Received: 3 August 2021 | Revised: 26 November 2021 | Accepted: 3 January 2022

DOI: 10.1002/ijc.33939

CANCER GENETICS AND EPIGENETICS

IJC INTERNATIONAL JOURNAL of CANCER

Differential Orthopedia Homeobox expression in pulmonary carcinoids is associated with changes in DNA methylation

Laura Moonen[1] | Lise Mangiante[2] | Daphne J. G. Leunissen[1] | Lisa M. V. Lap[1] | Aurelie Gabriel[2] | Lisa M. Hillen[1] | Guido M. Roemen[1] | Alexander Koch[1,3] | Manon van Engeland[1] | Anne-Marie C. Dingemans[4,5] | Matthieu Foll[2] | Nicolas Alcala[2] | Lynnette Fernandez-Cuesta[2] | Jules L. Derks[4] | Ernst-Jan M. Speel[1]

[1]Department of Pathology, GROW School for Oncology and Developmental Biology, Maastricht University Medical Centre, Maastricht, The Netherlands

[2]Rare Cancers Genomics Team (RCG), Genomic Epidemiology Branch (GEM), International Agency for Research on Cancer/World Health Organisation (IARC/WHO), Lyon, France

[3]Epify BV, Maastricht, The Netherlands

[4]Department of Pulmonary Diseases, GROW School for Oncology and Developmental Biology, Maastricht University Medical Centre, Maastricht, The Netherlands

[5]Department of Pulmonary Medicine, Erasmus MC Cancer Institute, University Medical Center, Rotterdam, The Netherlands

Correspondence
Ernst-Jan M. Speel, Department of Pathology, Maastricht University Medical Center, P. Debyelaan 25, P.O. Box 5800, 6202 AZ Maastricht, The Netherlands,
Email: ernstjan.speel@mumc.nl

Funding information
This work was supported by the Dutch Cancer Foundation [grant number 10956, 2017 to Ernst-Jan M. Speel], the French National Cancer Institute (INCa, PRT-K-17-047 to Lynnette Fernandez-Cuesta and Matthieu Foll), the Ligue Nationale Contre le Cancer (LNCC 2016 to Lynnette Fernandez-Cuesta), and the NeuroendocrineTumor Research Foundation (NETRF, Investigator Award 2019 to Lynnette Fernandez-Cuesta). Lise Mangiante has a fellowship from the LNCC. The results shown here are in part based upon data generated by the TCGA Research Network (https://www.cancer.gov/tcga) and the Rare Cancers Genomics initiative (http://rarecancersgenomics.com/lungnenomics/).

Abstract

Limited number of tumor types have been examined for Orthopedia Homeobox (OTP) expression. In pulmonary carcinoids, loss of expression is a strong indicator of poor prognosis. Here, we investigated *OTP* expression in 37 different tumor types, and the association between *OTP* expression and DNA methylation levels in lung neuroendocrine neoplasms. We analyzed publicly available multi-omics data (whole-exome-, whole-genome-, RNA sequencing and Epic 850K-methylation array) of 58 typical carcinoids, 27 atypical carcinoids, 69 large cell neuroendocrine carcinoma and 51 small cell lung cancer patients and TCGA (The Cancer Genome Atlas) data of 33 tumor types. 850K-methylation analysis was cross-validated using targeted pyrosequencing on 35 carcinoids. We report bimodality of *OTP* expression in carcinoids (OTP^{high} vs OTP^{low} group, likelihood-ratio test $P = 1.5 \times 10^{-2}$), with the OTP^{high} group specific to pulmonary carcinoids while absent from all other cohorts analyzed. Significantly different DNA methylation levels were observed between OTP^{high} and OTP^{low} carcinoids in 12/34 OTP infinium probes (FDR < 0.05 and β-value effect size > .2). OTP^{low} carcinoids harbor high DNA methylation levels as compared to OTP^{high} carcinoids. OTP^{low} carcinoids showed a significantly worse

Abbreviations: AC, atypical carcinoids; FDR, false discovery rate; FFPE, formalin-fixed paraffin-embedded; FPKM, fragments per kilobase of transcript per million mapped reads; GBM, glioblastoma; GO, gene ontology; GSEA, Gene-Set Enrichment Analysis; IHC, immunohistochemistry; IQR, interquartile range; LCNEC, large-cell neuroendocrine carcinoma; LGG, low-grade glioma; LNENs, lung neuroendocrine neoplasms; MOFA, multi-omics factor analysis; OTP, Orthopedia Homeobox; SCLC, small-cell lung carcinomas; TC, typical carcinoids; TCGA, The Cancer Genome Atlas; TNM, tumor node metastasis; WES, whole-exome sequencing; WHO, World Health Organization.

Laura Moonen and Lise Mangiante contributed equally to the first authorship.

Lynnette Fernandez-Cuesta, Jules L. Derks and Ernst-Jan M. Speel contributed equally to the last authorship.

overall survival (log-rank test $P = .0052$). Gene set enrichment analysis for somatically mutated genes associated with hallmarks of cancer showed robust enrichment of three hallmarks in the OTP^{low} group, that is, sustaining proliferative signaling, evading growth suppressor and genome instability and mutation. Together our data suggest that high *OTP* expression is a unique feature of pulmonary carcinoids with a favorable prognosis and that in poor prognostic patients, *OTP* expression is lost, most likely due to changes in DNA methylation levels.

KEYWORDS
epigenetics, methylation, neuroendocrine, Orthopedia Homeobox, pulmonary carcinoid

What's new?

Patients with pulmonary carcinoids that have low expression of the gene *OTP* (Orthopedia Homeobox) tend to have a worse prognosis. However, no mutations or genomic modifications have been found that explain the difference in *OTP* expression. Here, the authors analyzed DNA methylation data, and they found that changes in DNA methylation patterns were associated with the difference in *OTP* expression levels. These findings may suggest a role for epigenetic therapies in pulmonary carcinoid patients in the future.

1 | INTRODUCTION

Lung neuroendocrine neoplasms (LNENs) comprise a heterogeneous group of malignancies, mainly arising from pulmonary neuroendocrine cells, which account for approximately 20% of all primary lung cancers. According to the World Health Organization (WHO) 2021 classification, LNENs are subdivided into four different entities namely, typical carcinoids (TC, 1.8%), atypical carcinoids (AC, 0.2%), large-cell neuroendocrine carcinomas (LCNEC, 3%) and small-cell lung carcinomas (SCLC, 15%).[1,2] Even though pulmonary carcinoids only encompass 1% to 2% of all invasive lung malignancies, their occurrence has increased significantly over the past decades.[3,4]

LNENs share morphological and immunohistochemical features but they show a broad clinical-pathological spectrum with a different biological behavior. In contrast to the high-grade LNENs (eg, LCNEC and SCLC), carcinoids are characterized by a low mitotic frequency (≤10 mitosis/mm^2), none or punctuated necrosis, and a relatively favorable prognosis. Although carcinoids are considered as low- or intermediate grade tumors, they may show distant disease recurrence (TC: 1%-6% and AC: 14%-29%) in patients who initially underwent curative surgical resection.[5,6]

