1.66 \times 10^{-7}$ ; FDR = 0.039;  $OR_{\text{allelic}} = 2.01$ ) located at 6q24.3, in an intergenic region (tested allele frequency = 0.38, genotypic count detailed in Supplementary Table 1). At 10p12.2, rs28390243 had a p-value =  $2.39 \times 10^{-7}$  (FDR = 0.026), with an  $OR_{\text{allelic}} = 1.89$  (tested allele frequency = 0.38, genotypic count detailed in Supplementary Table 1). All the SNPs associated to DTC at 10p12.2 were near or within the *PIP4K2A* gene region.

A third susceptibility loci was observed at 17q21.32, which contain a large linkage disequilibrium region, mostly in the *SKAP1* gene. The minor allele of the lead SNP rs7211607 (tested allele frequency = 0.16, genotypic count detailed in Supplementary Table 1) in the *SKAP1* region had an  $OR_{\text{allelic}} = 2.37$ , (p-value of  $7.19 \times 10^{-7}$  and FDR of 0.026). Same associated regions were found to have the lowest p-values when excluding cases with non-invasive microcarcinoma from the analyses (Results not shown). The lead best SNP in this analysis was rs28390243, but not significant ( $OR_{\text{allelic}} = 2.04$ , FDR = 0.49).

Results from local imputation in each of the three regions on chromosomes 6, 10, and 17 found that the original lead SNPs in the three regions remained the most significant.

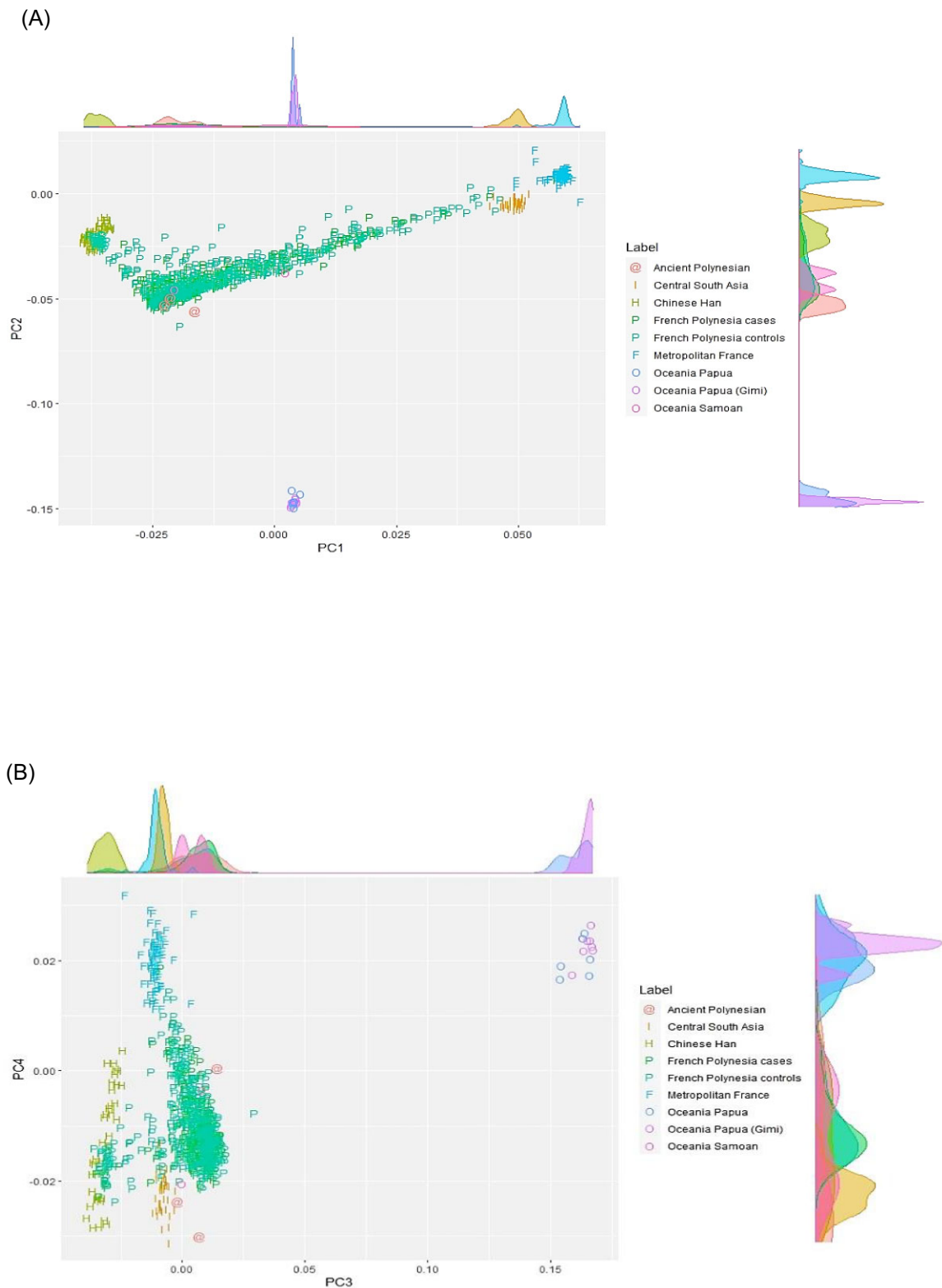
The statistical power to detect an association with rs1358902 (6q24.3) while considering the significance threshold at  $\alpha_1$  was 52% and 81% at  $\alpha_2$ . For rs28390243 (10p12.2), the power was 50% at  $\alpha_1$  and 80% at  $\alpha_2$ . This power was equal to 39% and 72% respectively at  $\alpha_1$  and  $\alpha_2$  for rs7211607 (17q21.32).

