Hairpin structure facilitates high-fidelity DNA amplification reactions in both qPCR and high-throughput sequencing

Kerou Zhang[1], Alessandro Pinto[3], Peng Dai[1], Michael Wang[1], Lauren Yuxuan Cheng[1], Ping Song[1], Luis Rodriguez[3], Cailin Weller[3], and David Yu Zhang[1,2,*]

[1] Department of Bioengineering, Rice University, Houston, TX, 77030, USA

[2] Systems, Synthetic, and Physical Biology, Rice University, Houston, TX, 77030, USA

[3] Nuprobe USA, 2575 W Bellfort Street, Houston, TX, 77054, USA

* To whom correspondence should be addressed. Tel: +1 713 348 2832; Fax: +1 713 348 5877; Email: genomic.dave@gmail.com

Present Address: David Yu Zhang, Nuprobe USA, 2575 W Bellfort Street, Houston, TX, 77054, USA

Abstract

Effective polymerase chain reactions (PCR) are important in bio-laboratories. It is essential to detect rare DNA-sequence variants for early cancer diagnosis or for drug-resistance mutations identification. Some of the common detection quantitative PCR (qPCR) methods are restricted in the limit of detection (LoD) because of the high polymerase misincorporation rate in Taq DNA polymerases. High-fidelity (HiFi) DNA polymerases have a 50- to 250-fold higher fidelity. Yet, there are currently no proper designs for multiplexed HiFi qPCR reactions. Moreover, the popularity of targeting highly multiplex DNA sequences requires minimizing PCR side products, as the potential of dimerization grows quadratically as the plexes of primers increases. Efforts tried before were either an add-on step, or technology-specific, or requiring high-level computing skills. There lacks an easy-to-apply and cost-effective method for dimerization reduction. Here, we presented the Occlusion System, composed of a 5'-overhanged primer and a probe with a short-stem hairpin. We demonstrated that it allowed multiplexing high-fidelity qPCR reaction, it was also compatible with the current variant-enrichment method to improve the LoD by 10-fold. Further, we found that the

Occlusion System reduced the dimerization up to 10-fold in highly multiplexed PCR. Thus, the Occlusion System satisfactorily improved both qPCR sensitivity and PCR efficiency.

Introduction

The Polymerase chain reaction (PCR) has been one of the most widely used methods in bio-laboratories. It could be used in conjunction with intercalating dyes or fluorophore-labeled probes to allow quantitative and real-time detection of nucleic acid targets using a fluorescence detection device, or it could be used upstream of a high-throughput sequencing reaction for target enrichment. Compared with the high-throughput sequencing methods, quantitative PCR (qPCR) reduced the total turnaround time with maintained or increased analytical sensitivity and reproducibility, thus benefiting the molecular diagnostics of diseases requiring timely diagnosis and treatment with clinical sensitivity and specificity, like clinical oncology and clinical microbiology[1–9].

Profiling nucleic acid variants with low allele frequencies, such as cancer mutations from liquid biopsy samples, drug-resistance mutations or pathogenic mutations from different types of clinical specimens[9], is clinically essential. Genetic alternations could result in a mixed population of mutant and normal cells with the tissue, which could progressively lead to oncogenesis[10]. Earlier diagnosis could be achieved with more sensitive detection of the low-abundance mutations for tumor indication, thus better prognosis and precision medicine[11,12]. Liquid biopsy has revealed an advantage over traditional gold standard tissue biopsy in its non-invasiveness, good concordance in genome profiling, and compatibility with different sample types[13,14]. However, somatic mutations detected from the liquid biopsy usually have lower variant allele frequency (VAF) than tissue biopsy[15], which added to the difficulty of sensitive and accurate detection of early-stage somatic mutations. Besides, liquid biopsy could offer a potential solution for longitudinal mutation profiling in diseases like minimal residual disease (MRD)[16,17], which usually requires sensitivity down to $1:10^4$ to $1:10^6$. Similarly, the need to detect low mutation rates within the drug-resistant *Mycobacterium tuberculosis* strains[18] or drug-resistance colorectal cancer[19] has been reported.

As the need to detect mutations with VAF low to 0.1% or even low to 0.01% has increased, the common qPCR diagnostic methods, which are mainly using polymerases lacking proofreading activity, are not satisfactory anymore. Current common qPCR diagnostics methods, like blocker displacement

amplification (BDA), allele-specific PCR (As-PCR)[20,21], ARMS-PCR[22,23], Cast-PCR[24,25], COLD-PCR[26] and etc., mainly have a limit of detection (LoD) around 0.1% to 1% VAF[27,28]. The high polymerase misincorporation rate would cause false-positive mutations thus obscure results[11] in qPCR method like BDA, restricting LoD (Figure S1A). Allele-specific blocker displacement amplification (As-BDA)[29] has been reported as an method capable of detecting and profiling multiplexed mutations down to 0.01% VAF within one qPCR reaction using Taq DNA polymerases. However, As-BDA relies much on the preliminary knowledge of target information, limiting its application in discovery of new biomarkers. Though there have been some bioinformatics designs[30] or methods, like incorporating unique molecular identifier[31,32] or single-molecule sequencing (SMRT-seq)[11,33], to minimize the effect of polymerase error in sequencing-based methods, there is not yet an effective solution in qPCR. qPCR usually has a simpler workflow, faster turnaround time, less cost, and more accessible instruments over the sequencing-based methods. High-fidelity DNA polymerases possess the 3' to 5' exonuclease activity known as proofreading features[34–37], leading to 50- to 250-fold increased fidelity[11]. Hence, high-fidelity DNA polymerases have the potential to achieve more accurate and sensitive detection in qPCR (Figure S1B). Besides intercalating dye as a fluorescence indicator, polymerases without proofreading activity are usually compatible with multiple methods, including molecular beacons[38], Sunrise primers[39], Scorpion primers[40], hydrolysis (TaqMan) probes[41] or fluorescence resonance energy transfer (FRET) probes[42], to indicate the production of PCR amplicons. TaqMan probes and etc. are generally favored over intercalating dyes because multiplex different spectrally distinct fluorophores can be used on different probes simultaneously to allow multiplexed detection and quantitation of 2-6 distinct DNA target species in a single reaction. However, high-fidelity DNA polymerases have only been reported to work with limited choice[43]. Designs like TaqMan probes and Sunrise primers have been demonstrated to be incompatible with high-fidelity DNA polymerase (Figure S2) for single target detection, let alone multiplexing. A key reason is that the 3' to 5' exonuclease activity of the high-fidelity polymerases would chew back and digest the 3' mismatched sequence or 3' chemical modification between the primer or probe and the template, and also high-fidelity polymerases are usually lack of 5' to 3' exonuclease activity to produce detectable signal with most of the probe designs. Thus, finding an appropriate probe design compatible with high-fidelity DNA polymerases, cost-efficient, easily accessible, and plex-scalable in qPCR has been challenging.

