Biosensors and Bioelectronics 179 (2021) 113079

Contents lists available at ScienceDirect

Biosensors and Bioelectronics

journal homepage: http://www.elsevier.com/locate/bios

Accurate genotyping of fragmented DNA using a toehold assisted padlock probe

Yanmin Gao [a,b], Hongyan Qiao [a,b], Victor Pan [c], Zhaoguan Wang [a,b], Jiaojiao Li [a,b], Yanan Wei [a,b], Yonggang Ke [c], Hao Qi [a,b,*]

[a] School of Chemical Engineering and Technology, Tianjin University, Tianjin, 300350, China
[b] Key Laboratory of Systems Bioengineering (Ministry of Education), Tianjin University, Tianjin, 300350, China
[c] Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology and Emory University, Atlanta, GA, 30322, United States

ARTICLE INFO

Keywords:
Cell-free DNA
Liquid biopsy
Toehold-assisted padlock probe
SNV
Broad-spectrum DNA enrichment

ABSTRACT

Fragmented DNA from blood plasma, i.e., cell-free DNA, has received great interest as a noninvasive diagnostic biomarker for "point-of-care" testing or liquid biopsy. Here, we present a new approach for accurate genotyping of highly fragmented DNA. Based on toehold-mediated strand displacement, a toehold-assisted padlock probe and toehold blocker were designed and demonstrated with new controllability in significantly suppressing undesired cross-reaction, promoting target recycling and point mutation detection by tuning the thermodynamic properties. Furthermore, toehold-assisted padlock probe systems were elaborately designed for 14 different single-nucleotide variants (SNVs) and were demonstrated to be able to detect low concentration of variant alleles (0.1%). In addition, a target, spanning a narrow sequence window of 29 nucleotides on average is sufficient for the toehold-assisted padlock probe system, which is valuable for the analysis of highly fragmented DNA molecules from clinical samples. We further demonstrated that the toehold-assisted padlock probe, in combination with a unique asymmetric PCR technique, could detect more target SNVs at low allele fractions (1%) in highly fragmented cfDNA. This allows accurate genotyping and provides a new commercial approach for high-resolution analysis of genetic variation.

1. Introduction

Highly fragmented DNA, especially cell-free DNA (cfDNA), has significant potential applications in biomedical diagnosis, including early cancer detection (Abbosh et al., 2018; Chen et al., 2020), determining the origin of metastatic relapse (Forshew et al., 2012), assessment of tumor staging and prognosis (Abbosh et al., 2018; Arko-Boham et al., 2019), selection of suitable treatment regimens (Peters et al., 2017; Soria et al., 2018), as well as monitoring therapy response (Knebel et al., 2019). The length distribution of fragmented DNA is typically in the range of 160–180 bp (Chae and Oh 2019), but can undergo dynamically changes according to the physiological and pathological conditions. Moreover, circulating tumor DNA (ctDNA) is shorter than cfDNA (Andersen et al., 2015; Li et al., 2020; Mouliere et al., 2018; Underhill et al., 2016), with especially small DNA fragments being significantly increase in some tumor types (Mouliere et al., 2018). In addition, the fragmentation pattern is another crucial feature, but the underlying mechanism is insufficiently understood (Heidary et al., 2014). Furthermore, human blood contains a very low proportion of ctDNA, especially in the early-stage cancer patients (Butler et al., 2017). Therefore, it is necessary to carefully consider all these features for accurate genotyping of ctDNA.

By current approaches, such as PCR-based method, it is difficult to cover all DNA fragments with different lengths and terminus. Different amplicon lengths lead to varied result in the detection of the same SNV, and even for an amplicon of the same length, different primer pairs do not generate consistent detection results (Andersen et al., 2015; Diehl et al., 2005). Therefore, for genotyping of fragmented DNA, one size does not fit all plasma-based DNA diagnostics (Heitzer and Speicher 2018; Mouliere et al., 2018), and new approaches are needed to achieve high-resolution analysis of short target sequence in fragmented DNA.

Since its development, the padlock probe has been widely applied in bioengineering and biomedical diagnosis (Ali et al., 2014). However, methods based on padlock probe are highly dependent on the fidelity of the utilized ligase, and most DNA ligases have a poor kinetic performance with imprecise substrate recognition (Kim and Mrksich 2010; Li

* Corresponding author. School of Chemical Engineering and Technology, Tianjin University, Tianjin, 300350, China.
E-mail address: haoq@tju.edu.cn (H. Qi).

https://doi.org/10.1016/j.bios.2021.113079
Received 20 November 2020; Received in revised form 25 January 2021; Accepted 3 February 2021
Available online 7 February 2021

et al., 2009; Wilson et al., 2013). Here, we present the accurate Genotyping of fragmented DNA using a toehold Padlock probe (GfDtP). For the first time, we demonstrate that an entropy-driven toehold-blocker can be designed to precisely regulate the padlock ligation reaction. Using this system, approximately 85% of non-desired cross-ligation was suppressed even at high probe concentration. Relying on a novel mechanism of toehold-driven ligation template recycling, our toehold-assisted padlock probe achieved a median 966-fold selectivity for 14 cancer-related DNA SNVs at very low allele fractions (0.1%). Furthermore, the short target sequence of only 29 nts necessary for the toehold probe could be enriched using single-primer asymmetric PCR amplification from cfDNA sample with various sizes and fragmentation patterns. By combining the broad-spectrum enrichment and high-resolution toehold padlock probe, we demonstrated that GfDtP could detect targeted SNVs even with a high background of total cfDNA, providing a suitable point-of-care or liquid biopsy approach for routine genotyping in therapy monitoring or disease screening at low cost.

2. Material and methods

2.1. Toehold-assisted padlock probe ligation

The ligation reaction was carried out in 15 μL reaction mixtures containing 1 × T4 ligation reaction buffer, T4 DNA ligase, the padlock probe, the blocker and the target (specified concentration). Before adding T4 DNA ligase, the reaction mixture was denatured at 95 °C for 3 min and cooled to 25 °C at a ramp of 0.1 °C/s using an Eppendorf Mastercycler. After annealing, T4 DNA ligase was added, incubated at 37 °C for 3 h, and inactivated by heating at 65 °C for 10 min. Then, exonuclease I and exonuclease III were added, incubated at 37 °C for 3 h and inactivated at 80 °C for 20 min.

