## **METHODOLOGY**



# qBiCo: a method to assess global DNA conversion performance in epigenetics via single-copy genes and repetitive elements



Faidra Karkala<sup>1</sup>, Roy B. Simons<sup>1</sup>, Floor Claessens<sup>1</sup>, Vivian Kalamara<sup>1</sup>, Manfred Kayser<sup>1</sup> and Athina Vidaki<sup>1\*</sup>

## Abstract

**Background** Human DNA methylation profiling offers great promises in various biomedical applications, including ageing, cancer and even forensics. So far, most DNA methylation techniques are based on a chemical process called sodium bisulfite conversion, which specifically converts non-methylated cytosines into uracils. However, despite the popularity of this approach, it is known to cause DNA fragmentation and loss affecting standardization, while incomplete conversion may result in potential misinterpretation of methylation-based outcomes.

**Results** To offer the community a solution, we developed qBiCo - a novel quality-control method to address the quantity and quality of bisulfite-converted DNA. qBiCo is a 5-plex, TaqMan<sup>®</sup> probe-based, quantitative (q)PCR assay that amplifies single- and multi-copy DNA fragments of converted and non-converted nature. It estimates four parameters: converted DNA concentration, fragmentation, global conversion efficiency, and potential PCR inhibition. We optimized qBiCo using synthetic DNA standards and assessed it using standard developmental validation criteria, showcasing that qBiCo is reliable, robust and sensitive down to picogram level. We also evaluated its performance by testing decreasing DNA amounts using several commercial bisulfite conversion efficiencies from 78 to 99.9%, while certain kits highly fragment DNA, demonstrating large variability in their performance. Towards building a prototype tool, we further optimized key functionalities, for example, by replacing the poorest performing single-plex assay and creating a more representative DNA standard. Aiming to scale-up and move towards implementation, we successfully transferred and validated our novel method in six different qPCR platforms from different major manufacturers.

**Conclusions** Overall, with the present study, we offer researchers in the epigenetic field a novel long-awaited QC tool that for the first time allows them to measure key quality and quantity parameters of the most popular DNA conversion process. The tool also enables standardization to prevent inconsistent data and false outcomes in the future, regardless of the downstream experimental analysis of DNA methylation-based research and applications across different fields of biology and biomedicine.

**Keywords** DNA methylation, Bisulfite conversion, Epigenetics, qPCR, Conversion efficiency, Quality control, Standardization

\*Correspondence: Athina Vidaki athina.vidaki@mumc.nl <sup>1</sup>Department of Genetic Identification, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands



© The Author(s) 2025. **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/.

#### Background

Epigenetics involves heritable modifications in gene expression that are not depicted by changes in DNA sequence, but rather appear as base and histone modifications [1]. DNA methylation is the most studied epigenetic modification, that chemically affects DNA by covalently adding a methyl group to the 5' carbon of cytosine. In humans, the majority of DNA methylation appears on cytosine-guanine dinucleotides (5'-CpG-3') [2]. As response to internal and external stimuli [3], DNA methylation regulates gene transcription by usually silencing a gene, while a non-methylated gene remains active and accessible to the transcription machinery [4]. Profusion of emerging evidence indicates the importance of DNA methylation as a molecular biomarker for detecting and monitoring phenotypes associated to lifestyle choices [5], ageing and age-related diseases [6] including cancer [7, 8] and cardiovascular disorders [9] as well as in neurodevelopmental disorders [10]. Furthermore, DNA methylation profiling can also be useful in more specialized applications, such as forensics [5, 11, 12].

To analyse genome-wide or targeted methylation, several high-throughput methods have been developed depending on the research question and downstream analysis. However, the vast majority require chemical treatment of DNA via the so-called bisulfite conversion (BC), a process where during DNA treatment with sodium bisulfite (NaHSO<sub>3</sub>), non-methylated cytosines are converted to uracils, whereas methylated cytosines remain unchanged [13]. Hence, follow-up PCR amplification of bisulfite-converted DNA (BC-DNA) replaces uracils with thymines, allowing the detection of methylation as an introduced C/T variant using standard downstream technologies, including microarrays, pyrosequencing and massively parallel sequencing [14]. Thus, due to its simplicity, cost-effectiveness and compatibility with various downstream genetic pipelines, BC has dominated the field over the last decades over enzymatic or immunoprecipitation approaches and is currently considered as the gold standard to study DNA methylation [14]. Additionally, it is the preferred method as it enables the analysis of any CpG of interest, without restrictions to specific enzyme recognition motifs and required enzyme-specific optimization [15].

Despite its popularity, however, BC requires aggressive chemical conditions such as low pH and elevated temperature, which can affect the integrity of DNA, resulting in damage and fragmentation. At the same time, choosing less aggressive conditions can affect the success and efficiency of conversion, resulting in follow-up overestimation of methylation and data misinterpretation [16]. When BC-DNA is too fragmented or 'contaminated' with carryover chemicals from the treatment, its detection and interpretation is also affected. Specifically, DNA incubation with sodium bisulfite results in depyrimidination and generation of abasic sites that are prone to DNA strand breaks, especially under heating or alkaline conditions [17]. Increased DNA fragmentation decreases the number of DNA molecules available for PCR amplification, longer targets are less likely to amplify, and eventually, the PCR results are no longer reproducible due to the large variability driven by stochastic events early during PCR [18].

Nowadays, epigenetic researchers perform BC via commercially available kits from various manufacturers, each coming with a slightly different treatment protocol. Generally, according to the manufacturers statements, commercial BC kits promise > 99% DNA conversion rate, at least 80% recovery rate of BC-DNA, and eluted DNA fragments of up to 2,000 bp in length. Thus far, researchers have attempted to empirically evaluate these parameters using small-scale and/or non-specific approaches. Specifically, the quantity of BC-DNA is usually assessed via Nanodrop spectrophotometer (Thermo Fisher Scientific - TFS), Qubit fluorometer (TFS), or Bioanalyzer (Agilent), while the degree of fragmentation of BC-DNA is via Bioanalyzer or agarose gel electrophoresis. When comparing the performance of different commercial BC kits using these methods, researchers concluded that key quantity and quality parameters strongly depend on the initial input amount of genomic DNA (gDNA) and the incubation time during BC [19-25]. Importantly, none of these methods is specific to BC-DNA, which is short, single/double-stranded hybrid, and contains 'contaminating' compounds and secondary structures that affect downstream analysis. Overall, lacking quality- and quantity-related indications does not only affect the accuracy, reproducibility and robustness of DNA methylation detection, but also the evaluation of epigenetic-based outcomes [23].

In an early effort, Ehrich et al. [24] developed multiple single-locus PCR assays for assessing post-conversion DNA quality in terms of fragmentation under increasing incubation temperature during bisulfite treatment. While this study showcased the impact of elevated temperature (>65 °C) on degrading BC-DNA fragments (<500 bp), no indication on BC efficiency or BC-DNA recovery could be devised. On the other hand, Sriraska et al. [25] aimed to verify the BC rate using a set of genomic and converted-specific primers, with incomplete conversion being detected when lower DNA amounts were used for conversion (<37.5 ng). However, their approach was based only on a single locus, therefore highly biased and unable to provide with a global DNA conversion measure. More recently, Kint et al. (2018) used a combination of agarose gel electrophoresis, quantitative PCR and digital PCR to measure BC-DNA fragmentation, spectroscopic measurements and digital PCR to evaluate BC-DNA

recovery, as well as sequencing to assess BC efficiency [19]. While this study revealed large performance differences among BC kits (for example, recovery ranging from 26.6 to 88.3%), their multi-level approach is not easy and practical to apply as a simple QC tool in experimental pipelines, as it also consumes large amounts of the oftenprecious BC-DNA sample.

So far, no BC-DNA-specific method exists that can thoroughly assess the BC process from both the qualitative and quantitative side, leading researchers to simply assume successful performance as stated per manufacturer to sell their products. Overall, the epigenetics community has largely overestimated the method's capabilities despite the warnings of the inventor of BC [26], Hikoya Hayatsu, who in 2008 wrote: "I now feel that the presently available data for DNA methylation need carefully be looked at, and that a scientifically sound, assured methodology free from any false positive or false negative assignments be established as soon as possible" [27].

In the present study, to offer the epigenetic community with a solution, we developed a novel QC method to simultaneously address the quantity and quality of BC-DNA (qBiCo). It is a 5-plex, TaqMan<sup>°</sup> probe-based quantitative (q)PCR assay, which amplifies single- and multi-copy fragments of converted and non-converted nature to estimate a sample's: (i) BC-DNA concentration, (ii) global BC efficiency, (iii) BC-DNA fragmentation and (iv) potential PCR inhibition. Here, we describe the initial development of the qBiCo assay, demonstrate its

 Table 1
 Information on bisulfite conversion (BC) kits used during the testing of gBiCo-v1 method

Kit No	Manufacturer	Kit name	Incubation	
			protocol	
1	Zymo Research	EZ DNA Methylation	14–16 h at 50 ℃	
2	Thermo Fisher Scientific	EpiJET BC	10 min at 98 °C / 150 min at 60 °C	
3	Diagenode	Premium Bisulfite	8 min at 98 ℃ / 60 min at 54 ℃	
4	Promega	MethylEdge® BC System	8 min at 98 ℃ / 60 min at 54 ℃	
5	Sigma Aldrich	Imprint® DNA Modification	90 min at 65 °C	
6	Qiagen	EpiTect Bisulfite	5 min at 95 °C / 25 min at 60 °C / 5 min at 95 °C / 85 min at 60 °C / 5 min at 95 °C / 175 min at 60 °C	
7	Abcam	Fast BC	20 min at 95 °C	
8	Analytik Jena	innuCONVERT	45 min at 85 °C	
9	Epigentek	Methylamp DNA modification	90 min at 65 °C	
10	Active Motif	BC	30 s at 95 ℃ / 20 min at 58 ℃	

applicability by evaluating several BC kits and describe our strategy moving towards a small-scale qBiCo technology prototype, which we validated using multiple qPCR platforms (technology readiness levels, TRL 1–4).