Known predictive factors for recurrence of disease after surgery for carcinoids are histopathological type (AC), tumor size and lymphatic involvement. Unfortunately, none of these factors enables the clinician to reliably predict patients at risk for recurrence up-front. Therefore, extensive follow-up is required for all patients with a carcinoid up to 15-year after surgical resection resulting in frequent over-surveillance.[1,7,8] A possible alternative is to subdivide carcinoids into prognostically relevant categories (at risk for recurrence vs no risk) using tumor-specific molecular features. Previously, we identified Orthopedia Homeobox (OTP) as a prominent molecular marker to accurately identify patients at risk for disease recurrence and showed that *OTP* transcription levels and nuclear protein expression levels were strongly correlated ($P = 5.9 \times 10^{-11}$, n = 60).[9] In addition, our research group and others have shown that OTP protein expression is absent in normal neuroendocrine cells of the lung, whereas OTP is highly expressed in carcinoid precursor lesions (ie, neuroendocrine cell hyperplasia and tumor lets), as well as in most pulmonary carcinoids.[9-14] Other NENs (ie, gastrointestinal/pancreatic) turned out to be OTP negative.[10] Intriguingly, pulmonary carcinoids without expression of OTP at the time of diagnosis develop more frequently metastases during follow-up.[11] These data suggest that OTP is a prognostic marker in pulmonary carcinoids and may play a role in carcinoid tumorigenesis. Although OTP has been described as a key player in the development of the hypothalamic neuroendocrine system of vertebrates, its specific role in tumorigenesis and carcinoid disease recurrence remains to be elucidated.

Here, we performed a comparative analysis on *OTP* expression throughout 37 different tumor types, using publicly available The Cancer Genome Atlas (TCGA) data (n = 33/37 tumor types) combined with publicly available LNEN data (n = 4/37 tumor types), to identify groups with similar expression profiles. Subsequently, we used multi-omics publicly available data on LNENs to study the molecular differences between OTP^{high} vs OTP^{low} LNENs, with a special focus on carcinoids.[15] In addition, we correlated transcriptomic and methylomic data to shed light on the regulatory mechanisms of OTP expression.

2 | MATERIALS AND METHODS

2.1 | Public LNEN and TCGA data

LNEN whole-exome sequencing (WES), transcriptome sequencing (RNAseq), Epic 850K Illumina arrays methylation data, mutational data and

corresponding clinical and histopathological annotations were available from the lungNENomics project within the Rare Cancers Genomics initiative http://rarecancersgenomics.com/lungnenomics/. Detailed information on data generation and quality controls can be found in Gabriel *et al* GigaScience 2020.[16] This dataset includes 88 carcinoids of which 58 TC, 27 AC and 3 unclassified carcinoids, 69 LCNEC and 51 SCLC.[15,17-19] All samples were collected from surgically resected tumors.

Transcriptomic data (In units of Fragment Per Kilobase of transcript per Million mapped read, FPKM) and corresponding clinical data of 33 different tumor types were gathered from the publicly available The Cancer Genome Atlas (TCGA) platform. Data were downloaded using the Bioconductor R package *TCGAbiolinks* (version 2.9.5). Duplicated samples were removed as well as one lung squamous cell carcinoma sample (TCGA-37-4129) because this sample was previously reported as misclassified.[20] TCGA *OTP* expression data was merged with LNEN *OTP* expression data to gather an overview of *OTP* expression patterns throughout 37 different tumor types.

2.2 | Data processing of expression data

Data processing was performed as described previously by Alcala et al,[15] providing quantification of expression at the gene level in two formats. (a) FPKM, one of the most popular formats for expression quantification, that facilitates comparisons across cohorts by mitigating technical batch effects through normalization based on gene length and sequencing library size; this is the format we used for comparing with TCGA cohorts, and for interpreting absolute expression levels. (b) Normalized read counts, obtained through the variance stabilization procedure described in Alcala et al using the *DESeq2* R package,[15] that facilitates downstream statistical analyses of the expression data within a cohort by reducing the relationship between mean and variance in expression level; this is the format we used for the statistical analysis of the expression (see determining an expression cut-off point procedure, and gene set enrichment analysis below).

2.3 | Clinical data

Clinical cohort data were retrieved from Alcala et al.[15] Data included, among others, age (in years), sex (male or female), smoking status (never smoker, former smoker, passive smoker and current smoker), Union for International Cancer Control/American Joint Committee on Cancer stage (Stage I-IV) and survival (calculated in months from surgery to last day of follow-up or death).

2.4 | Gene-set enrichment analysis of somatic mutations

Gene-set enrichment analysis of somatic mutations (GSEA) was performed as described previously.[15] In short, GSEA for somatic mutations was evaluated independently for each set of genes linked to a hallmark of cancer, taking into account genes with multiple mutations, using the Fisher's exact test (Fisher.test R function, *stats* package version 4.0.4).[21] *P* values for both the OTP^{high} and OTP^{low} group were adjusted for multiple testing. Altered hallmarks, including the mutated genes, are presented in Table S1.

In addition, we performed robustness analyses to assess the validity of the results, especially to outlier samples/genes that would have a high influence on the statistical results, e.g., that would alone drive the significance of a particular hallmark. First, we assessed the influence of each individual sample using a jackknife procedure (ie, for each sample, we performed the GSEA test after removing this sample). Second, we assessed the influence of each gene using a jackknife procedure (ie, for each gene, we performed the GSE test without this gene). A threshold adjusted *P*-value of .05 was fixed to select enriched pathway.

2.5 | Methylome analyses

DNA methylation levels at each OTP Infinium probe (publicly available 850K array data) were analyzed for 51 LNEN samples (n = 19 TC, n = 15 AC, n = 17 LCNEC) of the total cohort, combined with 528 publicly available TCGA normal lung adenocarcinoma (n = 274) and squamous cell carcinoma (n = 254) tissues (both 450K array data). Differential DNA methylation was tested using the Wilcoxon rank-sum test on β values, to allow comparisons between the nonnormally distributed distributions of β values. False discovery rate (FDR) was controlled using the Benjamini-Hochberg procedure to correct for multiple hypothesis testing. In addition, an overview of the DNA methylation data at all OTP probes together with the genomic location of these probes was presented. The exon and transcript annotations were retrieved from Ensembl using the BioMart tool (Ensemble Genes 75, Homo sapiens genes GRCh38.p12), CpG island locations were downloaded from the UCSC genome browser,[22] and the EPIC probe locations were retrieved from Zhou et al using the human genome version GRCh38.p12.

2.6 | Validation of the methylome analyses using pyrosequencing

To cross-validate the methylome analyses findings, we have performed pyrosequencing on n = 35/88 fresh frozen carcinoid samples of the sequencing cohort. DNA was isolated using the Gentra Puregene tissue kit 4g (Qiagen, Hilden, Germany), following the manufacturer's instructions. DNA samples were quantified by the fluorometric method (Quanti-iT PicoGreen dsDNA Assay, Life Technologies, CA) and integrity was checked by gel electrophoresis using a 1.3% agarose gel. Next, sodium bisulfite-modification, which converts unmethylated cytosine residues to uracil residues, was performed on 100 ng genomic DNA using the EpiTect Bisulfite Kit (Qiagen, Hilden, Germany), according to the manufacturer's instructions. Bisulfite modified DNA was amplified