The increasing demand for targeting highly multiplex DNA sequences in biomedical research has raised a necessity for minimizing the production of PCR side products, like primer-dimer and non-specific amplicons, to maximize usage of the sequencing reads. The potential of dimerization grows quadratically with the number of total primers increases[44]. The current multiplex PCR method AmpliSeq[45] is primarily composed of enzymatic digestion of excessive primer sequences post amplicon-generation PCR, followed by a ligation step of adaptor sequences. Other methods reported[44,46] used a second PCR step to append adaptor sequence on the amplicons, which saved hands-on time with higher efficiency as the ligation efficiency is usually less than 30%. However, both methods sophisticate the total sequencing workflow. Furthermore, the addition of adaptor sequences increases the possibility of another round of dimer formation and the non-specific binding. Efforts have been made to prevent or minimize primer-dimer formation during PCR reactions, like optimizing PCR condition[47], upgrading primer-design software[44], modifying the structure of primers[48] and etc. Alternatively, additional steps have been added in the library preparation steps, like size-selection purification, enzymatic clean-up and etc., to remove dimer and other PCR by-product. But they are either only case-specific, need repetitive optimization for new panels, or require computing skillsets, or add on extra hands-on time and overall cost. Therefore, there is an urgency to develop an easy-to-use and generalizable solution decreasing dimerization during PCR and streamlining the library preparation workflow.

Here, we present the Occlusion System, composed of Occlusion Primer and Occlusion Probe, provides a straightforward solution for detecting multiplexed targets in high-fidelity DNA polymerase qPCR reaction by functioning as a probe-based fluorescence indicator tool. Moreover, the Occlusion System, is compatible with other variant enrichment (diagnostic) methods like BDA[27], in both single-plex and multi-plex settings. We demonstrated that, in terms of analytical performance, the Occlusion System with BDA in high-fidelity qPCR reactions (HiFi BDA Occlusion System) was sensitive in VAF down to 0.01% using synthetic variants, which was improved by 10-fold compared with Taq-based BDA qPCR reactions, and accurate in VAF calling with a percentage error less than 23.1% on commercial reference samples. We applied the HiFi BDA Occlusion System in peripheral blood mononuclear cell (PBMC) samples from 27 acute myeloid leukemia (AML) patients and 15 healthy donors, and detected 6 FLT3 mutations, 2 IDH1 mutations in 27 patients. Comparative results between the HiFi BDA Occlusion qPCR and ddPCR in *FLT3*

gene showed 100% concordance in mutations over 0.1% VAF with the similar amount of DNA input. Further, the HiFi BDA Occlusion System was advantageous in detecting mutation with VAF down to 0.01% in clinical samples, by increasing the amount of DNA input to improve the LoD. We also found that the Occlusion System was promising in improving PCR efficiency and increasing the on-target rate of NGS libraries by reducing the formation of primer-dimer in the NGS. By applying the Occlusion System in two medium-sized NGS panels (80-plex and 179-plex), the overall on-target rate was increased up to 10-fold. Furthermore, we demonstrated that the Occlusion System is compatible with both Taq-based DNA polymerase and high-fidelity DNA polymerase in decreasing dimerization in sequencing library preparation. In brief, the Occlusion System improved clinical detection limit by 10-fold together in variant detection qPCR reactions by enabling previous variant enrichment method with multiplexed detection with high-fidelity DNA polymerases, and enhanced on-target rate up to 10-fold in high-throughput sequencing by reducing adaptor-associated non-specific products.

Materials and Methods

Study samples

Twenty-seven AML PBMC samples were purchased from Discovery Life Science (DLS) with written informed patient consent. Fifteen PBMC samples from healthy volunteers were purchased from Zen-Bio, who collected the samples with written informed patient consent. IRB approval was not required as these are Exemption 4 (commercial, de-identified samples). Detailed sample information is summarized in Tables S3 and S4.

Oligonucleotides and repository samples

Primers, BDA blockers, Occlusion primers, Occlusion probes and synthetic double-stranded DNA fragments (gBlocks) were purchased from Integrated DNA Technologies (IDT). Detailed design sequence was summarized in Tables S16 and S17. Primers, BDA blockers and gBlocks were purchased as standard desalting-purified, Occlusion primers and Occlusion probes as HPLC-purified; they were resuspended in 1× IDTE buffer (10 mM Tris, 0.1 mM EDTA) in 100 uM (gBlocks were resuspended in 10 ng/μL) as the original stock and stored at 4 °C. The dilution buffer composed of 100ng/μL carrier RNA (Qiagen, 1017647)

in 1× IDTE buffer with 0.1% TWEEN 20 (Sigma Aldrich), was used to serially dilute gBlocks. Human cell-line gDNA NA18562 and NA18537 repository samples (Coriell Biorepository), Commercial reference gDNA samples (Tru-Q 7 (1.3% Tier) Reference Standard (HD734) and Myeloid DNA Reference Standard (HD829)) (Horizon Discovery) were diluted with 1x IDTE buffer (IDT). All the dilutions and original stock were stored at 4 °C for short-term usage, −20 °C for long-term storage except Horizon reference gDNA samples (both dilutions and original stock were stored at 4 °C).

DNA extraction from PBMC samples

DNA was extracted from PBMC samples using the DNA Blood Mini kit (Qiagen, 51104). NanoDrop spectrophotometer was used to measure the yield of DNA. All DNA materials were stored at -20℃ until ready for analysis.

Reference material preparation

Synthetic gBlocks, used for mimicking natural mutations, were quantified by qPCR. The concentration of gBlocks were then adjusted by comparing cycle threshold (Ct) values with that of 40 ng/μL human gDNA NA18562. Then gBlocks were further diluted with gDNA NA18562 to prepare reference samples of 10%, 5%, 1%, 0.3%, 0.1%, 0.03%, and 0.01%VAF by serial dilution. 10% and 1% VAF reference samples were quantitated in NGS before further dilution to lower-VAF samples.

Occlusion qPCR protocol

All Occlusion qPCR assays were performed in the CFX96 Touch Real-Time PCR Detection System using 96-well plates. Phusion Hot Start Flex 2X Master Mix (Thermo Scientific) was used for each qPCR reaction. Roughly 200ng DNA in 5 μL were loaded as input into each reaction. The thermocycling program started with 30 s polymerase activation and initial denaturation at 98 ˚C, followed by 55 cycles of 10 s at 98 ˚C for DNA denaturing, 30 s at 63 ˚C for annealing and 30 s at 72 ˚C for extension (98 ˚C: 30s - (98 ˚C: 10 s - 63 ˚C: 30 s - 72 ˚C: 30 s) x 55). Fluorescence signal was collected at 63 ˚C per cycle. All the reactions were conducted in duplicate or triplicate. Details of composition concentration are summarized in Tables S13 to S15.

ddPCR quantitation protocol

ddPCR mutation assays from Bio-Rad were used to quantitate mutations in the *FLT3* gene [dHsaMDS2514588 for *FLT3* 2503G>A (D835N), dHsaMDS420144135 for *FLT3* 2508_2510del (I836del), dHsaMDV2010047 for *FLT3* 2503G>T (D835Y), dHsaMDV2510492 for *FLT3* 2503G>C (D835H)]. The whole workflow was conducted on a Q200 Droplet Digital PCR System (Bio-Rad). A 20 µL reaction containing 10 µL 2X ddPCR Supermix for Probes (No dUTP) (Bio-Rad, 1863024), 1 µL 20x target (FAM) and wildtype (HEX) primers/probe and roughly 20ng samples was prepared and loaded into the middle wells of DG8 cartridges (Bio-Rad, 1864008). 70 µL of the Droplet Generation Oil for Probes (Bio-Rad, 1963005) was then added to the top wells of the cartridges. The DG8 cartridges were put in the QX200 Droplet Generator (Bio-Rad, 10031907) for droplet generation. Then the thermo cycling, with all ramp speed set at 2 ˚C/sec, started with 10 min at 95 ˚C, followed by 40 repeated cycles of 30 s at 94 ˚C and 1 min at 55 ˚C. There is a 10 min incubation step at 98 ˚C following the repeated cycle before holding the plate at 4 ˚C until the next step. The plate was then loaded onto the QX200 Droplet Reader (Bio-Rad, 1864003) for droplet counts. QuantaSoft software was used to collect and analyze the data. Further data analysis on MATLAB used the code published previously[27].

