Posted on Authorea 17 May 2021 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.162128311.16782690/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

A tailored LNA clamping design principle: efficient, economized, specific and ultrasensitive for the detection of point mutations

Hao Yang[1], Mengqiu Yan[1], Gaolian Xu[1], Xiaohua Qian[1], Ruiying Zhao[1], Yuchen Han[1], Ling Zhang[1], Hongchen Gu[2], and Hong Xu[3]

[1]Affiliation not available
[2]Shanghai Jiao Tong University
[3]Shanghai Jiao Tong University, Shanghai, China

May 17, 2021

Abstract

In the development of personalized medicine, the ultrasensitive detection of point mutations that correlate with diseases is important to improve the efficacy of treatment and guide clinical medication. In this study, locked nucleic acid (LNA) was introduced as an amplification suppressor of a massive number of wild-type alleles in an amplification refractory mutation system (ARMS) to achieve the detection of low-abundance mutations with high specificity and sensitivity of at least 0.1%. By integrating the length of clamp, base type, number and position of LNA modifications, we have established a "shortest length with the fewest LNA bases" principle from which each LNA base would play a key role in the affinity and the ability of single base discrimination could be improve. Finally, based on this LNA design guideline, a series of the most important single point mutation sites of epidermal growth factor receptor (EGFR) was verified to achieve the optimal amplification state which as low as 0.1% mutation gene amplification was not affected under the wild gene amplification was completely inhibited, demonstrating that the proposed design principle has good applicability and versatility and is of great significance for the detection of circulating tumor DNA.

A tailored LNA clamping design principle: efficient, economized, specific and ultrasensitive for the detection of point mutations

Hao Yang++ [a], Mengqiu Yan++ [a] , Gaolian Xu [a], Xiaohua Qian [a], Ruiying Zhao [b], Yuchen Han [b], Lin Zhang [c], Hongchen Gu*[a], Hong Xu*[a]

[a]School of Biomedical Engineering/Med-X Research Institute, Shanghai Jiao Tong University, Shanghai, 200030, PR China

[b]Shanghai Chest Hospital, Shanghai Jiao Tong University, Shanghai, 200030, PR China

[c]Shanghai Mag-Gene Nano Tech Co., Ltd, Shanghai, 200030, PR China

++ To whom contribute equally to the article.

* To whom correspondence should be addressed.

Hong Xu, Add.: School of Biomedical Engineering/Med-X Research Institute, Shanghai Jiao Tong University,1954 Huashan Road, Xuhui District, Shanghai, China; Tel: 62933743; Fax: 62932907; Email: xuhong@sjtu.edu.cn

Abstract

Posted on Authorea 17 May 2021 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.162128311.16782690/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

In the development of personalized medicine, the ultrasensitive detection of point mutations that correlate with diseases is important to improve the efficacy of treatment and guide clinical medication. In this study, locked nucleic acid (LNA) was introduced as an amplification suppressor of a massive number of wild-type alleles in an amplification refractory mutation system (ARMS) to achieve the detection of low-abundance mutations with high specificity and sensitivity of at least 0.1%. By integrating the length of clamp, base type, number and position of LNA modifications, we have established a "shortest length with the fewest LNA bases" principle from which each LNA base would play a key role in the affinity and the ability of single base discrimination could be improve. Finally, based on this LNA design guideline, a series of the most important single point mutation sites of epidermal growth factor receptor (EGFR) was verified to achieve the optimal amplification state which as low as 0.1% mutation gene amplification was not affected under the wild gene amplification was completely inhibited, demonstrating that the proposed design principle has good applicability and versatility and is of great significance for the detection of circulating tumor DNA.

Key words: locked nucleic acid (LNA), LNA clamping modification principle, point mutant, ultrasensitive detection with high specificity, cell-free tumor DNA.

Introduction

Point mutants are a class of somatic mutants that serve as significant biomarkers for heritable human diseases, drug development, cancer detection, companion diagnostics and prognosis[1, 2] . The greatest challenge of point mutant detection is the identification of very low-abundance mutant targets against a background of large amounts of wild-type alleles. Since the ratio of mutant targets to wild-type genes is 1%-0.1% or even lower, detection methods with high sensitivity and specificity are urgently required[3, 4] . To meet these requirements, numerous methods have been developed, including amplification refractory mutation system (ARMS)[5] , next-generation sequencing (NGS)[6, 7] , high-resolution melting[8] and microarrays[9] . High-resolution melting and microarrays are quick and high throughput whereas usually lack sensitivity and specificity. Next-generation sequencing (NGS) can reach the limit of detection below 0.1% by enhancing the sequencing depth to more than 10000X[10] but is expensive and time-consuming.

ARMS is based on the application of allele-specific (AS) PCR primers that are perfectly matched to the mutant targets at 3' end[11, 12] . However, imperfect hybridization between the AS primer and wild-type gene may lead to false-positive results with relatively low efficiency[13, 14] . To further improve the specificity, peptide nucleic acid (PNA)[15] , a high-affinity DNA analogue with a polypeptide-like skeleton, was introduced to bind the wild-type gene during the PCR annealing stage, which impedes the extension of the AS primer and thus inhibits the amplification of wild-type genes[16] . Simultaneously, the PNA-DNA duplex provides excellent mismatch discriminatory power, and has already been used as a clamping agent for the detection of many mutant alleles[17-19] . But PNA also has some shortcomings, including poor solubility, high cost, a long synthesis period, and an inflexible design pattern (changing the length only)[20] . In particular, PNA could also inhibit the amplification of mutant templates with ultra-low concentrations and potentially producing false negatives[18] .

Locked nucleic acid (LNA), another DNA analogue with excellent water solubility[21, 22] , can be synthesized using the same process as conventional oligonucleotides with relative low cost. Moreover, LNA-modified oligonucleotides have great flexibility in terms of different numbers, positions and base types, thus they have been widely used as probes[23, 24] or primers for nucleic acid detection[25, 26] . Nevertheless, the principle of LNA-modified oligonucleotide design used as clamping systems remains rare [27-29] . The great challenge for the clamping system is to completely inhibit the amplification of massive numbers of wild-type gene, while with great sensitivity to mutation. Faced with this arduous task, there are few systematic studies on the thermodynamic parameters of LNA clamping[30, 31] . In the absence of effective design theory guidelines, people either designed LNA-modified clamping based on their own experience or obtained optimized LNA-modified clamping through many tedious repeated optimization experiments. In this study, based on the LNA-mediated PCR clamping system (LMPC), we proposed for the first time how to realize an optimized LNA clamping design strategy using AS primers for the detection of point mutations to achieve high sensitivity and specificity. First, LNA clamping are systemically investigated via different parameters,

Posted on Authorea 17 May 2021 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.162128311.16782690/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

including the number of, base type, and position of LNA modifications, ensuring that LNA clamping can completely inhibit the wild-type gene amplification signal but amplify mutant targets with high efficiency. And then, LNA clamping acquired by the new design guideline was successfully used as an amplification suppressor in ARMS PCR to achieve the ultrasensitive detection of the five most important mutant of epidermal growth factor receptor (EGFR) alleles with sensitivity of 0.1% against a background of 10,000 wild-type alleles.

Materials & methods

2.1 Oligonucleotides, templates and reagents

LNA-modified oligonucleotides (labeled with NH_2-C6 at the 3' end), primers, plasmids and TaqMan probes were all obtained from Sangon Biotech (Shanghai) Co. Ltd. with the sequences shown in Supplemental Table S1 and S3. All the TaqMan probes were designed using PrimerExpress® software (Applied Biosystems, Foster City, CA). Champagne TaqTM DNA polymerase and EvaGreen® Dye (20x stock solution in water) were purchased from Vazyme Biotech Co. Ltd. and Biotium, respectively. The EGFR Multiplex Cell Free Reference Standard Set was purchased from Horizon Discovery (HDx). AmoyDx® DNA and RNA Extraction Kits, the Silica-based Spin Column and the Human EGFR Gene Mutation Detection Kit (Multiplex real-time PCR) were purchased from Amoy Diagnostics Co., Ltd.

2.2 Preparation of PCR templates

The genomic DNA of A549 cells was extracted using a Genomic DNA Kit (TianGen Biotech Co. Ltd) and served as a wild-type background with the concentration quantified using digital PCR (QX200 Droplet Digital PCR, Bio-Rad). Plasmids containing different EGFR point mutation fragments (L858R, T790M, S768I, L861Q and G719A) were constructed and quantified by digital PCR.