#### Methods

#### DNA sample preparation

During the different phases of technology development, one or more commercially available DNA samples were used: the human high methylated gDNA standard (100ng/µl) (EpigenDX, USA) (qBiCo-v1 validation and BC kit testing), the human Methylated DNA (250 ng/ µl) (Zymo Research, USA) (qBiCo-v1 BC kit testing and qBiCo-v2 functionality improvement) and the Quantifiler<sup>™</sup> THP DNA Standard (100 ng/µl) (TFS, USA) (qBiCo-v1 BC kit testing). For the final phase of instrument-specific validation (qBiCo-v2), two commercially available biobank human blood samples (Sanquin, The Netherlands) were employed, together with a range of commercially available and already extracted non-human DNA samples from different species (Novagen, USA): Felis catus (cat), Canis lupus familiaris (dog), Mus musculus (mouse), Rattus (rat), Gallus gallus domesticus (chicken), Sus scrofa domesticus (pig), Bos Taurus (cow), and Rhesus macaque (monkey). gDNA isolation was performed using the QIAamp® DNA Investigator kit (Qiagen, Germany), following the whole blood protocol. All human DNA samples included in this study were quantified prior to BC to assess for their DNA quantity and quality, using the Quantifiler<sup>™</sup> DUO DNA quantification kit (TFS).

#### **BC-DNA** sample preparation

For qBiCo-v1 validation, 200 ng of the human high methylated gDNA (100 ng/µl) (EpigenDX, USA) were bisulfite-converted using the EZ DNA methylation kit (Zymo Research) using the manufacturer's instructions. During qBiCo-v1 testing, all three DNA standards mentioned above underwent BC in five different DNA amount – 200, 100, 50, 10, 1 ng using a total of ten commercially available BC kits to compare their performance (Table 1). The selected BC kits are popular and belong to ten different companies. All kits are optimized to convert DNA within the range of 100 ng – 2 µg, with a suggested optimal DNA amount set at 200 ng. BC protocols and incubation times were followed according to manufacturer's instructions and BC-DNA samples were eluted in 10 µl.

During functionality improvement towards qBiCo-v2, we prepared and employed a set of 'master' samples that realistically resemble BC-DNA quantity/quality, often seen in real-world analysis. For this, and to capture all possible variation across commercially available products, we used 16 different kits according to manufacturer's instructions: EpiTect Bisulfite kit (Qiagen), EpiTect

Fast bisulfite kit (Qiagen), EZ DNA Methylation kit (Zymo Research), EZ DNA Methylation Gold kit (Zymo Research), BC kit (Active motif, Belgium), Methylamp DNA modification kit (Epigentek, USA), EpiJET BC kit (TFS), MethylCode<sup>™</sup> BC kit (TFS), Premium Bisulfite kit (Diagenode, USA), Fast BC kit (Abcam, England), Epi-Mark<sup>®</sup> BC kit (New England Biolabs, USA), DNA BC kit I (Biovision, England), Imprint® DNA Modification kit (Sigma-Aldrich), CpGenome Direct Prep Bisulfite Modification kit (Merck Millipore, USA), innuCONVERT (Analytik Jena, Germany), MethylEdge® BC System (Promega, USA). To prepare the master samples, the Human Methylated DNA (250 ng/µl) (Zymo Research) was converted using different DNA inputs (200, 50, 20, 5 and 2 ng) per kit and eluted in 50  $\mu$ l. Then, for each amount, 30 µl of each BC-DNA sample was combined, to form a 480 μl master solution.

For the final phase of instrument-specific validation (qBiCo-v2), the whole blood DNA samples were bisulfite-converted using EZ DNA Methylation kit (Zymo Research). Each sample was converted 20 times using 50 ng and three times using 5 ng of DNA input. For all conversions, BC-DNA was eluted in 10 µl. To ensure the same quantity and quality across instruments, eluates of the same DNA input (e.g., 50 ng conversions) were pooled together in one tube and subsequently aliquoted to separate tubes. These samples were used to assess repeatability, sensitivity, and reproducibility. For sensitivity, the measured BC-DNA samples of 5 ng/µl were serially diluted (2X) up to nine times. For robustness, 5 ng/ µl of BC-DNA was UV treated for different time periods, while for specificity, 200 ng of the non-human samples described above were bisulfite-converted.

#### **BC-DNA evaluation using existing methods**

Two BC-DNA samples (initial gDNA input of 200 and 12.5 ng) were evaluated using other existing spectrophotometric and fluorometric methods often employed for BC-DNA quantity/quality assessment. For all reactions 1 µl was used, while for all methods reactions were performed in duplicate. First, we employed a NanoDrop Microvolume Spectrophotometer (TFS) to measure absorbance at 260 nm using the single-stranded DNA mode option. Next, we employed a 2100 Bioanalyzer instrument (Agilent Technologies, USA) using the RNA 6000 Pico kit (Agilent Technologies). Finally, we used the Qubit<sup>™</sup> 4 fluorometer (TFS) using the Qubit<sup>™</sup> singlestranded DNA Assay kit (TFS). Kit choices were driven by the fact that strands have lost their complementarity making BC-DNA mostly single-stranded and resembling the secondary structure of RNA. For all methods we used the manufacturer's instructions to calculate the BC-DNA concentration.

#### Synthetic DNA standard

Normally, for the quantification of human gDNA a commercially available human DNA standard is used. However, for the quantification of BC-DNA, this is not possible given that we cannot control neither the quantity nor the quality following BC. Hence, we opted for creating a synthetic DNA standard that resembles the composition of human BC-DNA. For this, we designed five artificial double-stranded gene fragments (gBlocks<sup>™</sup>, Integrated DNA Technologies, Coralville, IA, United States) with DNA sequences identical to the expected generated PCR product sequences when using the proposed assays in each qBiCo version. In qBiCo-v2, additional sequences were added on both the 5' and 3' ends of the gBlocks<sup>™</sup> to improve production and stability (Additional file 1: Table S4).

During qBiCo method formulation (qBiCo-v1), the four different gBlocks<sup>\*\*</sup> (except for IPC) were mixed in a ratio that resembles the one naturally found on human DNA based on our design. In other words, both the *hTERT* Short / *hTERT* Long assays as well as the *LINE1* Genomic / *LINE1* Converted assays were each mixed in a 1:1 ratio, while the single- / multi-copy elements were mixed in ~1 : 200 (GRCh37). Practically, the individual gBlocks<sup>\*\*</sup> were received in concentration of 10 ng/µl, and then mixed to resemble the expected copy numbers as in 50 ng/µl of human BC-DNA. Subsequently, serial dilutions by factor two were performed to generate the eight standards used in qPCR, with concentration range from 50 to 0.39 ng/µl.

During qBiCo prototype development (qBiCo-v2), we decided to simplify and improve the quantification process by directly working with copy numbers instead of concentration. To translate the given  $gBlock^{m}$  concentration in DNA copy numbers, we used the following equation:

Number of copies (molecules) = 
$$\frac{DNA(ng) \times N_A\left(\frac{molecules}{mal}\right)}{\frac{Molecular mass\left(\frac{m}{mal}\right) \circ f ragment}{Fragment length (bp)} \times Given in Avogadro\left(\frac{ng}{g}\right)}$$

where  $N_A$  (Avogadro's number) = 6.02214076 × 10<sup>23</sup> mol<sup>-1</sup>, and Given in Avogadro = 10<sup>9</sup>ng/g. Additionally, this time we aimed to create a more representative synthetic DNA standard able to quantify what is usually observed in human BC-DNA samples. Hence, the four different gBlocks<sup>™</sup> (except for IPC) were mixed in specific empirical ratios following systematic optimization, with a final volume of 100 µl. Practically, the individual gBlocks<sup>™</sup> were first diluted four to six times by factor ten and combined accordingly. Subsequently, the synthetic DNA standard mix was diluted by factor two to create the first standard used in qPCR. Four serial dilutions by factor three were performed to generate the rest of the standards. Information on the sequences of the synthetic DNA fragments, copy numbers per qBiCo version and dilution strategy are provided in Additional file 1: Tables S1, S4, S5 and Table 2.