using methylation-specific primers, which were designed using the Pyromark Assay design SW 2.0 (Qiagen, Hilden Germany; Table 1). PCR amplification was performed using the PyroMark PCR kit (Qiagen, Hilden Germany), according to the manufacturer's protocol, on a T100 thermal cycler (Bio-Rad). Afterward, PCR products were loaded on a 2% agarose gel and visualized using Proxima AQ-4 imaging software (Isogen, De Meern, The Netherlands). Next, the methylation status of cg26576712 (genomic location 77640156-77640158) and cg02493167 (genomic location 77639893-77639895), both located in the promoter region of *OTP*, were determined by pyrosequencing using PyroMark Q24 Advanced CpG reagents and a PyroMark Q24 Instrument upgraded with the PyroMark Q24 Advanced software (Qiagen, Hilden, Germany) according to the manufacturer's protocol. After pyrosequencing, the methylation percentage of each site was determined by analyzing the pyrograms using the PyroMark Q24 Advanced software (Qiagen, Hilden, Germany). Every PCR and pyrorun included a bisulfite converted and methylated DNA obtained from a RKO cell line, EpiTect unmethylated and bisulfite-converted controls (Qiagen, Hilden, Germany) and a H_2O control. In addition, an internal cytosine control was incorporated in the sequence to be analyzed. Pyrosequencing on normal lung samples (n = 3) was performed to determine the cut-off methylation level on cg26576712 and cg02493167 (mean positivity threshold 32.2%).

2.7 | Immunohistochemistry

Immunohistochemistry (IHC) on formalin-fixed paraffin-embedded (FFPE) tissue sections was performed using the rabbit anti-OTP primary polyclonal antibody (HPA039365; Atlas Antibodies). The primary antibody was diluted 1:3000 and incubated overnight at 4°C. Antibodies were detected by Bright Vision Poly-HRP-anti-mouse/rabbit/rat immunoglobulin G (IgG; Immunologic) followed by peroxidase-DAB (3,3′-diaminobenzidine) visualization. Expression of OTP IHC was scored as described previously.[9]

2.8 | Statistical analysis

Statistical analyses were conducted using SPSS for Mac version 25 (SPSS Inc., Chicago, IL) and RStudio for Mac version 3.6.1. The chi-squared test and Fisher's exact test were used to compare categorical data. One-sample *t*-test was performed to test the significance between measures in the Bland-Altman plots. Survival analysis for LNENs was performed using Cox's proportional hazard model (R package *survival* version 3.2-11), with statistical differences between groups performed using the Wald test and global fit of the model assessed using the log-rank test. Two-sided *P* values < .05 were considered significant. Bimodality in gene expression data was tested using the function normalmixEM from the R package *mixtools*. A likelihood ratio test was performed, using the chi-squared test to choose between a null model (unimodal Gaussian distribution) and the bimodal model, rejecting the null when *P*-value < .05. The optimal cut-off was then defined as the lowest density point of the two Gaussian mixture distributions (in our case, 8.7).

2.9 | Independent dataset of pulmonary carcinoids

To confirm our findings, we used the publicly available gene expression, mutation and Illumina 450K methylation array data from an independent pulmonary carcinoid cohort.[23] The cohort comprised of 30 primary lung carcinoids including 17 TC and 13 AC. Methylome data were available in 17 patients (n = 10 TC and n = 7 AC). The data were downloaded from Supporting Information data files (https://www.omicsdi.org/dataset/geo/GSE118133) reported in Reference 23. Data processing and subsequent analysis were performed as previously described above.

3 | RESULTS

3.1 | Patterns of *OTP* expression in pulmonary carcinoids

To gain insights into the expression patterns of *OTP* throughout 33 different TCGA tumor types and four LNEN subtypes, we performed comparative data analyses. Results showed that *OTP* was expressed in pulmonary carcinoids with higher levels in TC (median 126.4, interquartile range [IQR] 72.9-193.4 FPKM) than in AC (median 0.16, IQR 0.06-57.7 FPKM). However, both TC and AC groups include samples

TABLE 1 Overview of the primer combinations for both CpG sites within the promoter region of OTP and their specifications

Location	Primer	Sequence	Nt	T_m (°C)	% CG	Product length (bp)
OTP cg02493167	Forward	GGGAGTAGTAAATATTAGTTTTTATTGTGA	30	58.8	26.7	160
	Reverse	ATTCTATACCATTTCTAATCTACTCCTAAA	30	57.3	26.7	
	Pyrosequencing	ATGTTTTGTTATAAATATAATTG	23	39.2	13.0	
OTP cg26576712	Forward	GTTTTTAGTTAGTATTTTTAATGTTTTGTTTAAGT	35	57.2	17.1	116
	Reverse	CCTTCCACAAAAAAATAACCCAATAA	26	58.4	30.8	
	Pyrosequencing	ATGTTTTGTTTAAGTTAATTGG	22	44.6	22.7	

Abbreviations: bp, base pairs; CG, cytosine-guanine content; Nt, nucleotides; T_m, melting temperature.