The estimates of the radiation doses before the age of 15 years to the thyroid gland were low, with a maximal dose of 22 mGy among cases, and 13 mGy among controls.

In the interaction analysis, none of the SNPs reached the significance threshold. The most significant was observed in many SNPs (in LD:  $R^2$  between 0.85 and 1,  $D' = 1$ ) located in 21q21.1, all in the *MIR99AHG* region. These results did not change when we excluded individuals with a radiotherapy history and when non-invasive microcarcinoma cases were excluded.

We compared our findings with results from our previous study that were based on a subpopulation the current study (137 cases and 218 controls) (Maillard et al.<sup>11</sup>). We failed to find the same significant interaction in rs944289. However, the main effect of the genotype "T/T" had the same direction as in Maillard et al. findings<sup>11</sup> ( $OR_{T/T \text{ genotype}} = 1.45$ , 95% Confidence Interval (CI) = 0.26–7.98) (Allelic ORs are listed in Supplementary Table 2).

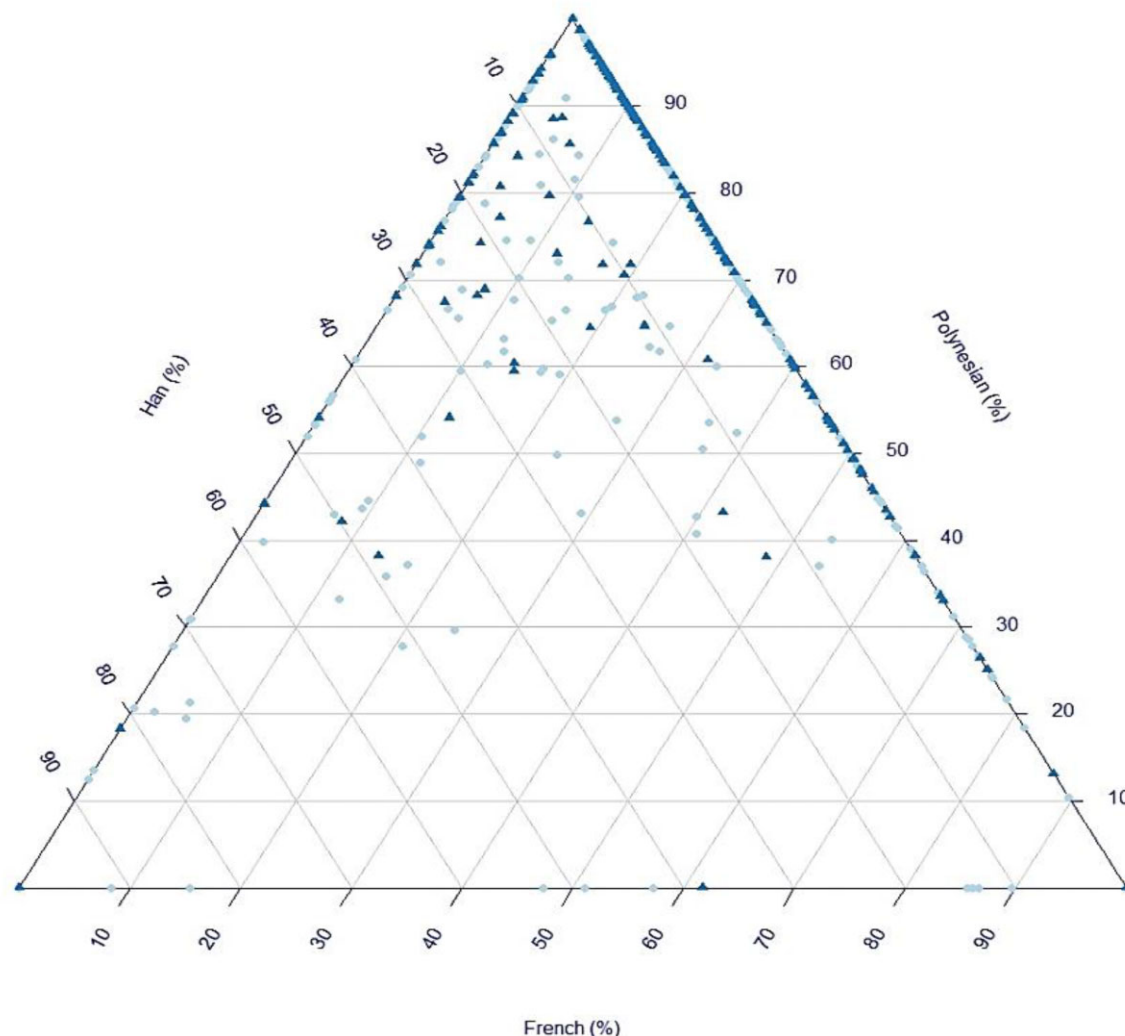
According to our significance threshold for the admixture mapping analysis, only two loci in chromosomes 14 and 15 were found to be significantly associated with DTC risk in a population of East-Asian ancestry. The locus 15q14 (34 830 119–34 903 911) was located within the *AQR* gene region, with a p-value =  $2.95 \times 10^{-5}$ . The second locus at 14q21.2 (43 863 839–43 935 211), with a p-value =  $4.20 \times 10^{-5}$ , was located in an intergenic region. Results from the analysis of haplotypes coded according to European



