

BRIEF REPORT

Open Access



# A validated restriction enzyme ddPCR cg05575921 (*AHRR*) assay to accurately assess smoking exposure

Sandra Fitzgerald<sup>1,2</sup>, Basharat Bhat<sup>3</sup>, Cristin Print<sup>1,2</sup> and Gregory T. Jones<sup>3\*</sup>

## Abstract

**Background & Methods** In this study, a novel restriction enzyme (RE) digestion-based droplet digital polymerase chain reaction (ddPCR) assay was designed for cg05575921 within the *AHRR* gene body and compared with matching results obtained by bisulfite conversion (BIS) ddPCR and Illumina DNA methylation array.

**Results** The RE ddPCR cg05575921 assay appeared concordant with BIS ddPCR ( $r^2 = 0.94$ ,  $P < 0.0001$ ) and, when compared with the Illumina array, had significantly better smoking status classification performance for current versus never smoked (AUC 0.96 versus 0.93,  $P < 0.04$ ) and current versus ex-smoker (AUC 0.88 versus 0.83,  $P < 0.04$ ) comparisons.

**Conclusions** The RE ddPCR cg05575921 assay accurately predicts smoking status and could be a useful component of 'precision-medicine' chronic disease risk screening tools.

**Keywords** Aryl hydrocarbon receptor repressor (*AHRR*), Bisulfite conversion, DNA methylation, Droplet digital PCR, Epigenetics, Restriction enzyme, Smoking exposure

## Introduction

Numerous epigenome-wide association studies have reported robust and reproducible DNA methylation changes correlated with smoking exposure, with the strongest association consistently being cg05575921, within the gene body of *AHRR* [1, 2]. Sensitive PCR methods have subsequently been developed to assess the methylation status of cg05575921 [3–5]. These techniques typically utilize bisulfite conversion, which is a

harsh chemical process resulting in significant DNA degradation and can be associated with issues regarding conversion efficiency if not carefully conducted [6]. Alternative PCR assays have been reported which avoid the need for bisulfite conversion by utilizing methylation-sensitive restriction enzyme digestion. Unfortunately, the pattern of methylation levels generated does not always appear to closely match those typically associated with bisulfite conversion and DNA methylation array-based measures of this specific CpG site [7]. In this study, we aimed to develop a restriction enzyme digestion-based droplet digital PCR (ddPCR) assay for cg05575921 (*AHRR*) that produces comparable results with that of bisulfite conversion-based methods. We further confirm its sensitive and specific association with smoking status in a large clinically relevant elderly cohort.

\*Correspondence:

Gregory T. Jones

greg.jones@otago.ac.nz

<sup>1</sup> Department of Molecular Medicine & Pathology, The University of Auckland, Auckland, New Zealand

<sup>2</sup> Maurice Wilkins Centre of Research Excellence, Auckland, New Zealand

<sup>3</sup> Vascular Research Group, Department of Surgical Sciences, Dunedin Medical School, University of Otago, PO Box 56, Dunedin 9054, New Zealand



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Methods

This observational cohort study examined a subset of participants recruited as part of the New Zealand Vascular Genetics Research Programme. In brief, consecutively recruited participants of an abdominal aortic aneurysm screening project and a healthy age matched control group were examined [8]. Each participant was over 50 years of age and had self-reported smoking status (never, ex-smoker, current smoker), smoking pack years, age and biological sex data collected at time of recruitment. One pack year was defined as twenty cigarettes (or tobacco equivalent) smoked every day for one year and was calculated using the publicly available online tool <https://www.smokingpackyears.com>.

For the initial ddPCR setup and validation experiments (Additional file 1: Sheets S1 and S4), separate subgroups of never, ex and current smokers (with high pack year consumption) were purposefully selected (without age or sex matching) selected, but sample status was blinded during ddPCR assessment.

A detailed description of the DNA extraction and the two droplet digital PCR methods utilized is provided in Additional file 2. In brief, genomic DNA was extracted from EDTA-anticoagulated whole blood using a modified salting-out method and stored at  $-20^{\circ}\text{C}$  until analysis. DNA concentration was determined using a nanophotometer (Implen P360).