2.2. Ligation reaction with target DNA (0.1% variant allele frequency)

The ligation reaction was carried out in 30 μL reaction mixtures containing 1 × T4 ligase reaction buffer, 25 U of T4 DNA ligase, 1 μM probe, 1 μM of blocker and 0 or 0.1% of the target (100 nM). The conditions were the same as for the toehold-assisted padlock probe ligation.

2.3. Real-time rolling cycle amplification (RCA)

The reaction was performed in 50 μL reaction mixtures containing 1 × Isothermal Amplification buffer, 6 mM $MgSO_4$, 100 nM primer, 0.2 mM dNTPs, 0.075 × SYBR Green I, 8 U of Bst2.0 WarmStart DNA polymerase and 10 μL of the ligation product. The reaction was incubated at 55 °C and fluorescence was measured every 30 s using a QuantStudio 6 Flex Real-Time PCR Systems (Thermo Fisher Scientific, USA).

2.4. EGFR L861Q detection

The asymmetric PCR was performed in 50 μL reaction mixtures containing 1 × Phusion HF buffer, 8 μM dNTPs, 192 μM ddNTPs, 4 μM forward primer, 30 ng phosphorylated target DNA and 1 U Phusion high-fidelity DNA polymerase. Thermocycling conditions were as follows: 30 s at 98 °C; 150 cycles of: 15 s at 98 °C, 30 s at 60 °C, 5 s at 72 °C. Then 5 U lambda exonuclease was added to digest the phosphorylated target DNA. The reaction mixture was incubated at 37 °C for 3 h, then inactivated at 75 °C for 10 min, yielding the ssDNA. The ligation reaction with wild type (WT) blocker and SNV blocker was carried out in 50 μL reaction mixtures containing 1 × T4 ligation buffer, 50 U T4 DNA ligase, 1 μM padlock probe, 1 μM WT blocker, 1 μM SNV blocker and 10 μL of the ssDNA. Except for the blocker, the ligation reaction system with only WT blocker or without blocker contained the same components. The conditions of the ligation reaction were the same as for the toehold-assisted padlock probe ligation.

3. Results and discussion

3.1. Principle and feasibility of the toehold-assisted padlock probe

In the toehold-assisted ligation reaction, the 5′ and 3’ termini of the probe have to displace the blocker successively and align with the template to form a circular ligation complex (Fig. 1A). Because toehold-mediated strand displacement is a reversible process, padlock probes were first designed to test whether the blocker strands could undergo discrete strand displacement. The designed blocker was composed of a toehold of 13 nt and a 6 nt sequence overlapping with the probe (Fig. S1A). The padlock probe was incubated with the template at an equimolar ratio in the presence or absence of blocker, and then treated with exonuclease I and exonuclease III to digest non-ligated linear DNA strands, leaving only circular DNA molecules (Fig. S1B). DNA polymers with various molecular weight were identified, with the smallest being the expected ligation product from a single padlock probe, and the larger circular DNA being the product of cross ligation of multiple padlock probe strands (Fig. 1B). The ligation yield of the padlock probes was quantified (Fig. 1C). This demonstrated that the branch migration of the blocker initiated by the toehold could tend to move continuously in one direction, allowing discrete strand displacement.

However, there were fewer ligation product and polymers in the presence of blocker. In the toehold-assisted ligation reaction, the blocker protects the major sequence from probe-template assembly, leaving a small segment for hybridization with the probe’s 3′ terminal toehold, after which the entire assembly sequence is exposed to the probe’s 5′ terminal sequence. Due to this spatial effect, a single padlock probe preferentially assembles on one template instead of undergoing cross-assembly of two probes on one template. Different ratios of the blocker, padlock probe, and template were used to verify the decrease of cross ligation. A high blocker concentration led to an almost 10-fold increase in the ratio of mono to cross-ligation products (Fig. S2).

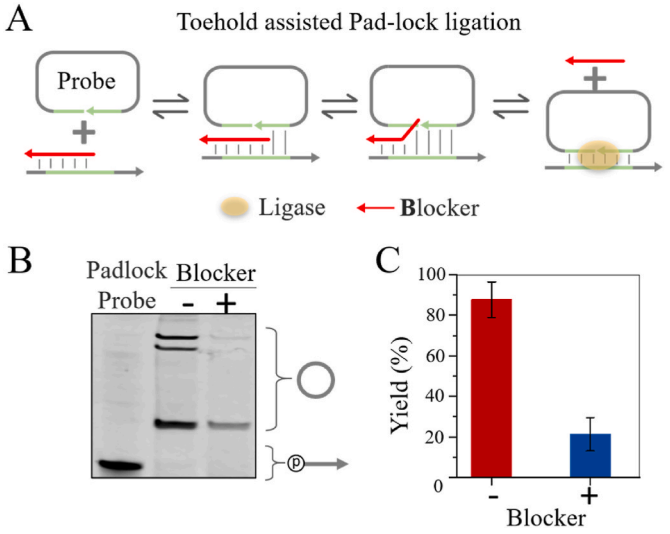

Fig. 1. Toehold-assisted ligation reaction. **(A)** Schematic presentation of the discrete strand displacement mechanism in the complex assembly of padlock probes in the presence of blocker (strand in red). **(B)** The padlock probe was ligated using T4 DNA ligase in the presence or absence of blocker. The product was digested with exonuclease I and exonuclease III and then analyzed in a denaturing PAGE gel. **(C)** Ligation yield in the absence or presence of blocker. Error bars represent the means ± SD, n = 3. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