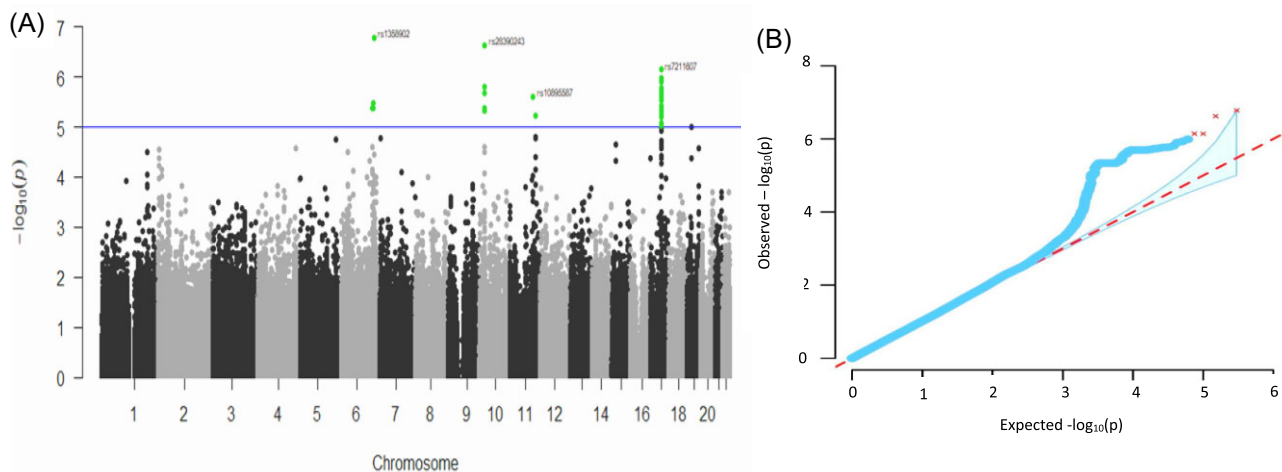
**Figure 1.** First four principal components with the corresponding proportion of explained variance in a worldwide PCA: (A) PCs 1 & 2; (B) PCs 3 & 4. We show a subset of the PCA with French Polynesians in a regional context.

ancestry showed an association with the *PIP4K2A* gene region in 10p12.2 (22 556 003–22 730 905), with a  $p$ -value =  $1.11 \times 10^{-6}$ , the very same result from the global GWAS. When this analysis was run with a haplotype coding according to the South-Asian ancestry, only one locus in the *RYBP* gene region (chromosome

3:72 374 593–72 446 623) was associated with DTC risk with a  $p$ -value =  $8.37 \times 10^{-5}$ . The other top associated loci were similar to those found in analysis with a haplotype coding according to European Ancestry. Results from this analysis are shown in Supplementary Table 3.



**Figure 2.** ADMIXTURE analysis showing ancestry percentage of individuals in our dataset. The three corners correspond to French, Han and ancient Polynesians used as reference groups. Cases are presented by dark blue triangles and controls by light blue dots.



**Figure 3.** GWAS results. (A) Manhattan plot of the 300 908 SNPs (represented underneath) for DTC risk. Color represents the various chromosomes; the blue dotted line indicates  $p$ -value  $< 10^{-5}$ . (B) The blue line represents the observed  $-\log_{10}(p)$ -values; the red dotted line represents the expected  $-\log_{10}(p)$ -values; the blue area represents the 95% confidence interval. The genomic inflation factor  $\lambda = 1.03$ .