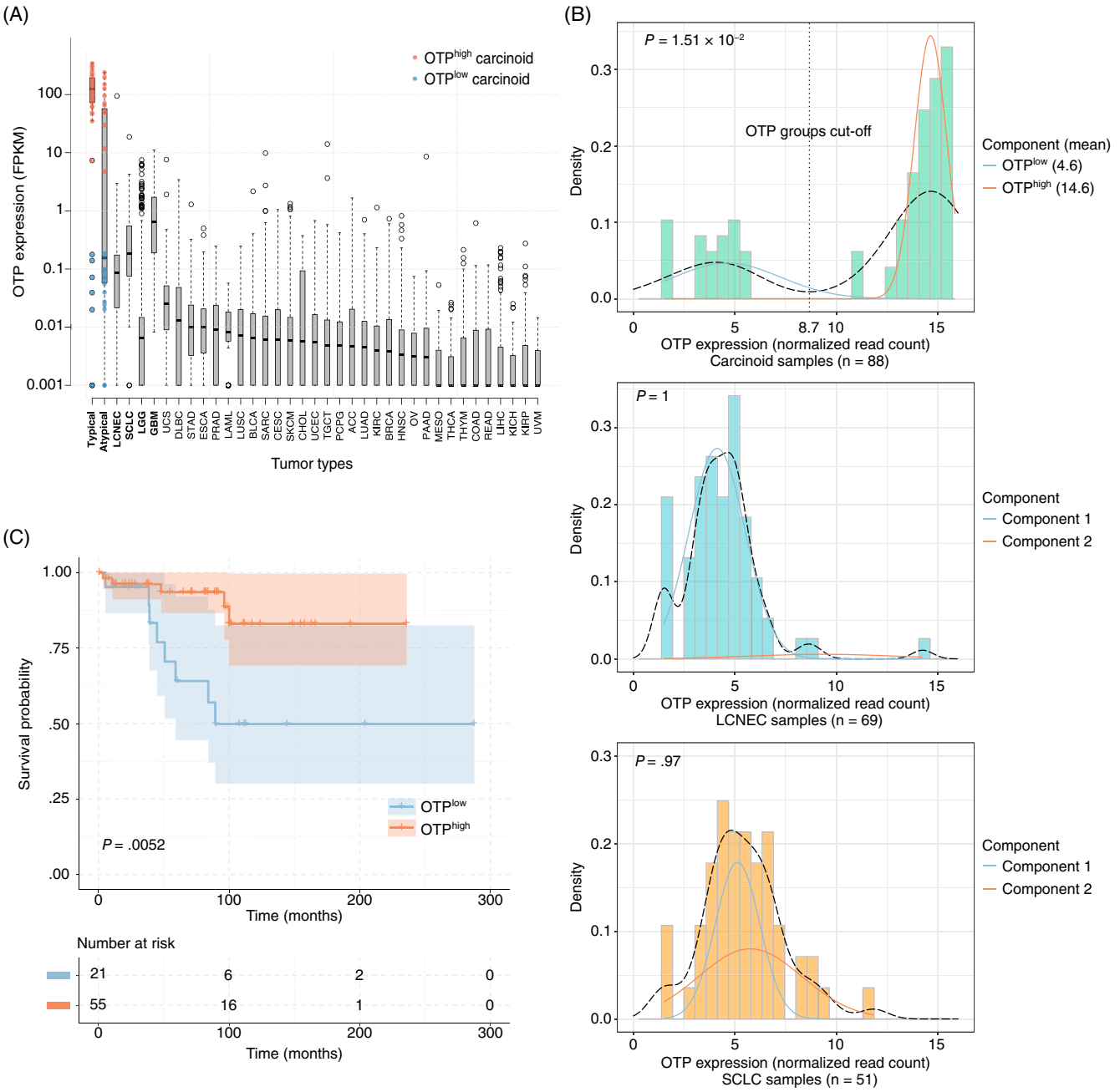


FIGURE 1 (A) RNA gene expression of *OTP* in 37 different tumor types highlighted using boxplots in fragment per kilobase million (FPKM). Center lines represent the median and box bounds represent the interquartile range (IQR). (See Table S4 for the abbreviation list of the tumor types.) (B) Histograms presenting the *OTP* expression pattern, in units of normalized read counts, in carcinoids (upper panel), LCNEC (middle panel) and SCLC (lower panel). The striped curve represents the distribution fit of the two Gaussian mixture distributions (component 1, in blue and component 2, in orange). The *OTP* cut-off is determined as the lowest density point of the two Gaussian mixture distributions (upper panel, $x = 8.7$). (C) Kaplan-Meier curves of overall survival probability for the OTP^{high} and OTP^{low} group in pulmonary carcinoid patients [Color figure can be viewed at wileyonlinelibrary.com]

with lower and higher *OTP* expression, respectively (Figure 1A). All other tumor types, including pulmonary LCNEC (median 0.09, IQR 0.02-0.17 FPKM) and SCLC (median 0.19, IQR 0.08-0.55 FPKM), showed very low to no *OTP* expression. Both the glioblastoma (GBM; median 0.64, IQR 0.19-1.7) and low-grade glioma (LGG; median 0.007, IQR 0.001-0.04) cohort show an overall low median expression of *OTP*, some samples within these groups display higher *OTP* expression levels (Figure 1A). Nevertheless, these data suggest that *OTP* is a highly specific marker for pulmonary carcinoids as compared to all other TCGA tumor types.

Considering that high- and low-*OTP* expression samples were detected in both TC and AC, we assessed the *OTP* gene expression of 88 sequenced carcinoids to determine the optimal cut-off point for OTP^{high}- and OTP^{low} -classification. A mixture model of two Gaussian distributions was fitted to the distribution of *OTP* expression (normalized read counts) in LNENs (Figure 1B). Data showed a clear bimodal distribution of *OTP* in carcinoids ($P = 1.5 \times 10^{-2}$) suggesting that within the carcinoid cohort, two distinct groups exist, which can be separated based on *OTP* expression levels (Figure 1B, upper panel). Next, we confirmed the absence of *OTP* bimodality in both the LCNEC- and SCLC sample cohort (Figure 1B, middle and bottom panels, respectively). By applying this cut-off point (lowest density point of the two Gaussian mixture distributions 8.7, in units of normalized read counts) to the high-grade LNENs, only one LCNEC sample and one SCLC sample were considered OTP^{high}, whereas all other high-grade LNENs were classified as OTP^{low}. In fact, Alcala et al. recently showed that this LCNEC sample (S00602) clustered with carcinoid tumors according to Multi-Omics Factor Analysis (MOFA).[15]

3.2 | Molecular and clinical characteristics of the OTP^{high} and OTP^{low} carcinoids

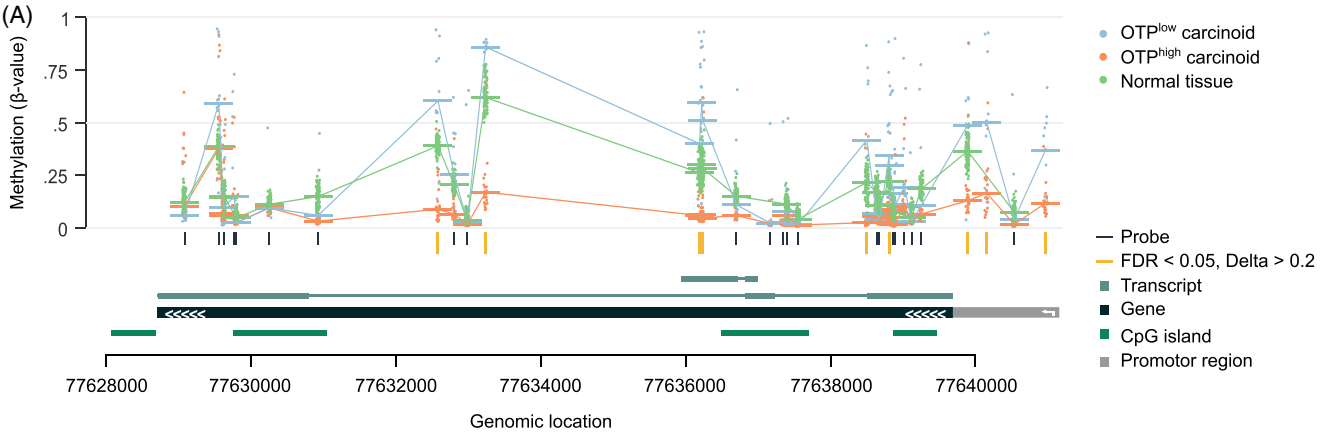
Baseline clinical characteristics are presented in Table 2. The OTP^{high} cohort was enriched for females (69%), while the OTP^{low} cohort contained relatively more males (79%, Fisher's exact test *P*-value $= 7.9 \times 10^{-5}$). Both age and smoking status were comparable within the two groups (Fisher's exact test both *P*-value = .16). In addition, patients in the OTP^{high} group were more frequently diagnosed as TC, whereas the OTP^{low} group was enriched for AC histopathology (Fisher's exact test *P*-value = .00013). Compared to OTP^{low}, patients in the OTP^{high} group were more frequently diagnosed in lower Tumor Node Metastasis (TNM) stages (Fisher's exact test of stage I-II vs stages III-IV *P*-value = .01).

Kaplan-Meier survival analysis was performed for the pulmonary carcinoids clustered in OTP^{high} and OTP^{low}. Results showed that carcinoids with low *OTP* expression harbor an unfavorable survival as compared to carcinoids with high *OTP* expression (log-rank test *P*-value = .0052, Figure 1C).