2.3 Real-time PCR

Standard real-time PCR was carried out in a total volume of 20 μl containing 0.15 μM each primer, 0.1 μM TaqMan MGB probe, 1.5 μM LNA oligonucleotides and 1x buffer (10 mM Tris-HCl, pH 8.3, 10 mM KCl, 0.3 M dNTP and 1.5 mM $MgCl_2$). Thermal cycling was performed using a LightCycler® 96 System (Roche) with two amplification steps: initial denaturation at 95°C for 2 min, followed by 15 cycles of 95°C for 1 s, 70°C for 20 s and 65°C for 20 s, then 30 cycles of 95°C for 1 s, 70°C for 20 s and 60°C for 20 s. The fluorescence signal was collected in the second step.

2.4 Melting curve experiments

The composition of the 20 μl melting reaction included 1.5 μM LNA clamping, 1.5 μM complementary segments (Supplemental Table S2), 1x Eva Green, and 1x PCR buffer. Melting curve assays were carried out on the LightCycler® 96 System (Roche) with following program: 55°C for 10 min, denaturation at 95°C for 10 s, and cooling to 55°C for 50 s, followed by finally increasing the temperature from 55°C to 97°C at a heating rate of 0.2°C/s, and the fluorescence was acquired at 0.1°C intervals. The melting curve was analyzed using LightCycler® software.

2.5 Agarose gel electrophoresis of PCR products

10 μl of PCR product was mixed with 1 μl of loading buffer, and electrophoretic analysis was carried out in 2% agarose gel stained with Tanon Nucleic Acid Dye (10000×). The sample was electrophoresed at a voltage of 150 mV for 30 min and then imaged on a Vilber Lourmat.

2.6 Cell-free DNA and clinical sample detection

The EGRF cfDNA point mutation reference standard samples were diluted in wild-type DNA (10000 copies) to obtain a series of mutation samples with concentrations of 1%, 0.4%, 0.2%, 0.1%, and 0% (negative). The genomic DNA of clinical formalin-fixed and paraffin-embedded (FFPE) samples was extracted using DNA Extraction Kits (Amoy Diagnostics) and tested by the Human EGFR Gene Mutation Detection Kit (Amoy Diagnostics) and LMCS we developed respectively.

Posted on Authorea 17 May 2021 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.162128311.16782690/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Results and discussion

3.1 Establishment of the LNA-mediated real-time PCR clamping system

The LNA-mediated PCR clamping system (LMPC) is schematically presented in Figure 1 Here, the AS primer is specifically complementary to mutant targets, while the LNA-modified amplification suppressor (LNA clamping) is perfectly complementary to wild-type with the mutant site in the middle. The completely complementary LNA-wild-type gene heteroduplex has a higher melting temperature (Tm) than that of AS primer-wild-type gene. Hence, during amplification, the formation of a heteroduplex between the LNA-modified amplification suppressor and wild-type DNA is favored over the formation of the primer-wild-type DNA duplex, thus leading to the suppression of AS primer hybridization and extension, which promises high specificity. However, due to the existence of the single base mismatch, the Tm for LNA-modified amplification suppressor hybridization with mutant-type DNA was sharply decreased, leaving the AS primer able to hybridize effectively with mutant targets and therefore to achieve the highly efficient amplification of rare mutant targets in the presence of a large number of wild-type alleles. For LMPC, the expected result is to completely inhibit the amplification of the massive number of wild-type gene (such as 10000 copies) while exerting as little effect as possible on the amplification efficiency of low-abundance mutant targets (e.g., lower than 10 copies), as shown in Figure 1B (optimal state).

Unfortunately, under real circumstances, the amplification of mutant targets would be suppressed in most clamping systems (as shown in Figure 1B, deficient state) even with the use of LNA oligonucleotides as clamping agents due to the lack of LMPC design theory. To obtain an LNA-modified amplification suppressor with excellent performance, the role of the difference in Tm between the LNA-mutant-type and LNA-wild-type heteroduplex was explored by investigating the effects of length, number of LNA bases and modification position. As illustrated in the schematic amplification curve of q-PCR (Figure 1B), although the wild-type gene could be inhibited by LNA clamping in both situations, there is a significant difference in mutant targets. In the optimal state, the presence or absence of LNA clamping had little effect on the amplification of rare mutant targets, while inhibition occurred in the deficient state. In general, to avoid limiting mutant amplification, the Tm of LNA clamping with a mutant-type DNA heteroduplex should be far lower than that of the AS primer, which makes the LNA clamping unable to recognize and hybridize with mutant targets during PCR. The premise is that as the suppression of mutant targets by LNA clamping weakens, the detection sensitivity should be gradually improved. Therefore, this research aims to determine what kind of LNA clamping is the most effective as well as to establish an LNA clamping design principle, which will greatly improve the performance of LMPC amplification for the highly sensitive detection of mutant targets.

3.2 The key factors in mismatch discrimination of LNA clamping

As shown in the Figure 1, the structure of the LNA clamping may be important for the sensitivity and specificity of the detection of the mutation targets, and our pre-experiments showed that the mismatch discrimination capability of the fully modified LNA oligonucleotides was not as good as that of the interval-modified LNA oligonucleotides (LNA modification at every second base, the data were not shown in this article), and the same result had also been reported in a previous study[27, 32] . It is precisely this interesting phenomenon to make us ponder whether there is a more rational design philosophy from which the clamping system can perform better with as few as LNA bases in mutant target detection. To investigate the design rules and obtain the optimal design parameter, the L861Q of the EGFR point mutant was chosen as the research model.

First, several interval LNA modification oligomers with different lengths (11-15 bp) were designed to explore the influence of LNA-modified clamping length on the detection performance. As the length of LNA increased, the Ct value of wild-type gene in real-time PCR increased gradually, indicating that the amplification efficiency of wild-type gene decreased gently (Figure S1). When the length of LNA clamped up to 14 bp, of which the Tm was more than 10 degrees higher than that of the AS primer, wild-type gene (10000 copies) were completely inhibited within 30 cycles of amplification plus ten cycles of preamplification. Similarly, the amplification of mutant targets (the ratio of mutant targets to wild-type gene was 1% and 0.1%, respectively)

Posted on Authorea 17 May 2021 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.162128311.16782690/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

maintained the same trend, but the inhibition efficiency was weaker than that of the wild-type gene owing to the existence of a mutant base (Figure S1D).

For LMPC, the formation of LNA/wild-type gene duplexes is the key factor to inhibit wild-type gene amplification. To evaluate the stability of LNA/DNA duplexes, the melting temperature (Tm) of the LNA clamping was measured using a melting curve (Figure S1) and predicted by IDT OligoAnalyzer (Table S3) respectively. Figure 2A illustrates that Tm (perfect match) is higher than Tm (mismatch) and both are increased with the length of LNA-modified clamping. In addition, the difference in Tm between matched and mismatched duplexes (denoted as ΔTm^a, ΔTm^a =Tm (perfect match)-Tm (mismatch)) of 861 LNA1-LNA5 is very close(9.6°C-11.2°C). The high Tm means high binding affinity, so if the Tm is high enough, the extension of AS primer on wild-type gene can be completely inhibited and the wild-type gene would be totally bound by the clamping during the annealing stage of PCR (such as LNA4 and LNA5). On the other hand, in such a high Tm based on above interval LNA modification clamping, the Tm (mismatch) between LNA clamping and mutant targets would be also high, which would reduce the amplification efficiency of mutant targets. It is clear in principle that Tm (perfect match) is a key factor in inhibiting the amplification of wild-type gene, but Tm (mismatch) determines the efficiency of mutant targets amplification indicating that the simplex LNA clamping modification pattern, which has no influence on ΔTm^a, would make it difficult to achieve high mismatch discrimination ability. Therefore, an LNA clamping design model needs to be explored in a much calmer and more rational manner.

To further reveal the effect of ΔTm^a, the mismatch discrimination ability of LNA-modified (LNA4) or unmodified oligomers (LNA6) at the mutant site was designed and studied. The Ct value of mutant targets are 16.56 (1%), 19.86 (0.1%) and 18.25 (1%), 20.91 (0.1%) for LNA4 and LNA6, respectively, suggesting that the inhibition of LNA4 for mutant targets is significantly lower (Figure S2). Then, the Tm values of the matched and mismatched states of both oligomers were analyzed according to the melting curves shown in Figure 2B. It is noticeable that the ΔTm^a of 861 LNA4 and LNA6 are 10.2°C and 6.96°C, respectively, indicating that the modification of LNA on the mutant site, which made ΔTm^a higher, would dramatically increase the mismatch discrimination ability, leading to higher amplification efficiency of mutant targets. Therefore, not only is high Tm of LNA-modified clamping needed to ensure the amplification suppression of wild-type gene, but more importantly, a high ΔTm^a is required to guarantee both effective suppression of wild-type gene amplification and the minimum impact on the mutation target amplification efficiency.