Occlusion NGS protocol

Occlusion NGS for 80-plex panel[46] and 179-plex[49] starts by mixing the sample with the adaptor primer mixes, possibly mixed with the Occlusion probes as noted, and Phusion Hot Start Flex 2X Master Mix. The forward and reverse adaptor primers were each at a final reaction concentration of 15 nM in a reaction volume of 50 µL.

The first thermocycling protocol for the 80-plex panel included an initial 98 °C denaturation for 30 s, followed by two cycles of 98°C for 10s, 60°C for 5min and 72°C for 2min, and a final extension at 72°C for 5min. The products were then purified using Ampure Beads (Beckman Coulter) at a ratio of 1.6X and eluted in water at 30 µL. Then the index PCR was performed. Each 50 µL reaction included 500 nM each Nextera index, 25 µL Phusion Hot Start Flex 2X Master Mix and 10 µL purified template. The thermocycling protocol is: 98 °C for 30 s, followed by repeated cycles of 98°C for 10s, 63°C for 1min and 72°C for 1min, and a final extension at 72°C for 5min. The number of repeated cycles for library without Occlusion probe

is 25, that for library with Occlusion probe is 30. The indexed libraries were size selected with ratios of 0.7× and 0.3× Ampure beads were used.

The thermocycling protocol for the179-plex panel included an initial 98 °C denaturation for 30 s, followed by two cycles of 98°C for 10s, 63°C for 30min and 72°C for 15s, and second two cycles of 98°C for 10s, 63°C for 15s and 72°C for 15s, then five cycles of 98°C for 10s, and 71°C for 30s. A mixture of two 1.5μM universal primers would be added into the reactions which are still kept inside the thermo cycler. The amplicon products were then bead purified at a ratio of 1.6X and eluted in 30 μL. Then the Truseq index PCR was performed with 250 nM each index in each 50 μL reaction using iTaq Universal Probes Supermix (Bio-Rad). The thermocycling protocol is: 95 °C for 3 min, followed by repeated cycles of 95°C for 10s, 65°C for 30s, and a final extension at 65°C for 2 min. The number repeated cycle for library without Occlusion probe is 15, that for library with Occlusion probe is 19.

Then the concentrations of libraries were quantified by Qubit, the length of libraries was measured in Bioanalyzer using high sensitivity DNA chips for QC. Libraries were sequenced in 2 x 150 bp pair-end sequencing in Illumina Miseq instrument.

Results

Hairpin-based Fluorescence indicator

To find a probe-based design as the fluorescence indicator in high-fidelity qPCR reactions, like the TaqMan probe, we previously attached a fluorophore modified overhang sequence to the 5' of the primer sequence, paired with a companion strand which is reverse complementary to the overhang sequence and attached with a quencher. However, the quencher was digested by the proofreading feature of high-fidelity polymerase and then generated an abnormal amplification curve from the initial cycle as shown from the experimental result (Figure 1A).

Inspired by RNA hairpin in intrinsic transcription termination[50–53], where the polymerase would fall off and end the transcription once meeting the hairpin termination structure, we empirically found that hairpin structure efficiently hindered the proofreading feature of high-fidelity DNA polymerases. Then we hypothesized that this hairpin structure could be further compatible with the fluorophore or quencher modification. Applying a hairpin structure in the companion strand, though the quencher was well protected,

raises a new problem that the polymerase could extend on the companion strand due to the inappropriate length of the hairpin stem (20nt in Figure 1B). As known, a 10 to 20nt priming region is for the initiation of nucleic acid synthesis[54]. By adjusting the length of the stem to 4nt, we finally found a promising hairpin structure, which could sufficiently protect the companion strand from quencher digestion and self-extension (Figure 1C).

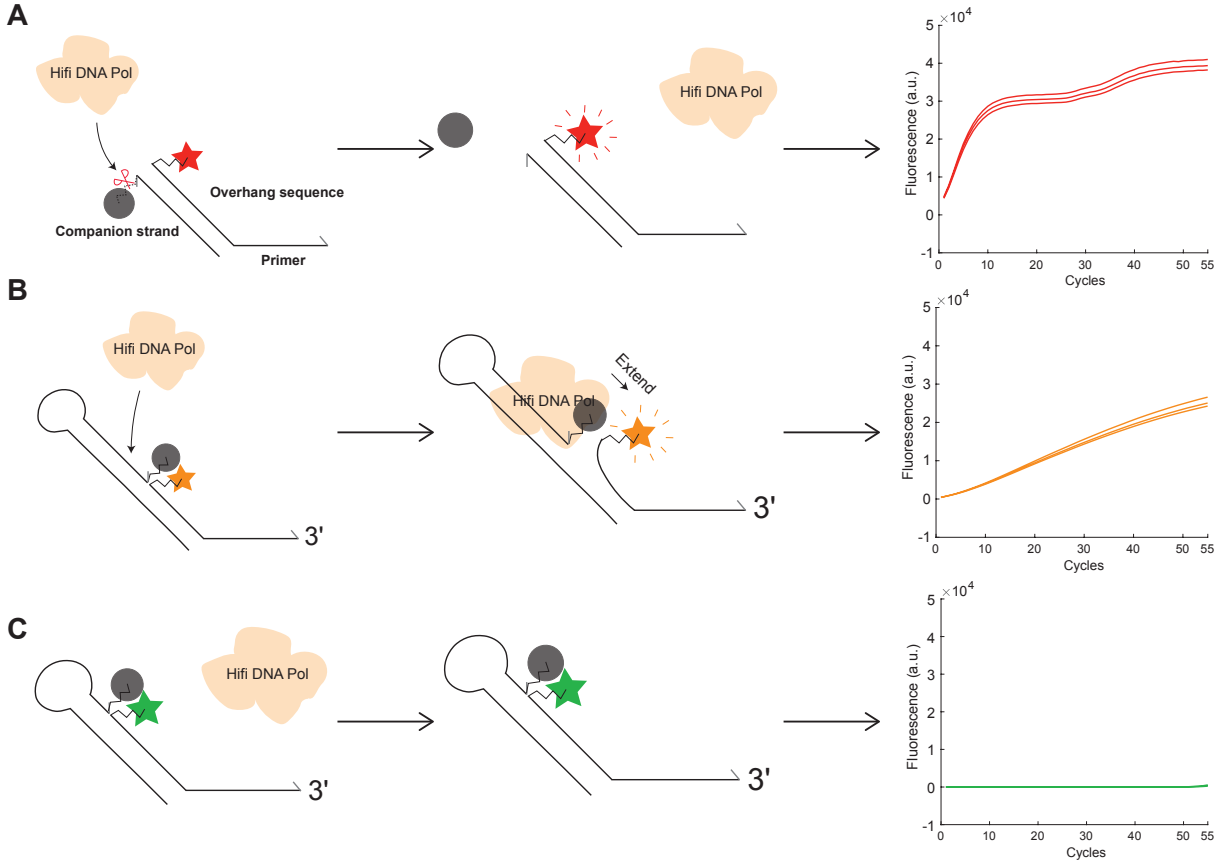

Figure 1. Hairpin structure effectively hindered the proofreading feature of high-fidelity (HIFI) DNA polymerases.