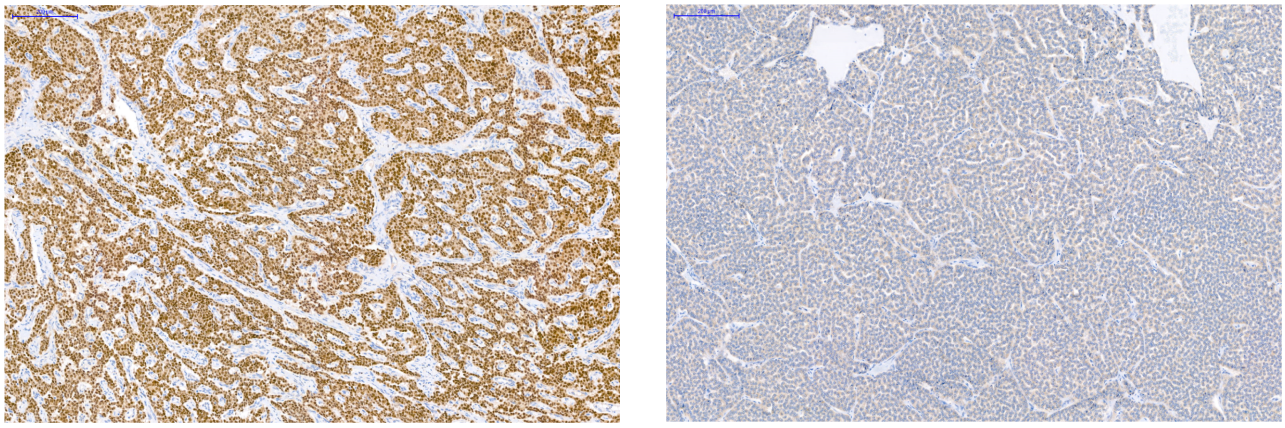
TABLE 2 Patient characteristics of the OTP^{high} and OTP^{low} group

Variable	Groups		*P*-value
	OTP^{high}	OTP^{low}	
Patients, n	64	24	
Age, years			
Mean ± SD	51.6 ± 18.3	58.3 ± 12.8	
Median (IQR)	54 (16-80)	58 (29-80)	.16
Gender			7.9×10^{-5}
Female	44 (68.7)	5 (20.8)	
Male	20 (31.3)	19 (79.2)	
Smoking status			.16
Current	14 (21.9)	3 (12.5)	
Former	13 (20.3)	8 (33.3)	
Never	23 (35.9)	6 (25.0)	
Passive		1 (4.2)	
Histopathological classification			1.30×10^{-4}
Typical	50 (78.1)	8 (33.3)	
Atypical	12 (18.8)	15 (62.5)	
Unclassified	2 (3.1)	1 (4.2)	
TNM stage			.01
I-II	60 (94)	18 (75.0)	
III-IV	3 (5)	6 (25.0)	
Unknown	1 (1)		
Survival censor			.03
Alive	50 (78.1)	14 (58.3)	
Death	6 (9.4)	8 (33.3)	
Unknown	8 (12.5)	2 (8.3)	
Median survival in months	79.3	59	

Data are presented as n (%) unless otherwise stated. Abbreviations: IQR, interquartile range; TNM, tumor node metastasis.

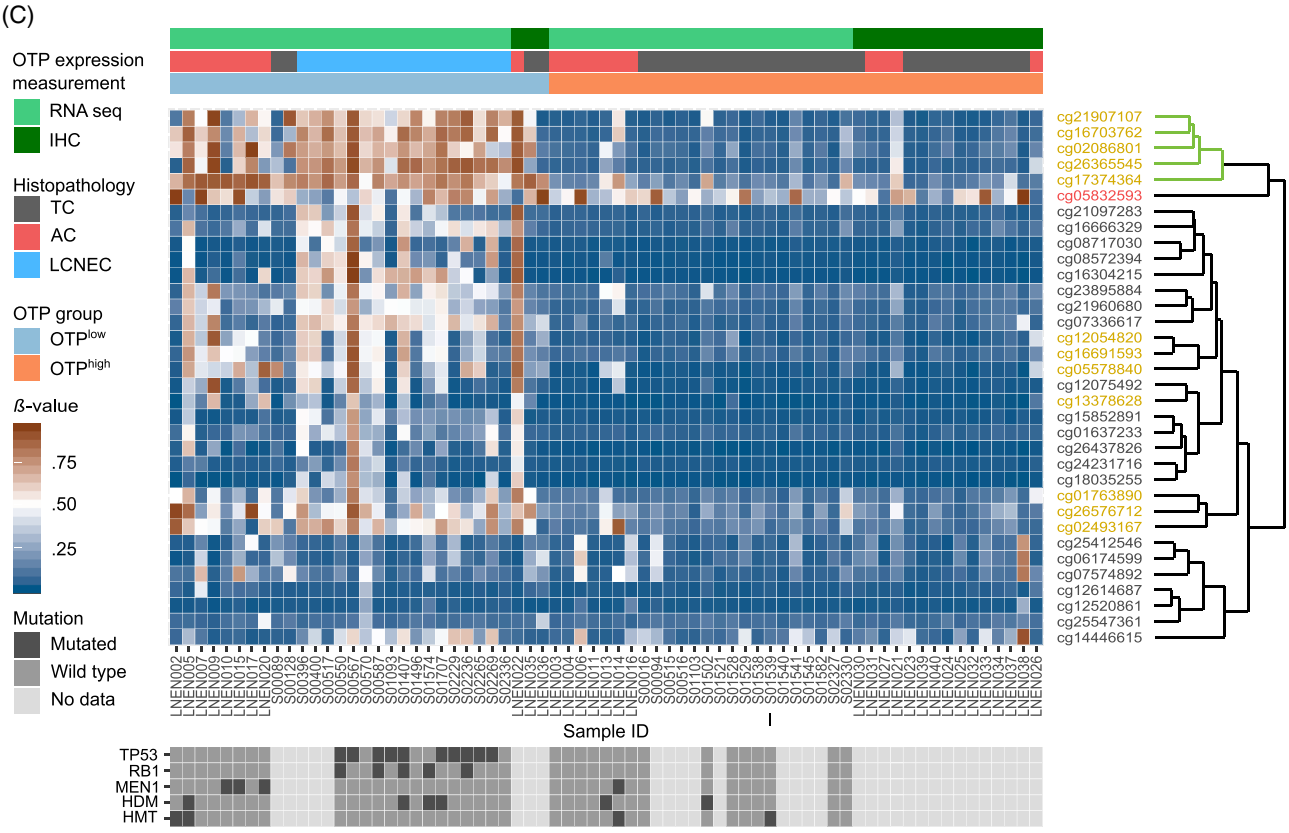


(B)



nOTP+/β-value: 0.07/pyrosequencing: 8%

nOTP-/β-value: 0.48/pyrosequencing: 42%



To investigate whether the survival difference between OTP^{high} and OTP^{low} carcinoids might be associated with somatic gene mutations, we performed a GSEA for mutated genes related to hallmarks of cancer.[21] The analysis showed that mutated genes in carcinoids clustered in the OTP^{low} group were strongly enriched in three hallmarks, that is, sustaining proliferative signaling ($P = 1.05 \times 10^{-3}$), evading growth suppressor ($P = 1.11 \times 10^{-8}$) and genome instability and mutation ($P = 6.53 \times 10^{-5}$, Figure S1, Table S1). Jackknife analyses of samples showed that OTP^{low} enrichments for somatic alterations were not influenced by a single sample (Table S1). In addition, jackknife analyses of mutated genes underlined that *MEN1* is necessary for the enrichment of genome instability and mutation and sustaining proliferative signaling hallmarks in the OTP^{low} group (Table S1). In the OTP^{high} cluster, three hallmark enrichments nearly reached significance including activating invasion and metastasis ($P = .03$), evading growth suppressor ($P = .02$) and sustaining proliferative signaling ($P = .04$, Table S1). However, it must be considered that jackknife analyses for both samples and mutated genes in the OTP^{high} group revealed poor stability of enrichments, indicating that the enrichments are driven by specific samples and/or mutated genes. Interestingly, *MEN1* mutations were found in 6 out of the 14 carcinoids clustered in the OTP^{low} group while only one case with a *MEN1* mutation was found in the OTP^{high} group of 40 patients. This is in agreement with several studies proving that *MEN1* mutations are associated with poor prognosis.[17,24,25] These data suggest that the clear and robust enrichment of hallmarks of cancers in the OTP^{low} group may explain tumor aggressiveness and thereby the difference in survival.