3.2 The influence of the number of LNA bases on the mismatch discrimination ability of LNA clamping

Although interval modification with LNA enhanced the sensitivity of the LNA clamping system to 0.1% under the background of 10000 copies of wild-type gene, the amplification efficiency of mutant targets was still suppressed unfortunately compared with the LNA-free group, as described above (the Ct value of real-time PCR was increased by at least 2.7, as shown in Figure S1). To find the essential factor of ΔTm^a in LMPC, a series of LNA clamps with different numbers of LNA bases but with similar Tm were designed (LNA7-10). The Tm of these clamping segments is high enough to ensure that the wild-type amplification could be completely restrained (shown in Table S3). In addition, the mutation location was selected preferentially for modification according to the result of Figure 2B. As shown in Figure 3, the signal of wild-type gene could be completely inhibited by all LNA clamping to avoid the interference on mutant signal. It is worth noting that the Ct value of mutant targets decreases as the LNA-modified bases increases (Figure 3A, 3B and 3D). So, at the same Tm, the amplification efficiency of rare mutant targets could be improved by adding LNA bases. And, more remarkable, the Ct value of LNA 10 is almost the same as that of LNA free (no LNA clamping in PCR mix), indicating that it is possible to obtain the optimized LNA clamping system as we predicted earlier.

To further reveal the factors that cause differentiated property of these LNA clamps, melting curve analysis was adopted to measure the Tm at perfect match and mismatch state respectively. As shown in Figure 3E and 3F, the melting curves of all matched duplexes for wild-type gene almost overlap, which is consistent with the predicted results. However, the Tm of mismatched duplexes for mutant targets decreased as the number of LNAs increased, improving ΔTm^a and ultimately the enhancement amplification efficiency of

Posted on Authorea 17 May 2021 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.162128311.16782690/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

mutant targets. According to the result of LNA7-LNA10 (Figure 3B and Figure 3C), it was obvious that with the addition of LNA-modified bases to a certain extent, the ΔTm^a and detection efficiency of LNA clamping systems was assuredly increased. However, it should be noted that it is not enough to focus only on ΔTm^a. For example, the ΔTm^avalues of LNA4 and LNA5 are close to LNA10 (10.02°C, 9.74°C and 9.02°C, respectively), but the amplification efficiency of mutant targets was suppressed in the systems with LNA4 and LNA5 (Figure 2A). This might be because even when ΔTm^a is high enough, if the number of LNA bases is too much, as in LNA4 and LNA5, the Tm of the mismatch state would also increase further, promoting the formation of the LNA clamping-mutant target heteroduplex and finally reducing the amplification efficiency. Hence, it was indicated that the values of Tm and ΔTm^a should be kept within a reasonable range.

Here, to highlight the effect of single base mismatch and the addition of LNA bases more deeply, the Tm of LNA-modified clamping (LNA 7-10) must be tailored and designed to maintain a consistent value. It can be seen from the results of LNA7-10, with the addition of the LNA bases, that although the length of the clamping segments was shortened from 21 bp to 18 bp, the high Tm could still be maintained, indicating that the binding affinity from LNA bases increased. As shown in Table S4 (LNA 7-10), as the number of LNA bases increased, the Tm contributed by each LNA base (denoted as $\Delta Tm^{b'}$, $\Delta Tm^{b'}$= (Tm (with LNA)-Tm(without LNA))/number of LNA bases) also increased, which means that each LNA base plays a key role in the stability of the LNA-DNA heteroduplex. Hence, the affinity with the mismatched templates might be dramatically decreased, further effectively improving the ΔTm^a. However, if there were too many LNA bases in the clamping segments, such as LNA4 or LNA5, even if the length of the clamping segments was only 14 bp or 15 bp, the $\Delta Tm^{b'}$ decreased (Table S3), finally reducing the mismatch discrimination property. Therefore, it is worth noting that a small amount of LNA modification not only improves the performance of LNA clamping systems but also reduces the cost.

3.4 Optimization of LNA Clamping system

To obtain the best performance of the LNA clamping system, the LNA modification characteristics have been further studied and optimized. Combining the previous results, there are several key factors for LNA clamping design: one is high Tm for perfectly matched wild-type gene, the second is high $\Delta Tm^a$when the single base mismatch occurs, and the third is high increased affinity per LNA ($\Delta Tm^{b'}$). Thus, by adjusting the base type, location and amount of LNA bases, an optimal design model would be obtained. As mentioned before, we think that the most effective and economical LNA clamping design strategy should have the shortest length with the fewest LNA-modified bases; meanwhile, the Tm is high enough (approximately 75°C according to our results).

Based on the high Tm, which can inhibit wild-type gene amplification, a simulation analysis of various LNA modification patterns was evaluated systematically by applying different lengths from 11 bp-16 bp and different numbers of LNA-modified bases from 1-7. It is worth noting that when the length of the LNA clamp is less than 14 bp, the number of LNA-modified bases must be at least 6 to achieve a relatively high Tm to inhibit wild-type gene amplification. Therefore, the length of LNA clamping should be more than 14 bp to minimize the number of LNA bases (according to the simulation analysis data). On the basis of the rule, we chose the shortest length with the fewest LNA bases, and three optimized LNA oligomers were screened out from a large number of candidate sequences that we simulated and calculated (LNA11-LNA13). In Figure 4, the sensitivity and specificity of the obtained three LNA clamps plus LNA 10 were tested again. As shown in Figure 4E and Figure 4F, ΔTm^a(LNA11: 11.4°C, LNA12: 11.3°C, LNA13: 8.3°C) and $\Delta Tm^{b'}$ (LNA11: 2.57°C, LNA12: 3.20°C, LNA13: 3.70°C) (shown in Table S3 and S4) of the four LNA clamps are extremely high, indicating that the optimal LNA clamps with high ΔTm^a and $\Delta Tm^{b'}$ could be obtained from the design principle. Furthermore, the detection performance at low abundance mutations is verified by real-time PCR. In addition, the amplification of wild-type gene by primers other than LNA 11, whose Tm is relatively low, can be inhibited (Figure 4A). The Ct values of these LNA clamping systems (Figure 4B, 4C and 4D) were almost the same as that of the LNA-free system when 1% and 0.1% mutant targets were detected respectively, demonstrating that the optimized amplification efficiency of mutant targets can be achieved successfully in these systems.

Posted on Authorea 17 May 2021 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.162128311.16782690/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Hence, combining the experimental observations with previous suggestions, we propose the following principles to obtain the optimal LNA clamping system for the ultrasensitive detection of single point mutant targets with low abundance:

1. The length of the LNA clamp should be short (14-16 bp), and the number of LNA-modified bases should be 4-6. LNA clamping with the shortest length and the fewest LNA bases should be selected as the most valuable.
2. The ΔTm^a of perfectly matched and mismatched LNA/DNA duplexes should be more than 10°C, and the increased affinity per LNA base ($\Delta Tm^{b'}$) should be as high as possible.
3. The Tm of the LNA oligonucleotides should be at least 10 degrees higher than that of the AS primer to prevent the binding of the AS primer to wild-type templates.
4. The mutant site of the oligonucleotides must be modified, and the mutant site should be placed in the middle of the clamping chain.

In addition, based on the design strategy proposed above, to realize the precise regulation of the LNA-modified base automatically, a homebrew analysis software for obtaining the optimal LNA clamping sequence based on the OligoAnalyzer interface has been developed in the Python language. Through a combination of local traversal and cloud computing, we first enter the sequence containing mutation information into the software and annotate the mutant site. Then, several optimal LNA clamps can be obtained automatically and quickly, avoiding tedious labor operations and significantly saving time (Figure S4).

