KAPABIOSYSTEMS

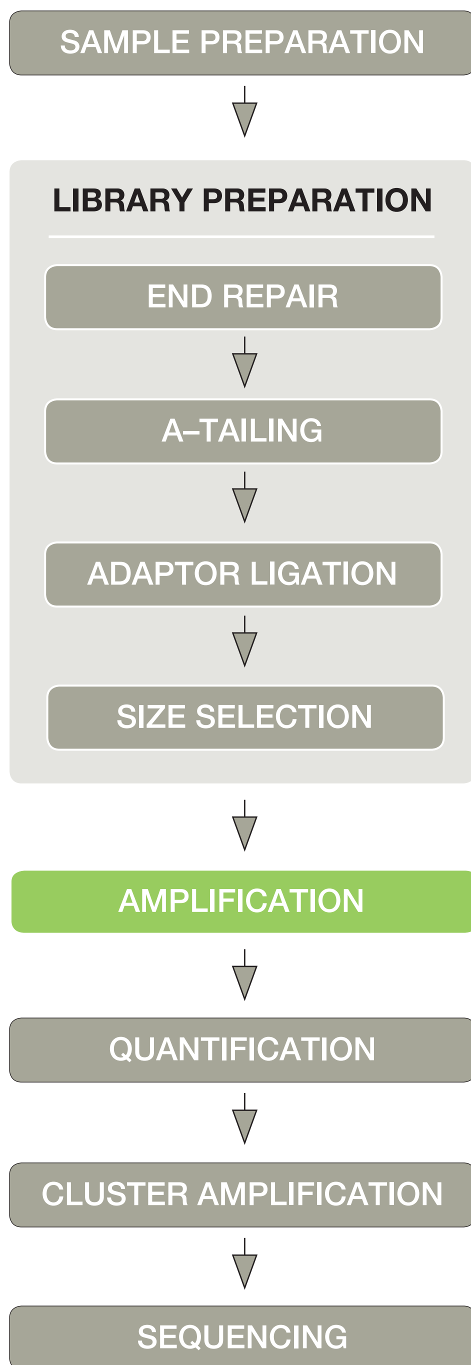
KAPA Library Amplification Kits

Reduce amplification bias and improve coverage.

KAPA HiFi Library Amplification Kits contain a novel DNA polymerase engineered for:

- Improved amplification of GC- and AT-rich genomic regions.
- Reduced enzyme bias resulting in improved sequencing coverage.
- Industry leading fidelity.

Illumina Sequencing Workflow

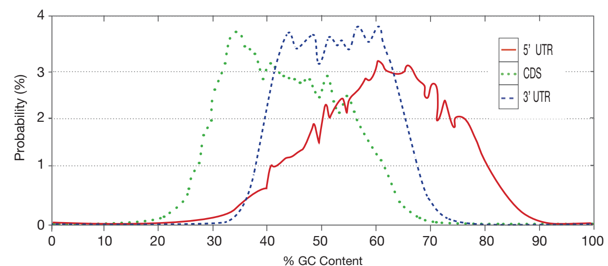


KAPA HiFi Library Amplification Kits

High fidelity PCR is used to selectively enrich library fragments carrying appropriate adaptor sequences and to amplify the amount of DNA prior to sequencing. During PCR enrichment, the polymerase does not synthesize all library fragments with equal efficiency. This amplification bias exacerbates uneven sequence coverage.

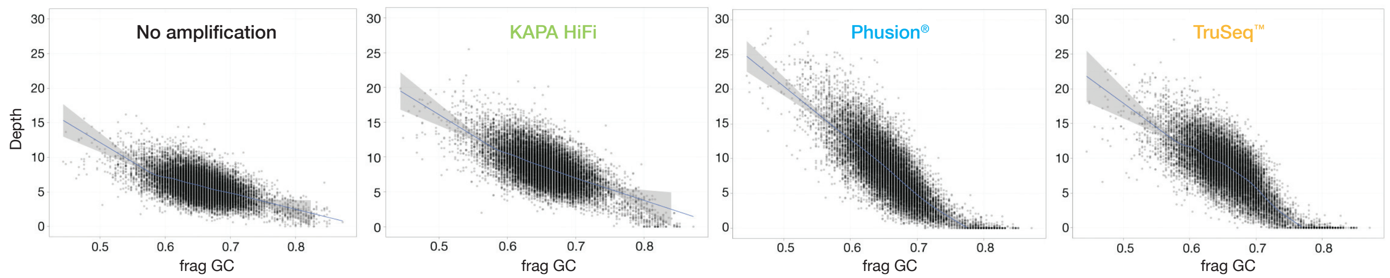
GC content is known to be an important factor in NGS library amplification bias (Aird *et al.* Genome Biology 2011, 12:R18). DNA polymerases commonly used for library amplification display significant bias against GC- and AT-rich regions. Genomes with extreme average GC content are particularly susceptible to amplification bias. High-GC content elements such as 5' UTRs, CpG islands and first exons, contained within the balanced human genome, are also vulnerable to amplification bias (Zhang L *et al.* PNAS 2004; 101:16855-16860).