A custom ddPCR assay was designed to quantify the proportion of methylated copies of cg05575921. This assay utilized the methylation-sensitive nature of the restriction enzyme HpaII (RE ddPCR). HpaII digestion was performed on purified gDNA and the digested DNA used as template for the subsequent ddPCR. Quantification of the number of methylated cg05575921 CpG sites was performed using a FAM-labeled probe. To allow for normalization of the RE ddPCR assay, a HEX-labeled probe was designed to a region of the gene *KIT* that did not contain any HpaII cut sites. The per cent methylation of the cg05575921 *AHRR* CpG site was determined by dividing FAM (methylated cg05575921) by HEX (copies of *KIT*), each being calculated by Poisson distribution (cp/ $\mu\text{L}$ ) according to manufacturer's guidelines.

A second ddPCR assay, based on the bisulfite (BIS) ddPCR assay published by Arroyo et al. [3], was used for comparison with the RE ddPCR approach. The primers and probes were as previously published [3] with the exception of the forward primer which was redesigned in Primer3Plus [9] to achieve a smaller amplification product of 126bp. Neither primers nor probes contained polymorphic bases with global population minor allele frequencies greater than 0.1%. Bisulfite-converted (EZ DNA Methylation Kit (Zymo)) gDNA was used as template for subsequent ddPCR. In this assay, cg05575921

methylation was assessed using a methylated-site-specific FAM probe and an unmethylated-site-specific HEX probe. Because this assay assessed FAM and HEX signal, assessing the same amplicon quantification was performed using fraction abundance (per cent methylated = FAM divided by the total of FAM and HEX, each probe being measured as cp/ $\mu\text{L}$ ).

Illumina Infinium MethylationEPIC (v1) array data were generated as previously reported [8]. Raw idat files were processed using the GenomeStudio software (Illumina) and Bioconductor packages ChAMP [10] and minifi (v1.46) [11]. Several different iterations of normalization schema were applied to the array data, including Noob, SWAN, BMIQ and GenomeStudio [12]. Normalized beta-values for cg05575921 were then extracted for paired comparisons to ddPCR results generated using matching DNA samples. Results for all techniques were reported as a percentage of methylation for ease of comparison.

Statistical analysis was conducted by assessing paired correlations between assays using Fischer Z-transformation of the Pearson correlation coefficient. The agreement between assays was tested using Bland–Altman analysis. The smoking status classification performance of the assays was assessed by receiver operating characteristic (ROC) curve analysis, with the method of DeLong et al. [13] being used to calculate of the standard error of the area-under-the-curve (AUC). Pairwise comparisons were conducted between the respective AUC generated by the EPIC array and RE ddPCR assays. Optimized smoking status group association thresholds were calculated using the Youden J statistic and tabulated with their corresponding sensitivity and specificity measures. Confusion matrices were used to present the positive and negative predictive performance of the dichotomized threshold values. Statistical tests were performed using MedCalc version 19.2 (MedCalc Software, Ostend, Belgium).

## Results

The overall cohort examined in this study ( $n=1227$ , Table 1) had a mean age of 70 years, 71% were male, and 37, 51 and 12% were never, ex or current smokers, respectively.

It has been previously demonstrated that a bisulfite-converted DNA ddPCR assay can accurately determine the CpG methylation status of cg05575921 within *AHRR* [3, 4, 14]. Furthermore, strong concordance between BIS ddPCR and Illumina Epic array measures of cg05575921 has been reported [4, 14]. In this study we wanted to investigate whether a restriction digestion approach coupled with droplet digital PCR could also be used to accurately measure the level of methylation of the cg05575921 CpG site. Initial investigations were conducted on gDNA

**Table 1** Demographics of cohort comparing ddPCR and EPIC array cg05575921 values

	Total	Never smoked	Ex-smoker	Current smoker
Number	1227	450	630	147
Sex, <i>n</i> of males (%)	871 (71.0)	300 (66.7)	473 (75.1)	98 (66.9)
Age (years)	70.1 ± 8.2	70.8 ± 7.6	70.6 ± 8.2	65.6 ± 8.9
Smoking pack years	7 (0–30)	0	20 (8–38)	34 (21–48)
History of cardiovascular disease*	708 (57.7)	214 (47.6)	367 (58.3)	127 (86.4)
cg05575921, Illumina array % DNA methylation	78.3 ± 12.2	84.7 ± 7.3	77.6 ± 10.8	61.4 ± 12.8
cg05575921, RE ddPCR % DNA methylation	75.8 ± 16.2*	84.7 ± 8.3	75.4 ± 14.2*	50.5 ± 15.3*

