KAPA HIFI URACIL+: A NOVEL HIGH FIDELITY POLYMERASE FOR IMPROVED BISULFITE SEQUENCING

ROSS I.M. WADSWORTH[1], ERIC VAN DER WALT[1], CHRISTINE CLARK[2], MARTIN T. SWAIN[3], GAVIN J. RUSH[1], JOHN F. FOSKETT III[1], and PAUL J. MCEWAN[1]

1 KAPA BIOSYSTEMS, WOBURN, MA, USA | 2 WELLCOME TRUST SANGER INSTITUTE, CAMBRIDGE, UK | 3 UNIVERSITY OF ABERYSTWYTH, UK

INTRODUCTION

DNA methylation plays an important role in modulating gene expression in a wide variety of organisms. While capture techniques allow the enrichment of highly methylated genomic regions, methods of choice for studying methylation at base-pair resolution utilize bisulfite-treated DNA. Bisulfite (BS) converts unmethylated cytosines to uracil - which are read as thymines in subsequent sequencing reactions - while 5'-methylated cytosines (5-mC) remain unconverted.

High fidelity "B-family" DNA polymerases commonly used for NGS library amplification are unable to amplify BS-converted DNA due to an inherent "uracil read-ahead" function that detects promutagenic uracil in the template strand and stalls DNA synthesis.

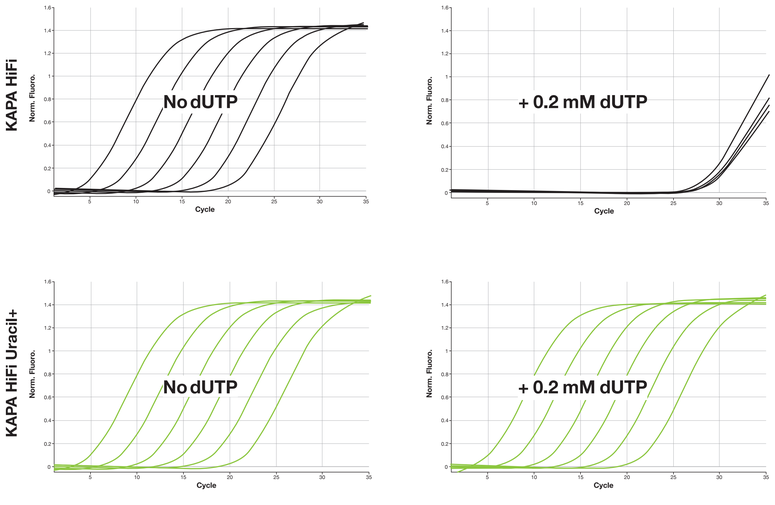
We previously developed KAPA HiFi, a novel engineered high fidelity DNA polymerase which shows reduced amplification bias and increased amplification efficiency compared with standard library amplification reagents, yielding sequencing reads with more even coverage depth and more complete representation across reference sequences.

We have subsequently developed a new version of this polymerase which is tolerant of uracil in the template strand. Here we demonstrate the utility of KAPA HiFi Uracil+ DNA Polymerase for the high-fidelity amplification of BS-converted NGS libraries.

REAL-TIME AMPLIFICATION WITH URACIL

dUTP inhibits high fidelity PCR -- dUTP is readily incorporated during strand extension, but the uracil read-ahead function of Type B polymerases prevents DNA containing uracil from acting as template in subsequent rounds of amplification.

We used real-time PCR to compare the performance of KAPA HiFi and KAPA HiFi Uracil+ in the presence or absence of dUTP.



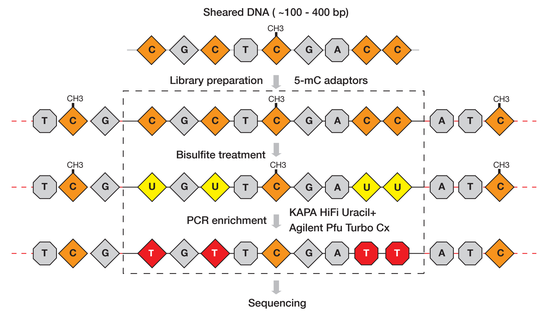
KAPA HiFi Uracil+ is not inhibited by dUTP.