3.2. Toehold-assisted controllability of the padlock probe

Next, we explored the thermodynamic principles of blocker design to improve the yield while still suppressing cross ligation. In the assembled DNA complex (Fig. 2A), blocker and probe overlap at a sequence segment necessary for assembly, whereby both probe and blocker have their own respective toehold sequences. The toehold-assisted assembly of the DNA ligation complex is a reversible process (Fig. 2B), whereby the toeholds of both probe and blocker can be tuned to achieve a desired reaction yield and quality. The difference in Gibbs free energy between $\Delta G_{T\text{-}p}$ and $\Delta G_{T\text{-}b}$ was designated as $\Delta\Delta G_{toe}$. A group of blockers with various lengths of the blocker toehold and probe toehold were designed (Fig. 2C and Table S1) and designated as B (number of nucleotide bases in the probe toehold, number of nucleotide bases in the blocker toehold). The yields of mono- and cross-ligation products were quantified in reaction with a 1:1:10 molar ratio of template, probe and blocker (Fig. 2D and Fig. S3). In the absence of blocker, the total ligation yield was about 90%, but the cross-ligation product accounted for about 50% of the total. Although the total ligation yield decreased in most of the blocker-modulated reactions, all of the blockers significantly decreased the yield of cross-ligation, to 6.5% on average. The cross-ligation yield was reduced by 85% compared to the 44% in the absence of blocker. However, only few of the blockers afforded very good ligation. The blocker B (14, 15) generated a total ligation yield of about 84% while cross-ligation accounted for only about 11%.

It was observed that the $\Delta\Delta G_{toe}$ is related to the ligation yield (Fig. 2E). Overall, the yield of ligation product in the presence of type II blockers were higher than with type I blockers. In type I blockers, $\Delta\Delta G_{toe}$ values below 0 improved the ligation efficiency. Interestingly, even type II blockers with similar $\Delta\Delta G_{toe}$ (sequence not covering the joined point in the probe) generated higher ligation yields than type I blockers (sequence covering the joined point in the probe). This may be due to the long probe toehold increasing the binding efficiency, which resulted in a high ligation yield with slightly increased cross-ligation. In addition, a short blocker sequence decreased the binding stability, resulting in a high ratio of cross-ligation products.

3.3. High sensitivity of SNV detection at low allele fractions

To test the sensitivity of the toehold-assisted padlock probe, a single nucleotide variant (SNV), comprising G in the original template as target X, and T in the variant template as target S (Fig. 3A), was designed. The difference in Gibbs free energy for blocker binding ($\Delta\Delta G_{(x\text{-}s)}$) was approximately −1.21 kcal/mol for the G to T mutation. Padlock probes were ligated on target X or S in the presence of one set of blockers at a molar ratio of P: B: T = 100:100:10 nM, and the yields were quantified (Fig. 3B and Fig. S4). As designed, the padlock probes ligated on target X or S at almost the same efficiency in the absence of blocker, while all blockers led to higher ligation yields on target S (Fig. 3C). This indicated that the toehold blocker was able to protect target X much better from probe binding than target S. The blocker B (11, 13) showed the highest Q value, describing the variant discrimination ability (Zhang et al., 2012). Except for the G/T mutation, a series of mutations including substitutions and indels (insertions or deletions) (Fig. S5A), with ΔΔG(x-s) values ranging from −0.55 to −6.30 kcal/mol, were used to test the sensitivity of B (11, 13). High Q values were obtained for most of these mutations (Figs. S5B and S5C), demonstrating the robustness and flexibility of the toehold-assisted padlock probe.

Notably, B (11, 13) gave a much high ligation yield on target S than on target X (Fig. 3D), clearly demonstrating an efficient template recycling mechanism in B (11, 13). It is possible that toehold blockers can introduce a template recycling function to remove the ligated product