3.5 Validation and application of the LNA clamping design model

To further verify the reliability and universality of the LNA-modified design rules we proposed, other point mutation types in EGFR, such as L858R, T790M, S768I, and G719A, which are the most common mutation sites of EGFR and are closely related to the guidance of targeted drug treatment of cancer, were chosen to validate the LNA clamping design principle. Wild-type gene (10000 copies) and plasmids containing different ratios of EGFR point mutation templates were used to explore the sensitivity and specificity of LMPC. The optimal LNA clamping of different mutant sites were designed via homebrew design software, and the detection performance was evaluated using real-time PCR and agarose gel electrophoresis. Because the amplification of wild-type gene cannot be completely inhibited in the LNA-free group, the amplification signal of ultra-low-concentration mutation targets (mutant ratio of 0.1%) would be overlapped by the amplification of wild-type gene, which would finally interfere with the assessment of LMPC (Figure S3). Therefore, to highlight whether LNA clamp would affect the amplification of mutant targets before and after addition, wild-type genes were detected separately rather than added to the mutant targets in each group (LNA and LNA free). As shown in Figure 5, according to the amplification curves for all point mutant sites, the wild-type amplification in each mutant site can be totally suppressed by LNA clamp, and further, the amplification efficiency on mutant targets (100 and 10 copies) of LNA and LNA-free groups are almost identical except L858R, because the binding affinity of FP in this site was decreased by the second structure. Then, the amplicons were analyzed by agarose gel electrophoresis (Figure 5E). There is no target band of wild-type gene in the LNA clamping group, indicating that LNA clamping can totally inhibit the amplification of wild-type gene. Moreover, the brightness of the mutant band in the electrophoretogram was almost the same between the LNA and LNA-free groups, which is consistent with the real-time PCR results. Hence, the results indicated that our LNA clamping design principles are universal for point mutation at different sites.

Recently, many studies have shown the clinical utility of circulating cell-free DNA (cfDNA) as the most important target of liquid biopsy in cancer detection. However, ultra-low-abundance mutant target detection in variable amounts of cfDNA could be extremely challenging due to the demand for ultra-sensitivity [33] . To test the performance of the LNA clamping system in cfDNA samples, the EGRF cfDNA reference standard samples were diluted in wild-type DNA (10000 copies) to obtain a series of cfDNA samples with concentrations of 1%, 0.4%, 0.2%, 0.1%, and 0% (In addition, 500 copies of mutant targets as a positive control). The Ct values gradually increased with decreasing cfDNA concentration, and as little as 0.1% cfDNA could be detected correctly at all single mutant sites (Figure 6A). In general, in the field of nucleic

Posted on Authorea 17 May 2021 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.162128311.16782690/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

acid detection, the limit of detection of PCR is usually defined as measurand concentration that produces at least 95% positive replicates. To make the sensitivity of 0.1% in these systems more statistically significant, at least 16 replicates were performed in each site and as shown in Figure S5, 100% positive replicates could be obtained at S768I, L858R, G719X and T790M, except for L861Q (93.7%). The results show that our LNA clamping system can perform well and reach a sensitivity of 0.1% (10 copies of mutant targets in 10000 wild-type gene) in cfDNA, which is comparable to the level of digital PCR and achieves the limit sensitivity of real-time fluorescence quantitative PCR technology. In addition, LMPC we proposed here has excellent specificity. This is because that the templates in cfDNA sample contain all mutant sites, but only one site can be amplified in specific LMPC under such a complex sample, indicating that this system has the determination of different sites with high selectivity.

Furthermore, there are still some differences between the real samples and the standard samples. Therefore, it is crucial to further explore the properties of our system in clinical samples. The commercial EGFR Mutation Detection Kit was selected as a reference, and 24 clinical FFPE samples were tested using commercial kits and our LNA clamping system. As shown in Figure 6B, all the mutant results of EGFR were identified, and the detection results of the commercial kit test and our clamping system are exactly the same as the 100% compliance rate in these 24 samples, demonstrating that the method has definitely excellent performance in actual clinical sample detection and prospects for clinical application

Conclusions

In this study, we reported novel, universal and efficient LNA clamping design rules for ultra-low abundance point mutant detection. A series of LNA clamps with high Tm, ΔTm^a and increased affinity per LNA base (ΔTm^b') can easily be obtained through this system, which could completely suppress the amplification of massive wild-type gene but with almost no influence on a few copies of mutant targets. Therefore, it can overcome the bottleneck problem in existing clamping systems that the amplification efficiency of the mutant targets must be decreased to inhibit the wild-type gene amplification signal. Then, the sensitivity of rare mutant target detection could be greatly improved, which is extremely important for the detection of ctDNA. Furthermore, the design pattern can dramatically decrease the length of oligonucleotides and the number of LNA-modified bases, which is suitable for multiplex PCR. In addition, optimal LNA clamping could be obtained efficiently and easily by homebrew analysis software. Moreover, the LNA clamping system was applied in the most common point mutation of EGFR and can achieve a highly sensitive and specific result in both cfDNA standard samples and clinical samples, providing an efficient strategy for mutant target detection in ctDNA and good prospects for clinical application.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (Grant No. 31927803); Exploratory Projects of Natural Science Foundation of Shanghai (Grant No. 19ZR1476000); Innovation Research Plan supported by Shanghai Municipal Education Commission (Grant No. ZXWF082101) and SJTU funding (ZH2018ZDA35 and 19X150010008); Cross Research Fund of Biomedical Engineering of Shanghai Jiao Tong University (Grant No. ZH2018QNA69).

References

1. Siravegna, G., Marsoni, S., Siena, S., and Bardelli, A. (2017) Integrating liquid biopsies into the management of cancer, *Nature reviews. Clinical oncology* .

2. Cescon, D. W., Bratman, S. V., Chan, S. M., and Siu, L. L. (2020) Circulating tumor DNA and liquid biopsy in oncology, *Nature Cancer, 1* , 276-290.

3. Diehl, F., Li, M., Dressman, D., He, Y., Shen, D., Szabo, S., Diaz, L. A., Jr., Goodman, S. N., David, K. A., Juhl, H., Kinzler, K. W., and Vogelstein, B. (2005) Detection and quantification of mutations in the

Posted on Authorea 17 May 2021 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.162128311.16782690/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

plasma of patients with colorectal tumors, *Proc Natl Acad Sci U S A, 102* , 16368-16373.

4. Holdhoff, M., Schmidt, K., Donehower, R., and Diaz, L. A., Jr. (2009) Analysis of circulating tumor DNA to confirm somatic KRAS mutations,*J Natl Cancer Inst, 101* , 1284-1285.

5. Vannucchi, A. M., Pancrazzi, A., Bogani, C., Antonioli, E., and Guglielmelli, P. (2006) A quantitative assay for JAK2(V617F) mutation in myeloproliferative disorders by ARMS-PCR and capillary electrophoresis,*Leukemia, 20* , 1055-1060.

6. Newman, A. M., Bratman, S. V., To, J., Wynne, J. F., Eclov, N. C., Modlin, L. A., Liu, C. L., Neal, J. W., Wakelee, H. A., Merritt, R. E., Shrager, J. B., Loo, B. W., Jr., Alizadeh, A. A., and Diehn, M. (2014) An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage, *Nature medicine, 20* , 548-554.

7. Razavi, P., Li, B. T., Brown, D. N., Jung, B., Hubbell, E., Shen, R., Abida, W., Juluru, K., De Bruijn, I., Hou, C., Venn, O., Lim, R., Anand, A., Maddala, T., Gnerre, S., Vijaya Satya, R., Liu, Q., Shen, L., Eattock, N., Yue, J., Blocker, A. W., Lee, M., Sehnert, A., Xu, H., Hall, M. P., Santiago-Zayas, A., Novotny, W. F., Isbell, J. M., Rusch, V. W., Plitas, G., Heerdt, A. S., Ladanyi, M., Hyman, D. M., Jones, D. R., Morrow, M., Riely, G. J., Scher, H. I., Rudin, C. M., Robson, M. E., Diaz, L. A., Jr., Solit, D. B., Aravanis, A. M., and Reis-Filho, J. S. (2019) High-intensity sequencing reveals the sources of plasma circulating cell-free DNA variants, *Nat Med, 25* , 1928-1937.

8. Ladas, I., Fitarelli-Kiehl, M., Song, C., Adalsteinsson, V. A., Parsons, H. A., Lin, N. U., Wagle, N., and Makrigiorgos, G. M. (2017) Multiplexed Elimination of Wild-Type DNA and High-Resolution Melting Prior to Targeted Resequencing of Liquid Biopsies, *Clin Chem,63* , 1605-1613.

9. Kawamata, N., Ogawa, S., Seeger, K., Kirschner-Schwabe, R., Huynh, T., Chen, J., Megrabian, N., Harbott, J., Zimmermann, M., Henze, G., Schrappe, M., Bartram, C. R., and Koeffler, H. P. (2009) Molecular allelokaryotyping of relapsed pediatric acute lymphoblastic leukemia,*Int J Oncol, 34* , 1603-1612.

10. 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., and Rosenfeld, N. (2012) Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA, *Sci Transl Med,4* , 136ra168.

11. Little, S. (2001) Amplification-refractory mutation system (ARMS) analysis of point mutations, *Curr Protoc Hum Genet, Chapter 9* , Unit 9 8.