## Discussion

The main purpose of this study was to explore the role of genetic factors in DTC occurrence in the FP population. Compared to the GLOBOCAN statistics, the sex ratio in thyroid cancer cases worldwide is 3.87, which is similar to our study.<sup>41</sup> Smoking habits are more frequent in our study population compared to the general population in FP (36.5% of men and 41.4% of women smokers).<sup>42</sup>

Our results from the GWAS suggest potential DTC susceptibility loci at 6q24.3, 10p12.2, and 17q21.32. While all these regions have been previously linked to cancer, none of them was previously associated with DTC risk.

At locus 6q24.3, all variants were in non-coding regions, the closest genes were *GRM1*, *RAB32*, and *ADGB*. These genes have been associated with the risk and prognosis of ovarian cancer,<sup>43</sup> gastric cancer<sup>44</sup> and breast cancer.<sup>45</sup> In addition, this region is also associated with autoimmune thyroiditis<sup>46</sup> and in the regulation of proliferation and invasion of thyroid cancer cells<sup>47</sup> and with familial lung cancer risk.<sup>48</sup>

The lead SNP (rs1358902) in chromosome 10 is located near a promoter for the *PIP4K2A* gene in the 10p12.2 locus. This locus has been associated with childhood B-cell acute lymphoblastic leukemia occurrence by a previous GWAS,<sup>49</sup> in the European ancestry population, and in multiple other populations including Hispanic, African, and Chinese.<sup>50,51</sup> Variants in this gene were also found to be associated with acute lymphoblastic leukemia prognosis.<sup>52,53</sup> The results from a cross-ancestry meta-analysis identified an association between a variant in the *PIP4K2A* gene region and breast cancer susceptibility in European and Asian populations.<sup>54</sup> Although the role of *PIP4K2A* is not well-defined and mainly depends on the biological context, this gene has recently been suggested to be involved in oxidative stress, cellular senescence, and tumor growth.<sup>55,56</sup> It has also been suggested to play an oncogenic role and was found to be frequently overexpressed in multiple types of leukemia and some types of solid cancers in cancer cohorts including TCGA.<sup>55,56</sup>

Another signal was identified on chromosome 17 in the region of *SKAP1* (SRC kinase-associated phosphoprotein 1 gene (Fig. 3A)) which had been previously reported to be associated with ovarian cancer in the European and Chinese Han populations.<sup>57,58</sup> The expression levels of these genes were found to be associated with ovarian, endometrial, and colorectal cancer risk.<sup>59–61</sup> The 17q21.32 locus is additionally associated with endometrial, prostatic, and renal cancers.<sup>62–65</sup> The biological function of *SKAP1* related to these cancer types is not yet well elucidated.

The observed power under the  $\alpha_1$  assumption varied from 39% to 51% for the three leading SNPs of the region. This power is between 78% and 81% if we consider  $\alpha_2$ . In view of the small size of the general FP population (almost 300 000), the present study included an important proportion of the total DTC cases occurred in FP.

In radiation-related DTC studies, only two variants in *FOXE1* and *ATM* were associated with the disease in two different populations.<sup>66</sup> In *FOXE1*, only the association of rs965513 with radiation-related DTC could be replicated in a population exposed to higher levels of radiation caused by Chernobyl disaster.<sup>67</sup> Notably, the same variant rs1801516 in *ATM* could have an opposite effect between the FP population<sup>11</sup> and the Belarusian population.<sup>68</sup> This could highlight the importance of taking into account ancestry when considering the genetic susceptibility to DTC, and the importance of further exploration of the concerned genetic regions, as the causal variant could be a different one in the same region.

We did not observe a significant interaction between thyroid radiation dose due to atmospheric nuclear tests and genetic factors, but this result was expected given the very low level of exposure resulting in a limited statistical power for testing this interaction.

The interaction found by a former study with different doses estimation and genotyping technology<sup>11</sup> was not replicated in the current study. This difference could be caused by the different thyroid IR doses between both studies, the partially shared study population, and the difference between genotyping technologies, as about 10% of the genotyping results are different between the first and the second studies.

A more recent study focused on the genomic profile of radiation-related papillary thyroid cancer after the Chernobyl accident, which included normal and tumor tissues, with whole-genome, mRNA, and microRNA sequencing; DNA methylation profiling; and genotyping arrays. The germline DNA analyses yielded three possible associations with the radiation dose of variants in *SMAD3*, *TERT*, and 9q22.33.<sup>69</sup> However, none of these genes/loci were found in our study. Thus, our results from genetic susceptibility analysis identified novel genes related to DTC with many of them linked to hormonal-dependent cancers. Results from admixture mapping highlighted two regions: 15q14 (34 830 119–34 903 911) and 14q21.2 (43 863 839–43 935 211) to be associated with DTC risk in a population with East Asian ancestry. These results should be taken with precaution because of the limited sample size of the study and the high admixture of this population. Altogether, results from this analysis show that the outline of the genetic regions associated with DTC risk is relatively different depending on the ethnic origins.