#### qPCR assay design

qBiCo is based on the well-established real-time TaqMan<sup>®</sup>-based quantitative PCR technology. It comprises five assays in a multiplex reaction, each of which amplifies either a single- or multi-copy locus to calculate different BC-DNA quality/quantity parameters: global conversion efficiency, concentration, fragmentation and inhibition. Initially, the Ensembl genome browser (https:/ /grch37.ensembl.org/index.html) was used to locate and extract the DNA sequences of the regions of interest in the human genome (GRCh37/GRCh38). Next, the expected BC-DNA sequences were obtained using MethPrimer (https://www.urogene.org/cgi-bin/methp rimer/methprimer.cgi) [28]. BC-DNA-specific primers were designed using BiSearch following standard parameters (including no CpG in the primer sequence when possible) [29]. Fluorescently labelled TaqMan<sup>®</sup> probes were manually designed to target converted C-rich areas and considering instrument-specific requirements. Lastly, Autodimer was employed to check for hairpin and primer-dimer formation [30]. Details on primer/probe sequences, fluorophore labelling for each probe, and length of each fragment per qBiCo version are provided in Additional file 1: Tables S1 and S4.

#### qPCR protocol

The five assays were amplified simultaneously in a qPCR reaction using the oligo sequences and conditions described in Additional file 1: Table S1 (v1) and Table S4 (v2). Unless mentioned differently, all reactions were performed in Hard-Shell<sup>®</sup>, thin-wall PCR 96-well plates (Bio-Rad, USA) in a CFX96 Touch<sup>™</sup> Real-Time PCR Detection system (BioRad).

For qBiCo-v1 the PCR reaction was optimized in a final volume of 20 µl containing: 10 µl of 2x EpiTect<sup>®</sup> Methy-Light qPCR reagent (Qiagen), 2 µl of 25 mM MgCl<sub>2</sub> (Applied Biosystems, USA), 0.8 µl of 20 mg/ml BSA (Bio-Labs, USA), 3.2 µl of nuclease-free water, 2 µl of primer/ probe mix, 1 µl of diluted IPC-gBlock<sup>™</sup> (1 ng), and 1 µl of BC-DNA template. The cycling conditions for this qPCR consisted of polymerase activation and denaturation

**Table 2** Composition of the updated synthetic DNA standard(STD) in qBiCo-v2 method

qPCR assay	Number of copies (DNA molecules)					
	StdA	StdB	StdC	StdD	StdE	
hTERT Short	6,000	2,000	667	222	74	
TPT1 Long	6,000	2,000	667	222	74	
LINE1 Converted	384,000	128,000	42,667	14,222	4,741	
LINE1 Genomic	11,400	3,800	1,267	422	141	
Synthetic IPC	3,000	3,000	3,000	3,000	3,000	

(95 °C, 5 min), followed by 33 cycles of denaturation (95 °C, 15 s), annealing (56 °C, 30 s) and extension (60 °C, 70 s). Nuclease-free water was used as negative control. In most cases, qPCR reactions were performed in triplicate.

For qBiCo-v2 the PCR reaction volume was decreased to 10 µl, containing 5 µl of 2x EpiTect<sup>®</sup> MethyLight qPCR reagent (Qiagen), 1 µl of 25 mM MgCl<sub>2</sub> (Applied Biosystems), 0.2 µl of 20 mg/ml BSA (BioLabs), 1.3 µl of nuclease-free water, 1 µl of primer/probe mix, 0.5 µl of diluted IPC-gBlock<sup>™</sup> (Table 2), and 1 µl of BC-DNA template. For experiments performed using QS 7 Flex (TFS) and BioRad CFX384 (BioRad), Hard-Shell°, thin-wall PCR 384-well plates (BioRad) were used. For experiments performed using QS 5 (TFS) and 7500 Fast (TFS), MicroAmp<sup>®</sup> Fast 96-well reaction plates (0.1 ml) were used. For qPCR runs performed using the RotorGene Q instrument (Qiagen), the reaction volume was doubled and contained: 10 µl of 2X EpiTect° MethyLight qPCR reagent (Qiagen), 2 µl of 25mM MgCl<sub>2</sub> (Applied Biosystems), 0.4 µl of 20 mg/ml BSA (BioLabs), 4.1 µl of nuclease-free water, 2 µl of primer/probe mix, 0.5 µl of diluted IPC-gBlock<sup>™</sup> (3,000 copies, STD5, Additional file 1: Table S5), and 1  $\mu$ l of BC-DNA template. The reactions were performed in 0.1 ml Strip Tubes (Qiagen), which were placed in a Rotor-Disc 72 Rotor (Qiagen). All qPCR reactions were performed in triplicate, except for assessing repeatability (five replicates) and specificity (duplicates).

#### Data analysis

For qBiCo-v1, firstly the PCR efficiency and  $R^2$  of the standard curve of each assay was calculated to verify the success of the qPCR run. The baseline threshold of each assay was determined in a qualitative manner, by setting it at the beginning of the exponential (geometric) phase. The equation of each standard curve is obtained by the BioRad CFX Maestro software using the following formula:

$$\mathbf{CT} = \mathbf{m} \left[ \mathbf{log}(\mathbf{Qty}) \right] + \mathbf{b}$$

where m is the slope, b the y-intercept, and Qty the starting DNA quantity of each individual standard. According to this equation, the software provides information about the  $R^2$ , slope and PCR efficiency. When  $R^2 \ge 0.985$ ,  $-3.3 \le \text{slope} \le 3.6$ , and  $90\% \le PCR$  efficiency  $\le 110\%$ , the qPCR run was indicated as successful. Regarding the qBiCo indices, we followed the following formulas:

BC DNA concentration (ng) = [hTERT Short]

$$BC \ efficiency \ (\%) = \frac{[LINE1 \ Converted]}{[LINE1 \ Converted] + [LINE1 \ Genomic]} \times 100$$

$$BC DNA fragmentation index = \frac{[hTERT Short]}{[hTERT Long]}$$

Naturally, the intact BC-DNA ratio is calculated as 1 minus the fragmented BC-DNA amount. The presence of PCR inhibition was indicated qualitatively if the average IPC-Cq of the sample was higher than the average plus standard deviation (SD) of the average IPC-Cq of the synthetic standards.

For qBiCo-v2, a semi-automated pipeline using a spreadsheet macro file (Microsoft Excel) was developed to provide with a quicker and more standardized approach. Because of the different set-up among instruments, we created three files in total: Additional file 2 for 96-well, Additional file 3 for 72-well and Additional file 4 for 384-well set-up. Here is a short step-by-step description of the analysis: First, results including the well position, targeted assay, fluorophore, sample name, and Cq values are manually extracted from the instrument. On the "Sample Data" tab of the data analysis file, the plate layout is imported, while the extracted results are imported on the "Raw Results" tab of the file. Only in the case of 72-well setup, results are first imported on the "Rotorgene72\_FormatCorrection" tab to be converted to a suitable form before importing them on the "Raw Results" tab. Unless mentioned differently, all samples and standards were tested in triplicate. To exclude any individual qPCR reaction outliers, we examine whether the Cq value of a technical replicate of a sample lies within two SDs of the mean Cq of the triplicates. Hence, when a replicate lies outside this range, it is excluded from analysis and automatically flagged in the pipeline. In case of duplicates, a data point is considered outlier, when: SD / average Cq>0.02. Possible PCR inhibition is now indicated when the IPC-Cq value has more than one qPCR cycle difference than the IPC-Cq values of the synthetic standards. Following these steps, the equation and metrics of each standard curve are calculated and evaluated as described previously. Finally, the Cq values are then translated to copy numbers to calculate the qBiCo indices using the formulas mentioned above.

QC-ed qPCR and qBiCo parameter data were analysed using R (version 3.6.3, 29th February 2020) with R-packages: beanplot, cowplot, data.table, dplyr, ggbreak, ggplot2, ggpubr, ggrepel, grid, gridExtra, gtable, plyr, reshape2, rio, rstatix and writexl. During instrumentspecific qBiCo-v2 validation, we employed various methods to assess differences with statistical significance depending on the type of data. Particularly, we used the non-parametric Wilcoxon signed-rank test to assess differences between two gDNA inputs (repeatability). On the other hand, we employed the non-parametric Kruskal-Wallis test (one-way ANOVA) to assess differences between decreasing BC-DNA inputs (reproducibility/ sensitivity) and among differently treated BC-DNA samples (robustness). Finally, we used student's t-test to assess pairwise differences between the optimal and changed conditions for all experiments. We assessed significance based on difference levels: \*: p < 0.05, \*\*: p < 0.01, \*\*\*\*: p < 0.001.

#### Results

## Method formulation based on single- and multi-copy targets

Our envisioned method is based on co-amplifying five DNA regions with five sets of primers and probes labelled with different fluorophores based on TaqMan<sup>®</sup>based qPCR technology. The five assays were designed to assess the conversion efficiency, quantity, fragmentation and inhibition of a BC-DNA sample (Additional file 1: Table S1). Hence, it was important that they target as many non-CpG cytosines with known non-methylated levels as possible, and no CpG sites with unknown methylation levels. The first assay targets a single-copy gene (*hTERT*) via a short PCR fragment corresponding to the BC-DNA concentration index. The second assay targets a longer version of the same single-copy gene (*hTERT*) allowing us to calculate the BC-DNA fragmentation index. The choice of hTERT gene was arbitrary, motivated by the use of this gene in existing commercial assays for assessing genomic DNA quantity/quality. The third and fourth assays target the genomic and converted versions of a multi-copy repetitive element (LINE1) allowing to estimate a representative, genome-wide conversion efficiency index. The choice of LINE1 element is also arbitrary, motivated by its wide spread across the human genome (17%). The fifth assay targets a spiked-in, short, artificial DNA fragment acting as an internal positive control (IPC) that can give clues on potential PCR inhibition. Importantly, given the difficulty and uncertainty involved with using BC-DNA as standards, we opted for synthesizing artificial DNA fragments for this purpose.