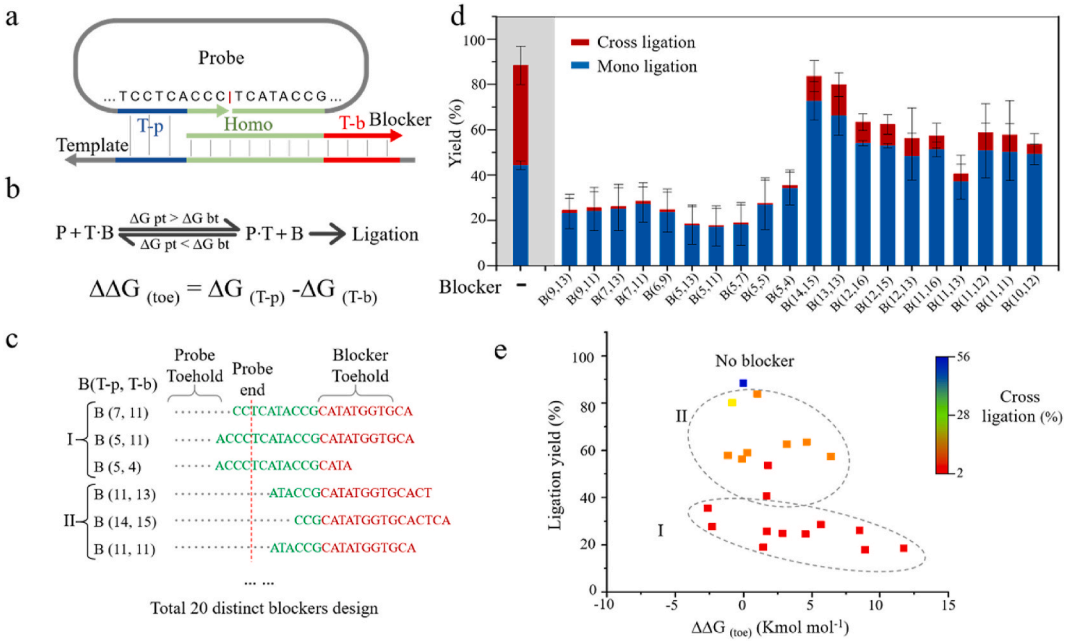

Fig. 2. Toehold-modulated controllability. **(A)** Schematic of the assembled padlock complex modulated by the toehold blocker. The toehold of the probe is shown in blue, the blocker toehold in red, and the overlapped sequence in green. **(B)** The biochemical reaction in toehold-assisted ligation reaction. The Gibbs free energy of the probe toehold ($\Delta G_{(T\text{-}p)}$) and blocker toehold ($\Delta G_{(T\text{-}b)}$) was calculated using NUPACK (Supplementary information 2). The difference between $\Delta G_{(T\text{-}p)}$ and $\Delta G_{(T\text{-}b)}$ was designated as $\Delta\Delta G_{(toe)}$. **(C)** Blockers consisting of a homology segment (green) and toehold segment (red) were designed. The corresponding sequence of the probe toehold is represented by gray dots and the blocker is denoted as B (number of nucleotide base in the probe toehold, number of nucleotide bases in the blocker toehold). A total of 20 distinct blockers were designed, with those covering the probe joining point (red dashed line) grouped as type I and others as type II. **(D)** Cross- and mono ligation yields were quantified densitometrically based on denaturing gels. The padlock probe (0.5 μM) was ligated with 0.5 μM template in the presence of 5 μM blocker. **(E)** The reaction yields of various blocker-modulated ligations were plotted against the corresponding $\Delta\Delta G_{(toe)}$ and the heatmap color represents the percentage of cross ligation in the total ligated circular ssDNA. Error bars represent the means ± SD, n = 3. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

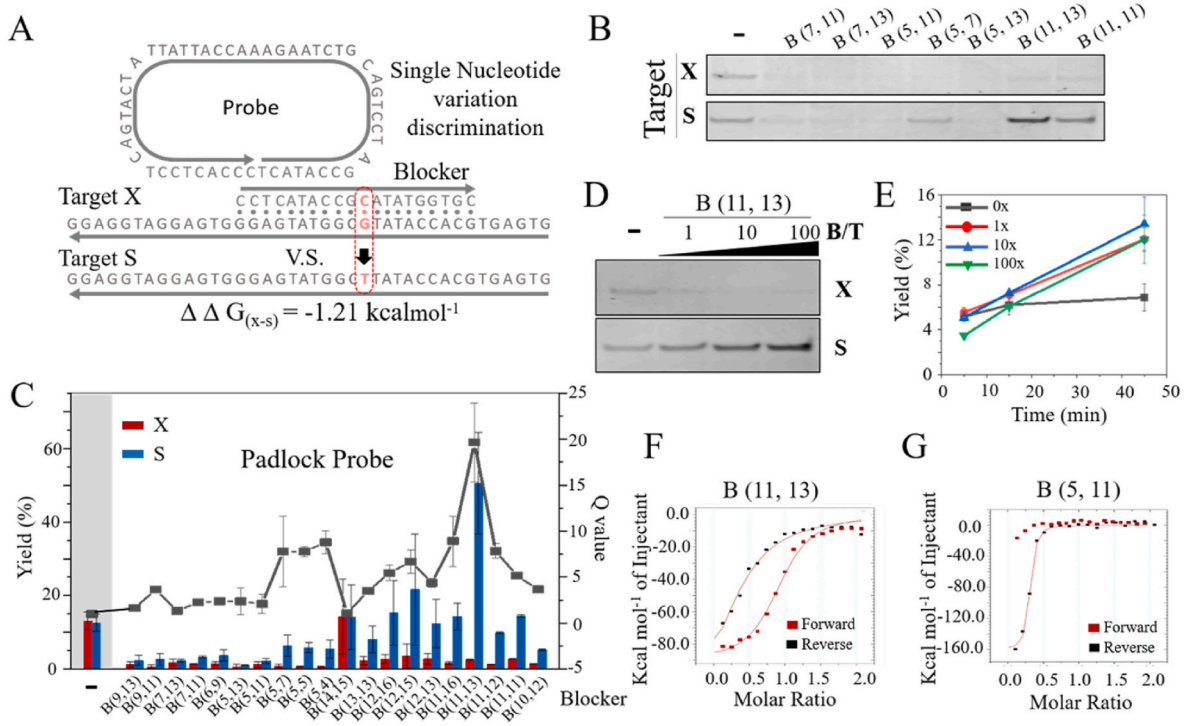

Fig. 3. Ultra-high sensitivity of the toehold-assisted padlock probe. **(A)** Illustration of binding to templates with single-nucleotide variation, G in target X which is perfectly complementary to the blocker, and T in target S. The shift of Gibbs free energy for binding to targets X and S was −1.21 kcal/mol according to NUPACK. **(B)** The padlock probe (100 nM) was ligated on the template (10 nM) by T4 DNA ligase in the presence of various blockers (100 nM). After exonuclease digestion, the ligation products were analyzed on a denaturing PAGE gel. **(C)** The discrimination factor Q (black square) was calculated from the ligation yield of the padlock probe on target X (red column) or S (blue column) in the presence of various blockers. **(D)** Ligation product of the padlock probe (100 nM) on 10 nM target X or S in the presence of 0, 10, 100, or 1000 nM blocker B (11, 13). **(E)** The ligation yield on target S in the presence of various concentrations of blocker was quantified after various reaction times. The ligation process modulated by B (11, 13) **(F)** or B (5, 11) **(G)** was monitored by isothermal titration calorimetry. The data were corrected by subtraction of blank experiments and then fitted via nonlinear regression. The values of △H, Ka, and N can be found in Fig. S6. Error bars represent the means ± SD, n = 3. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

from the template by toehold invasion. Although ligated products can still react with the blocker and template complex, at least non-ligated probe will have an almost equal chance to compete. The target S was recycled about 5 times with the aid of B (11,13), while the ligation reaction with target X was tightly suppressed at molar ratio of 10:1 between blocker and template. This template recycling mechanism is mainly responsible for the discrimination sensitivity. Additionally, template recycling was assessed by a time course of the ligation reaction at a molar ratio of P:T = 10:1 and various concentrations of blocker, from 0 to 100-fold of the template. The ligation reaction stopped at around 15 min in the absence of blocker, but B (11, 13) pushed the ligation reaction on target S further, and the ligation yield did not stop increasing even at the terminal time point of 45 min (Fig. 3E). Moreover, no obvious cross-ligation was observed in the template recycling assay. Therefore, the designed toehold blocker successfully established a new enzyme-free template recycling function for *in vitro* biochemical ligation reactions, allowing high-fidelity discrimination between target sequences with single-nucleotide mutations.