A). Chemical modification on the companion strand would be digested by HIFI DNA polymerases. The companion strand is reverse complementary to the overhang sequence on the primer, and the 5' overhang sequence is modified with a fluorophore modification. The companion strand has a 3' quencher modification. This design failed to function as a fluorescence indicator, because the HIFI DNA polymerases would digest the 3' quencher modification, resulting in abnormal increased signal even with the absence of any template sequences.

B). Hairpin structure on the companion strand protect the 3' chemical modification. By applying a hairpin structure with 20nt stem to the 3' of the companion strand, the digestion of HIFI DNA polymerases

was effectively prohibited. However, the long stem structure of the hairpin provided a binding site for the HIFI DNA polymerases to extend on, resulting in linear increase of fluorescent signal without any template sequences.

C). Hairpin with a short stem reveals optimal performance for preventing self-extension and 3' digestion. Adjusted to a 4nt-stem structure, the hairpin on the 3' of the companion strand promisingly functions as a fluorescence indicator. When there is no template present, there is not any abnormal fluorescence generated.

Overview of the Occlusion System

The Occlusion System is composed of an Occlusion Primer and an Occlusion Probe, Occlusion Probe is originated from the companion strand as described above. As shown in Figure 2A, during the polymerase extension on the reverse primer, the Occlusion probe would be strand-displaced, the separation of fluorophore and quencher led to the generation of fluorescence. We demonstrated the feasibility of the Occlusion System in detecting single target (Figure S3) and multiplexed targets (Table S1). Moreover, we set a design criterion to introduce mismatch(es) among the overhang binding region to weaken the occupancy of the DNA polymerases, thus improving the PCR efficiency (Figure S4).

Then we applied the Occlusion System to the BDA system, as the schematic was shown in Figure S5. In the absence of any other fluorescence indicator, the Occlusion System is compatible with diagnostic methods like the BDA. The result reveals non-overlapping fluorescence separation between samples with different VAFs and clear background with imperceptible noise in the fluorescent channel (Figure 2B, Figure S6). We plotted Ct values of each VAF sample as each data points versus logarithmic VAF values, to generate a fitted curve from linear regression. The fitted curves exhibited a strong linear correlation with r-squares ranging from 0.896 to 0.999 (Figure 2C and Figure S7). Each 10% and 1% VAF sample was NGS quantified (Table S2), and each 0.01% VAF sample was Sanger sequenced post reaction (Figure S8 and S9). To see if the Occlusion System allows multiplexed readout in qPCR reactions with BDA, we designed different BDA Occlusion sets to each of the two target genes of interest (Figure 2D). Both channels for each pair could achieve effective variant enrichment. In this way, we showed the Occlusion System is capable of multiplexing qPCR reactions, and compatible with diagnostic methods like BDA, achieving ultra-sensitive variant enrichment with VAF down to 0.01% in high-fidelity qPCR.

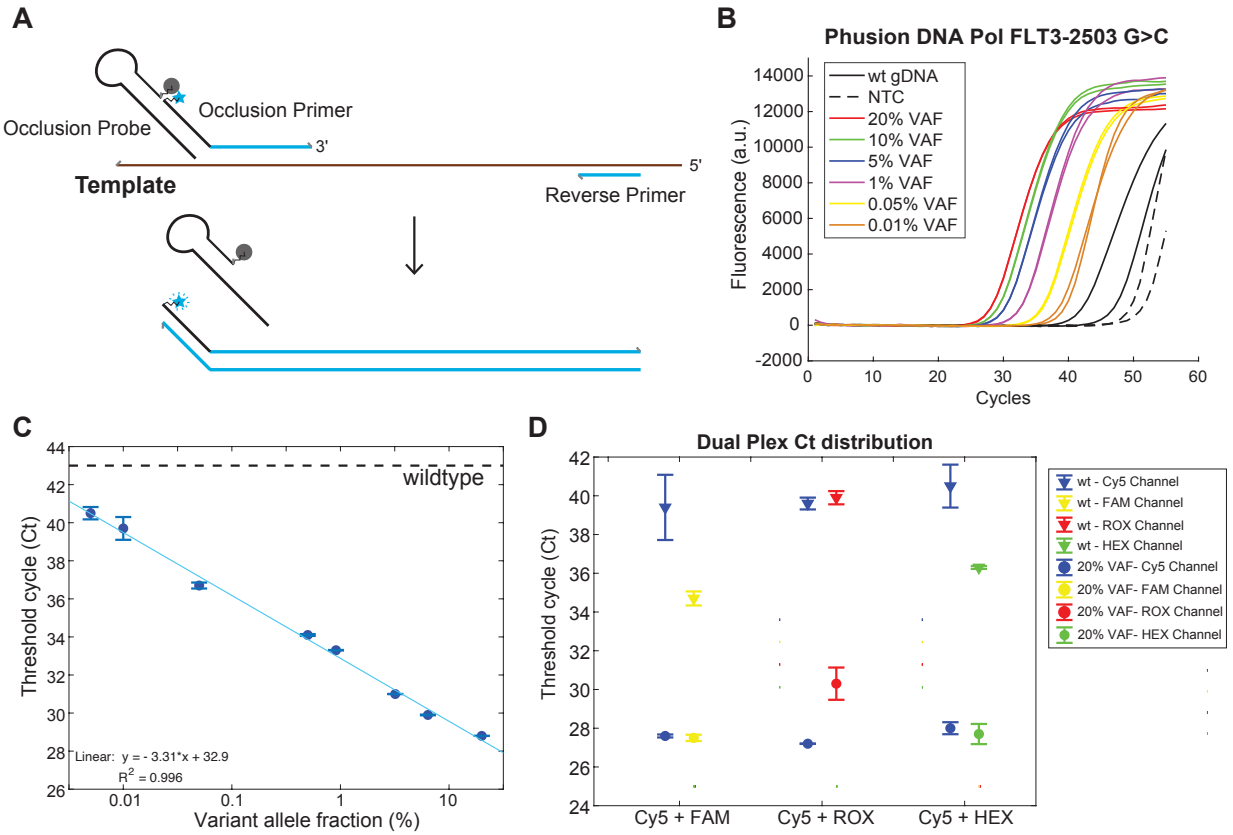

Figure 2. Overview of the Occlusion System.

A). Schematic of the Occlusion System. The Occlusion System mainly comprises an Occlusion Probe and an Occlusion Primer. The Occlusion Probe is originated from the companion strand described in Figure **1C**. The Occlusion Primer has a primer region and a 5' overhang region modified with fluorophore. The Occlusion Primer could bind to the template sequence and extend on it. As the Reverse Primer extends on the strand generated from the Occlusion Primer, it would encounter and then displace the Occlusion Probe, causing the separation of the fluorophore and quencher, thus, the signal would increase, indicating the generation of amplicons.

B). The Occlusion System with blocker displacement amplification (BDA) technology. The compatibility of the Occlusion System with the BDA technology by applying the Occlusion System to the reverse primer of the BDA design, as shown in Figure **S5**. Duplicate qPCR results for various mixtures of synthetic gBlock variants and genomic DNA with different VAFs. Input 200 ng or equivalent to 60,000 copies of molecules per reaction.

C). Threshold cycle (Ct) values summary for the results observed in b. The error bar shows one standard deviation.

D). Multiplexing of the Occlusion System with BDA. By combining two BDA Occlusion sets together inside one qPCR reaction, we could demonstrate the multiplexing ability of the Occlusion System with BDA.

The genes of interest here are *FLT3* and *DNMT3A*. Amplification of *FLT3* region was indicated by Cy5 channel. The FAM, ROX, and HEX channels stand for amplicons from *DNMT3A* gene. The variant templates used are the mixture of 20% VAF FLT3 2503G>C and 20% VAF DNMT3A 2644C>T.