#### Development and evaluation using synthetic standards

During the first phase of development, we aimed to create the proof-of-concept version of our novel method (qBiCo-v1). First, we optimized each qPCR assay separately, which resulted in achieving similar performance for all five assays (Fig. 1A). Moving into a multiplex qPCR reaction, Fig. 1B indicates successful achievement of harmonized amplification of all assays in a single reaction, where single-copy assays (*hTERT* Short & Long) were detected in the same quantification cycle (Cq), as did the multi-copy assays (*LINE1* Converted & Genomic). Finally, we showed that a maximum of 32 cycles were preferable to avoid false-positive detection driven by lowlevel background signal, particularly for the very sensitive



Fig. 1 qBiCo-v1 assay development using synthetic DNA standards. (A) Amplification curves of the single qPCR assays for the eight serially diluted DNA standards, each labelled with a different fluorescent dye; (B) Amplification curves of an example 5-plex qBiCo assay (STD3, in duplicate); (C) Performance of negative controls during 16 qPCR runs to determine the threshold where no false positive signals are obtained; (D) Standard curves of the four assays (50–0,39 ng/µl) used to calculate the concentration of each fragment in a sample; (E) Heatmap of the detection limit of each individual assay based on the serially diluted DNA standards; (F) Spike-in control performance in the serially diluted DNA standards during six qPCR runs; (G) Assay efficiency using the same DNA standard dilution series within five days. RFU: Relative fluorescent units; STD: Synthetic standard; IPC: Internal positive control; Cq: Quantification cycle

*LINE1 Genomic* assay that can also amplified in low levels of contaminating non-human DNA (Fig. 1C).

Subsequently, the performance of qBiCo was evaluated by studying the PCR efficiency and accuracy of each individual qPCR assay depending on initial input DNA template. First, to assess amplification efficiency, the pooled synthetic DNA standard was serially diluted by a factor of two, covering a resulting qPCR input DNA range from 50 down to 0.39 ng. Based on each assay's standard curve (triplicate analysis), we obtained a wellperforming PCR efficiency (98% - hTERT Short, 99% - hTERT Long, 94% - LINE1 Converted, 90% - LINE1 Genomic) and  $R^2 > 0.98$  in all assays (Fig. 1D). Importantly, the single- and multi-copy assays were performing comparatively, even at the lowest input. Yet, as expected the accuracy of DNA copy quantification decreased with decreasing DNA input (Fig. 1E). More specifically, we provide the first indication that quantification is sensitive and possible even down to 0.78 ng input DNA without remarkably sacrificing accuracy (81% - hTERT Short, 80% - hTERT Long, 75% - LINE1 Converted, 90% - LINE1 Genomic). Additionally, we assessed the performance of the IPC using data from six experiments, demonstrating that its detection is expected after the 27th PCR cycle in the absence of any inhibition (Fig. 1F). Finally, we showed that the synthetic DNA is stable and can be used for up to two days after preparation without compromising performance (Fig. 1G), which can be practical for highthroughput analysis.

#### Initial method validation using standard criteria

During the next phase of development, we aimed to test the above described qBiCo-v1 assay in terms of standard method performance criteria based on the three main parameters it measures: (i) global BC efficiency (Additional file 5: Fig S1), (ii) BC-DNA concentration (Additional file 5: Fig S2), and (iii) BC-DNA fragmentation (Additional file 5: Fig S3). To this end, we employed one of the most popular BC kits (EZ DNA methylation kit, Zymo Research) based on optimal conditions (200



**Fig. 2** Comparison of BC-DNA measurement capabilities between qBiCo-v1 and other currently used but non-specific methods. Measurements were performed in duplicate. Dotted lines show the expected DNA amounts. Information on protocols used are found in the Methods. BC: Bisulfite conversion; gDNA: Genomic DNA

ng gDNA input) to avoid introducing further bias. All analyses were performed at the qPCR quantification level and the associated data can be found in Additional file 1: Table S2.

Global BC efficiency measurements were highly repeatable based on six replicates (average SD = 0.37%) at least for highly converted samples (BC > 85%) (Additional file 5: Fig S1A). When measurements were replicated in two additional qPCR experiments, the reproducibility was very high, but we noticed possible small run-to-run bias (3-4% lower for one sample) (Additional file 5: Fig S1B). Additionally, we could successfully measure the expected BC efficiency even down to at least 0.24 ng of BC-DNA input (Additional file 5: Fig S1C). Strikingly, the quantification of this parameter was very robust and did not significantly change even in the presence of PCR inhibitors at extremely high levels (800 µM hematin) (Additional file 5: Fig S1D). On the other hand, artificial initial gDNA degradation using UV exposure, even for 1 min, strongly impacted BC efficiency measurement (>10% lower) (Additional file 5: Fig S1E). Importantly, we can accurately and reproducibly determine the right portion of BC-DNA in converted/non-converted mixtures, indicating strong linearity of our detection ( $R^2 = 0.994$ ) (Additional file 5: Fig S1F).

Results were very similar regarding these performance parameters for the measurement of BC-DNA concentration (Additional file 5: Fig S2). Our approach is able to relatively accurately detect BC-DNA down to 0.195 ng/  $\mu$ l (Additional file 5: Fig S2C), based on a linear fashion  $(R^2 = 0.989, Additional file 5: Fig S2F)$ . Additionally, given that this is the only parameter that other researchers have previously attempted to quantify based on existing methods such as fragment analysis (Bioanalyzer), spectrophotometry (Nanodrop), fluorometry (Qubit) (see details in Materials and Methods), we were also interested in comparing qBiCo's performance against these previously used three methods, particularly at two different gDNA amounts (optimal: 200 ng and suboptimal/16 times less: 12.5 ng). As shown in Fig. 2, our qBiCo-v1 assay was the only one resulting in the detection of the expected recovered BC-DNA amount for both samples, assuming 90-100% efficiency in BC-DNA recovery using the employed BC kit. Based on these first results, qBiCo surpasses in performance and quantification abilities all three other methods, which all tend to underestimate the BC-DNA quantity with limited reproducibility.

However, measuring BC-DNA fragmentation levels turned out more challenging (Additional file 5: Fig S3A-E). Not only was there large variation (>25%) in the detected BC-DNA fragmentation ratio between technical replicates of the same sample within the same experiment, but also across qPCR runs, making this parameter less stable and reproducible. Given that the BC-DNA

fragmentation levels we detected in our samples even under optimal conditions were very high (60–90%), the effect of PCR inhibitors and DNA degradation was more severe. Despite these difficulties, we found the performance of qBiCo-v1 during this initial validation process promising.

#### Performance evaluation of commercial BC kits

Next, we aimed to test its applicability by assessing popular BC protocols used by epigenetic researchers. To achieve this aim, we employed our method to assess the performance of ten different commercially available BC kits using five decreasing gDNA inputs, (200 ng to 1 ng) using three independent commercial DNA samples. Details on employed BC kits can be found in Table 1, while associated data are shared in Additional file 5: Table S1.

Overall, there was striking evidence that the BC process can differ substantially between commercial kits and their protocols, despite the manufacturers' promises of >99% BC effiency even at minute amounts (<1 ng). Regarding BC efficiency (Fig. 3A), all kits except kit 10 (Active Motif) resulted in levels > 90% using optimal gDNA inputs (200 ng). Kit 10 performed really poorly for all samples and DNA amounts, with BC rate ranging from 2 to 80% indicating a systemic issue in the BC process. As expected, the efficiency of BC is more variable and reduces with decreasing amounts of input DNA. However, there are certain kits that based on our work performed extremenly well (>95%) even at the lowest amounts of input DNA (1-10 ng), particularly kit 3 (Diagenode), kit 5 (Sigma Aldrich) and kit 6 (Qiagen). Moreover, it was common to observe that often only one of the three DNA samples in each input resulted in decreased BC performance, which impacted the observed variance for a particular kit; for example, for kit 4 (Promega).

Regarding BC-DNA recovery, the results were even more negatively surprising, particularly because most kit manufacturers promise at least 80% recovery. Figure 3B showcase a remarkable variability among kits in terms of the amount of eluted BC-DNA, regardless of the initial gDNA input. While there are certain kits (1-3): Zymo Research, TFS, Diagenode) that on average suffered from 0 to 30% DNA loss at optimal amounts, there are others (5-7: Sigma Aldrich, Qiagen, Abcam) that constantly cause ~ 50% DNA loss. Similarly, certain kits (9-10: Epigentek, Active Motif) degrade or lose>70% of the available DNA. We also noted that measured BC-DNA recovery is highly variable among samples within a kit, but the trend per kit is similar across DNA amounts, indicating a systematic performance. It is worth noting that no data were obtained when only 1 ng of gDNA was converted.

Finally, it is well-known that the BC process causes DNA fragmentation, but it has been challenging so far to measure it in a quantitiative way. Here, we measured detection differences between the *hTERT* Short (85 bp) and *hTERT* Long (235 bp) assays. Even at this short length, we noticed > 50% less fluorescent signal for the longer fragments, even under optimal conditions (200 ng input) (Fig. 3C). Observations and trends among kits were similar down to 50 ng, with kit 1 (Zymo Research) outperforming the other kits. Given the severe fragmentation, we failed to detect any signal for the longer fragment even at the 10 ng input. In combination with our previous observations during initial validation, this qBiCo-v1 parameter is by far the least sensitive.