The kinetics of branch migration were further studied by isothermal titration calorimetry, which is considered a standard method for measurement the energy of binding between biomolecules (Gourishankar et al., 2004; Vander Meulen and Butcher 2012; Velazquez-Campoy et al., 2004). The blockers B (5, 11) with very weak ligation yield on target X, and B (11, 13) with a high ligation yield, were assessed in forward titration (probe titrated against blocker and target complex) and reverse titration (blocker titrated against probe and target complex) (Fig. S6). Interestingly, B (11, 13) exhibited efficient strand exchange in both the forward and reverse directions, whereby the forward was slightly favored (Fig. 3F). By contrast, B (5, 11) showed a completely different behavior (Fig. 3G), which may explain the difference in ligation yield.

3.4. Generality applicability of the toehold-assisted padlock probe for identifying SNV at low allele fractions

Based on SNV discrimination via a specific toehold blocker, we further designed a genetic variation detection system by combining the sensitivity of the toehold padlock with isothermal rolling circle amplification (RCA) (Fig. 4A). First, a padlock probe was ligated on a target strand mix (0, 0.1, 1, 10, and 100% of target S in target X) in the presence of a designed toehold blocker, after which the ligated circular ssDNA was isolated and subjected to RCA. The results demonstrated that the toehold-assisted padlock probe can detect SNV with ultra-high sensitivity at a variant/wild type ratio of 1/1000 (Fig. 4B). Blockers for 14 well-known human nonpathogenic SNVs across five genes including EGFR, TP53, KRAS, GNAS, and STK11, which are widely utilized as biomarkers for cancer diagnosis, were designed and tested on synthetic DNA strands. To quantify the detection at a SNV/WT ratio of 1/1000, the normalized fold-change β was calculated as previously reported (Wang and Zhang 2015). An average β value of about 966 was obtained, whereby the highest value reached 3731 for the R273C variant of TP53 (Fig. 4C and Fig. S7). Notably, the β value varied significantly for different genes, indicating that the secondary structure of the target may have an effect on probe binding. In addition, when G or C is substituted by T (e.g., R213*, R273C and G12V), a higher β value can be achieved. Furthermore, when G or C base is substituted by T in the same target (e.g., G12C and G12V), this also leads to a higher β. These results provide a valuable reference for designing toehold blockers and padlock probes for other SNVs.

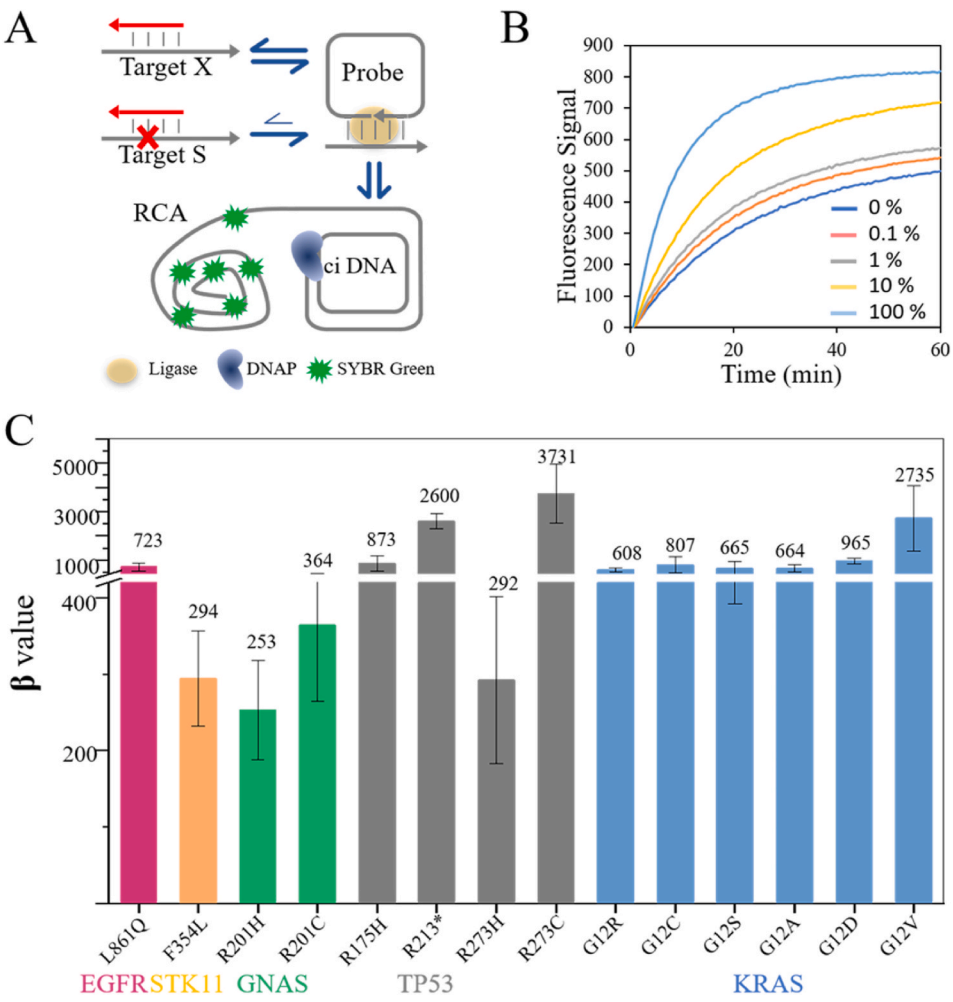

Fig. 4. Detection of genomic SNVs at a low allele fraction of 0.1%. **(A)** Schematic of genetic SNV detection based on the toehold-assisted padlock probe combined with rolling circle amplification (RCA). **(B)** Real-time monitoring of the amplification of the ligated products of padlock probes at various ratios of target S diluted by target X in the presence of B (11, 13). **(C)** The normalized fold-change β for detecting synthetic DNA of 14 non-pathogenic SNVs diluted to 0.1% by wild-type DNA molecules. Error bars represent the means ± SD, n = 3.