The significant associations identified in our study need to be replicated and functionally validated to understand the biological mechanisms involved. An increase in the sample size in the future will permit the exploration of further orders of interaction such as the interaction between gene sets.

In addition to genetic susceptibility results, our study identified significant admixture in FP related to Asian and European populations. However, the dominating ancestry in FP is related to ancient local French Polynesians. Our study suggests that our results were not biased by population structure or admixture, especially because of the use of a mixed model. Future exploration of ancestry in FP could enable the study of the potential correlation between geographic regions (archipelago and atolls), ancestry, and DTC. A larger study population could help to better understand these potential interconnections, especially comparing to classic GWAS study population sizes, though increasing the sample size may be difficult due to the small size of the French Polynesian population and therefore the relatively small number of DTC cases in this population. Besides, the use of non-ethnically adapted microarrays strongly reduced the number of analyzable SNPs and genomic regions, as these microarrays are designed for European population and doesn't cover the same variants as in FP population. Exploring genetic factors of DTC in non-European populations will allow to propose better polygenic risk scores by identifying other variants.

## Conclusion

The French Polynesian population has genetic, geographic, and cultural specificities. Here, we attempted to study genetic DTC risk factors while accounting for genetic structure in an admixed population. Our results are suggestive of a role of three loci: 6q24.3, 10p12.2, and 17q21.32. The genes present at these loci were not previously associated with DTC, but rather with other



hormone-dependent cancers or psychiatric disorders. Given that the microarray used for genotyping in the present study was designed for the population of European ancestry, it would be more adapted to perform whole genome sequencing in the Polynesian population.

## Supplementary data

Supplementary materials are available at [PCMED](#) online.

## Acknowledgements

This work was supported by Institut National du Cancer (Grant No. 9533) and Fondation ARC (Grant No. PGA120150202302). We would like to thank all volunteers who participated in this study. We also thank the CEPH-Biobank team for their involvement in the preparation of DNA samples and the CNRGH teams for their involvement in DNA genotyping.

## Conflict of interest statement

The authors have no conflicts of interest to declare. All co-authors have seen and agree with the contents of the manuscript and there is no financial interest to report.

None declared. Besides, as the Editorial Board Member of *Precision Clinical Medicine*, the author Jean-Baptiste Cazier was blinded from reviewing and making decision on this manuscript.