Restriction enzyme (RE) digestion ddPCR cg05575921 (*AHRR*) values were significantly lower compared to array derived values (\*paired *t* test,  $P < 0.0001$ ) in both ex and current smokers. There was no significant difference ( $p = 0.28$ ) between values in never smokers. Cardiovascular disease included documented medical history of ischemic stroke or transient ischemic attack, coronary heart disease or peripheral artery disease [8]. Continuous measures are represented as either mean ± 1 standard deviation or median and interquartile range

isolated from 52 cohort participants (23 current smokers, 6 ex-smokers and 23 never smokers) which were assessed using both the cg05575921 RE and BIS ddPCR assays. There was a strong correlation between measures ( $r^2 = 0.94$ ,  $P < 0.0001$ ), with an overall coefficient of variance between the two assay types of 5.2% (Additional file 1: Sheet S1). To investigate the quantitative consistency of the restriction enzyme approach, the RE ddPCR assay was repeated for 16 separate samples (2–6 replicates), as well as a technical replicate included in each 96-well batch. The RE ddPCR assay had good technical reproducibility in the 16 individual replicates, with a mean coefficient of variation of 6.0% (Additional file 1: Sheet S1). The plate batch replicate sample had a coefficient of variation of 3.2% (mean methylation  $91.4 \pm 2.9\%$ ,  $n = 12$  batch replicates).

Because the RE ddPCR assay involved the amplification of two separate PCR products (*AHRR* and *KIT* amplicons), we performed an additional analysis to confirm that there was no PCR amplification bias between the two primer/probes sets. This involved comparison of FAM and HEX copies/ $\mu\text{L}$  in ddPCR reactions which had not undergone restriction enzyme digest. (as described in the supplementary methods, with results in Additional file 1: Sheet S4). There was no strong bias in amplification between the two PCR reactions, and therefore, inclusion of no enzyme normalization made no significant difference to the smoking status predictive performance of the assay (Additional file 1: Sheet S4 Table 2).

We then examined the effect of different normalization schema on Illumina Infinium MethylationEPIC (v1) array derived cg05575921 values. Although there did appear to be a subtle influence on cg05575921 values across smoking status groups (Additional file 1: Sheet S2), this did not significantly (all pairwise ROC-curve comparisons  $P > 0.05$ ) influence the smoking status classification performance of array derived values and, given this equipoise and its wide

use in the literature, subsequent results are reported using Illumina Genome Studio normalization.

Next, we examined the full set of 1227 samples, matching cg05575921 RE ddPCR results with their respective Illumina EPIC array data. The comparative quantitation of DNA methylation at cg05575921 is shown in Table 1 and Additional file 1: Sheet S3.