#### Scaling-up towards a small-scale technology prototype

Based on our experience so far and aiming to advance the accuracy, sensitivity and applicability of our method, we followed several optimization steps to improve its functionalities and to make it more user-friendly and cost-effective. Our overall aim was to create an improved version (qBiCo-v2) towards a small-scale technology prototype that can form the basis of a future commercial product.

First, we aimed to improve the individual qPCR assay design. Due to its poor performance, we decided to redesign the *hTERT* Long assay used to obtain the BC-DNA fragmentation index. However, despite redesigning to amplify a slightly shorter fragment at another region of the *hTERT* gene and further adjusting the primer/ probe ration in the multiplex qPCR reaction, the performance of the redesigned assay was not satisfactory (efficiency of 61%) (Additional file 5: Fig S4A). Therefore, we decided to completely replace this assay with a similar one (222 bp) using another single-copy gene (TPT1) that resulted in a more stable and efficient amplification (efficiency of 91.8%) (Additional file 5: Fig S4B). Additionally, to boost the amplification of both our single-copy assays, we tested the effect of reducing the copies of the IPC assay (originally 6,000 copies per reaction). As a result, we obtained robust IPC detection down to 3,000 copies without compromising the measurement of PCR inhibition (Additional file 5: Fig S4C).

Moreover, to optimize the amplification efficiency of all assays, we examined multiple parameters. For example, we further adjusted the primer/probe mix concentrations, the amplification program by changing the denaturation time and temperature, and we examined the effect of adding betaine or DMSO in the reaction, besides  $MgCl_2$  and BSA that are already present in the assay. However, the extra additives led to instability of the *TPT1* Long fragment (data not shown) and therefore, we did not proceed further. Finally, to increase costeffectiveness, we also reduced the qPCR reaction volume



**Fig. 3** Testing of ten commercial BC kits based on qBiCo-v1 indices. (**A**) BC efficiency, (**B**) BC-DNA recovery and (**C**) BC-DNA fragmentation levels ( $\geq$  235 bp) depending on the input DNA amount used for the BC of three commercial DNA standards. Red circles indicate outliers > ± 2SD. Information on kits used are found in Table 1. BC: Bisulfite conversion

by 50% (10  $\mu$ l) without compromising performance. The optimized qPCR conditions and updated oligo sequences (qBiCo-v2) are presented in the Additional file 1: Table S4. Importantly, we aimed to adjust the composition and

preparation strategy of our synthetic DNA standard to make it more representative, considering our observations during the BC kit testing and the usual gDNA input range researchers use. Our goal was to create a synthetic standard that resembled as much as possible the detection and assay detection ratio of samples from the 'field'. To capture the entire spectrum of expected eluted BC-DNA variability, we created a testing 'master' sample by converting a commercial methylated DNA standard using five different amounts and a much larger range of BC kits (n = 16). Following a thorough tuning strategy and multiple optimization qPCR runs, we settled with a new synthetic DNA standard mix (Table 2) and dilution strategy (Additional file 1: Table S5) with higher robustness and less variability among technical replicates (Additional file 5: Fig S5). More specifically, we increased the number of dilutions and pipetted volumes during preparation using TE buffer (pH = 8.0) instead of water. These adjustments increased the stability of the fragments upon storage and subsequently improved the metrics of the standard curves.

Beyond the wet lab, we also aimed to improve and streamline the data analysis part. First, instead of originally estimating the nanograms of DNA involved, we switched our approach to calculating the actual copy numbers for each detected gene fragment. Second, to enable large-scale analysis in shorter time, we developed a user-friendly spreadsheet file based on formulas and macros for a semi-automatic way of analysing the qPCR data (Additional file 2). The data analysis pipeline also includes the automatic calculation of both qPCR performance metrics and qBiCo parameters per sample. Third, we developed a strategy to systematically detect data outliers, applicable to both duplicate and triplicate analysis.

#### Applicability to other real-time qPCR platforms

During the next phase of development, to enable uniform implementation in other laboratories we aimed to transfer and test the qBiCo-v2 method across real-time qPCR platforms of several manufacturers. To achieve our goal, beyond using our own BioRad CFX96 instrument, we collaborated with other researchers to employ five additional commonly encountered qPCR instruments: Bio-Rad CFX384, TFS 7500 Fast, TFS QuantStudio 5 (QS5), TFS QuantStudio 7 Flex (QS7 Flex), and Qiagen Rotor-Gene Q.

Overall, we successfully adopted our qPCR protocol in all five platforms, by simply adjusting the dyes to fit each instrument's spectra and the qPCR reaction volume (Additional file 1: Table S6) and our analysis pipeline to fit the different software used for the experiment set-up and data extraction (Additional files 3 and 4). Additionally, we evaluated its performance based on the obtained Cq values and relative fluorescent units (RFU) when running the same qPCR run on all instruments (Additional file 1: Table S7, Additional file 5: Fig S6 and S7). As expected, there were substantial differences in terms of fluorescence intensity, which can be explained by instrument-to-instrument specifications as each instrument works with different (log) scales. Nevertheless, Cq values were comparable across instruments for each qPCR assay (Additional file 1: Table S7). Importantly, we compared the PCR efficiency and goodness-of-fit of the obtained standard curves for each of the four qBiCo-v2 qPCR assays (without IPC) (Fig. 4). We detected subtle



Fig. 4 Comparative qBiCo-v2 assay performance on six qPCR instruments. (A) PCR efficiency (%) and (B) goodness-of-fit values (R<sup>2</sup>), of each qPCR assay of qBiCo-v2. Data are included from the initial method transfer runs on six qPCR systems from three biotechnology manufacturers (BioRad, Thermo Fisher Scientific, Qiagen). Dashed red lines indicate ideal performance measurements, while the grey box indicates acceptable range by standard qPCR guidelines (PCR efficiency between 90–110%)

differences in PCR efficiency, particularly evident for the *TPT1* Long assay for most instruments, and for the *LINE1* assay for the TFS instruments, the latter of which is detected with a proprietary dye (ABY<sup> $\infty$ </sup>) (Fig. 4A, Additional file 1: Table S6).

#### Instrument-specific developmental validation

To further establish the capacity, usability, and possible instrument-specific limitations of our qBiCo-v2 method, we decided to conduct a small-scale validation study per qPCR platform. Our methodology to assess different performance evaluation parameters is described in Materials and Methods, while all data per instrument and statistical evaluation can be found in Additional file 1: Tables S8 and S9, respectively.

First, we evaluated the ability of accurately assessing BC-DNA at two different gDNA inputs: 50 and 5 ng, with an expected maximum of 5 and 0.5 ng/µl input into qBiCo, respectively. We chose these amounts based on our previous experience optimizing this method, but also aiming to go lower with the gDNA input and evaluating the applicability of qBiCo-v2. For BC efficiency, we observed concordant results between instruments: average BC between replicates ranging 0.987-0.999 at 50 ng and 0.940-0.994 at 5 ng (Fig. 5A). Nevertheless, for BC-DNA concentration, we observed a larger variation between instruments for both input DNA amounts: average BC-DNA recovery between replicates ranging 0.493-2.252 at 50 ng and 0.443-1.409 at 5 ng (Fig. 5B). For three instruments (7500 Fast, CFX384 and CFX96), there were statistically significant differences between the two amounts (p < 0.01, Additional file 1: Table S9). It was also clear that the 7500 Fast instrument overestimated the concentration of BC-DNA compared to the other instruments used. Similarly, for BC-DNA fragmentation, we observed a similar trend with BC-DNA recovery; average BC-DNA fragmentation index between replicates ranging 0.861-1.116 at 50 ng and 1.026-6.412 at 5 ng (Fig. 5C). In this case, the CFX96 instrument seems to have overestimated the fragmentation of BC-DNA compared to the other instruments applied.

Next, we evaluated the ability of accurately assessing BC-DNA at decreasing qPCR DNA inputs ranging from 5 down to 0.0195 ng. To this end, we employed the BC-DNA samples based on 50 ng of gDNA input, to minimize the effects of the BC process itself. Overall, results and trends among instruments were similar; however, it was evident that adding less BC-DNA into the qBiCo reaction can significantly alter the obtained indices. This was particularly clear for measuring BC efficiency (p < 0.01) and BC-DNA concentration (p < 0.0001), for most instruments (Additional file 1: Table S9). Specifically for BC efficiency, measurements were significantly affected below 0.1563 ng BC-DNA (Fig. 5D). Similarly,

BC-DNA concentration measurements were significantly impacted at amounts lower than 0.08 ng BC-DNA, for most instruments (Fig. 5E, Additional file 1: Table S9). In line with our observations described above, the 7500 Fast instrument overestimates the BC-DNA concentration. Finally, for BC-DNA fragmentation, and as expected from our previous experience, we failed to obtain sufficient data below 0.1563 ng BC-DNA, except for Rotor-Gene Q which seems to be the most sensitive for this qBiCo index (Fig. 5F). Overall, considering all parameters, we can conclude that 150 pg input DNA is the limit of detection of qBiCo-v2.