## References

1. Wilmschurst JM, Hunt TL, Lipo CP, et al. High-precision radiocarbon dating shows recent and rapid initial human colonization of East Polynesia. *Proc Natl Acad Sci U S A* 2011;**108**(5):1815–20. doi: 10.1073/pnas.1015876108
2. Villedieu-Liou M. *Résultats du recensement général de la population de la Polynésie Française du 3 septembre 1996*. Paris, France: EP-SILON. INSEE, 1998.
3. Ribbe J, Tomczak M. An impact assessment for the French nuclear weapon test sites in French Polynesia. *Mar Pollut Bull* 1990;**21**(11):536–42. doi: 10.1016/0025-326X(90)90302-O.
4. Drozdovitch V, Bouville A, Doyon F, et al. Reconstruction of individual radiation doses for a case-control study of thyroid cancer in French Polynesia. *Health Phys* 2008;**94**(5):418–33. doi: 10.1097/01.HP.0000299293.06218.88
5. De Vathaire F, Le Vu B. Cancer mortality in french polynesia between 1984 and 1992. *Br J Cancer* 1996;**74**(10):1680–1. doi: 10.1038/bjc.1996.609.
6. De Vathaire F, Le Vu B, De Vathaire CC. Thyroid cancer in French Polynesia between 1985 and 1995: Influence of atmospheric nuclear bomb tests performed at Mururoa and Fangataufa between 1966 and 1974. *Cancer Causes Control* 2000;**11**(1):59–63. doi: 10.1023/a:1008961503506.
7. Le Vu B, De Vathaire F, De Vathaire CC, et al. Cancer incidence in French Polynesia 1985–95. *Trop Med Int Health* 2000;**5**(10):722–31. doi: 10.1046/j.1365-3156.2000.00624.x
8. Simon SL, Bouville A. Radiation doses to local populations near nuclear weapons test sites worldwide. *Health Phys* 2002;**82**(5):706–25. doi: 10.1097/00004032-200205000-00016.
9. Brindel P, Doyon F, Rachédi F, et al. Menstrual and reproductive factors in the risk of differentiated thyroid carcinoma in native women in French Polynesia: A population-based case-control study. *Am J Epidemiol* 2008;**167**(2):219–29. doi: 10.1093/aje/kwm288
10. Cléro I, Doyon F, Chungue V, et al. Dietary iodine and thyroid cancer risk in French Polynesia: A case-control study. *Thyroid* 2012;**22**(4):422–9. doi: 10.1089/thy.2011.0173
11. Maillard S, Damiola F, Clero E, et al. Common variants at 9q22.33, 14q13.3, and ATM loci, and risk of differentiated thyroid cancer in the French Polynesian population. *PLoS One* 2015;**10**(4):e0123700. doi: 10.1371/journal.pone.0123700
12. Brindel P, Doyon F, Bourgain C, et al. Family history of thyroid cancer and the risk of differentiated thyroid cancer in French Polynesia. *Thyroid* 2010;**20**(4):393–400. doi: 10.1089/thy.2009.0350
13. Hińcza K, Kowalik A, Kowalska A. Current knowledge of germline genetic risk factors for the development of non-medullary thyroid cancer. *Genes* 2019;**10**(7):482. doi: 10.3390/genes10070482.
14. Drozdovitch V, Bouville A, Tetuanui T, et al. Behavior and food consumption pattern of the french polynesian population in the 1960s -1970s. *Asian Pac J Cancer Prev* 2019;**20**(12):3667–77. doi: 10.31557/APJCP.2019.20.12.3667
15. Gudmundsson J, Thorleifsson G, Sigurdsson JK, et al. A genome-wide association study yields five novel thyroid cancer risk loci. *Nat Commun* 2017;**8**:1, 14517. doi: 10.1038/ncomms14517
16. Truong T, Lesueur F, Sugier P, et al. Multiethnic genome-wide association study of differentiated thyroid cancer in the EPITHYR consortium. *Int J Cancer* 2021;**148**(12):2935–46. doi: 10.1002/ijc.33488
17. Guibon J, Sugier PE, Kulkarni O, et al. Fine-mapping of two differentiated thyroid carcinoma susceptibility loci at 2q35 and 8p12 in Europeans, Melanesians and Polynesians. *Oncotarget* 2021;**12**(5):493–506. doi: 10.18632/oncotarget.27888
18. De Vathaire F, Zidane M, Xhaard C, et al. Assessment of differentiated thyroid carcinomas in French Polynesia after atmospheric nuclear tests performed by France. *JAMA Netw Open* 2023;**6**(5):e2311908. doi: 10.1001/jamanetworkopen.2023.11908
19. De Vathaire F, Drozdovitch V, Brindel P, et al. Thyroid cancer following nuclear tests in French Polynesia. *Br J Cancer* 2010;**103**(7):1115–21.
20. Drozdovitch V, de Vathaire F, Bouville A. Ground deposition of radionuclides in French Polynesia resulting from atmospheric nuclear weapons tests at Mururoa and Fangataufa atolls. *J Environ Radioact* 2020;**214-215**:106176. doi: 10.1016/j.jenvrad.2020.106176.
21. Drozdovitch V, de Vathaire F, Bouville A. Radiological Impact of atmospheric nuclear weapons tests at Mururoa and Fangataufa Atolls to populations in Oceania, South America and Africa: Comparison with French Polynesia. *Asian Pac J Cancer Prev* 2021;**22**(3):801–9. doi: 10.31557/APJCP.2021.22.03.801
22. Drozdovitch V, Bouville A, Taquet M, et al. Thyroid doses to french polynesians resulting from atmospheric nuclear weapons tests: estimates based on radiation measurements and population lifestyle data. *Health Phys* 2021;**120**(1):34–55. doi: 10.1097/HP.0000000000001262
23. Amos CI, Dennis J, Wang Z, et al. The oncoarray consortium: A network for understanding the genetic architecture of common cancers. *Cancer Epidemiol Biomarkers Prev* 2017;**26**(1):126–35. doi: 10.1158/1055-9965.EPI-16-0106
24. Chang CC, Chow CC, Tellier LCAM, et al. Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaSci* 2015;**4**:7. doi: 10.1186/s13742-015-0047-8.
25. Kent WJ, Sugnet CW, Furey TS, et al. The human genome browser at UCSC. *Genome Res* 2002;**12**(6):996–1006. doi: 10.1101/gr.229102