Overall, the toehold padlock can be robustly applied to ultra-sensitive genetic variation analysis at single-nucleotide resolution. Compared with other approaches, there is a notable competitive advantage of this method, because only a very narrow detection window of 29 nts is sufficient for analysis. This feature is valuable for the analysis of highly fragmented DNA molecules from clinical samples.

3.5. Accurate genotyping of a target SNV in cell-free DNA

The cfDNA found in blood at low concentrations is highly fragmented, with an average size of 160–180 bp. To detect more DNA fragments containing the target SNV, single primer asymmetric amplification was developed to enrich the target fragment in the presence of ddNTPs, in which a short single-stranded DNA with an inactivated 3' terminus is efficiently amplified (Fig. 5A). Theoretically, the fragmented DNA can be enriched irrespective of its length, as long as it contains the primer binding. Furthermore, ddNTP termination to a large extent avoids the cross reaction between these amplified ssDNA. According to the size assessment of the original cfDNA (Fig. 5B top), the mean length is 173 bp, with 98.2% over 50 bp in length, 90.5% over 80 bp, and 54.1% over 150 bp. By contrast, the ssDNA from asymmetric amplification is mainly in the size range of 70–80 nts (Fig. 5B bottom), the optimal size for the toehold-assisted padlock probe. Because the necessary sequence for GfDtP is approximately 50 nts in length, the broad-spectrum enrichment of fragmented DNA and highly sensitive toehold padlock probe could be combined to construct the highly sensitive system for the genotyping of fragmented DNA (Fig. S8).

For quantitative assessment, the total target DNA and the SNV were measured in the absence of blocker ($Ct_{no\text{-}b}$ measured from the RCA amplification curve) and in the presence of WT toehold blocker ($Ct_{one\text{-}b}$). The signal-to-noise ratio of the sample was measured in the presence of both WT and SNV toehold blockers ($Ct_{two\text{-}b}$), in which all the target DNA fragment were blocked (Figs. S9A-C). Then, the fraction of SNV in total DNA fragments, i.e., the observed mutation ratio, was calculated as: ($Ct_{two\text{-}b}$-$Ct_{one\text{-}b}$) × 100%/($Ct_{two\text{-}b}$-$Ct_{no\text{-}b}$). When the synthesized 100% WT and 100% SNV DNA fragment, the corresponding observed mutation ratios were 0 and 100%, with the $Ct_{one\text{-}b}$ almost equal to $Ct_{two\text{-}b}$ and $Ct_{no\text{-}b}$, respectively (Figs. S9D and S9E).

For further validation of the fragmented DNA detection, we synthesized 90 nts DNA carrying EGFR WT and the L861Q mutation (c.2582 T > A; (Harvey et al., 2020)) were analyzed. The mixed samples with variant allele fractions of 5, 1 and 0.1%, were quantified and yielded clearly different observed mutation ratios (Fig. 5C). The observed mutation ratios and the allele fractions fitted well to the regression equation $y = 2.1933 \times e^{(-x/0.79298)} + 0.12527$ with $R^2 = 0.9992$, which could be used as a standard reference for absolute quantification. Then, samples comprising 1% SNV in the form of DNA fragments with fixed length of 50 bp, 90 bp, 150 bp and the mixture of these were quantified and the observed mutation ratios were calculated as 31 ± 2% for all samples with various sizes (Fig. 5D).

Finally, the EGFR SNV L861Q in standard cfDNA (Horizon Discovery Ltd.) was analyzed by GfDtP. In addition to the DNA fragments enriched by asymmetric PCR (Fig. S10A), DNA fragments with fixed lengths of 80 and 150 bp were amplified from the 1% standard EGFR cfDNA by conventional PCR (Fig. S10B). Interestingly, the observed mutation ratio was quantified as respectively 1.02% and 0.59% for the 80 bp and 150