Clinical validation of the HiFi BDA Occlusion System

To further validate whether the fitted curve generated from the mimicking VAF samples is accurate enough for calculating VAF from clinical specimens, we designed BDA Occlusion sets for *IDH1*, *FLT3*, *DNMT3A* gene targets and ran commercial reference samples on these genes (Figure 3A and 3B, Figure S7A). The closeness of the reference samples (shown as the light blue star in Figure 3A, orange star in 3B, and magenta star in S7A) to the fitting curve exhibits the accuracy of the fitting result. The percentage error between the called VAF and claimed VAF is 1% for 5% reference sample, 23.1% and 20.8% for 1.30% reference sample. We expect the sample with lower VAF would have more randomness with the variation from the number of input variant sequences based on Poisson distribution. To demonstrate the compatibility with clinical samples, we performed HiFi BDA Occlusion qPCR reactions on the DNA extracted from PBMC samples of AML patients and healthy individuals. We used the *GAPDH* housekeeping gene to quantify the input of samples, a summary of Ct value was shown in Table S5 to S6. Then we applied *FLT3* and *IDH1* BDA Occlusion designs in the DNA samples extracted from all AML patients and healthy donors, with 6 *FLT3* mutations, 2 *IDH1* mutations reported from 7 individual AML samples (detailed Ct values were collected and summarized in Table S7 to S10. By confirming the mutation pattern through Sanger sequencing (Figure 3D and S10), we successfully called the VAFs of these samples based on the Ct value and the fitted curve of the according gene (Table S11). More, we validated those *FLT3* mutations on ddPCR platform (Figure S11 to S14 and Table S12). As shown in Figure 3C, there was 100% concordance between the VAFs called by the ddPCR and the HiFi BDA Occlusion System in the *FLT3* mutations with VAF higher than 0.1% using 20ng DNA input. The HiFi BDA Occlusion System successfully reported mutations from two more patients with VAF between 0.01% to 0.05% by increasing the total amount of DNA input to 200ng. Thus, we concluded that the HiFi BDA Occlusion System was accurate and sensitive in identifying VAF from unknown samples, showed advantages in detecting mutations down to 0.01% VAF with flexible and

scalable DNA input, and was fully compatible with clinical specimens, indicating a promising application in facilitating clinical diagnosis with a fast turnaround time.

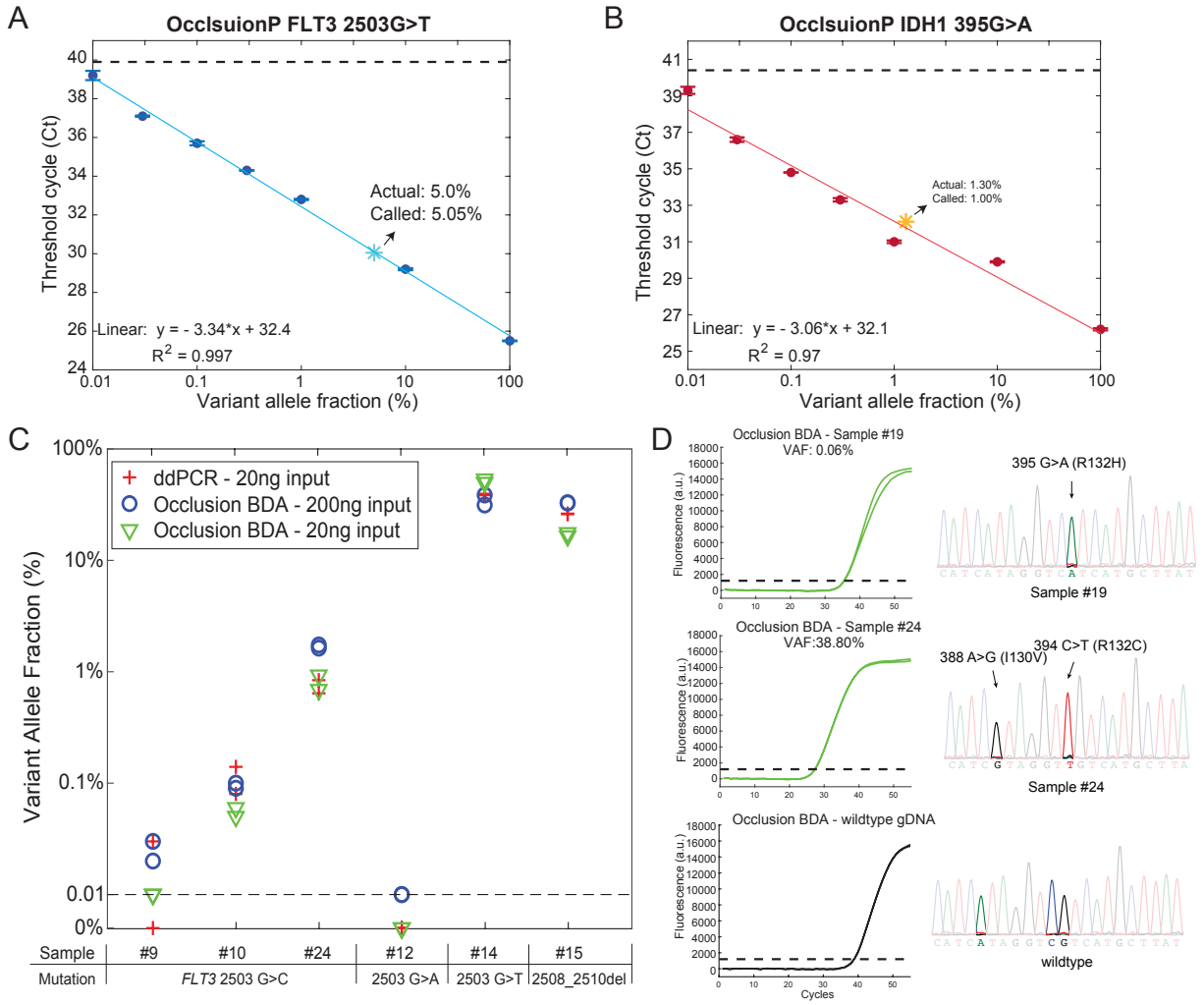

Figure 3. Ct determination validation with commercial reference samples

A). Quantitation validation from standard VAF curve of the *FLT3* 2503G>T mutation with a 5.0% reference sample. Ct value of each VAF sample was plotted as each individual blue dots with one standard deviation as error bar. The solid blue line was the fitted line, showing a strong linear correlation between log(VAF) and Ct values with an r-square of 0.997. The 5.0% VAF reference sample was marked as a light blue star, its coordinate was determined by the claimed VAF and the obtained Ct value from the Occlusion qPCR reaction.

B). Quantitation validation from standard VAF curve of the *IDH1* 395G>A mutation with a 1.30% reference sample. The VAF samples were presented as individual red dots, with the fitted line shown as solid red. The linear correlation in *IDH1* 395G>A mutation has an r-square around 0.97. The 1.30% VAF

reference sample was marked as a yellow star, its closeness to the fitted line revealed the accuracy of VAF calling.

C). Summary of comparative results between ddPCR and HiFi BDA Occlusion qPCR in *FLT3* mutations.

D). HiFi BDA Occlusion qPCR results and Sanger traces of AML samples in *IDH1* gene.