3.3 | Mechanisms of OTP inactivation

3.3.1 | Genomic OTP inactivation

To examine which underlying regulatory mechanism may cause the expression differences between OTP^{high} and OTP^{low} carcinoids, we investigated publicly available gene mutation data.[15] Results, to date, have shown no gene-inactivating somatic mutations, alterations by chimeric transcripts or genomic rearrangements in the *OTP* gene (Supplementary data, Alcala *et al*[15]).[17]

To further investigate the regulatory mechanism underlying *OTP* expression, we evaluated a possible association between DNA methylation and *OTP* expression using both transcriptome and methylome data (Illumina Infinium 850K for tumor samples and 450K for normal samples). Transcriptomic and methylomic data were available in 51 samples including 24 OTP^{high}, 10 OTP^{low} and 17 LCNEC samples. Results showed 12/34 OTP Infinium probes harboring a significantly different methylation level (FDR < 0.05 and delta > 0.2) between OTP^{high} and OTP^{low} carcinoids, of which three probes were located in the promotor region (cg02493167, cg26576712 and cg01763890; Figure 2A, Table S2). At these 12 loci, OTP^{high} carcinoids harbor a lower methylation level (based on β values) as compared to OTP^{low} carcinoids. In addition, normal TCGA lung samples showed a baseline methylation level, which was greater than OTP^{high} and lower than OTP^{low} carcinoids (Figure 2A, specified in green). This is interesting since OTP expression have not yet been reported in normal lung tissues. Evaluation of the LCNEC samples showed a median RNA expression level of 0.09 FPKM and high methylation levels (Figure S2). Of note, the methylation level of LCNEC samples was higher as compared to OTP^{low} carcinoids. Hence, we hypothesized that epigenetically hypermethylation of *OTP* could lead to gene-silencing and subsequent loss of protein expression.

To cross-validate the *OTP* methylation within the groups of pulmonary carcinoids, we performed pyrosequencing targeting two single promoter probes present in the Infinium 850K assay (cg02493167 and cg26576712). We have targeted these two CpGs since it has been frequently described that methylation within the promoter region of genes is associated with transcriptional silencing. Results revealed that 26% (n = 9/35) of patients harbored an increased methylation status while 74% (n = 26/35) showed no or low methylation percentages on cg02493167, using the methylation percentage of normal lung samples as the methylation threshold (average methylation percentage of 32.3%). Targeting cg26576712 showed an increased methylation status in 8/23 (35%) of the pulmonary carcinoids while 15/23 (65%) showed a low methylation percentage. Twelve cases failed due to technical issues. Additional IHC analysis (n = 14) confirmed loss of OTP expression in OTP-methylated carcinoids while expression was present in unmethylated carcinoids (Figure 2B). Bland-Altman plots showed no proportional difference between pyrosequencing and 850K arrays (Figure S3; $P = .562$ for cg26576712 and $P = .069$ for cg02493167).

The identified association between DNA methylation and *OTP* expression raised the question whether one or multiple CpGs are associated with expression. For this purpose, we have generated a heatmap of the β values of all cg-sites for both carcinoid clusters (ie, OTP^{high} and OTP^{low}) and LCNEC (Figure 2C). Complete-linkage

FIGURE 2 (A) Plot showing the DNA methylation levels at each OTP infinium probe (850K) of all carcinoid samples (OTP^{high} and OTP^{low}) combined with TCGA normal lung adenocarcinoma and squamous cell carcinoma tissues (450K). The *y*-axis on the right shows the β values; a horizontal bar was drawn at the median β-value for each probe. Differential DNA methylation between OTP^{high} and OTP^{low} carcinoids was calculated using the Wilcoxon rank-sum test (significant different cg-sites are presented in yellow). (B) Representative images illustrating OTP IHC of a pulmonary carcinoid patient showing nuclear OTP positivity and a low methylation level (left panel) and a pulmonary carcinoid patient showing absence of OTP protein expression and a high methylation percentage (right panel). (C) Heatmap of the methylation level (in β values) for the OTP^{high} and OTP^{low} group (*x*-axis) for each cg-site (*y*-axis). The cg-sites which harbor a significantly different methylation level between the groups are presented in yellow. The upper green legend bar represents the OTP level measurement, the middle bar represents the histopathological diagnosis of each sample and the lower bar indicates the OTP group. Somatic mutations are represented in the lower rectangle

clustering using the Euclidean distance metric revealed a cluster with cgs that strongly correlates when applying a tree height cut-off of three (Figure 2C, cg21907107, cg16703762, cg02086801, cg26365545, cg17374364, specified in green). These cgs might possibly explain together the regulation of OTP expression. Noteworthy is cg17374364, which shows a β-value above .5 in all OTP^{low} samples (Figure 2C). While OTP^{high} samples show overall a low methylation level, data reveal that some samples (eg, LNEN013, LNEN014, S01502 and S01539) tend to show a higher methylation level (β-value > .5). Somatic mutation analysis revealed that these samples harbored mutations in genes associated with Gene Ontology (GO) terms related to histone demethylase activity (HDM that is, *KDM4A*, *KDM5C* and *FBXL19*) and histone methylation (HMT that is, *DOT1L*; Figure 2C). In addition, all cases tested with OTP IHC expression consistently showed that the OTP expression groups matched methylation groups (Figure S4).

3.4 | Independent confirmation of our findings

We verified our findings using published pulmonary carcinoid data from Laddha et al,[23] which included gene expression, mutation and Illumina 450K methylation array data of 30 samples (17 TCs and 13 ACs). Bimodality testing confirmed our previous findings showing a clear bimodal distribution of *OTP* expression within the group of carcinoids (Figure S5A, $P = 6.9 \times 10^{-9}$). After the OTP cut-off (lowest density point of the two Gaussian mixture distributions 10.2, in units of normalized read counts), 7 patients were allocated to the OTP^{low} group and 23 patients to the OTP^{high} group. To investigate whether *OTP* expression is also associated with methylation in this cohort, we have generated a heatmap using the Illumina 450K methylation array data. Results showed that the highest methylation levels, in β values, were observed in carcinoids clustered in the OTP^{low} group (Figure S5B). Results showed that, in line with our findings described above, all OTP^{low} samples harbor a β-value above .5 on cg17374364, while OTP^{high} samples show overall a low methylation level (i.e., a β-value below .5). Furthermore, we observed a variable methylation level throughout all samples for cg05832593 (Figure S5B, specified in red), also consistent with our data (Figure 2C, specified in red).