12. Newton, C. R., Graham, A., Heptinstall, L. E., Powell, S. J., Summers, C., Kalsheker, N., Smith, J. C., and Markham, A. F. (1989) Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS), *Nucleic Acids Res, 17* , 2503-2516.

13. Kwok, S., Kellogg, D. E., McKinney, N., Spasic, D., Goda, L., Levenson, C., and Sninsky, J. J. (1990) Effects of primer-template mismatches on the polymerase chain reaction: human immunodeficiency virus type 1 model studies, *Nucleic Acids Res, 18* , 999-1005.

14. Latorra, D., Campbell, K., Wolter, A., and Hurley, J. M. (2003) Enhanced allele-specific PCR discrimination in SNP genotyping using 3' locked nucleic acid (LNA) primers, *Hum Mutat, 22* , 79-85.

15. Kim, H. R., Lee, S. Y., Hyun, D. S., Lee, M. K., Lee, H. K., Choi, C. M., Yang, S. H., Kim, Y. C., Lee, Y. C., Kim, S. Y., Jang, S. H., Lee, J. C., and Lee, K. Y. (2013) Detection of EGFR mutations in circulating free DNA by PNA-mediated PCR clamping, *J Exp Clin Cancer Res, 32* , 50.

16. Hyrup, B., and Nielsen, P. E. (1996) Peptide nucleic acids (PNA): synthesis, properties and potential applications, *Bioorg Med Chem, 4* , 5-23.

17. Liu, Y., Lei, T., Liu, Z., Kuang, Y., Lyu, J., and Wang, Q. (2016) A Novel Technique to Detect EGFR Mutations in Lung Cancer, *Int J Mol Sci, 17* , 792.

Posted on Authorea 17 May 2021 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.162128311.16782690/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

18. Machnik, G., Skudrzyk, E., Buldak, L., Labuzek, K., Ruczynski, J., Alenowicz, M., Rekowski, P., Nowak, P. J., and Okopien, B. (2015) A Novel, Highly Selective RT-QPCR Method for Quantification of MSRV Using PNA Clamping Syncytin-1 (ERVWE1), *Mol Biotechnol, 57* , 801-813.

19. Yoshida, H., Kim, Y. H., Ozasa, H., Nagai, H., and Hirai, T. (2017) EGFR T790M detection in circulating tumor DNA from non-small cell lung cancer patients using the PNA-LNA clamp method, *Anticancer Research, 37* , 2721-2725.

20. Nikiforov, T. T., and Jeong, S., . (1999) Detection of hybrid formation between peptide nucleic acids and DNA by fluorescence polarization in the presence of polylysine, *Analytical Biochemistry, 275* , 248.

21. Koshkin, A. A., Nielsen, P., Meldgaard, M., Rajwanshi, V. K., Singh, S. K., and Wengel, J. (1998) LNA (locked nucleic acid): an RNA mimic forming exceedingly stable LNA: LNA duplexes, *Journal of the American Chemical Society, 120* , 13252-13253.

22. Singh, S. K., Koshkin, A. A., Wengel, J., and Nielsen, P. (1998) LNA (locked nucleic acids): synthesis and high-affinity nucleic acid recognition, *Chemical Communications* , 455-456.

23. Johnson, M. P., Haupt, L. M., and Griffiths, L. R. (2004) Locked nucleic acid (LNA) single nucleotide polymorphism (SNP) genotype analysis and validation using real-time PCR, *Nucleic Acids Research, 32* , e55.

24. Suzawa, K., Yamamoto, H., Ohashi, K., Hashida, S., Tomida, S., Kubo, T., Maki, Y., Soh, J., Tsukuda, K., Kiura, K., Miyoshi, S., and Toyooka, S. (2017) Optimal method for quantitative detection of plasma EGFR T790M mutation using droplet digital PCR system, *Oncol Rep, 37* , 3100-3106.

25. Ballantyne, K. N., van Oorschot, R. A., and Mitchell, R. J. (2008) Locked nucleic acids in PCR primers increase sensitivity and performance, *Genomics, 91* , 301-305.

26. Levin, J. D., Fiala, D., Samala, M. F., Kahn, J. D., and Peterson, R. J. (2006) Position-dependent effects of locked nucleic acid (LNA) on DNA sequencing and PCR primers, *Nucleic Acids Res, 34* , e142.

27. Ishige, T., Itoga, S., and Matsushita, K. (2018) Locked Nucleic Acid Technology for Highly Sensitive Detection of Somatic Mutations in Cancer, *Adv Clin Chem, 83* , 53-72.

28. Nafa, K., Hameed, M., and Arcila, M. E. (2015) Locked Nucleic Acid Probes (LNA) for Enhanced Detection of Low-Level, Clinically Significant Mutations, *Methods Mol Biol, 1392* , 71-82.

29. Zhang, S., Chen, Z., Huang, C., Ding, C., Li, C., Chen, J., Zhao, J., and Miao, L. (2019) Ultrasensitive and quantitative detection ofEGFRmutations in plasma samples from patients with non-small-cell lung cancer using a dual PNA clamping-mediated LNA-PNA PCR clamp, *The Analyst, 144* , 1718-1724.

30. McTigue, P. M., Peterson, R. J., and Kahn, J. D. (2004) Sequence-dependent thermodynamic parameters for locked nucleic acid (LNA)-DNA duplex formation, *Biochemistry, 43* , 5388-5405.

31. Kurreck, J., Wyszko, E., Gillen, C., and Erdmann, V. A. (2002) Design of antisense oligonucleotides stabilized by locked nucleic acids,*Nucleic Acids Res, 30* , 1911-1918.

32. Strling, Z. M., and Koch, T. (2008) *Locked Nucleic Acids* , John Wiley & Sons, Inc.

33. Diaz, L. A., Jr., and Bardelli, A. (2014) Liquid biopsies: genotyping circulating tumor DNA, *Journal of clinical oncology : official journal of the American Society of Clinical Oncology,32* , 579-586.

Figure legend

Figure 1. The LNA-mediated real-time PCR clamping system (LMPC). (A) The LNA oligomer was designed to perfectly match the wild-type sequence. The forward primer partially overlapped the LNA binding position, and the 3' end of the primer was designed to bind the point mutation site. During PCR, an LNA/DNA hybrid with a perfect match inhibited annealing of the primer, thus suppressing the amplification of wild-type DNA. However, an LNA/DNA hybrid with a single base-pair mismatch did not restrain the

amplification of mutant alleles. (B) Various LNA modification modes based on length, base number, and position were evaluated. LNA modification has a great impact on clamping performance. The optimal amplification state of LMPC is that the amplification of wild-type gene can be inhibited completely with no influence on mutant targets. But in most clamping system, the amplification of mutant targets would be suppressed as deficiency state.

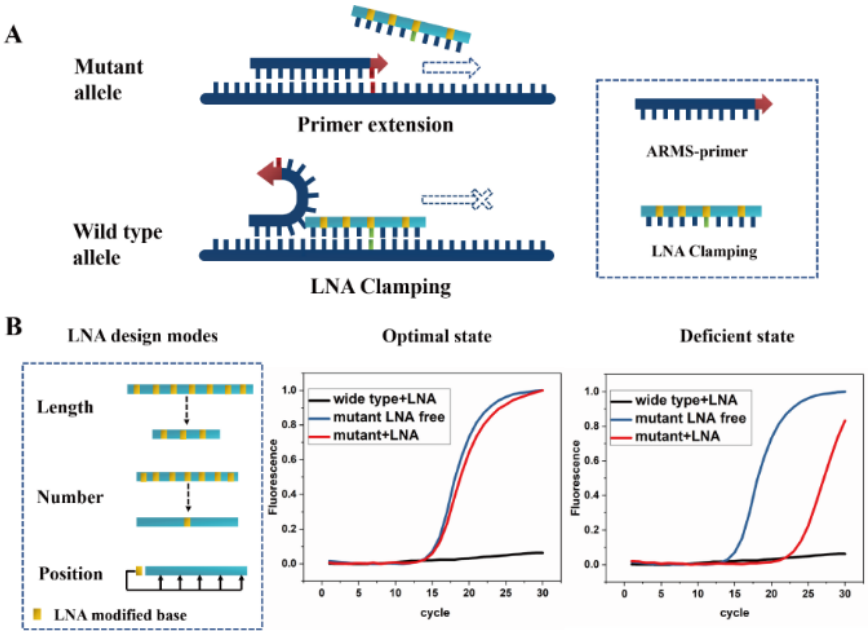


Figure 2. (A) The melting temperatures of perfectly matched (WT) and mismatched (MUT) LNA/DNA duplexes and the differences in Tm (ΔTm^a) between the wild type and mutant sequences. (B) The melting curves of different LNA/DNA duplexes. "M" represents the mismatch, and "W" represents the perfect match.