Importantly, using this sensitivity threshold, we tested the reproducibility of the qBiCo-v2 method across qPCR platforms. The results were very promising for all decreasing amounts (Additional file 5: Fig S8, Additional file 1: Table S9). Strong statistically significant differences were only detected for the CFX96 instrument for all three qBiCo indices (p < 0.001) and for the 7500 Fast instrument in terms of BC efficiency (p < 0.0001). Moreover, we evaluated the ability of accurately assessing BC-DNA in compromised, artificially degraded samples using UV treatment. As expected, the measurements using all instruments were significantly affected (Fig. 5G-I, Additional file 1: Table S9), particularly for longer UV exposure times (60-120 s). Finally, we evaluated the specificity of BC-DNA evaluation using qBiCo-v2 using a range of non-human DNA samples (Additional file 1: Table S8). Unsurprisingly, we obtained false positive results for *Rhesus macaque* (monkey), particularly when measuring global BC efficiency, which can be explained since we target the evolutionally conserved region of the LINE1 repeat.

#### Discussion

BC has been the golden standard for decades, allowing researchers to easily translate methylation differences into sequence variation that can be detected using standard genetic technologies [14], each with their own strengths and weaknesses [31]. So far, these techniques have been used for methylation biomarker discovery and validation, uncovering method-specific methylation biases among targeted techniques [32] and between genome-wide and PCR-based approaches [33], introduced either experimentally or during data analysis [34]. Currently, as we move towards larger global research efforts, more difficult and scarce templates, and implementation in the clinic, the need for methylation assay standardization becomes prominent. Currently, the lack of proper standardization stands out as the most important contributor to DNA methylation variability [35]. Yet, while laboratories validate methylation assays post-BC [36], the performance of BC remains a black box. In genome-wide analysis, BC control probes are only



Fig. 5 qBiCo-v2 assay validation using synthetic DNA standards on six qPCR instruments. Testing (A-C) repeatability at two DNA inputs, (D-F) sensitivity down to 0.02 ng/µl, and (G-I) robustness at different UV exposure times, per qBiCo index and qPCR instrument (colour-coded). All qBiCo indices are presented in ratios. BC: Bisulfite conversion

assessed post-analysis after sample processing [37], not as QC for downstream decision making to exclude inappropriately converted samples. In targeted analysis, BC-DNA is enriched by using converted-specific primers or probes, making some think a QC is not necessary. However, different BC-DNA inputs in the PCR and storage conditions that can cause further fragmentation, were among the most significant factors for large methylation variation [35]. Measuring the BC-DNA input is also essential in statistically interpretating methylation data, to map stochastic events that are particularly relevant in low amounts (<1 ng) of BC-DNA [38].

While usually scientists assume its successful performance, there are a few studies evaluating BC kits [19– 25], already highlighting great variance using targeted, non-specific approaches. Motivated by these small-scale efforts, we aimed to offer the community an innovative and complete technological solution for global BC performance assessment, which can also potentially form the basis of a future commercial product. Our goal was to establish a method and prototype tool that can simultaneously assess both the quality and quantity of BC-DNA prior to downstream analysis using a bench-top instrument commonly accessible in molecular biology laboratories. It is important that the BC-QC solution is simple, cost-effective, reliable and robust, hence we opted for a method based on real-time qPCR, which is also often used for the QC of gDNA. Additionally, qPCR enables the simultaneous amplification of multiple small DNA regions in one reaction, hence the assessment of several QC parameters at the same reaction.

Compared to previous studies that have assessed BC-DNA recovery and fragmentation, to the best of our knowledge we are the first to build a QC solution that measures a global BC efficiency index. We did this by analysing the genomic and converted versions of an evolutionally conserved, (non-CpG) C-rich region of a repetitive element (LINE1). Recently, Hong and Shin (2021) developed another multiplex qPCR-based system (BisQuE) with a similar goal in mind, but their BC efficiency measurement is based only on one cytosine of a single gene intron [39], which can highly bias the assessment. In contrast, via the LINE1 Converted probe we target five cytosines across ~180-200 different LINE1 regions in the genome per sample, making our BC efficiency measurement much more representative compared to targeting a single-copy locus. While we expect that each human genome might display unique genetic variants in some of these LINE1 copies, we do not envision this as problematic since the BC efficiency measure is based on a ratio of hundreds of targets, hence is quite robust to individual copy effects. On the other hand, the choice of targets for our single-copy assays was not as straightforward. We opted for the hTERT gene, since we also employ this target for gDNA QC assessment in our laboratory (Quantifier<sup>™</sup> Duo DNA quantification kit, TFS). This allows us for avoiding locus-specific biases when comparing quantities obtained before and after BC. Of course, *LINE1* and hTERT are just viable options we could have used, and in principle the use of any common repetitive element and single-copy gene respectively, would work as long as the number of included non-CpG cytosines is similar and sufficient. While we do not expect that the choice of locus would significantly impact the results, we cannot exclude small effects from sequence differences, local secondary structures and/or unknown methylation levels.

Altogether, qBiCo-v1 showed promising performance, particularly in comparison to other non-specific methods commonly used as QC. While these methods are not BC-DNA-specific, they take advantage of its nature being mostly single-stranded for their quantification. Despite having only preliminary data, we successfully highlighted qBiCo's superior performance. However, Page 14 of 17

there were indications of low sensitivity and robustness issues with the *hTERT* Long assay, which drove the overall limit of detection of qBiCo-v1 at 780 pg BC-DNA per µl reaction. At first glance this looks very little, but if we consider that BC-DNA is usually eluted in 10  $\mu$ l and that BC-DNA recovery is rather low (often < 50%), this would translate into at least 15 ng gDNA into BC. On the other hand, likely due to the multi-copy nature of *LINE1* assays, the measurement of BC efficiency was very robust even at extreme inhibition and fragmentation levels. However, caution is required as we observed low specificity with the LINE1 Genomic assay, which was often amplified earlier than anticipated. Finally, from our initial validation efforts it was clear that the fragmentation index was the weakest assay in qBiCo, where we often obtained no data, even if we only target a 235 bp long fragment.

Eventually, we were able to highlight qBiCo's need and potential in the field, when using a wide range of BC kits as part of research and diagnostics assays. The employed kits were selected from ten different manufacturers. They display similar protocols concerning the basic structure and steps in the BC process - sodium bisulfite incubation, DNA binding, desulfonation, washing and elution of BC-DNA, which allowed for a fair comparison. Nevertheless, there is variance within some steps i.e., incubation length and temperature, spin column width and size. Overall, our findings were striking. First, it is clear that a promised BC efficiency of >99% is not always achieved. While we uncovered systematic issues with some kits, others achieved consistent successful performance even down to 10 ng. Yet, qBiCo could pick out individual sample failures and differences. One question that remains to be determined by the community is the threshold one should choose for successful BC efficiency assessment. Should it be 90, 95 or 99%? Lastly, it was clear that results were worse when treating 1 ng of BC-DNA. Detected BC efficiencies were much lower, which can be driven not only by the BC process itself, but also by the capabilities of qBiCo (in the best-case scenario amounts translate to 100 pg input into each reaction).

Additionally, our assessment in terms of BC-DNA recovery and fragmentation is in line with recent work by others. Using Qubit and testing 12 BC kits, average BC-DNA recovery was measured between 26.6 and 88.3% using kit-specific optimal gDNA inputs (135 ng to 2  $\mu$ g) [19]. On the other hand, using qPCR-based BisQuE Hong and Shin previously showed that average BC-DNA recovery at 50 ng input could range between 18.2% and 50.6% [39]. Results are similar also in terms of BC-DNA fragmentation, even though we acknowledge that it is difficult to compare quantitatively as previous methods use different approaches, indices and fragment lengths. Here, we show that even with the best performing kit, the intact BC-DNA portion (at least at 235 bp level) is no more

than 50-60% of the sample. Previous observations using qPCR have reported several Cq differences between their short and long fragments [19]. Also, in another previous study, BC-DNA degradation indices were measured between 1.5 and 2.5 [40], which agree with our findings. We cannot exclude locus-specific biases in these assessments, driven for example by DNA accessibility, 3D DNA structures or binding proteins, but overall BC damages DNA substantially. This might be extra challenging when applying qBiCo to low-quality/quantity BC-DNA samples like cell-free or forensic-type DNA. For such applications, the right choice of BC kit is critical. In a recent study on cell-free tumour DNA, the EpiTect kit (QIAGEN) performed best in terms of BC-DNA recovery and fragmentation based on digital PCR and Bioanalyzer, respectively [41]. Based on these and our findings, every kit performs differently having their own strengths and weaknesses. Therefore, there is a need for every lab to thoroughly validate their BC kit of choice for their own application, for which qBiCo offers a complete solution. This concerns not only well-established commercial BC kits, but also newly developed, not-yet-commercial BC methods like rapid ones combining DNA extraction and conversion [42] or based on microfluidics [43] as well as ones using enzymatic approaches [44].

To enable widespread adoption of qBiCo, we continued our technology development towards building a smallscale prototype. Following a set of functionality improvements, we built an improved our method (qBiCo-v2). Despite replacing our long assay and targeting a new gene (TPT1), we still found its performance weak and not robust enough. Future work could explore alternative solutions in a larger screening approach. For example, an idea is to offer a more sensitive and robust approach for BC-DNA recovery and fragmentation also via repetitive elements [45]. Still, caution should be taken in the chosen length of the long assay to make it meaningful enough for downstream applications. Moreover, we updated our approach of creating the synthetic standard from mixing to preparation and improved it more in terms of quantitative power and applicability. We ensured that the synthetic standards were amplified and detected at similar time points with BC-DNA samples considering 16 different BC kits and 5 different amounts. Nevertheless, we still observed substantial batch effects, so it is important to find ways to standardize the production of the synthetic standard in the future. Importantly, we semi-automated the analysis from the raw qPCR data to the qBiCo indices to enable for a faster and normalized approach.