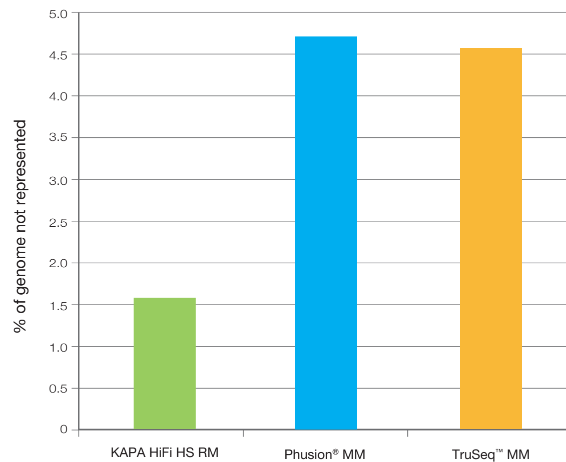
Distributions of GC content in 5' UTRs, CDS, and 3' UTRs of human genes.



KAPA HiFi Library Amplification Kits have been designed to address PCR-induced bias. Kits contain the novel KAPA HiFi DNA Polymerase, engineered for high fidelity and processivity and capable of balanced amplification of complex library DNA. Kits are supplied as a ready-to-use master mix (2X) containing all components for PCR, except primers and template.

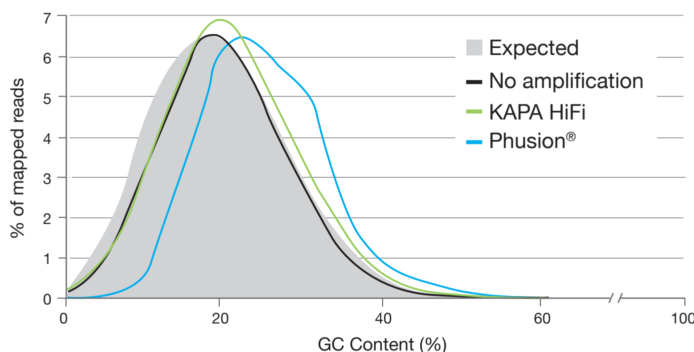
Effect of high-GC content on coverage depth for libraries amplified using common proof-reading (B-family) polymerases

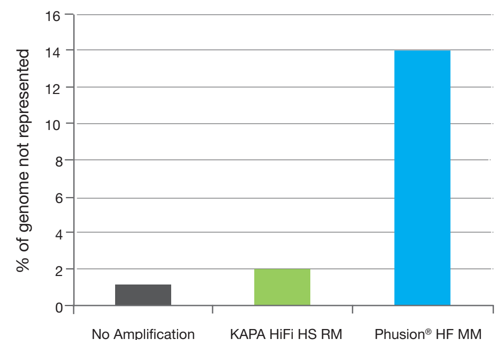




Indexed Illumina TruSeq™ libraries prepared from identical sheared *M. tuberculosis* (65% GC) gDNA were amplified using the indicated PCR reagents, and compared to an equivalent unamplified library by paired-end sequencing (2 x 75 bp). After filtering and aligning read pairs to reference sequences, 250 000 read pairs were randomly sampled for each genome, and scatter plots of mean sequence coverage depth vs. GC content were generated by analyzing 250 bp windows. GC-rich *M. tuberculosis* sequences were under-represented following library amplification using either Phusion® HF Master Mix or Illumina TruSeq™ PCR Master Mix. In contrast, library amplification with KAPA HiFi HotStart Master Mix resulted in coverage distribution across the range of GC-content that is almost indistinguishable from that of the unamplified control.