Posted on Authorea 17 May 2021 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.162128311.16782690/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Posted on Authorea 17 May 2021 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.162128311.16782690/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

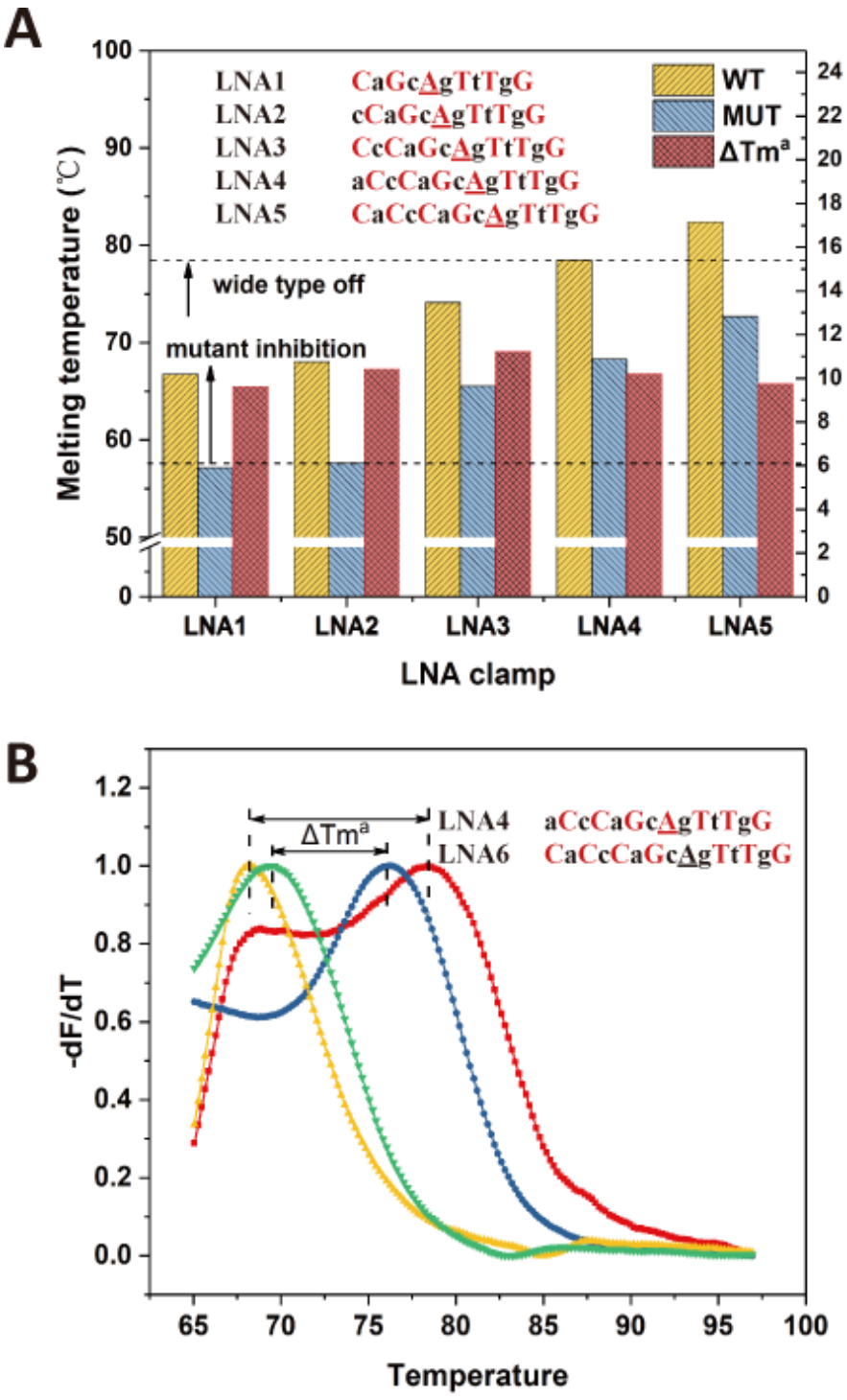


Figure 3. The amplification efficiency of different LNA-mediated real-time PCR systems and the blank control (without LNA) were evaluated. Serially diluted standard templates at different mutation proportions of 1%, 0.1% and 0% (in 10000 copies of wild-type gene) were tested. Real-time PCR plots of different LNA systems at 0% (A), 1% (B), and 0.1% (C) mutant DNA. (D) Line chart of the mean Ct (threshold cycles) value calculated from three individual replicates. (E) The melting curves of different LNA clamps with

mutant sequences and (F) wild-type sequence.

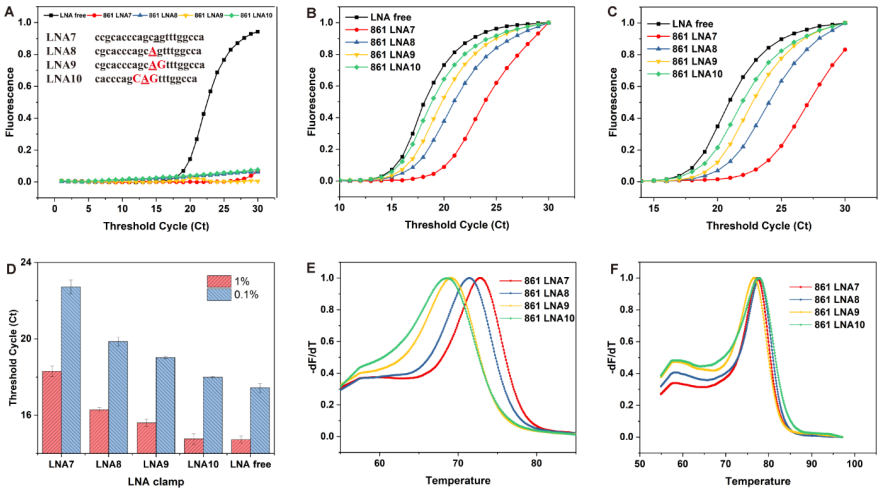


Figure 4. Evaluation of four optimized LNA oligonucleotides. Real-time PCR plots of different LNA systems at 0% (A), 1% (B) and 0.1% (C) mutant templates. (D) Line chart of the mean Ct (threshold cycles) value calculated from three individual replicates. (E) The melting curves of different LNA oligomers with mutant sequences and (F) wild-type sequence.

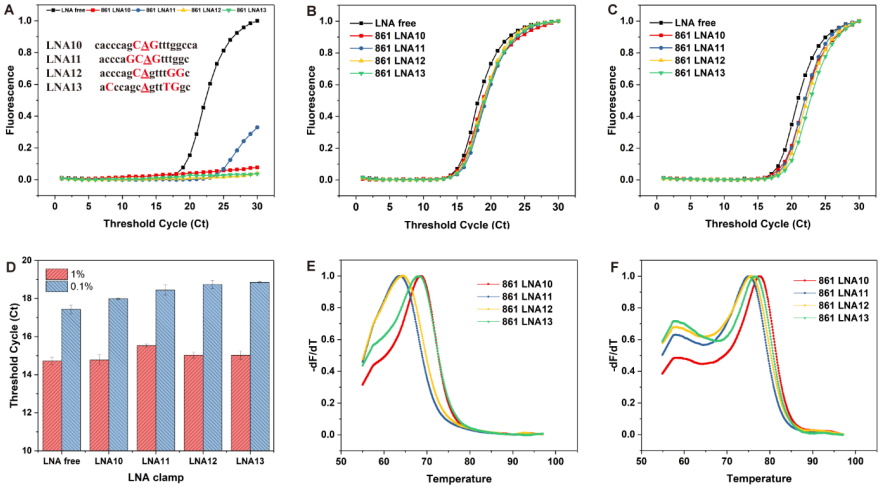


Figure 5. To avoid interference from wild-type gene signals, no wild-type gene were added to the mutant targets to verify the universality of LNA-mediated real-time PCR for the detection of EGFR point mutations. The amplification comparison between the LNA system and the normal PCR system (without LNA clamp) was assessed in the samples that contained 10000 copies of wild-type DNA (from A549 cells) and 100 and 10 copies of mutant DNA, respectively. Plasmids harboring (A) the G719A, (B) T790M, (C) S768I, and (D) L858R mutant sequence was directly amplified in each system. (E) Agarose gel electrophoresis of amplicons in both systems, 1(LNA free-10000 wt), 2(LNA-10000 wt), 3(LNA free-100 mut), 4(LNA-100 mut), 5(LNA free-10 mut), 6(LNA-10 mut), 7 (negative control).