In addition, to further enable implementation, we tested the qBiCo-v2 protocol in several commonly used instruments from three manufacturers. We were pleased by how smooth this transfer was as the only necessary adjustments were driven by instrument-specific

limitations in terms of compatible dyes (TFS), number of channels (BioRad) and minimum reaction volume (Qiagen). In the future, using our approach we are confident that scientists can easily extend the transfer to additional instruments, for example the LightCycler<sup>®</sup> platforms (Roche). Generally, our instrument-specific validation results aligned with qBiCo-v1, with detectable variance within and between BC-DNAs, qPCR runs, qBiCo indices and qPCR instruments still being observed. We can explain these by simple errors during sample preparation and storage as well as micro-pipetting, but systematic, instrument-specific limitations were also uncovered. Nevertheless, our statistical evaluation should be considered cautiously due to the small sample size and missing data particularly in the case of the TPT1 Long fragment and robustness experiments. Overall, based on our empirical observations, we provide useful guidance to researchers that possess these instruments and wish to perform and/or extend these validation procedures in their own laboratory. Future implementation of qBiCo in the (high throughput) end-user environment will further highlight its true potential and usefulness.

#### Conclusion

In 2017, Lind and van Engeland highlighted: "Only when including quality and standardization at every level of DNA methylation analyses, will we be able to achieve the robustness to independently validate DNA methylation analyses and to compare multiple methylation studies in systematic reviews. This is the only way to more efficiently develop future DNA methylation-based biomarkers." [46]. This is particularly true in specialized applications like forensics, where data standardization and scrutiny are of paramount importance, currently hindered by the lack of BC assessment [47]. In this study, we built the first-of-its-kind method and prototype tool for global BC performance assessment based on qPCR. qBiCo can detect several parameters of a BC-DNA sample: efficiency, recovery, fragmentation and inhibition. Here, we demonstrated our technology development: from basic method formulation (TRL-1) to technology concept (TRL-2) to critically improved, established proof-of-concept (TRL-3) to testing and validation of a small-scale prototype (TRL-4). We conclude that we achieved our aim of offering a BC-QC solution to the epigenetic community, providing evidence of its performance across different samples, BC-DNA inputs, BC kits, qPCR instruments and other conditions. We also present a thorough critical method assessment and suggestions for further improvement and development. Motivated by our findings, we call epigenetic researchers to shift their perspective from assumption to empirical assessment with regards to BC kit performance and integrate qBiCo as a QC step in their methylation assays regardless

of downstream analysis. We also hope to inspire others to develop similar methods for other types of data, i.e. BC-RNA data in epitranscriptomics [48].

#### Abbreviations

BC	bisulfite conversion
BC-DNA	bisulfite-converted DNA
CpG	cytosine-guanine dinucleotides
Cq	quantification cycle
DNA	deoxyribonucleic acid
gDNA	genomic DNA
hTERT	human telomerase reverse transcriptase
IPC	internal positive control
LINE1	long interspersed nuclear element 1
PCR	polymerase chain reaction
RFU	relative fluorescent units
qBiCo	qualification/quantitation of bisulfite-converted DNA
QC	quality control
qPCR	quantitative PCR
SD	standard deviation
STD	standard
TFS	Thermo Fisher Scientific
TPT1	tumour protein, translationally-controlled 1
TRL	technology readiness level

#### Supplementary Information

The online version contains supplementary material available at https://doi.or g/10.1186/s43682-025-00033-3.

Supplementary Material 1	
Supplementary Material 2	
Supplementary Material 3	
Supplementary Material 4	
Supplementary Material 5	

#### Acknowledgements

We would like to thank previous members of the Department of Genetic Identification (Erasmus MC) for their technical assistance and contributions to this project, in particular Rochelle Chotkan during qPCR primer design and development of single-plex reactions, Diego Montiel González during LINE1 primer design and Benjamin Planterose Jiménez during implementation of R for data analysis. Floor Claessens contributed to this study as part of her bachelor's education (Forensic Research) at Hogeschool van Amsterdam, with work carried out during her research internship at the Department of Genetic Identification at Erasmus MC. Furthermore, we would also like to thank the multiple collaborators for enabling us access to different qPCR instruments: Isabel Chu (Department of Haematology, Erasmus MC, TFS QuantStudio 7 Flex), Claudia Erpelinck (Department of Haematology, Reamus MC, TFS 7500 Fast), and Cathleen van der Lee (QIAGEN, RotorGene Q).

#### Author contributions

Conceptualization: AV; Methodology: FK, AV; Investigation: FK, FC, VK; Data curation: FK, RS, AV; Formal analysis: FK, RS, AV; Visualization: RS, AV; Funding acquisition: AV; Resources: MK; Supervision: AV; Writing – original draft: AV; Writing – review & editing: FK, RS, FC, VK, MK.

#### Funding

This work was financed partly by Erasmus MC and partly by the Dutch Research Council (NWO) via a Demonstrator grant awarded to A.V. by the Applied and Engineering Sciences (TTW) domain (project number 18560).

#### Data availability

All data generated or analysed during this study are included in this published article and its supplementary information files.

#### Declarations

#### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

#### **Competing interests**

Athina Vidaki was the sole inventor of the presented qBiCo technology on a filed patent application (Publication numbers: CN114761578A; EP4028551A1; US2022372574A1; WO2021048410A1), which, however, has been discontinued by Erasmus MC for financial reasons. The remaining authors have declared that no competing interests exist.

#### Received: 26 November 2024 / Accepted: 5 January 2025 Published online: 13 January 2025

#### References

- 1. Eccleston A, DeWitt N, Gunter C, Marte B, Nath D. Epigenetics Nat. 2007;447:395. https://doi.org/10.1038/447395a.
- Moore LD, Le T, Fan G. DNA methylation and its basic function. Neuropsychopharmacology. 2013;38(1):23–38. https://doi.org/10.1038/npp.2012.112.
- Mazzio EA, Soliman KFA. Basic concepts of epigenetics. Epigenetics. 2012;7(2):119–30. https://doi.org/10.4161/epi.7.2.18764.
- Newell-Price J, Clark AJ, King P. DNA methylation and silencing of gene expression. Trends Endocrinol Metab. 2000;11(4):142–8. https://doi.org/10.10 16/s1043-2760(00)00248-4.
- Vidaki A, Kayser M. Recent progress, methods and perspectives in forensic epigenetics. Forensic Sci Int Genet. 2018;37:180–95. https://doi.org/10.1016/j. fsigen.2018.08.008.
- Reale A, Tagliatesta S, Zardo G, Zampieri M. Counteracting aged DNA methylation states to combat ageing and age-related diseases. Mech Ageing Dev. 2022;206:111695. https://doi.org/10.1016/j.mad.2022.111695.
- Perez RF, Tejedor JR, Fernandez AF, Fraga MF. Aging and cancer epigenetics: where do the paths fork? Aging Cell. 2022;21(10):e13709. https://doi.org/10.1 111/acel.13709.
- Locke WJ, Guanzon D, Ma C, Liew YJ, Duesing KR, Fung KYC, et al. DNA methylation Cancer biomarkers: translation to the clinic. Front Genet. 2019;10:1150. https://doi.org/10.3389/fgene.2019.01150.
- Perna L, Zhang Y, Mons U, Holleczek B, Saum K-U, Brenner H. Epigenetic age acceleration predicts cancer, cardiovascular, and all-cause mortality in a German case cohort. Clin Epigenetics. 2016;8(1):64. https://doi.org/10.1186/s131 48-016-0228-z.
- Haghshenas S, Bhai P, Aref-Eshghi E, Sadikovic B. Diagnostic utility of genome-wide DNA methylation analysis in mendelian neurodevelopmental disorders. Int J Mol Sci. 2020;21(23):9303. https://doi.org/10.3390/ijms212393 03.
- Planterose Jimenez B, Liu F, Caliebe A, Montiel Gonzalez D, Bell JT, Kayser M, et al. Equivalent DNA methylation variation between monozygotic co-twins and unrelated individuals reveals universal epigenetic inter-individual dissimilarity. Genome Biol. 2021;22(1):18. https://doi.org/10.1186/s13059-020-02 223-9.
- Ambroa-Conde A, Giron-Santamaria L, Mosquera-Miguel A, Phillips C, Casares de Cal MA, Gomez-Tato A, et al. Epigenetic age estimation in saliva and in buccal cells. Forensic Sci Int Genet. 2022;61:102770. https://doi.org/10.1016/j. fsigen.2022.102770.
- Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, et al. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proc Natl Acad Sci U S A. 1992;89(5):1827–31. https://doi.org/10.1073/pnas.89.5.1827.
- Pajares MJ, Palanca-Ballester C, Urtasun R, Alemany-Cosme E, Lahoz A, Sandoval J. Methods for analysis of specific DNA methylation status. Methods. 2021;187:3–12. https://doi.org/10.1016/j.ymeth.2020.06.021.
- Kurdyukov S, Bullock M. DNA methylation analysis: choosing the right method. Biology (Basel). 2016;5(1). https://doi.org/10.3390/biology5010003.
- Genereux DP, Johnson WC, Burden AF, Stoger R, Laird CD. Errors in the bisulfite conversion of DNA: modulating inappropriate- and failed-conversion frequencies. Nucleic Acids Res. 2008;36(22):e150. https://doi.org/10.1093/nar/ gkn691.