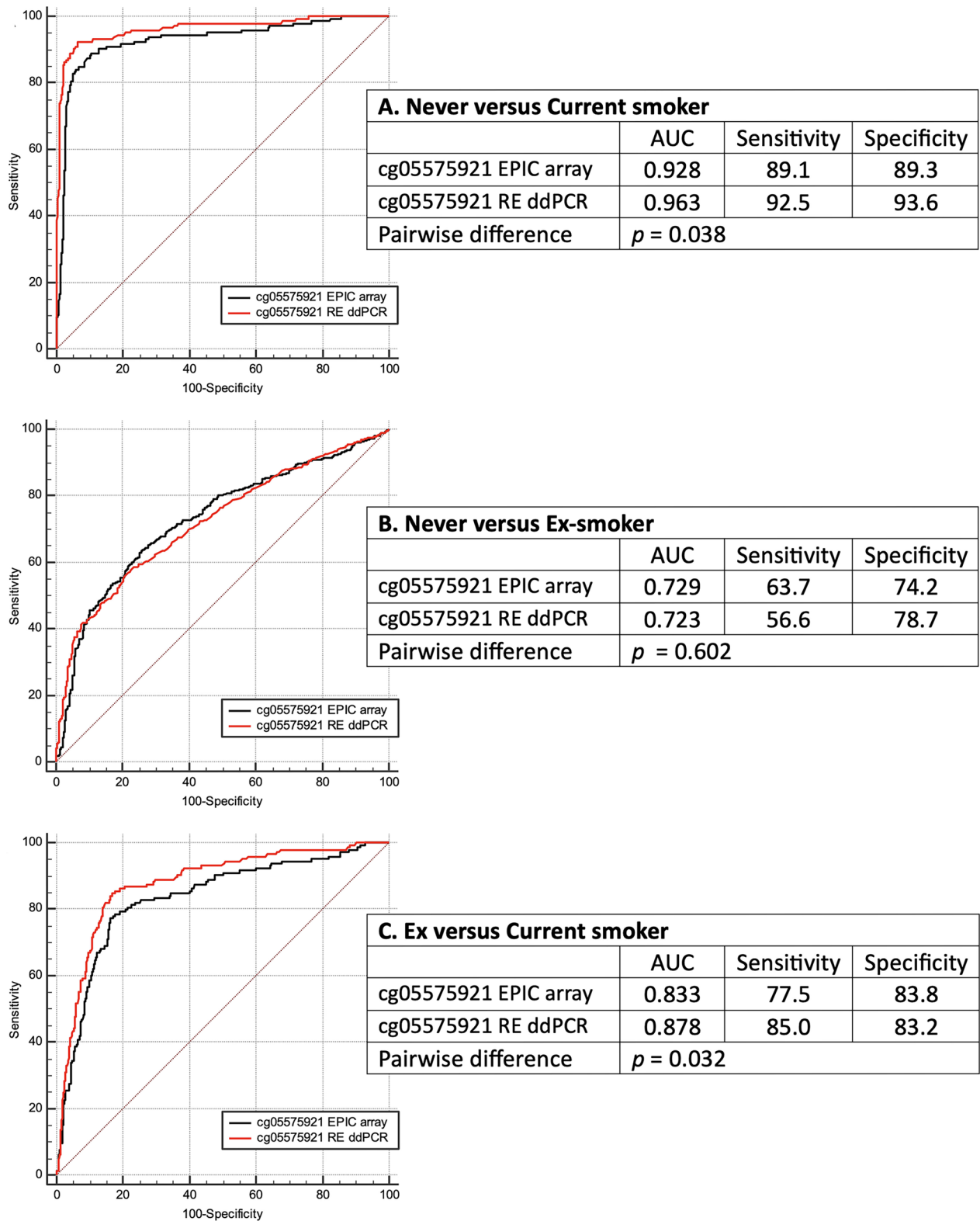
Within the never smoked group, there was no statistically significant difference ( $p = 0.28$ ) between RE ddPCR and Illumina array cg05575921 methylation values. However, the RE ddPCR values were significantly lower compared to their matching Illumina array values in both the ex ( $t$ -value 4.7,  $P < 0.0001$ ) and current ( $t$ -value 8.5,  $P < 0.0001$ ) smoker groups, with this difference being most apparent in current smokers (Table 1 and Additional file 1: Sheet S3B).

Bland–Altman analysis also demonstrated that disagreement between paired test measures was not uniform (Additional file 1: Sheet S3C), with good agreement in values greater than 80% methylation (the majority of whom were never smokers), but increasingly poorer agreement within values 65–80% (mostly ex-smokers) and less than 65% (mostly current smokers). This was also evident in the nonlinear correlation between measures (Additional file 1: Sheet S3D).

Finally, the smoking status classification performance of cg05575921 RE ddPCR was compared with that of EPIC array using ROC-curve analysis. The RE ddPCR values had statistically better predictive performance, as assessed by pairwise differences in AUC, as well as positive and negative predictive values (confusion matrices in Additional file 1: Sheet S3 Table 2) for the current versus never smoked (Fig. 1A) and current versus ex-smoker (Fig. 1C) comparisons.