4 | DISCUSSION

Pulmonary carcinoids are rare lung tumors with a relatively indolent course of disease although a subgroup shows a more aggressive disease course. Previously, we have shown that low *OTP* gene and protein expression is associated with a poor prognosis and others have shown that OTP expression may be utilized to identify patients at risk for disease recurrence.[11] However, thus far, the mechanisms underlying the regulation of OTP expression have not been clarified. Here, we evaluated *OTP* expression and methylation levels within 208 LNEN samples, 33 other tumor subtype cohorts and normal lung tissue using publicly available transcriptomics and methylomics data and identified a unique and bimodal expression of *OTP* in lung carcinoids. To date, no mutations or other genomic modifications (ie, chimeric transcripts and/or genomic rearrangements) have been reported in the *OTP* gene. Therefore, we comprehensively analyzed epigenomic data, revealing, for the first time, that differential *OTP* expression patterns could be explained by epigenetic modifications. Our findings were verified in 30 additional pulmonary carcinoids samples from Laddha *et al*.

Previously, Alcala et al correlated gene expression and promoter methylation in pulmonary carcinoids to identify genes, which expression can be explained by methylation of CpG islands.[15] Results highlighted several top candidate genes including two homeobox genes (*HNF1A* and *HNF4A*). However, *OTP*, although methylated, was not among these top candidate genes. This might be the result of the fact that DNA methylation does not occur exclusively at CpG islands. Most of the tissue-specific DNA methylation seems to occur at CpG island shores (region of lower CpG density that lies in close proximity [~2 kb] of CpG island)[26-29] as analyzed in-depth here.

One of the cg sites, cg17374364, harbored a β-value above .5 in all OTP^{low} samples. Nevertheless, also five of the 39 OTP^{high} samples showed a β-value above .5 (LNEN013, LNEN014, S01502, S02330 and LNEN021). Some of these samples carried mutations in lysine demethylases of histone methyltransferases, which might have impacted the methylation status of the above-mentioned cg site.

In relation to the spectrum of LNENs (low-grade to LCNEC), we and others have shown that downregulation of OTP expression is correlated with poor prognosis. Both a subgroup of carcinoids as well as LCNEC are OTP negative. It is interesting to investigate whether this observation matches with other observations in these tumors suggesting a temporal transition from low-grade to high-grade NE carcinomas.[30,31] Alcala et al and Laddha et al identified through different analyses and in a different dataset, three equivalent molecular groups of carcinoids (i.e., A1, A2 and B *vs* LC1, LC2 and LC3, with different clinical features).[15,16,23] Importantly, in both the A1-A2 and LC1-LC3 groups *OTP* was generally highly expressed, while *OTP* was downregulated in the B and LC2 group. Moreover, they showed that within cluster A1 a subcohort existed, also referred to as "supra"-carcinoids, with molecular and clinical characteristics most similar to LCNEC; these supra-carcinoids also showed low *OTP* expression. Alcala et al showed that these supra-carcinoids are, albeit their shared low *OTP* expression, vastly different from the carcinoid B cluster on the molecular level.[15] Despite these distinct genomic features based on both genome-wide expression and methylation profiles, we here observe that these supra-carcinoids showed all high methylation levels similar to the OTP^{low} carcinoids. Most remarkable is LNEN021, because this supra-carcinoid clustered as OTP^{high} based on IHC analysis, but harbored high methylation levels on cg17374364 in contrast to the other OTP^{low} samples. This suggests that despite the high genomic difference between OTP^{low} and supra-carcinoids, they share this common feature of high methylation levels on a specific CpG in the *OTP* gene. It remains to be investigated whether these clusters of carcinoids evolve or

whether they are distinct entities from the beginning and progress as a result of punctuated tumor evolution.[32]

The high expression levels in pulmonary carcinoids as compared to all other cancers, the correlation between loss of expression and poor prognosis and the simplicity of IHC to detect expression mark OTP as a highly suitable diagnostic and prognostic marker for daily clinical practice. However, different studies have reported additional genes which expression correlated with overall patient survival.[9,15,23,33,34] It remains to be studied whether a panel of markers can further improve the prognostication of pulmonary carcinoids.

Our study implicates DNA methylation as possible regulatory mechanism of *OTP*; nevertheless, our study has several limitations. First, it is a retrospective study with a rather small cohort. Albeit pulmonary carcinoid is an orphan disease, future studies should validate our findings using larger cohorts containing extensive clinical data (ie, relapse-free survival). Second, our study did not contain data on the methylation level of normal neuroendocrine cells of the lung. It would be very informative to investigate the methylation levels in these specific cells and in other neuroendocrine tumor types in future studies. Furthermore, we should investigate whether methylation may occur as the result of temporal tumor evolution by analyzing the methylation levels in tissue of primary carcinoid tumors at diagnosis that later developed disease recurrence as well as the corresponding metastatic lesion. Third, even though we did not observe significant differential expression levels in the TET (TET1, TET2 and TET3) and DNMT family (DNMT1, DNMT3A and DNTM3B), two enzyme families that play a major role in DNA methylation, between OTP^{high} and OTP^{low} carcinoids, future studies should investigate the members of the methylation machinery in more detail (Table S3). Finally, future studies might investigate the role of miRNAs as additional regulators of OTP expression.[35]

5 | CONCLUSION

We show that within the group of pulmonary carcinoid patients, two distinct groups exist which can be separated based on *OTP* expression. To the best of our knowledge, we are the first to prove that the differential *OTP* expression within pulmonary carcinoids is associated with changes in DNA methylation. These findings arouse curiosity about whether epigenetic therapies might be useful for pulmonary carcinoid patients in the future.

CONFLICT OF INTEREST

Dr Speel reports personal fees from Amgen, Eli Lilly and Novartis, outside the submitted work. Nonfinancial support from Biocartis and Abbvie and grants from AstraZeneca, Pfizer, Novartis and Bayer, outside the submitted work. Dr Dingemans reports financial conflicts (ad board, paid to institute) of Eli Lilly, Roche, Amgen, Jansen, Pfizer, Astra-Zeneca, Bayer, Boehringer, Ingelheim, Sanofi, outside the submitted work. Dr van Engeland is co-founder and shareholder of Epify BV and MLA Diagnostics BV, companies which develop cancer DNA methylation markers. All other authors declare no potential conflicts of interest. Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer/World Health Organization.

DATA AVAILABILITY STATEMENT

Data sources and handling of the publicly available datasets used in our study are described in the Materials and Methods. Further details and other data that support the findings of our study are available from the corresponding author upon request.

ETHICS STATEMENT

All new specimens were collected from surgically resected tumors, applying local regulations and rules at the collecting site, and including patient consent for molecular analyses as well as collection of de-identified data, with approval of the IARC Ethics Committee.