Improvement of PCR efficiency - Reduction of primer dimer formation by the Occlusion System

We further hypothesized that the Occlusion System could also potentially improve the PCR efficiency by reducing primer-dimer formation in the sequencing library preparation methods. Usually, reads of an Amplicon or Target Sequencing library are from on-target amplicons, dimers, and off-target amplicons (non-specific amplification). A typical protocol for Amplicon Sequencing usually includes an adaptor-appending step, by either ligation or PCR. The ligation step is more economical but may lead to more dimer from the self-ligation of the adaptors[55]. Adaptor PCR generally has a higher efficiency over ligation, but it could introduce a higher possibility of dimer formation and off-target amplification due to the extended length of primers, resulting in a lower percentage of reads from target amplicons and, thus, a lower on-target rate (Figure 4A). The influence of dimerization and non-specific amplification on the total quality of the sequencing library and sequencing becomes more severe as the plex number of the sequencing panel significantly increases[44].

The Occlusion System can bind to the 5' overhang region of the adaptor primers, thus restricting their non-specific binding to either genomic regions or other primers for non-specific amplification and dimerization (Figure 4B). What is more, the Occlusion System could provide an easy, ready-to-use design solution to any panels with universal or specific 5' overhang sequences during the PCR step. We validated our hypothesis by applying the Occlusion System to two previously designed NGS panels, one is a 80-plex panel[46]; the other is 179-plex[49]. In the presence of the Occlusion Probe, the percentage of reads of 80-plex target amplicons increased by 100% (from 40% to 80%), the 179-plex target reads increased by 700% (Figure 4C). We continued to explore if the Occlusion System was compatible with both Taq-based DNA polymerases and high-fidelity DNA polymerases. By modifying the protocol of the 179-plex panel (protocol details in Supplementary Section S6), we showed that the Occlusion System could improve the sequencing efficiency in libraries using either Taq-based or high-fidelity DNA polymerase (Figure S15). The Occlusion

System also revealed potential in reducing the non-specific amplification with genomic DNA as input (Figure S15B). In this way, we concluded that the Occlusion System effectively improves PCR efficiency and increases the on-target rate of sequencing libraries, especially in large-scale NGS panels.

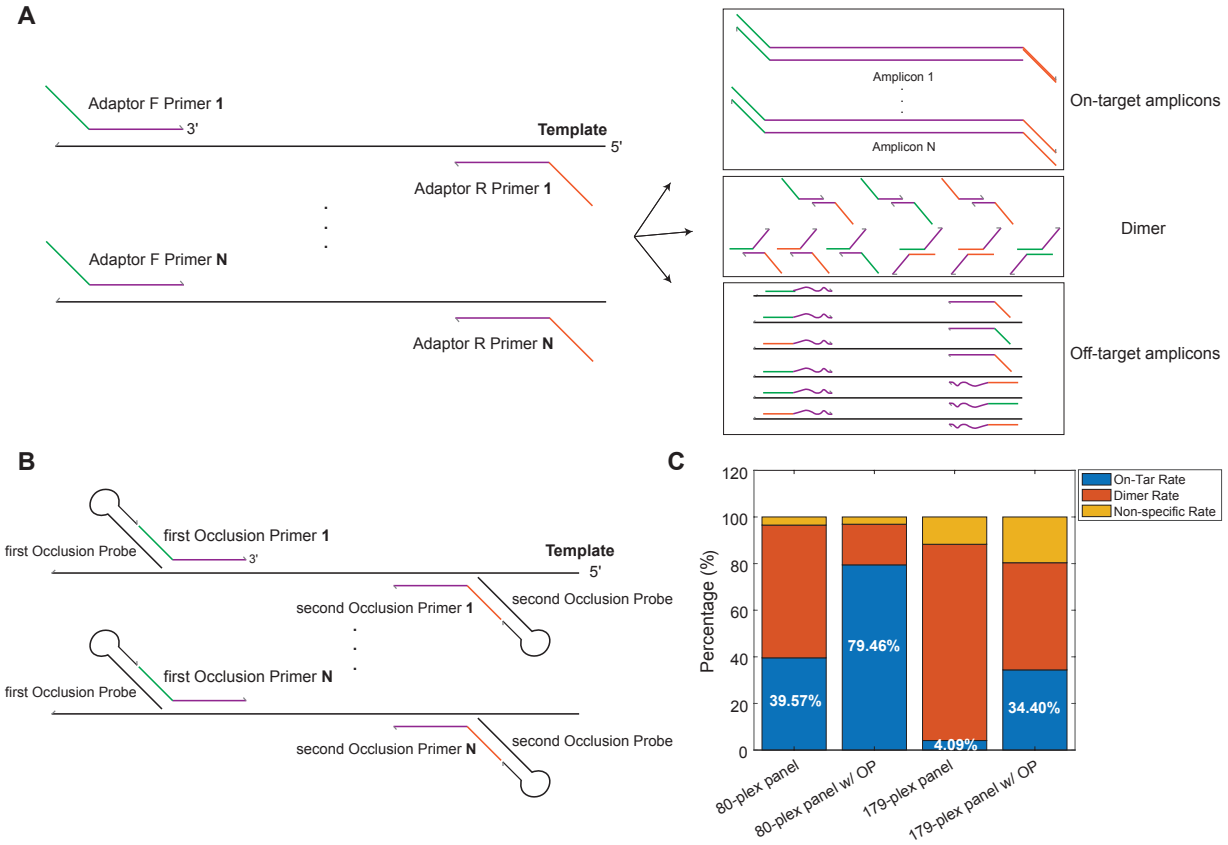

Figure 4. The Occlusion System with Next-generation sequencing (NGS) to reduce dimer generation.

A). Schematic of the formation of On-target amplicons, Dimer and Off-target amplicons in NGS adaptor PCR.

B). The Occlusion System in adaptor PCR. The Occlusion System could be designed for each adaptor primer (Occlusion Primer), wherein each Occlusion Primer would be paired with an Occlusion Probe. The reverse complementarity of the Occlusion Probe with the overhang region of the Occlusion primer would make the Occlusion Probe fully bind to the overhang region, thus effectively reduce the chance of non-specific amplification and dimerization.

C). The Occlusion System productively increased the on-target rate of large-scale NGS panels. By applying the Occlusion System in an 80-plex and a 179-plex NGS panels, we found the Occlusion System could improve the on-target rate by 100% to 700%.

DISCUSSION

The high polymerase misincorporation rate has been a restriction in the LoD of variant detection methods. The appearance of high-fidelity DNA polymerases, with proof-reading activity, offered a potential solution. We found that the Occlusion System, with rationally designed short-stem hairpin structures, could hinder the 3' to 5' exonuclease activity of high-fidelity DNA polymerases, and function as a probe-based fluorescence indicator in multiplexed high-fidelity qPCR reactions. This Occlusion System presented here further provides a solution for multiplexed variant enrichment reactions with high-fidelity DNA polymerases. It allows convenient transition for other qPCR technologies, from previously used non-high-fidelity DNA polymerases to high-fidelity DNA polymerases. This convenience translates into practical advantages for molecular diagnostics and genomics research, allowing previous variant detection methods to achieve better sensitivity and specificity, and enabling detection of multiple targets within the same reaction. We demonstrated the compatibility of the Occlusion system with the BDA technology could achieve the detection of multiplexed variants with VAF low to 0.01%. BDA has been mostly reported to have an LoD only around 0.1% VAF, except As-BDA[29] which could detect mutation low to 0.01% VAF. However, As-BDA requires preliminary knowledge of mutation or small insertion/deletion information, thus limited in the total number of mutations detected within one reaction and restricted in the discovery of novel mutations. The HiFi BDA Occlusion system offset this limitation by achieving a hypothesis-free variant enrichment within the target region with ultra-sensitivity, and the specific mutation identity could be confirmed by Sanger sequencing or other sequencing technologies. Furthermore, the same sequence of Occlusion probe could be re-used in designs targeting different templates within the same fluorescent channel, reducing the overall cost in oligo synthesis. Thus, we believe the Occlusion System provides a cost-efficient, user-friendly, and plex-scalable solution in high-fidelity DNA qPCR reactions, which is not achievable by other methods.