Posted on Authorea 17 May 2021 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.162128311.16782690/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Posted on Authorea 17 May 2021 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.162128311.16782690/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

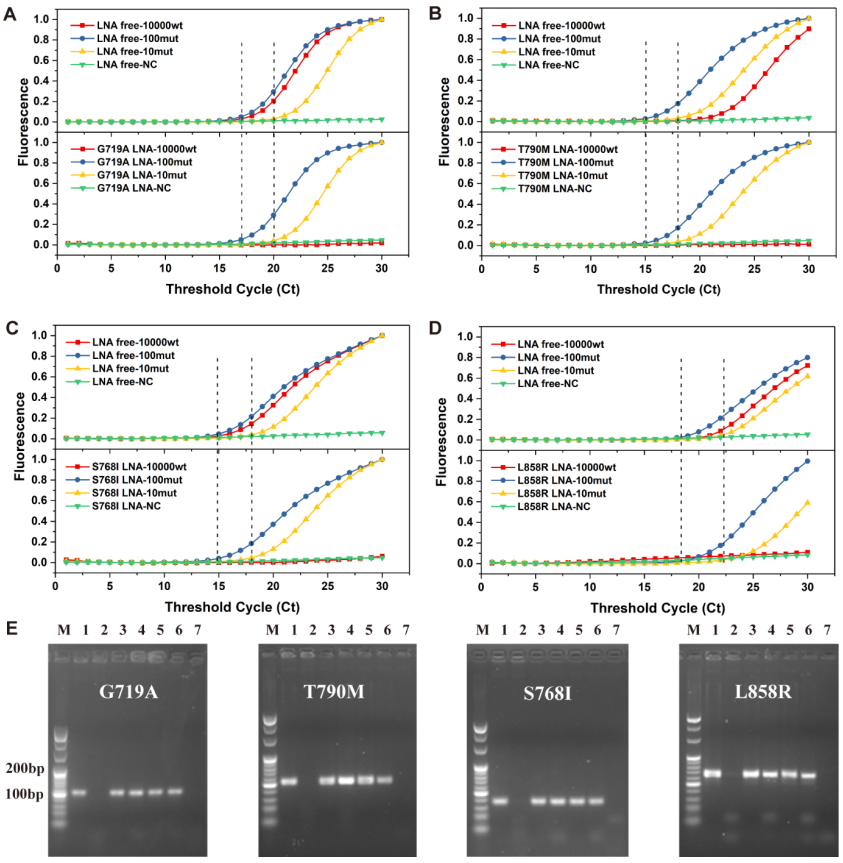
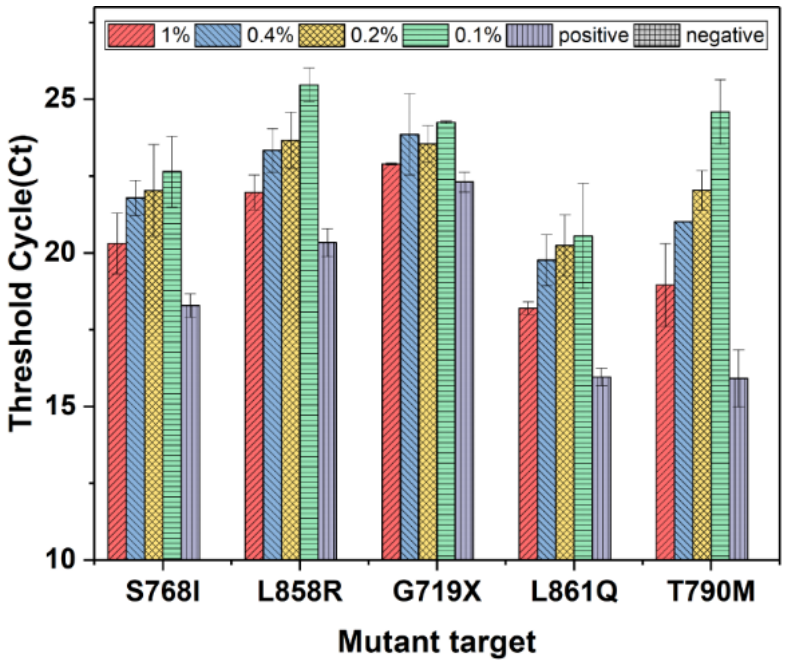


Figure 6. The Ct values of different ratios of EGFR cell-free DNA reference standard samples at different single point mutant sites (A). Comparison of clinical sample results between commercial kits and LNA clamping systems (B).

Posted on Authorea 17 May 2021 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.162128311.16782690/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