26. Loh PR, Palamara PF, Price AL. Fast and accurate long-range phasing in a UK Biobank cohort. *Nat Genet* 2016; **48**(7):811–6. doi: 10.1038/ng.3571.
27. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* 2009; **5**(6):e1000529. doi: 10.1371/journal.pgen.1000529.
28. Howie B, Fuchsberger C, Stephens M, et al. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet* 2012; **44**(8):955–9. doi: 10.1038/ng.2354
29. Auton A, Abecasis GR, Altshuler DM, et al. A global reference for human genetic variation. *Nature* 2015; **526**(7571):68–74. doi: 10.1038/nature15393
30. Bergström A, McCarthy SA, Hui R, et al. Insights into human genetic variation and population history from 929 diverse genomes. *Science* 2020; **367**(6484):eaay5012. doi: 10.1126/science.aay5012
31. Posth C, Nägele K, Collieran H, et al. Language continuity despite population replacement in Remote Oceania. *Nat Ecol Evol* 2018; **2**(4):731–40. doi: 10.1038/s41559-018-0498-2
32. Price AL, Patterson NJ, Plenge RM, et al. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006; **38**(8):904–9. doi: 10.1038/ng1847
33. Alexander DH, Lange K. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinf* 2011; **12**:1, 246. doi: 10.1186/1471-2105-12-246.
34. Conomos MP, Miller MB, Thornton TA. Robust inference of population structure for ancestry prediction and correction of stratification in the presence of relatedness. *Genet Epidemiol* 2015; **39**(4):276–93. doi: 10.1002/gepi.21896.
35. Benjamini Y, Hochberg Y. On the adaptive control of the false discovery rate in multiple testing with independent statistics. *J Educ Behav Stat* 2000; **25**(1):60–83. doi: 10.3102/10769986025001060
36. Maples BK, Gravel S, Kenny EE, et al. RFMix: A discriminative modeling approach for rapid and robust local-ancestry inference. *Am J Hum Genet* 2013; **93**(2):278–88. doi: 10.1016/j.ajhg.2013.06.020
37. Gao X, Starmer J, Martin ER. A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genet Epidemiol* 2008; **32**(4):361–9. doi: 10.1002/gepi.20310.
38. Bland JM, Altman DG. Multiple significance tests: The Bonferroni method. *BMJ* 1995; **310**(6973):170. doi: 10.1136/bmj.310.6973.170.
39. Daya M, Rafaels N, Brunetti TM, et al. Association study in African-admixed populations across the Americas recapitulates asthma risk loci in non-African populations. *Nat Commun* 2019; **10**(1):880. doi: 10.1038/s41467-019-08469-7
40. Zaitlen N, Kraft P. Heritability in the genome-wide association era. *Hum Genet* 2012; **131**(10): 1655–64. doi: 10.1007/s00439-012-1199-6.
41. International Agency for Research on Cancer. Cancer over time. Available from: [https://gco.iarc.fr/overtime/en/dataviz/bars?sexes=1\\_2&sort\\_by=value2&mode=cancer](https://gco.iarc.fr/overtime/en/dataviz/bars?sexes=1_2&sort_by=value2&mode=cancer) [cited 2023 Jan 3].
42. Kessaram T, McKenzie J, Girin N, et al. Tobacco smoking in islands of the pacific region, 2001–2013. *Prev Chronic Dis* 2015; **12**:E212. doi: 10.5888/pcd12.150155.
43. Kamieniak MM, Rico D, Milne RL, et al. Deletion at 6q24.2–26 predicts longer survival of high-grade serous epithelial ovarian cancer patients. *Mol Oncol* 2015; **9**(2):422–36. doi: 10.1016/j.molonc.2014.09.010
44. Zhou N, Liu C, Wang X, et al. Downregulated SASH1 expression indicates poor clinical prognosis in gastric cancer. *Hum Pathol* 2018; **74**:83–91. doi: 10.1016/j.humpath.2018.01.008
45. Zeller C, Hinzmann B, Seitz S, et al. SASH1: A candidate tumor suppressor gene on chromosome 6q24.3 is downregulated in breast cancer. *Oncogene* 2003; **22**(19):2972–83. doi: 10.1038/sj.onc.1206474
46. Yoshida N, Nishikori M, Izumi T, et al. Primary peripheral T-cell lymphoma, not otherwise specified of the thyroid with autoimmune thyroiditis. *Br J Haematol* 2013; **161**(2):214–23. doi: 10.1111/bjh.12255
47. Sun D, Zhou R, Liu H, et al. SASH1 inhibits proliferation and invasion of thyroid cancer cells through PI3K/Akt signaling pathway. *Int J Clin Exp Pathol* 2015; **8**(10):12276–83.
48. Musolf AM, Simpson CL, Moiz BA, et al. Genetic variation and recurrent haplotypes on chromosome 6q23-25 risk locus in familial lung cancer. *Cancer Res* 2021; **81**(12):3162–73. doi: 10.1158/0008-5472.CAN-20-3196
49. Migliorini G, Fiege B, Hosking FJ, et al. Variation at 10p12.2 and 10p14 influences risk of childhood B-cell acute lymphoblastic leukemia and phenotype. *Blood* 2013; **122**(19):3298–307. doi: 10.1182/blood-2013-03-491316
50. Xu H, Yang W, Perez-Andreu V, et al. Novel susceptibility variants at 10p12.31-12.2 for childhood acute lymphoblastic leukemia in ethnically diverse populations. *J Natl Cancer Inst* 2013; **105**(10):733–42. doi: 10.1093/jnci/djt042
51. Deng J, Xue Y, Wang Y, et al. 10p12.31-12.2 variants and risk of childhood acute lymphoblastic leukemia in a Chinese population. *Leuk Lymphoma* 2015; **56**(1):175–8. doi: 10.3109/10428194.2014.910775
52. Jude JG, Spencer GJ, Huang X, et al. A targeted knockdown screen of genes coding for phosphoinositide modulators identifies PIP4K2A as required for acute myeloid leukemia cell proliferation and survival. *Oncogene* 2015; **34**(10):1253–62. doi: 10.1038/nc.2014.77
53. Liao F, Yin D, Zhang Y, et al. Association between PIP4K2A polymorphisms and acute lymphoblastic leukemia susceptibility. *Medicine (Baltimore)* 2016; **95**(18):e3542. doi: 10.1097/MD.0000000000003542
54. Shu X, Long J, Cai Q, et al. Identification of novel breast cancer susceptibility loci in meta-analyses conducted among Asian and European descendants. *Nat Commun* 2020; **11**(1):1217. doi: 10.1038/s41467-020-15046-w
55. Emerling BM, Hurov JB, Poulogiannis G, et al. Depletion of a putatively druggable class of phosphatidylinositol kinases inhibits growth of p53-Null tumors. *Cell* 2013; **155**(4): 844–57. doi: 10.1016/j.cell.2013.09.057
56. Zhang S, Li Z, Yan X, et al. Regulatory network and prognostic effect investigation of PIP4K2A in leukemia and solid cancers. *Front Genet* 2019; **9**:721. doi: 10.3389/fgene.2018.00721
57. Mostowska A, Sajdak S, Pawlik P, et al. Replication study for the association of seven genome-wide-identified loci with susceptibility to ovarian cancer in the polish population. *Pathol Oncol Res* 2015; **21**(2):307–13. doi: 10.1007/s12253-014-9822-6
58. Chen H, Ma H, Li L, et al. Genome-wide association study identifies new susceptibility loci for epithelial ovarian cancer in Han Chinese women. *Nat Commun* 2014; **5**:4682. doi: 10.1038/ncomms5682
59. Pu Y, Wei J, Wu Y, et al. THUMP3-AS1 facilitates cell growth and aggressiveness by the miR-218-5p/SKAP1 axis in colorectal cancer. *Cell Biochem Biophys* 2022; **80**(3):483–94. doi: 10.1007/s12013-022-01074-4