As the need for higher coverage of sequencing increases, the severity of primer dimer formation boosts as the total number of primers goes up. Most of the existing efforts to reduce primer dimer formation either focus on specific panel designs and require specialized coding skills or supplement additional steps to the workflow with extra cost and hands-on time. Thus, there lacks a method that could be easily adopted to a newly built sequencing panel with little computing effort or reagent cost. The Occlusion system showed competence in reducing dimerization over the current efforts by providing an easy-to-adapt hairpin design in the upstream PCR reactions during sequencing library preparation, leading to up to 10-fold improvement

of on-target rate in the sequencing results. Considering that the adaptor sequence is usually universal to the sequencing platforms, the Occlusion probe could be implemented economically and effortlessly in many types of sequencing libraries with various primer panels included.

We expect the Occlusion system could promote the diagnosis of early-stage cancer clinically. In this work, we primarily designed the HiFi BDA Occlusion System to detect mutations occurred within codon 126 to codon 133 in *IDH1* gene and codon 832 to codon 838 in *FLT3* gene. By applying the HiFi BDA Occlusion qPCR reactions in PBMC samples from 27 AML patients and 15 healthy donors, we successfully detected mutations in 7 AML samples and confirmed the results by Sanger sequencing. When compared with ddPCR, results are 100% concordant for mutations with VAF $\geq$ 0.1% with the similar amount of DNA input. When it comes to lower VAF mutations, the ddPCR either failed to report the mutation or showed inconsistency within a technical replicate. Unlike ddPCR which is less dymamic in the input with the concern of multiple molecules within the same droplet, the HiFi BDA Occlusion system successfully detected two samples with VAF below 0.05% using 10-fold more DNA input. The HiFi BDA Occlusion system has demonstrated its reliable capability, robustness and competitiveness in detecting VAF low to 0.01% in clinical PBMC samples, which reveals the potential of widespread applications in many other clinical specimen types. We also believed that the Occlusion system was a general method that could be used with many other variant detection methods in high-fidelity PCR reactions.

The current version of the Occlusion System still needs additional technologies or methods to determine the detailed mutation information. The multiplex-ability of the Occlusion system could be empirically improved further.

DATA AVAILABILITY

The sequences of the DNA oligos used for experiments are included in the supplementary file accompanying this manuscript. The detailed concentration of the oligos used for experiment are described in the Method and Materials section. The main data supporting the results in this study are available within the paper and its Supplementary Information. The datasets collected and/or analyzed during the current study available from the corresponding author on reasonable request.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR online.

ACKNOWLEDGEMENTS

Correspondence may be addressed to DYZ (genomic.dave@gmail.com).

The authors thank Quoc-Khanh Pham and Dr. Bao for use of their ddPCR instrument.

The authors thank Xuwen Li, Gavin Jiaming Li and Nina Guanyi Xie for editorial assistance.

FUNDING

This work was funded by NCI grant 5U01CA233364 to DYZ.

AUTHOR CONTRIBUTIONS

K.Z and D.Y.Z conceived the project. K.Z, A.P, L.R, C.W and R.W performed the experiments. K.Z, A.P, M.W analyzed the data. K.Z, A.P and L.Y.C conducted Occlusion qPCR design. K.Z, P.D and P.S conducted Occlusion NGS design. K.Z wrote the manuscript with input from all authors. D.Y.Z. revised the manuscript.

CONFLICT OF INTEREST

There is a patent pending on the Occlusion system method used in this work. K.Z, P.S, P.D, M.W and L.Y.C declare competing interests in the form of consulting for Nuprobe USA. D.Y.Z declares a competing interest in the form of consulting for and significant equity ownership in NuProbe Global, Torus Biosystems and Pana Bio.

REFERENCE

1. Kralik, P. & Ricchi, M. A Basic Guide to Real Time PCR in Microbial Diagnostics: Definitions, Parameters, and Everything. *Front. Microbiol.* **8**, 108 (2017).
2. DeBiasi, R. L. & Tyler, K. L. Polymerase Chain Reaction in the Diagnosis and Management of Central Nervous System Infections. *Arch. Neurol.* **56**, 1215–1219 (1999).
3. Wong, M. L. & Medrano, J. F. Real-time PCR for mRNA quantitation. *Biotechniques* **39**, 75–85 (2005).
4. Wang, D. *et al.* Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus–

Infected Pneumonia in Wuhan, China. *JAMA* **323**, 1061–1069 (2020).
5. Wang, W. *et al.* Detection of SARS-CoV-2 in Different Types of Clinical Specimens. *JAMA* **323**, 1843–1844 (2020).
6. Mackay, I. M. Real-time PCR in the microbiology laboratory. *Clin. Microbiol. Infect.* **10**, 190–212 (2004).
7. Mocellin, S., Rossi, C. R., Pilati, P., Nitti, D. & Marincola, F. M. Quantitative real-time PCR: a powerful ally in cancer research. *Trends Mol. Med.* **9**, 189–195 (2003).
8. Bernard, P. S. & Wittwer, C. T. Real-Time PCR Technology for Cancer Diagnostics. *Clin. Chem.* **48**, 1178–1185 (2002).
9. Espy, M. J. *et al.* Real-time PCR in clinical microbiology: applications for routine laboratory testing. *Clin. Microbiol. Rev.* **19**, 165–256 (2006).
10. Loeb, K. R. & Loeb, L. A. Significance of multiple mutations in cancer. *Carcinogenesis* **21**, 379–385 (2000).
11. Potapov, V. & Ong, J. L. Examining Sources of Error in PCR by Single-Molecule Sequencing. *PLoS One* **12**, e0169774 (2017).
12. Martincorena, I. & Campbell, P. J. Somatic mutation in cancer and normal cells. *Science (80-.).* **349**, 1483 LP – 1489 (2015).
13. Remon, J. *et al.* Liquid biopsy in oncology: a consensus statement of the Spanish Society of Pathology and the Spanish Society of Medical Oncology. *Clin. Transl. Oncol.* **22**, 823–834 (2020).
14. Palmirotta, R. *et al.* Liquid biopsy of cancer: a multimodal diagnostic tool in clinical oncology. *Ther. Adv. Med. Oncol.* **10**, 1758835918794630–1758835918794630 (2018).
15. Li, S. *et al.* Sensitive detection of tumor mutations from blood and its application to immunotherapy prognosis. *Nat. Commun.* **12**, 4172 (2021).
16. Schuurhuis, G. J. *et al.* Minimal/measurable residual disease in AML: a consensus document from the European LeukemiaNet MRD Working Party. *Blood* **131**, 1275–1291 (2018).
17. Jongen-Lavrencic, M. *et al.* Molecular Minimal Residual Disease in Acute Myeloid Leukemia. *N. Engl. J. Med.* **378**, 1189–1199 (2018).
18. Ford, C. B. *et al.* Mycobacterium tuberculosis mutation rate estimates from different lineages predict substantial differences in the emergence of drug-resistant tuberculosis. *Nat. Genet.* **45**, 784–790 (2013).
19. Misale, S. *et al.* Emergence of KRAS mutations and acquired resistance to anti-EGFR therapy in colorectal cancer. *Nature* **486**, 532–536 (2012).
20. Suzuki, Y., Sekiya, T. & Hayashi, K. Allele-specific polymerase chain reaction: A method for amplification and sequence determination of a single component among a mixture of sequence variants. *Anal. Biochem.* **192**, 82–84 (1991).
21. Billard, A. *et al.* The allele-specific probe and primer amplification assay, a new real-time PCR method for fine quantification of single-nucleotide polymorphisms in pooled DNA. *Appl. Environ. Microbiol.* **78**, 1063–1068 (2012).
22. Newton, C. R. *et al.* Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS). *Nucleic Acids Res.* **17**, 2503–2516 (1989).
23. Zhu, Y. *et al.* A novel ARMS-based assay for the quantification of EGFR mutations in patients with lung adenocarcinoma. *Oncol. Lett.* **15**, 2905–2912 (2018).
24. Didelot, A. *et al.* Competitive allele specific TaqMan PCR for KRAS, BRAF and EGFR mutation detection in clinical formalin fixed paraffin embedded samples. *Exp. Mol. Pathol.* **92**, 275–280 (2012).
25. Barbano, R. *et al.* Competitive allele-specific TaqMan PCR (Cast-PCR) is a sensitive, specific and fast method for BRAF V600 mutation detection in Melanoma patients. *Sci. Rep.* **5**, 18592 (2015).
26. Li, J. *et al.* Replacing PCR with COLD-PCR enriches variant DNA sequences and redefines the sensitivity of genetic testing. *Nat. Med.* **14**, 579–584 (2008).
27. Wu, L. R., Chen, S. X., Wu, Y., Patel, A. A. & Zhang, D. Y. Multiplexed enrichment of rare DNA variants via sequence-selective and temperature-robust amplification. *Nat. Biomed. Eng.* **1**, 714–723 (2017).
28. Bohers, E., Viailly, P.-J. & Jardin, F. cfDNA Sequencing: Technological Approaches and Bioinformatic Issues. *Pharmaceuticals* vol. 14 (2021).
29. Zhang, K., Rodriguez, L., Cheng, L. Y. & Zhang, D. Y. Single Tube qPCR detection and quantitation of hotspot mutations down to 0.01% VAF. *bioRxiv* 2021.08.12.456178 (2021)