## Discussion

Smoking has long been recognized as a leading risk factor for a wide range of chronic conditions, including cardiovascular diseases and cancers [15]. Current



**Fig. 1** Relative sensitivity and specificity of cg05575921 assays detection of smoking status. DNA methylation values for cg05575921 derived from Illumina EPIC assays (GS normalised) and RE ddPCR were significantly predictive of smoking status (all AUCs  $P < 0.0001$ ), however, the RE ddPCR assay was significantly more sensitive and specific in differentiating between **A** never and current smokers and **C** Ex and current smokers. **B** There was no significant difference in the predictive performance of either assay to distinguish between never and ex-smokers. Confusion matrices, tabulating the positive and negative predictive performances, are shown in Additional file 1: Sheet S3, Table 2.



clinical practice incorporates patient self-reported measures of smoking exposure, typically by assessing smoking status (never, ex or current smoker). Limitations and potential inaccuracies of such self-reported data have been suggested. For example, there is likely to be a wide range of smoking-related risk among ex-smokers depending on their volume of exposure (pack years) and duration of cessation. Similarly, second-hand exposure is not readily quantified with such measures.

Epigenetic markers, such as DNA methylation, are emerging as ‘precision medicine’ biomarkers that can surpass the performance of such ‘conventional’ markers in risk prediction [16]. This is particularly true for smoking, for which numerous epigenome-wide association studies have demonstrated robust and reproducible DNA methylation changes correlated with smoking exposure, with the strongest association consistently being cg05575921, within the gene body of *AHRR* [1, 2]. Consequently, sensitive PCR methods have been developed to assess the methylation status of this specific CpG site [5, 17]. The use of cg05575921 (*AHRR*) assays have been suggested to have the potential to improve risk prediction utility by a range of effects including assessment of second-hand smoke exposure [18], correcting for false or inaccurate self-reported measures [19] and by virtue of strong morbidity and mortality associations, even after adjusting for self-reported smoking history [5].

In this study, we report a technically accurate restriction enzyme digestion ddPCR assay, which appears to have similar performance to that of previously reported ddPCR assays [3, 4] without the need for the chemically harsh bisulfite conversion process [6], which can be an issue particularly when working with samples containing limited amounts of input DNA. However, previous methylation-sensitive restriction enzyme-based techniques do provide some reason for caution regarding this approach. Bravo-Gutierrez and colleagues reported a *HpaII* restriction enzyme digest method, but using qPCR rather than ddPCR as the detection platform. Unfortunately, the methylation patterns reported for the Bravo-Gutierrez assay do not match those typically associated with bisulfite conversion PCR and Illumina array-based measures of this specific CpG site [7]. Specifically, there was a substantial overlap in cg05575921 values when comparing smoking and non-smoking groups, and many current smokers were reported as having 100% methylation. This is contrary to the well-established literature, which consistently reports cg05575921 methylation values of 50–60% in the vast majority of current smokers [2, 17]. Nevertheless, other groups have convincingly demonstrated that some RE-based ddPCR assays can be highly sensitive and specific [20]. In addition, RE-based

approaches are substantially less time-consuming than that of bisulfite conversion (30–60 min versus 12–16 h).

In contrast to the qPCR-based observations of the Bravo-Gutierrez assay, the RE ddPCR assay developed in this study produced results that are consistent with the prior literature [2, 5, 17]. Regardless of which of the three (RE or BIS ddPCR or Illumina array) assessing techniques used, we observed that the hypomethylation of the cg05575921 DNA methylation site was statistically significantly, sensitively and specifically associated with smoking status. Interestingly, using a large cohort of 1227 individuals, we were also able to confirm that the RE ddPCR assay had significantly better predictive power for smoking status than measures derived from Illumina arrays. We do, however, note a slight potential difference with a previous report by Philibert et al. [4], who showed a strong and linear ( $R^2=0.98$ ) correlation between EPIC array and BIS ddPCR cg05575921 results (in 92 sample comparisons) which did not match the nonlinear correlation in our current RE ddPCR versus EPIC array comparison (Additional file 1: Sheet S3). The quantification of the BIS and RE ddPCR differed, with fractional abundance assessment being utilized in the (single amplicon) BIS ddPCR assay and, in contrast, the RE ddPCR method involved comparison of two separate amplicons (*AHRR* and *KIT*), therefore necessitating comparison via a simpler ratio of FAM/ HEX signal. Nevertheless, we still observed a strong linear correlation between the two (BIS and RE) ddPCR methods, and the reason for the nonlinear effect observed between the RE ddPCR and EPIC array paired comparisons remains unclear. One possible explanation that we considered was the effect of array normalization. This is an essential step in the analysis of array data that is needed to reduce the size of batch effects, correct for probe type bias and adjust for background noise and dye bias [12, 21]. It is worth noting that ddPCR assays do not require such correction, as they are internally calibrated, in the case of the restriction enzyme assay to the number of copies of a reference gene (*KIT*). While we did observe subtle differences due to the normalization method that was applied to the array data (Additional file 1: Sheet S2), the lower RE ddPCR values observed in current smokers (Additional file 1: Sheet S3B) was a much larger effect and an explanation for this particular observation therefore remains elusive. Nevertheless, an apparent consequence of this ‘skewing’ of the hypomethylated signal appeared to be associated with a more sensitive and specific smoking group classification performance for the RE ddPCR assay compared to array derived measures. Indeed, as shown in Additional file 1: Sheet S3A, we noted several (10 of 147, 6.8%) known current smokers that, while having