Low GC-content libraries result in variable bias depending on the polymerase used for amplification

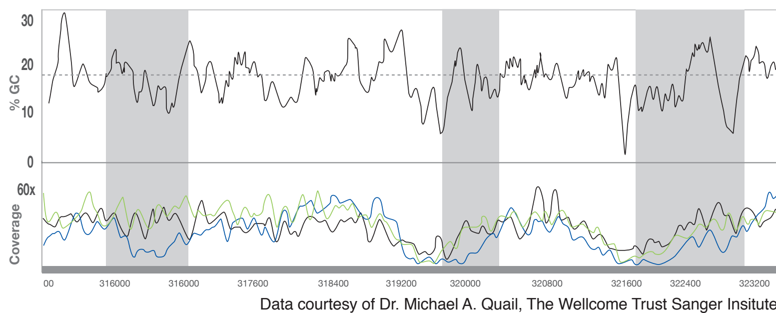




Libraries prepared from identical sheared *P. falciparum* (19% GC) gDNA were amplified using the indicated PCR reagents, and compared to an equivalent unamplified library. Observed frequencies of GC-content for reads are plotted for each condition tested (black = unamplified; green = KAPA HiFi HotStart Master Mix; blue = Phusion® HF Master Mix). The expected frequency distribution of reads is indicated by the grey shaded area. The unamplified library tracked the expected frequency distribution. Amplification with KAPA HiFi showed minimal bias while amplification with Phusion® resulted in a dramatic bias against reads with low GC-content. Average coverage depth for each library was 16.0x (unamplified control); 16.5x (KAPA HiFi); 18.8x (Phusion®). Data courtesy of Dr. Michael A. Quail, The Wellcome Trust Sanger Institute.

Library amplification can dramatically affect coverage uniformity

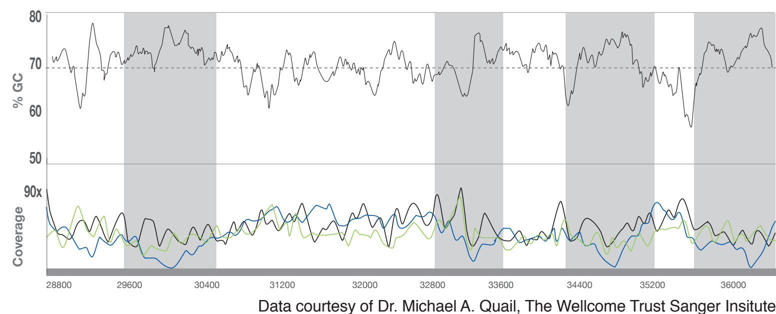
The following Artemis screen captures depict examples of coverage bias in libraries amplified with either KAPA HiFi HotStart Master Mix or Phusion® HF Master Mix, compared to an unamplified control library. In short stretches of either high or low-GC content, the degree of coverage bias varies according to the method used to amplify the library.



Data courtesy of Dr. Michael A. Quail, The Wellcome Trust Sanger Insitute

Coverage depth and GC content across a ~7kb region of the *P. falciparum* genome.

Within this region of the genome there are 3 locations where high-AT sequences (>80%) lead to coverage bias (grey bars). In all three regions coverage depth drops significantly after amplification with Phusion® (blue), while the library amplified using KAPA HiFi (green) shows more uniform coverage depth which tracks that of the unamplified control library (black).

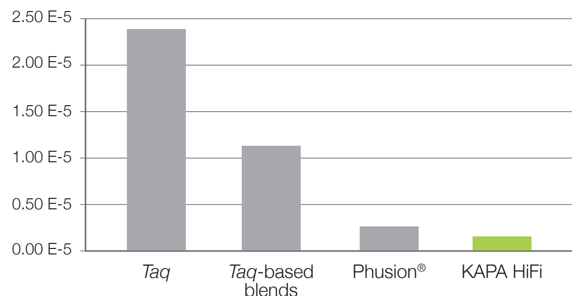


Data courtesy of Dr. Michael A. Quail, The Wellcome Trust Sanger Insitute

Coverage depth and GC content across a ~7kb region of the *B. pertussis* genome.

Within this region of the genome there are 4 distinct locations of high-GC sequence (>75%) that lead to sequence coverage bias (grey bars). in these regions the library amplified using Phusion® (blue) exhibits lower depth of coverage compared to the unamplified control. In contrast, the library amplified with KAPA HiFi (green) exhibits more even coverage depth, similar to the control library (black).

Superior accuracy for high fidelity library amplification



Error rates of DNA polymerases and blends.

The error rate of KAPA HiFi is calculated at 1 error in 3.54 x 10^6 bases covered (2.82 x 10^{-7}). The error rate of KAPA HiFi is 100X lower than *Taq* polymerase, 40X lower than polymerase blends and 2X lower than Phusion®.

Phusion® is a registered trademark of Thermo Fisher Scientific Inc. TruSeq™ is a trademark of Illumina Inc.

For sales please contact
sales@kapabiosystems.com
or call your local representative.

For more information on our complete range of products please visit **www.kapabiosystems.com**

ORDERING INFORMATION

Description	Code	Kit contents
KAPA HiFi Library Amplification Kit	KK2611	50 rxn
KAPA HiFi Library Amplification Kit	KK2612	250 rxn