SYBR Green® real-time PCR was used to monitor amplification of a specific 452 bp amplicon from ten-fold template dilutions (80 pM – 0.8 fM) without dUTP or with 0.2 mM dUTP. KAPA HiFi shows typical inhibition by uracil, while PCR amplification with KAPA HiFi Uracil+ is not inhibited. Amplification with Agilent Pfu Turbo Cx was completely inhibited by SYBR Green®, even in the absence of dUTP (data not shown), and could therefore not be compared in this assay.

PCR AMPLIFICATION OF WGBS LIBRARIES

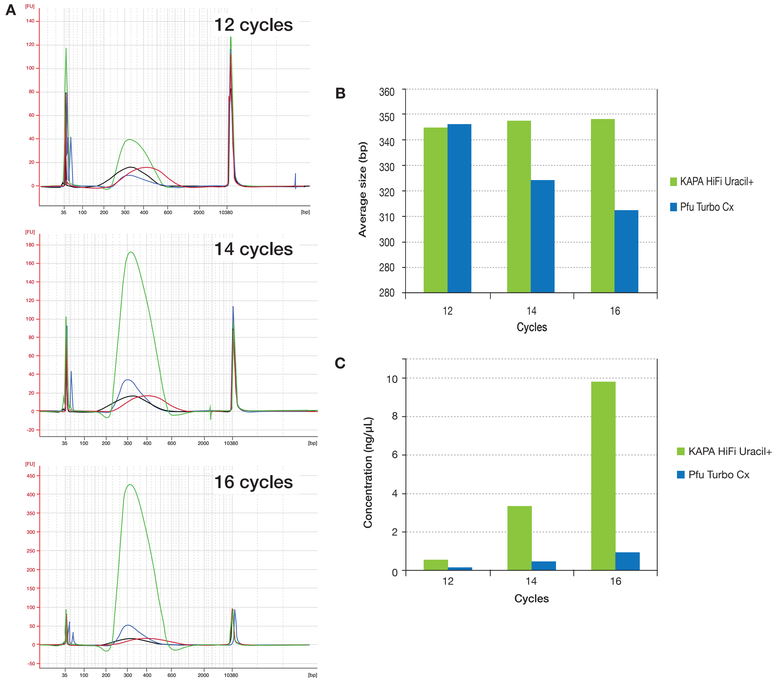
Complete bisulfite conversion of DNA results in DNA damage and low DNA recovery. During bisulfite treatment, cytosine is converted to uracil, which base-pairs with adenine, and subsequent PCR amplification replaces unmethylated G:C base pairs with A:T base pairs. Inhibitor carry-over, DNA damage, and high AT-content typically result in low PCR yields and detrimental size- and GC- biases during library amplification.

To compare KAPA HiFi Uracil+ with Agilent Pfu Turbo Cx, we amplified an Illumina TruSeq human whole-genome BS sequencing (WGBS) library and examined PCR yields and size bias.



A library containing 200 – 400 bp inserts was prepared from 5 µg of human gDNA using the KAPA Library Preparation Kit. Following BS conversion, the library was quantified using the KAPA Library Quantification Kit, which indicated that ~2.7% of amplifiable library material remained.

Undiluted BS-treated library DNA was amplified using KAPA HiFi Uracil+ or Agilent Pfu Turbo Cx and analysed using a Bioanalyzer 2100 High Sensitivity DNA chip.



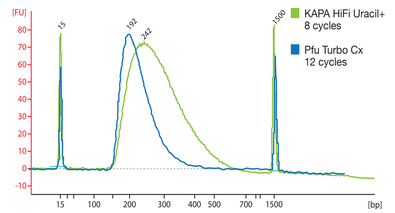
KAPA HiFi Uracil+ provides higher yields and minimal size bias in comparison with Agilent Pfu Turbo Cx.