RE ddPCR values (below 60%) consistent with their smoking status, had Illumina array values (over 80%) which were more consistent with never smoked status.

It has also been suggested that normalization, using a separate no RE reaction, may be required for RE ddPCR assays [20]. In a sub-analysis, we demonstrated that inclusion of a no enzyme normalization step did not significantly alter the high smoking status sensitivity and specificity of this particular DNA methylation ddPCR assay (Additional file 1: Sheet S4).

In conclusion, we report the development of a restriction enzyme digestion-based droplet digital PCR assay for cg05575921 (*AHRR*) that produces at least comparable results to that of previously reported ddPCR and Illumina array-based methods without the need for bisulfite conversion of the DNA sample. We further confirm that this RE ddPCR method is able to sensitively and specifically predict smoking status in a large clinically relevant elderly cohort. This bisulfite conversion-free ddPCR assay appears to have potential clinical utility as a precise molecular diagnostic method for the evaluation of smoking exposure.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13148-024-01659-1>.

**Additional file 1.** Supplementary file 2: Comparison of cg05575921 restriction enzyme versus bisulfite conversion ddPCR assays. Supplementary file 3: Effect of Illumina array normalisation on cg05575921 values. Supplementary file 4: Comparison of cg05575921 restriction enzyme ddPCR versus Illumina EPIC array values. Supplementary file 5: Effect of inclusion of a no enzyme normalisation to the cg05575921 RE ddPCR assay.

**Additional file 2.** Supplementary file 1: ddPCR methods.

## Acknowledgements

The authors wish to thank Mrs Vicky Phillips for her technical assistance.

## Author contributions

SF and GJ write the main manuscript text. SF designed the ddPCR assay. GJ performed the statistical analysis. BB analyzed the Illumina dataset, and CP was involved with ddPCR design and analysis. All authors reviewed the manuscript.

## Funding

This work was supported by grants from the Healthier Lives, National Science Challenge, the Health Research Council of New Zealand (14/155, 17/402, 20/144) and Genomics Aotearoa (a New Zealand Ministry of Business, Innovation and Employment funded research platform). SF is supported by the Cancer Research Trust John Gavin Fellowship.

## Availability of data and materials

Genome-wide datasets analyzed during the current study are not publicly available due to ethics approval restrictions. The ddPCR datasets analyzed during the current study are available from the corresponding author on reasonable request.

## Declarations

### Ethics approval and consent to participate

All participants gave written informed consent, and the study was approved by the National Health and Disability Ethics Committees of New Zealand (16/NTB/219 & URA/11/11/072).