ORCID

Laura Moonen https://orcid.org/0000-0003-3699-2534

REFERENCES

1. Caplin M, Baudin E, Ferolla P, et al. Pulmonary neuroendocrine (carcinoid) tumors: European neuroendocrine tumor society expert consensus and recommendations for best practice for typical and atypical pulmonary carcinoids. *Ann Oncol*. 2015;26:1604-1620.
2. Derks JL, Hendriks LE, Buikhuisen WA, et al. Clinical features of large cell neuroendocrine carcinoma: a population-based overview. *Eur Respir J*. 2016;47:615-624.
3. Korse CM, Taal BG, van Velthuysen M-LF, Visser O. Incidence and survival of neuroendocrine tumours in The Netherlands according to histological grade: experience of two decades of cancer registry. *Eur J Cancer*. 2013;49:1975-1983.
4. Dasari A, Shen C, Halperin D, et al. Trends in the incidence, prevalence, and survival outcomes in patients with neuroendocrine tumors in the United States. *JAMA Oncol*. 2017;3:1335-1342.
5. Hendifar AE, Marchevsky AM, Tuli R. Neuroendocrine tumors of the lung: current challenges and advances in the diagnosis and management of well-differentiated disease. *J Thorac Oncol*. 2017;12:425-436.
6. Reuling E, Dickhoff C, Plaisier P, Bonjer H, Daniels J. Endobronchial and surgical treatment of pulmonary carcinoid tumors: a systematic literature review. *Lung Cancer*. 2019;134:85-95.
7. Singh S, Bergsland EK, Card CM, et al. Commonwealth neuroendocrine tumour research collaboration and the North American neuroendocrine tumor society guidelines for the diagnosis and management of patients with lung neuroendocrine tumors: an international collaborative endorsement and update of the 2015 European neuroendocrine tumor society expert consensus guidelines. *J Thorac Oncol*. 2020;15:1577-1598.
8. Baudin E, Caplin M, Garcia-Carbonero R, et al. Lung and thymic carcinoids: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2021;32:439-451.
9. Swarts DR, Henfling ME, Van Neste L, et al. CD44 and OTP are strong prognostic markers for pulmonary carcinoids. *Clin Cancer Res*. 2013;19:2197-2207.
10. Moonen L, Derks J, Dingemans A-M, Speel E-J. Orthopedia Homeobox (OTP) in pulmonary neuroendocrine tumors: the diagnostic value and possible molecular interactions. *Cancer*. 2019;11:1508.

11. Papaxoinis G, Nonaka D, O'Brien C, Sanderson B, Krysiak P, Mansoor W. Prognostic significance of CD44 and orthopedia homeobox protein (OTP) expression in pulmonary carcinoid tumours. *Endocr Pathol*. 2017;28:60-70.
12. Roy M, Buehler DG, Zhang R, et al. Expression of insulinoma-associated protein 1 (INSM1) and orthopedia homeobox (OTP) in tumors with neuroendocrine differentiation at rare sites. *Endocr Pathol*. 2019;30:35-42.
13. Hanley KZ, Dureau ZJ, Cohen C, Shin DM, Owonikoko TK, Sica GL. Orthopedia homeobox is preferentially expressed in typical carcinoids of the lung. *Cancer Cytopathol*. 2018;126:236-242.
14. Viswanathan K, Borczuk AC, Siddiqui MT. Orthopedia homeobox protein (OTP) is a sensitive and specific marker for primary pulmonary carcinoid tumors in cytologic and surgical specimens. *J Am Soc Cytopathol*. 2019;8:39-46.
15. Alcala N, Leblay N, Gabriel A, et al. Integrative and comparative genomic analyses identify clinically relevant pulmonary carcinoid groups and unveil the supra-carcinoids. *Nat Commun*. 2019;10:1-21.
16. Gabriel AA, Mathian E, Mangiante L, et al. A molecular map of lung neuroendocrine neoplasms. *GigaScience*. 2020;9:giaa112.
17. Fernandez-Cuesta L, Peifer M, Lu X, et al. Frequent mutations in chromatin-remodelling genes in pulmonary carcinoids. *Nat Commun*. 2014;5:1-7.
18. Peifer M, Fernández-Cuesta L, Sos ML, et al. Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer. *Nat Genet*. 2012;44:1104-1110.
19. George J, Walter V, Peifer M, et al. Integrative genomic profiling of large-cell neuroendocrine carcinomas reveals distinct subtypes of high-grade neuroendocrine lung tumors. *Nat Commun*. 2018;9:1-13.
20. Balanis NG, Sheu KM, Esedebe FN, et al. Pan-cancer convergence to a small-cell neuroendocrine phenotype that shares susceptibilities with hematological malignancies. *Cancer Cell*. 2019;36:17-34.e7.
21. Kiefer J, Nasser S, Graf J, et al. A systematic approach toward gene annotation of the hallmarks of cancer. Paper presented at: AACR Annual Meeting 2017; April 1-5, 2017; Washington, DC; 2017.
22. Kent WJ, Sugnet CW, Furey TS, et al. The human genome browser at UCSC. *Genome Res*. 2002;12:996-1006.
23. Laddha SV, Da Silva EM, Robzyk K, et al. Integrative genomic characterization identifies molecular subtypes of lung carcinoids. *Cancer Res*. 2019;79:4339-4347.
24. Swarts DR, Scarpa A, Corbo V, et al. MEN1 gene mutation and reduced expression are associated with poor prognosis in pulmonary carcinoids. *J Clin Endocrinol Metabol*. 2014;99:E374-E378.
25. Simbolo M, Mafficini A, Sikora KO, et al. Lung neuroendocrine tumours: deep sequencing of the four World Health Organization histotypes reveals chromatin-remodelling genes as major players and a prognostic role for TERT, RB1, MEN1 and KMT2D. *J Pathol*. 2017;241:488-500.
26. Portela A, Esteller M. Epigenetic modifications and human disease. *Nat Biotechnol*. 2010;28:1057-1068.
27. Doi A, Park I-H, Wen B, et al. Differential methylation of tissue- and cancer-specific CpG Island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nat Genet*. 2009;41:1350-1353.
28. Irizarry RA, Ladd-Acosta C, Wen B, et al. The human colon cancer methylome shows similar hypo-and hypermethylation at conserved tissue-specific CpG Island shores. *Nat Genet*. 2009;41:178-186.
29. Van Vlodrop IJ, Niessen HE, Derks S, et al. Analysis of promoter CpG Island hypermethylation in cancer: location, location, location! *Clin Cancer Res*. 2011;17:4225-4231.
30. Derks JL, Leblay N, Lantuejoul S, Dingemans A-MC, Speel E-JM, Fernandez-Cuesta L. New insights into the molecular characteristics of pulmonary carcinoids and large cell neuroendocrine carcinomas, and the impact on their clinical management. *J Thorac Oncol*. 2018;13:752-766.
31. Volante M, Mete O, Pelosi G, Roden AC, Speel EJM, Uccella S. Molecular pathology of well-differentiated pulmonary and thymic neuroendocrine tumors: what do pathologists need to know? *Endocr Pathol*. 2021;32:1-15.
32. Baca SC, Prandi D, Lawrence MS, et al. Punctuated evolution of prostate cancer genomes. *Cell*. 2013;153:666-677.
33. Reuling EM, Naves DD, Thunnissen E, et al. A multimodal biomarker predicts dissemination of bronchial carcinoid. *medRxiv*. 2021. https://www.medrxiv.org/content/10.1101/2021.05.17.21257308v1
34. Pelosi G, Rindi G, Travis WD, Papotti M. Ki-67 antigen in lung neuroendocrine tumors: unraveling a role in clinical practice. *J Thorac Oncol*. 2014;9:273-284.
35. Rapa I, Votta A, Felice B, et al. Identification of microRNAs differentially expressed in lung carcinoid subtypes and progression. *Neuroendocrinology*. 2015;101:246-255.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Moonen L, Mangiante L, Leunissen DJG, et al. Differential Orthopedia Homeobox expression in pulmonary carcinoids is associated with changes in DNA methylation. *Int. J. Cancer*. 2022;150(12):1987-1997. doi:10.1002/ijc.33939