doi:10.1101/2021.08.12.456178.
30. Zagordi, O., Klein, R., Däumer, M. & Beerenwinkel, N. Error correction of next-generation sequencing data and reliable estimation of HIV quasispecies. *Nucleic Acids Res.* **38**, 7400–7409 (2010).
31. Filges, S., Yamada, E., Ståhlberg, A. & Godfrey, T. E. Impact of Polymerase Fidelity on Background Error Rates in Next-Generation Sequencing with Unique Molecular Identifiers/Barcodes. *Sci. Rep.* **9**, 3503 (2019).
32. Smith, T., Heger, A. & Sudbery, I. UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res.* **27**, 491–499 (2017).
33. Eid, J. *et al.* Real-Time DNA Sequencing from Single Polymerase Molecules. *Science (80-.).* **323**, 133 LP – 138 (2009).
34. Johnson, K. A. The kinetic and chemical mechanism of high-fidelity DNA polymerases. *Biochim. Biophys. Acta - Proteins Proteomics* **1804**, 1041–1048 (2010).
35. Joyce, C. M. & Benkovic, S. J. DNA Polymerase Fidelity: Kinetics, Structure, and Checkpoints. *Biochemistry* **43**, 14317–14324 (2004).
36. Eckert, K. A. & Kunkel, T. A. DNA polymerase fidelity and the polymerase chain reaction. *Genome Res.* **1**, 17–24 (1991).
37. Keohavong, P. & Thilly, W. G. Fidelity of DNA polymerases in DNA amplification. *Proc. Natl. Acad. Sci.* **86**, 9253 LP – 9257 (1989).
38. Chen, W., Martinez, G. & Mulchandani, A. Molecular Beacons: A Real-Time Polymerase Chain Reaction Assay for Detecting Salmonella. *Anal. Biochem.* **280**, 166–172 (2000).
39. Nazarenko, I. A., Bhatnagar, S. K. & Hohman, R. J. A closed tube format for amplification and detection of DNA based on energy transfer. *Nucleic Acids Res.* **25**, 2516–2521 (1997).
40. Thelwell, N., Millington, S., Solinas, A., Booth, J. & Brown, T. Mode of action and application of Scorpion primers to mutation detection. *Nucleic Acids Res.* **28**, 3752–3761 (2000).
41. Holland, P. M., Abramson, R. D., Watson, R. & Gelfand, D. H. Detection of specific polymerase chain reaction product by utilizing the 5'----3' exonuclease activity of Thermus aquaticus DNA polymerase. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 7276–7280 (1991).
42. Didenko, V. V. DNA probes using fluorescence resonance energy transfer (FRET): designs and applications. *Biotechniques* **31**, 1106–1121 (2001).
43. Zhang, M. *et al.* A novel quantitative PCR mediated by high-fidelity DNA polymerase. *Sci. Rep.* **7**, 10365 (2017).
44. Zhang, D. *et al.* Designing Highly Multiplex PCR Primer Sets with Simulated Annealing Design using Dimer Likelihood Estimation (SADDLE). *Nat. Portf.* (2021) doi:10.21203/rs.3.rs-491811/v1.
45. Sun, J.-M. *et al.* Small-cell lung cancer detection in never-smokers: clinical characteristics and multigene mutation profiling using targeted next-generation sequencing. *Ann. Oncol.* **26**, 161–166 (2015).
46. Song, P. *et al.* Selective multiplexed enrichment for the detection and quantitation of low-fraction DNA variants via low-depth sequencing. *Nat. Biomed. Eng.* (2021) doi:10.1038/s41551-021-00713-0.
47. Park, M., Won, J., Choi, B. Y. & Lee, C. J. Optimization of primer sets and detection protocols for SARS-CoV-2 of coronavirus disease 2019 (COVID-19) using PCR and real-time PCR. *Exp. Mol. Med.* **52**, 963–977 (2020).
48. Kawano, M. *et al.* Reduction of non-insert sequence reads by dimer eliminator LNA oligonucleotide for small RNA deep sequencing. *Biotechniques* **49**, 751–755 (2010).
49. Wu, L. *et al.* Ensemble of Nucleic Acid Absolute Quantitation Modules for Accurate Copy Number Variation Detection and Targeted RNA Profiling. *Nat. Portf.* (2021) doi:10.21203/rs.3.rs-923491/v1.
50. Wilson, K. S. & von Hippel, P. H. Transcription termination at intrinsic terminators: the role of the RNA hairpin. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 8793–8797 (1995).
51. Farnham, P. J. & Platt, T. Rho-independent termination: dyad symmetry in DNA causes RNA polymerase to pause during transcription in vitro. *Nucleic Acids Res.* **9**, 563–577 (1981).
52. Roberts, J. W. Mechanisms of Bacterial Transcription Termination. *J. Mol. Biol.* **431**, 4030–4039 (2019).
53. Gusarov, I. & Nudler, E. The Mechanism of Intrinsic Transcription Termination. *Mol. Cell* **3**, 495–504 (1999).
54. Michael M Cox, Jennifer A Doudna, M. O. *Molecular Biology Principles and Practice*. (WH

Freeman).
55. Head, S. R. *et al.* Library construction for next-generation sequencing: overviews and challenges. *Biotechniques* **56**, 61-passim (2014).