Human WGBS libraries were amplified using standard protocols with either 12, 14 or 16 cycles, and the amplified libraries analysed using a Bioanalyzer 2100 High Sensitivity DNA chip (A). In comparison with Agilent Pfu Turbo Cx, KAPA HiFi Uracil+ produced much higher yields (B) with very little size bias (C).

WGBS OF AN AT-RICH EUKARYOTIC PARASITE

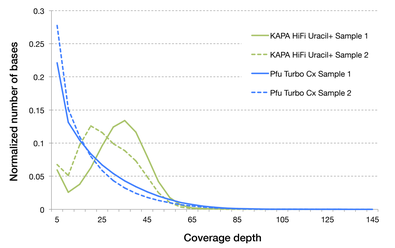
Researchers at the Wellcome Trust Sanger Institute and The University of Aberystwyth are collaborating in an ongoing effort to generate whole-genome bisulfite sequence (WGBS) data for a small, AT-rich (~65% AT), eukaryotic parasite genome.

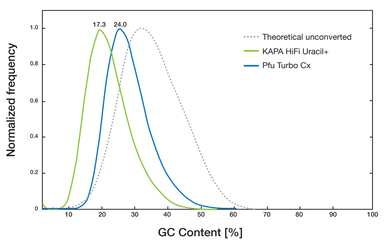
KAPA HiFi Uracil+ was compared with Pfu Turbo Cx for the construction of WGBS libraries, and was found to produce higher yields of amplified library material after fewer PCR cycles. Library fragments amplified with KAPA HiFi Uracil+ also displayed a wider size distribution, comprising a greater proportion of longer fragments.



Parasite WGBS libraries amplified using KAPA HiFi Uracil+ show higher yields and larger average fragment sizes.

A parasite WGBS library was amplified with KAPA HiFi Uracil+ or Pfu Turbo Cx, adjusting the number of PCR cycles for similar yields (8 cycles or 12 cycles, respectively). Amplified libraries were analysed using a Bioanalyzer 2100 High Sensitivity DNA chip.





KAPA HiFi Uracil+ provides greater coverage depth uniformity.

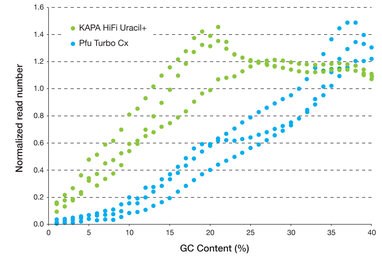
Replicate WGBS libraries were prepared from parasite gDNA, bisulfite-treated, and amplified using KAPA HiFi Uracil+ or Pfu Turbo Cx before paired-end sequencing on Illumina GAIIx.

KAPA HiFi Uracil+ provides improved representation of AT-rich sequences.

Replicate WGBS libraries were prepared from parasite gDNA, bisulfite-treated, and amplified using KAPA HiFi Uracil+ or Pfu Turbo Cx before paired-end sequencing on Illumina GAIIx.

IMPROVED HUMAN BISULFITE SEQUENCING

KAPA HiFi Uracil+ was assessed for BS sequencing of human DNA at the USC Epigenome Center Data Production Facility.



KAPA HiFi Uracil+ provides improved representation of AT-rich sequences for bisulfite-converted human DNA libraries.

100 ng of bisulfite treated library DNA was amplified (8 cycles) with Agilent Pfu Turbo Cx or KAPA HiFi Uracil+.

We would like to acknowledge the following people for their contributions to this work:

M Berriman[2], J Burton[2], RC Clark[2], AJ Coffey[2], NE Holroyd[2], A Palotie[2,4], CE Scott[2], JI Tsar[2], KK Geyer[3], KF Hoffmann[3], MT Swain[3] and C Nicolet[5]

2 Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, United Kingdom.

3 Institute of Biological, Environmental and Rural Sciences (IBERS), Aberystwyth University, Penglais Campus, Aberystwyth SY23 3DA, UK.

4 Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland.

5 Epigenome Center Data Production Facility, University of Southern California, Los Angeles, USA.