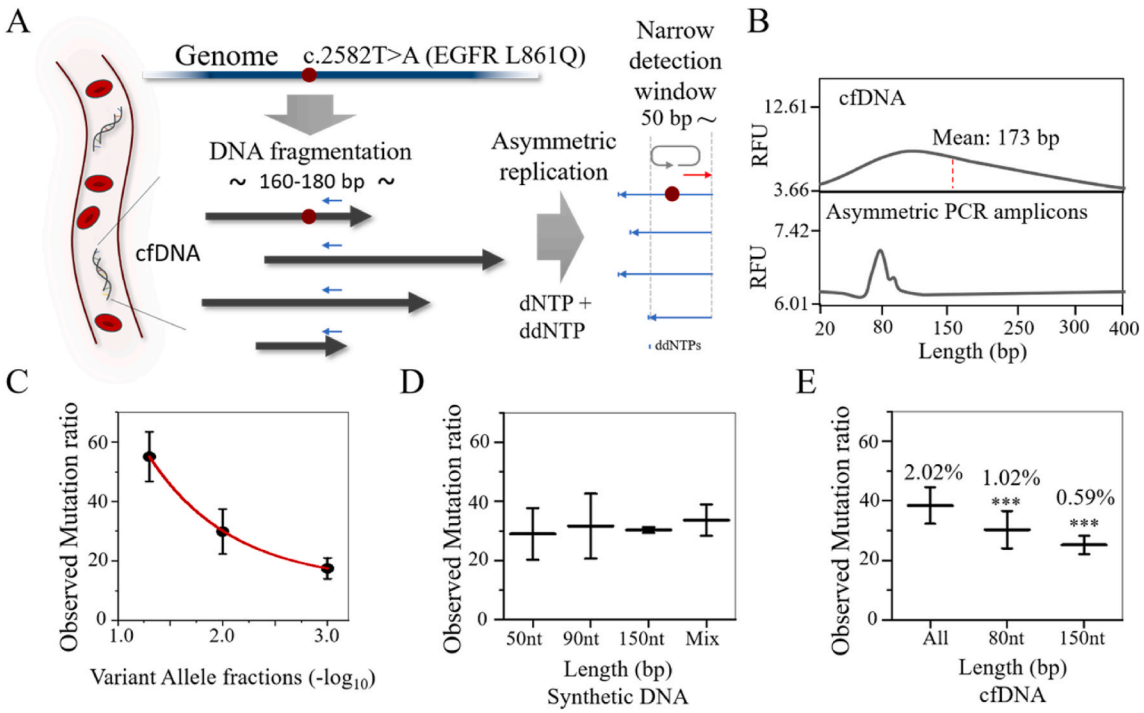

Fig. 5. Detection of a EGFR SNV at low variant allele fraction in reference cfDNA. (**A**) Scheme of EGFR mutation detection. (**B**) The fragment size distribution range of cfDNA (upper right) and asymmetric PCR amplicons (bottom right) was analyzed by capillary electrophoresis. (**C**) The regression curve of the observed mutation ratio and the inverse of the logarithmic variant allele fractions. The synthetic 90 nts WT/variant mix was tested at 5, 1, and 0.1%. (**D**) DNA of 50 nts, 90 nts, 150 nts, and their mixtures comprising 1% synthetic variant fragment among the WT yielded very similar observed mutation ratios. (**E**) The L861Q mutation from 1% reference EGFR cfDNA was detected by GfDtP in comparison to sample enriched by standard PCR (80 and 150 bp amplicons). The variant allele fractions were calculated using the regression equation from Fig. 5C. Error bars represent the means ± SD, n = 3.

bp DNA samples. However, GfDtP yielded a value of up to 2.02% (Fig. 5E), which clearly indicated that GfDtP fit well with the principle for detection of highly fragmented DNA and could detect a larger proportion of target SNVs among cfDNA fragments. Furthermore, by GfDtP, EGFR L861Q at a low molar fraction was identified in clinical cfDNA samples from cancer patients (data not shown). Therefore, the results demonstrate that GfDtP has evident advantages for the detection of SNVs in highly fragmented DNA in a fast and accurate fashion.

4. Conclusion

We successfully developed a novel method for precise analysis of highly fragmented DNA is developed by combining broad-spectrum enrichment method with a high-resolution toehold-assisted padlock probe. Interestingly, toehold-assisted padlock probe presents a very efficient solution for the design of robust and versatile controllability in template dependent ligations. Even at a high molar ratio of probe over template, padlock ligation was achieved while efficiently suppressing undesired cross-ligation by 85% on average. Moreover, the developed probe was able to detect 0.1% SNV in human genomic DNA, while needing only an average sequence window of 29 nts. Toehold-assisted padlock probe designed for 14 different target SNVs were tested and showed promising results with all target sequences. Although this method is still time consuming (around 6 h after optimization), further optimization and integration into electrochemical devices can pave a way for practical routine liquid biopsy and “point-of-care” systems.

CRediT authorship contribution statement

Yanmin Gao: Conceptualization, Methodology, Data curation, Writing - original draft. **Hongyan Qiao:** Investigation, Validation. **Victor Pan:** Writing - review & editing. **Zhaoguan Wang:** Investigation. **Jiaojiao Li:** Investigation. **Yanan Wei:** Validation. **Yonggang Ke:** Writing - review & editing. **Hao Qi:** Project administration, Funding acquisition, Supervision, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the National Key R&D Program of China (Grant No. 2020YFA0712104) and National Natural Science Foundation of China (Grant No. 21778039).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bios.2021.113079.

References

Abbosh, C., Birkbak, N.J., Swanton, C., 2018. Early stage NSCLC - challenges to implementing ctDNA-based screening and MRD detection. Nat. Rev. Clin. Oncol. 15 (9), 577–586.

Ali, M.M., Li, F., Zhang, Z., Zhang, K., Kang, D.K., Ankrum, J.A., Le, X.C., Zhao, W., 2014. Rolling circle amplification: a versatile tool for chemical biology, materials science and medicine. Chem. Soc. Rev. 43 (10), 3324–3341.

Andersen, R.F., Spindler, K.L., Brandslund, I., Jakobsen, A., Pallisgaard, N., 2015. Improved sensitivity of circulating tumor DNA measurement using short PCR amplicons. Clin. Chim. Acta 439, 97–101.

Arko-Boham, B., Aryee, N.A., Blay, R.M., Owusu, E.D.A., Tagoe, E.A., Doris Shackie, E.S., Debrah, A.B., Adu-Aryee, N.A., 2019. Circulating cell-free DNA integrity as a diagnostic and prognostic marker for breast and prostate cancers. Canc. Genet. 235–236, 65–71.