60. Kho PF, Wang X, Cuéllar-Partida G, et al. Multi-tissue transcriptome-wide association study identifies eight candidate genes and tissue-specific gene expression underlying endometrial cancer susceptibility. *Commun Biol* 2021;**4**(1):1211. doi: 10.1038/s42003-021-02745-3
61. Jareid M, Snapkov I, Holden M, et al. The blood transcriptome prior to ovarian cancer diagnosis: A case-control study in the NOWAC postgenome cohort. *PLoS One* 2021;**16**(8):e0256442. doi: 10.1371/journal.pone.0256442
62. Breyer JP, Avritt TG, McReynolds KM, et al. Confirmation of the HOXB13 G84E germline mutation in familial prostate cancer. *Cancer Epidemiol Biomarkers Prev* 2012;**21**(8):1348–53. doi: 10.1158/1055-9965.EPI-12-0495
63. Balint I, Fischer J, Ljungberg B, et al. Mapping the papillary renal cell carcinoma gene between loci D17S787 and D17S1799 on chromosome 17q21.32. *Lab Invest* 1999;**79**(12):1713–8.
64. Tsuruta T, Kozaki KI, Uesugi A, et al. miR-152 is a tumor suppressor microRNA that is silenced by DNA hypermethylation in endometrial cancer. *Cancer Res* 2011;**71**(20):6450–62. doi: 10.1158/0008-5472.CAN-11-0364
65. Dupont WD, Breyer JP, Johnson SH, et al. Prostate cancer risk variants of the HOXB genetic locus. *Sci Rep* 2021;**11**(1):11385. doi: 10.1038/s41598-021-89399-7
66. Zidane M, Cazier JB, Chevillard S, et al. Genetic susceptibility to radiation-related differentiated thyroid cancers: a systematic review of literature. *Endocr Relat Cancer* 2019;**26**(10):R583–96. doi: 10.1530/ERC-19-0321
67. Takahashi M, Saenko VA, Rogounovitch TI, et al. The FOXE1 locus is a major genetic determinant for radiation-related thyroid carcinoma in Chernobyl. *Hum Mol Genet* 2010;**19**(12):2516–23. doi: 10.1093/hmg/ddq123
68. Damiola F, Byrnes G, Moissonnier M, et al. Contribution of ATM and FOXE1 ( TTF2 ) to risk of papillary thyroid carcinoma in Belarusian children exposed to radiation. *Int J Cancer* 2014;**134**(7):1659–68. doi: 10.1002/ijc.28483
69. Morton LM, Karyadi DM, Stewart C, et al. Radiation-related genomic profile of papillary thyroid carcinoma after the Chernobyl accident. *Science* 2021;**372**(6543):eabg2538. doi: 10.1126/science.abg2538