- Tanaka K, Okamoto A. Degradation of DNA by bisulfite treatment. Bioorg Med Chem Lett. 2007;17(7):1912–5. S0960-894X(07)00091-1 [pii]. https://doi. org/10.1016/j.bmcl.2007.01.040
- Leontiou CA, Hadjidaniel MD, Mina P, Antoniou P, Ioannides M, Patsalis PC. Bisulfite Conversion of DNA: performance comparison of different kits and Methylation Quantitation of Epigenetic Biomarkers that have the potential to be used in non-invasive prenatal testing. PLoS ONE. 2015;10(8):e0135058. doi: PONE-D-15-06170 [pii]. https://doi.org/10.1371/journal.pone.0135058
- Kint S, De Spiegelaere W, De Kesel J, Vandekerckhove L, Van Criekinge W. Evaluation of bisulfite kits for DNA methylation profiling in terms of DNA fragmentation and DNA recovery using digital PCR. PLoS ONE. 2018;13(6):e0199091. doi: PONE-D-18-04302 [pii]. https://doi.org/10.1371/jou rnal.pone.0199091
- Tierling S, Schmitt B, Walter J. Comprehensive evaluation of Commercial Bisulfite-based DNA methylation kits and development of an alternative protocol with Improved Conversion performance. Genet Epigenet. 2018;10:1179237X18766097. https://doi.org/10.1177/1179237X18766097.
- 21. Worm Orntoft MB, Jensen SO, Hansen TB, Bramsen JB, Andersen CL. Comparative analysis of 12 different kits for bisulfite conversion of circulating cell-free DNA. Epigenetics. 2017;12(8):626–36. https://doi.org/10.1080/15592 294.2017.1334024.
- Holmes EE, Jung M, Meller S, Leisse A, Sailer V, Zech J, et al. Performance evaluation of kits for bisulfite-conversion of DNA from tissues, cell lines, FFPE tissues, aspirates, lavages, effusions, plasma, serum, and urine. PLoS ONE. 2014;9(4):e93933. PONE-D-14-02521 [pii]. https://doi.org/10.1371/journal.pon e.0093933
- Plongthongkum N, Diep DH, Zhang K. Advances in the profiling of DNA modifications: cytosine methylation and beyond. Nat Rev Genet. 2014;15(10):647–61. https://doi.org/10.1038/nrg3772.
- Ehrich M, Zoll S, Sur S, van den Boom D. A new method for accurate assessment of DNA quality after bisulfite treatment. Nucleic Acids Res. 2007;35(5):e29. https://doi.org/10.1093/nar/gkl1134.
- Sriraksa R, Chaopatchayakul P, Jearanaikoon P, Leelayuwat C, Limpaiboon T. Verification of complete bisulfite modification using calponin-specific primer sets. Clin Biochem. 2010;43(4–5):528–30. https://doi.org/10.1016/j.clinbioche m.2009.11.005.
- Hayatsu H, Wataya Y, Kazushige K. The addition of sodium bisulfite to uracil and to cytosine. J Am Chem Soc. 1970;92(3):724–6. https://doi.org/10.1021/ja 00706a062.
- Hayatsu H. Discovery of bisulfite-mediated cytosine conversion to uracil, the key reaction for DNA methylation analysis–a personal account. Proc Jpn Acad Ser B Phys Biol Sci. 2008;84(8):321–30. https://doi.org/10.2183/pjab.84.321.
- Li LC, Dahiya R. MethPrimer: designing primers for methylation PCRs. Bioinformatics. 2002;18(11):1427–31. https://doi.org/10.1093/bioinformatics/18.11.14 27.
- Tusnády GE, Simon I, Váradi A, Arányi T. BiSearch: primer-design and search tool for PCR on bisulfite-treated genomes. Nucleic Acids Res. 2005;33(1):e9. https://doi.org/10.1093/nar/gni012.
- Vallone PM, Butler JM. AutoDimer: a screening tool for primer-dimer and hairpin structures. Biotechniques. 2004;37(2):226–31. https://doi.org/10.2144/ 04372ST03.
- Jeddi F, Faghfuri E, Mehranfar S, Soozangar N. The common bisulfite-conversion-based techniques to analyze DNA methylation in human cancers. Cancer Cell Int. 2024;24(1):240. https://doi.org/10.1186/s12935-024-03405-2.
- De Chiara L, Leiro-Fernandez V, Rodriguez-Girondo M, Valverde D, Botana-Rial MI, Fernandez-Villar A. Comparison of Bisulfite pyrosequencing and methylation-specific qPCR for methylation Assessment. Int J Mol Sci. 2020;21(23). https://doi.org/10.3390/ijms21239242.
- 33. Chatterjee A, Macaulay EC, Ahn A, Ludgate JL, Stockwell PA, Weeks RJ, et al. Comparative assessment of DNA methylation patterns between reduced

representation bisulfite sequencing and Sequenom EpiTyper methylation analysis. Epigenomics. 2017;9(6):823–32. https://doi.org/10.2217/epi-2016-01 76.

- Piao Y, Xu W, Park KH, Ryu KH, Xiang R. Comprehensive Evaluation of Differential Methylation Analysis Methods for Bisulfite Sequencing Data. Int J Environ Res Public Health. 2021;18(15). https://doi.org/10.3390/ijerph18157975.
- Pharo HD, Honne H, Vedeld HM, Dahl C, Andresen K, Liestøl K, et al. Experimental factors affecting the robustness of DNA methylation analysis. Sci Rep. 2016;6:33936. https://doi.org/10.1038/srep33936.
- Sestakova S, Salek C, Remesova H. DNA methylation validation methods: a coherent review with practical comparison. Biol Proced Online. 2019;21:19. https://doi.org/10.1186/s12575-019-0107-z.
- Leung ML, Abdullaev Z, Santana-Santos L, Skaugen JM, Moore S, Ji J, Microarray-Based DNA. Methylation profiling: validation considerations for clinical testing. J Mol Diagn. 2024;26(6):447–55. https://doi.org/10.1016/j.jmol dx.2024.02.001.
- Naue J, Hoefsloot HCJ, Kloosterman AD, Verschure PJ. Forensic DNA methylation profiling from minimal traces: how low can we go? Forensic Sci Int Genet. 2018;33:17–23. https://doi.org/10.1016/j.fsigen.2017.11.004.
- 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. https://doi.org/10.3389/fgene.2021.618955.
- Shiga M, Asari M, Takahashi Y, Isozaki S, Hoshina C, Mori K, et al. DNA methylation-based age estimation and quantification of the degradation levels of bisulfite-converted DNA. Leg Med (Tokyo). 2024;67:102336. https://doi.org/1 0.1016/j.legalmed.2023.102336.
- Kresse SH, Brandt-Winge S, Pharo H, Flatin BTB, Jeanmougin M, Vedeld HM, et al. Evaluation of commercial kits for isolation and bisulfite conversion of circulating cell-free tumor DNA from blood. Clin Epigenetics. 2023;15(1):151. https://doi.org/10.1186/s13148-023-01563-0.
- Zamuner FT, Ramos-Lopez A, Garcia-Negron A, Purcell-Wiltz A, Cortes-Ortiz A, Cuevas AR, et al. Evaluation of silica spin–column and magnetic bead formats for rapid DNA methylation analysis in clinical and point–of–care settings. Biomed Rep. 2024;21(2):112. https://doi.org/10.3892/br.2024.1800.
- Turiello R, Nouwairi RL, Keller J, Cunha LL, Dignan LM, Landers JP. A rotationally-driven dynamic solid phase sodium bisulfite conversion disc for forensic epigenetic sample preparation. Lab Chip. 2023;24(1):97–112. https://doi.org/ 10.1039/d3lc00867c.
- Vaisvila R, Ponnaluri VKC, Sun Z, Langhorst BW, Saleh L, Guan S, et al. Enzymatic methyl sequencing detects DNA methylation at single-base resolution from picograms of DNA. Genome Res. 2021;31(7):1280–9. https://doi.org/10.1 101/gr.266551.120.
- Pham DAT, Le SD, Doan TM, Luu PT, Nguyen UQ, Ho SV, et al. Standardization of DNA amount for bisulfite conversion for analyzing the methylation status of LINE-1 in lung cancer. PLoS ONE. 2021;16(8):e0256254. https://doi.org/10.1 371/journal.pone.0256254.
- Lind GE, van Engeland M. Details matter: the role of genomic location and assay standardization in DNA methylation analyses. Epigenomics. 2017;9(7):933–5. https://doi.org/10.2217/epi-2017-0053.
- Naue J, Lee HY. Considerations for the need of recommendations for the research and publication of DNA methylation results. Forensic Sci Int Genet. 2018;37:e12–4. https://doi.org/10.1016/j.fsigen.2018.08.003.
- Johnson Z, Xu X, Pacholec C, Xie H. Systematic evaluation of parameters in RNA bisulfite sequencing data generation and analysis. NAR Genom Bioinform. 2022;4(2):lqac045. https://doi.org/10.1093/nargab/lqac045.

#### Publisher's note

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