Butler, T.M., Spellman, P.T., Gray, J., 2017. Circulating-tumor DNA as an early detection and diagnostic tool. Curr. Opin. Genet. Dev. 42, 14–21.
Chae, Y.K., Oh, M.S., 2019. Detection of minimal residual disease using ctDNA in lung cancer: current evidence and future directions. J. Thorac. Oncol. 14 (1), 16–24.
Chen, X., Gole, J., Gore, A., He, Q., Lu, M., Min, J., Yuan, Z., Yang, X., Jiang, Y., Zhang, T., Suo, C., Li, X., Cheng, L., Zhang, Z., Niu, H., Li, Z., Xie, Z., Shi, H., Zhang, X., Fan, M., Wang, X., Yang, Y., Dang, J., McConnell, C., Zhang, J., Wang, J., Yu, S., Ye, W., Gao, Y., Zhang, K., Liu, R., Jin, L., 2020. Non-invasive early detection of cancer four years before conventional diagnosis using a blood test. Nat. Commun. 11 (1), 3475.
Diehl, F., Li, M., Dressman, D., He, Y., Shen, D., Szabo, S., Diaz Jr., L.A., Goodman, S.N., David, K.A., Juhl, H., Kinzler, K.W., Vogelstein, B., 2005. Detection and quantification of mutations in the plasma of patients with colorectal tumors. Proc. Natl. Acad. Sci. U.S.A. 102 (45), 16368–16373.
Forshew, T., Murtaza, M., Parkinson, C., Gale, D., Tsui, D.W., Kaper, F., Dawson, S.J., Piskorz, A.M., Jimenez-Linan, M., Bentley, D., Hadfield, J., May, A.P., Caldas, C., Brenton, J.D., Rosenfeld, N., 2012. Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. Sci. Transl. Med. 4 (136), 136ra168.
Gourishankar, A., Shukla, S., Ganesh, K.N., Sastry, M., 2004. Isothermal titration calorimetry studies on the binding of DNA bases and PNA base monomers to gold nanoparticles. J. Am. Chem. Soc. 126 (41), 13186–13187.
Harvey, R.D., Adams, V.R., Beardslee, T., Medina, P., 2020. Afatinib for the treatment of EGFR mutation-positive NSCLC: a review of clinical findings. J. Oncol. Pharm. Pract. 26 (6), 1461–1474.
Heidary, M., Auer, M., Ulz, P., Heitzer, E., Petru, E., Gasch, C., Riethdorf, S., Mauermann, O., Lafer, I., Pristauz, G., Lax, S., Pantel, K., Geigl, J.B., Speicher, M.R., 2014. The dynamic range of circulating tumor DNA in metastatic breast cancer. Breast Cancer Res. 16 (4), 421.
Heitzer, E., Speicher, M.R., 2018. One size does not fit all: size-based plasma DNA diagnostics. Sci. Transl. Med. 10 (466).
Kim, J., Mrksich, M., 2010. Profiling the selectivity of DNA ligases in an array format with mass spectrometry. Nucleic Acids Res. 38 (1), e2.
Knebel, F.H., Bettoni, F., da Fonseca, L.G., Camargo, A.A., Sabbaga, J., Jardim, D.L., 2019. Circulating tumor DNA detection in the management of anti-EGFR therapy for advanced colorectal cancer. Front. Oncol. 9, 170.
Li, J., Yao, B., Huang, H., Wang, Z., Sun, C., Fan, Y., Chang, Q., Li, S., Wang, X., Xi, J., 2009. Real-time polymerase chain reaction microRNA detection based on enzymatic stem-loop probes ligation. Anal. Chem. 81 (13), 5446–5451.
Li, Y., Li, T., Liu, B.F., Hu, R., Zhu, J., He, T., Zhou, X., Li, C., Yang, Y., Liu, M., 2020. CRISPR-Cas12a trans-cleaves DNA G-quadruplexes. Chem. Commun. (Camb.) 56 (83), 12526–12529.
Mouliere, F., Chandrananda, D., Piskorz, A.M., Moore, E.K., Morris, J., Ahlborn, L.B., Mair, R., Goranova, T., Marass, F., Heider, K., Wan, J.C.M., Supernat, A., Hudecova, I., Gounaris, I., Ros, S., Jimenez-Linan, M., Garcia-Corbacho, J., Patel, K., Ostrup, O., Murphy, S., Eldridge, M.D., Gale, D., Stewart, G.D., Burge, J., Cooper, W. N., van der Heijden, M.S., Massie, C.E., Watts, C., Corrie, P., Pacey, S., Brindle, K.M., Baird, R.D., Mau-Sorensen, M., Parkinson, C.A., Smith, C.G., Brenton, J.D., Rosenfeld, N., 2018. Enhanced detection of circulating tumor DNA by fragment size analysis. Sci. Transl. Med. 10 (466).
Peters, S., Camidge, D.R., Shaw, A.T., Gadgeel, S., Ahn, J.S., Kim, D.W., Ou, S.I., Perol, M., Dziadziuszko, R., Rosell, R., Zeaiter, A., Mitry, E., Golding, S., Balas, B., Noe, J., Morcos, P.N., Mok, T., Investigators, A.T., 2017. Alectinib versus crizotinib in untreated ALK-positive non-small-cell lung cancer. N. Engl. J. Med. 377 (9), 829–838.
Soria, J.C., Ohe, Y., Vansteenkiste, J., Reungwetwattana, T., Chewaskulyong, B., Lee, K. H., Dechaphunkul, A., Imamura, F., Nogami, N., Kurata, T., Okamoto, I., Zhou, C., Cho, B.C., Cheng, Y., Cho, E.K., Voon, P.J., Planchard, D., Su, W.C., Gray, J.E., Lee, S. M., Hodge, R., Marotti, M., Rukazenkov, Y., Ramalingam, S.S., Investigators, F., 2018. Osimertinib in untreated EGFR-mutated advanced non-small-cell lung cancer. N. Engl. J. Med. 378 (2), 113–125.
Underhill, H.R., Kitzman, J.O., Hellwig, S., Welker, N.C., Daza, R., Baker, D.N., Gligorich, K.M., Rostomily, R.C., Bronner, M.P., Shendure, J., 2016. Fragment length of circulating tumor DNA. PLoS Genet. 12 (7), e1006162.
Vander Meulen, K.A., Butcher, S.E., 2012. Characterization of the kinetic and thermodynamic landscape of RNA folding using a novel application of isothermal titration calorimetry. Nucleic Acids Res. 40 (5), 2140–2151.
Velazquez-Campoy, A., Ohtaka, H., Nezami, A., Muzammil, S., Freire, E., 2004. Isothermal titration calorimetry. Curr. Protoc. Cell Biol. https://doi.org/10.1002/0471143030.cb1708s23. Chapter 17, Unit 17 18.
Wang, J.S., Zhang, D.Y., 2015. Simulation-guided DNA probe design for consistently ultraspecific hybridization. Nat. Chem. 7 (7), 545–553.
Wilson, R.H., Morton, S.K., Deiderick, H., Gerth, M.L., Paul, H.A., Gerber, I., Patel, A., Ellington, A.D., Hunicke-Smith, S.P., Patrick, W.M., 2013. Engineered DNA ligases with improved activities in vitro. Protein Eng. Des. Sel. 26 (7), 471–478.
Zhang, D.Y., Chen, S.X., Yin, P., 2012. Optimizing the specificity of nucleic acid hybridization. Nat. Chem. 4 (3), 208–214.