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

Received: 16 January 2024 Accepted: 15 March 2024

Published online: 25 March 2024

## References

- Elliott HR, Tillin T, McArdle WL, Ho K, Duggirala A, Frayling TM, et al. Differences in smoking associated DNA methylation patterns in South Asians and Europeans. *Clin Epigenetics*. 2014;6(1):4.
- Gao X, Jia M, Zhang Y, Breitling LP, Brenner H. DNA methylation changes of whole blood cells in response to active smoking exposure in adults: a systematic review of DNA methylation studies. *Clin Epigenetics*. 2015;7:113.
- Arroyo K, Nargizyan A, Andrade FG, Myint SS, Lu S, Pandey P, et al. Development of a Droplet Digital PCR DNA methylation detection and quantification assay of prenatal tobacco exposure. *Biotechniques*. 2022;72(4):121–33.
- Philibert R, Dogan M, Noel A, Miller S, Krukow B, Papworth E, et al. Dose response and prediction characteristics of a methylation sensitive digital PCR assay for cigarette consumption in adults. *Front Genet*. 2018;9:137.
- Bojesen SE, Timpson N, Relton C, Davey Smith G, Nordestgaard BG. *AHRR* (cg05575921) hypomethylation marks smoking behaviour, morbidity and mortality. *Thorax*. 2017;72(7):646–53.
- Hong SR, Shin KJ. Bisulfite-converted DNA quantity evaluation: a multiplex quantitative real-time PCR system for evaluation of bisulfite conversion. *Front Genet*. 2021;12: 618955.
- Bravo-Gutierrez OA, Falfan-Valencia R, Ramirez-Venegas A, Sansores RH, Hernandez-Zenteno RJ, Hernandez-Perez A, et al. Hypomethylation of *AHRR* (cg05575921) is related to smoking status in the Mexican Mestizo population. *Genes*. 2021;12(8):1276.
- Jones GT, Marsman J, Bhat B, Phillips VL, Chatterjee A, Rodger EJ, et al. DNA methylation profiling identifies a high effect genetic variant for lipoprotein(a) levels. *Epigenetics*. 2020;15(9):949–58.
- Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen JA. Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Res*. 2007;35(Web server issue):W71–4.
- Tian Y, Morris TJ, Webster AP, Yang Z, Beck S, Feber A, Teschendorff AE. ChAMP: updated methylation analysis pipeline for Illumina BeadChips. *Bioinformatics*. 2017;33(24):3982–4.
- Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, Irizarry RA. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*. 2014;30(10):1363–9.
- Bhat B, Jones GT. Data analysis of DNA methylation epigenome-wide association studies (EWAS): a guide to the principles of best practice. *Methods Mol Biol*. 2022;2458:23–45.
- DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a non-parametric approach. *Biometrics*. 1988;44(3):837–45.
- Takeuchi F, Takano K, Yamamoto M, Isono M, Miyake W, Mori K, et al. Clinical implication of smoking-related aryl-hydrocarbon receptor repressor (*AHRR*) hypomethylation in Japanese adults. *Circ J*. 2022;86(6):986–92.
- Bartecchi CE, MacKenzie TD, Schrier RW. The human costs of tobacco use. *N Engl J Med*. 1994;330(13):907–12.

16. Zhong J, Agha G, Baccarelli AA. The role of DNA methylation in cardiovascular risk and disease: methodological aspects, study design, and data analysis for epidemiological studies. *Circ Res.* 2016;118(1):119–31.
17. Philibert R, Dogan M, Beach SRH, Mills JA, Long JD. AHRR methylation predicts smoking status and smoking intensity in both saliva and blood DNA. *Am J Med Genet B Neuropsychiatr Genet.* 2020;183(1):51–60.
18. Reynolds LM, Magid HS, Chi GC, Lohman K, Barr RG, Kaufman JD, et al. Secondhand tobacco smoke exposure associations with DNA methylation of the aryl hydrocarbon receptor repressor. *Nicotine Tob Res.* 2017;19(4):442–51.
19. Andersen AM, Philibert RA, Gibbons FX, Simons RL, Long J. Accuracy and utility of an epigenetic biomarker for smoking in populations with varying rates of false self-report. *Am J Med Genet B Neuropsychiatr Genet.* 2017;174(6):641–50.
20. van de Leemkolk FEM, Nell RJ, Versluis M, de Koning EJP, Huurman VAL, Alwayn IPJ, et al. Quantification of unmethylated insulin DNA using methylation sensitive restriction enzyme digital polymerase chain reaction. *Transpl Int.* 2022;35:10167.
21. Welsh H, Batalha C, Li W, Mpye KL, Souza-Pinto NC, Naslavsky MS, Parra EJ. A systematic evaluation of normalization methods and probe replicability using Infinium EPIC methylation data. *Clin Epigenetics.* 2023;15(1):41.

### **Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.