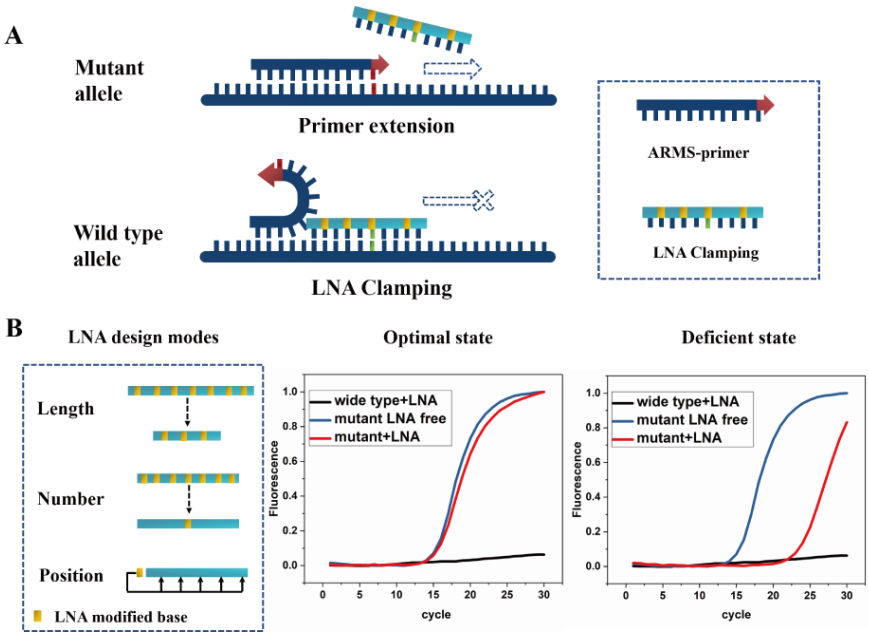
A



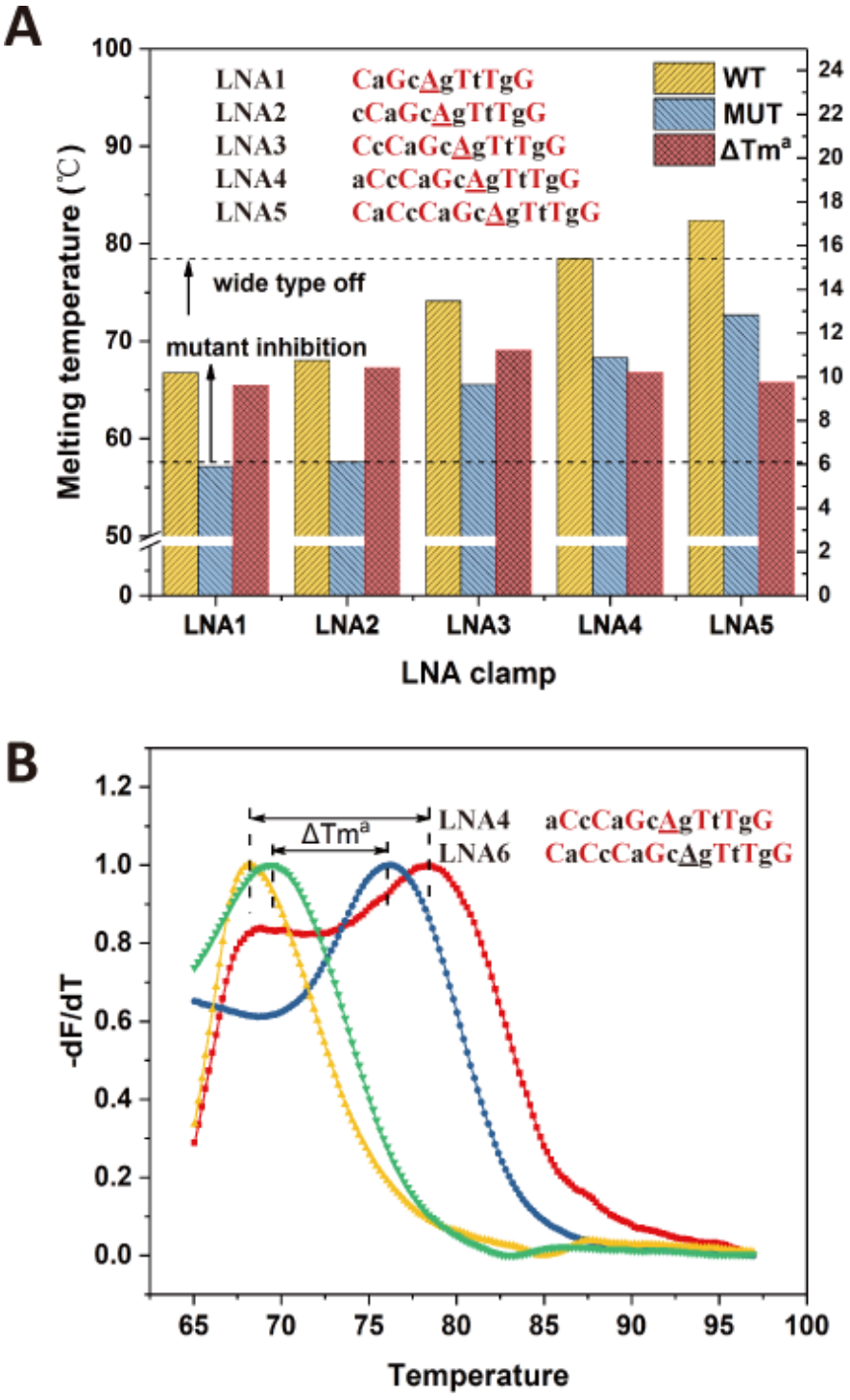
B

Sample Tag Number	Commercial kit	LNA clamping system (Ct)	Correlation
1	G719X/S768I	G719X（13.64/13.61） S768I（13.44/13.54）	match
2	G719X	G719X（17.11/16.82）	match
3	G719X	G719X（18.36/18.87）	match
4	S768I/L858R	S768I（18.15/18.29） L858R（18.33/18.60）	match
5	S768I	S768I（20.54/22.92）	match
6	L861Q	L861Q（17.05/16.99）	match
7	L861Q	L861Q（17.94/17.99）	match
8	L861Q	L861Q（13.93/13.93）	match
9	L861Q	L861Q（14.08/14.30）	match
10	L858R	L858R（17.32/17.60）	match
11	L858R	L858R（17.96/17.73）	match
12	L858R/T790M	L858R（19.97/19.70） T790M（18.23/19.38）	match
13	G719X	G719X（18.42/17.89）	match
14	G719X	G719X（17.90/17.00）	match
15	L861Q	L861Q（15.31/15.15）	match
16	negative	negative	match
17	negative	negative	match
18	L858R	L858R（14.85/15.21）	match
19	L858R	L858R（17.54/17.92）	match
20	L858R	L858R（22.30/21.42）	match
21	G719X	G719X（16.48/16.92）	match
22	negative	negative	match
23	negative	negative	match
24	negative	negative	match

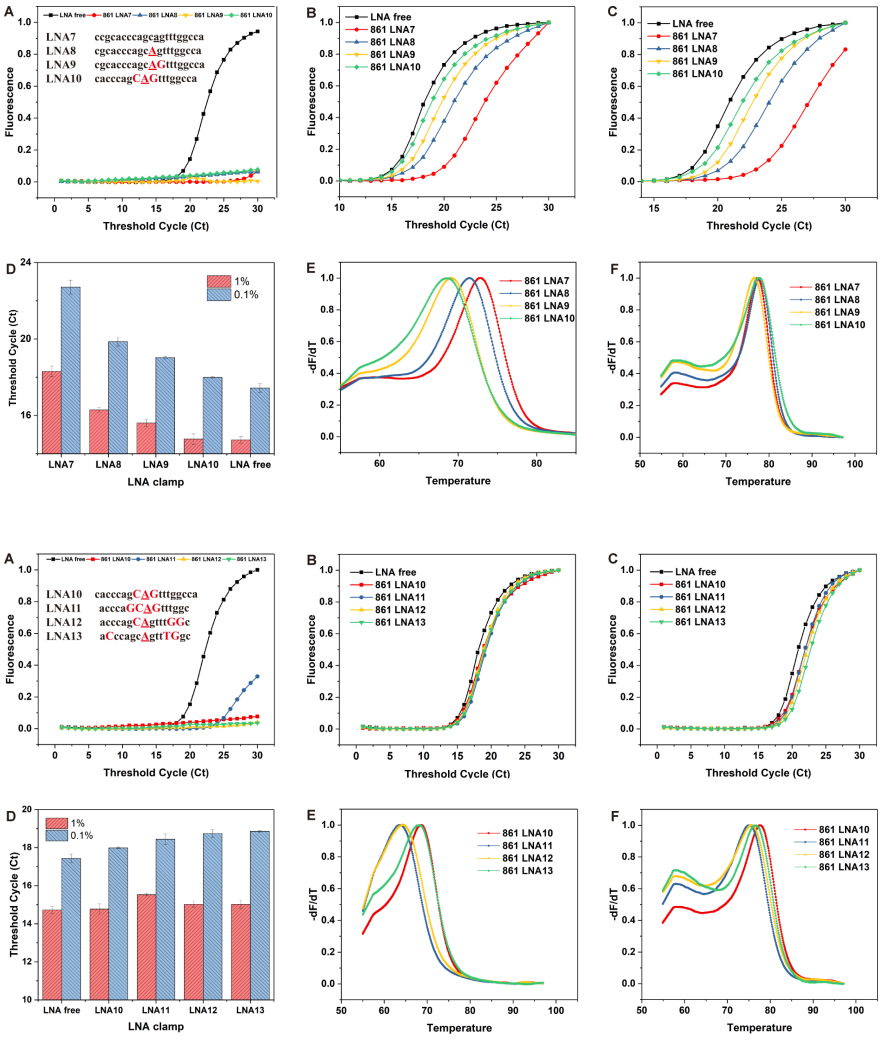
Posted on Authorea 17 May 2021 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.162128311.16782690/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.



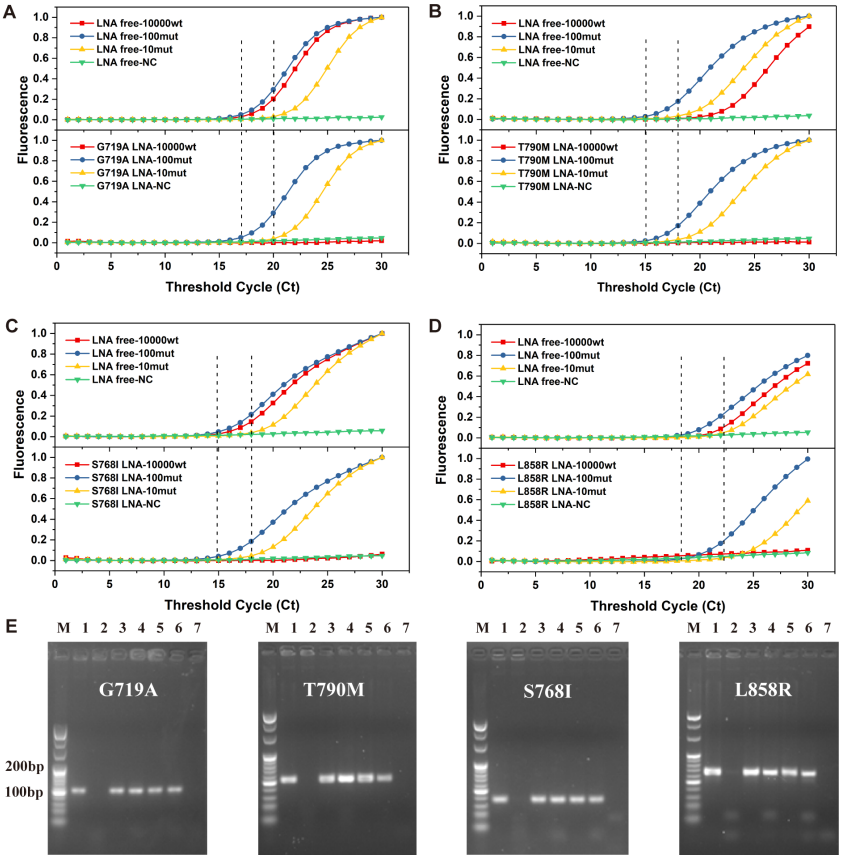
Posted on Authorea 17 May 2021 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.162128311.16782690/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.



Posted on Authorea 17 May 2021 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.162128311.16782690/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

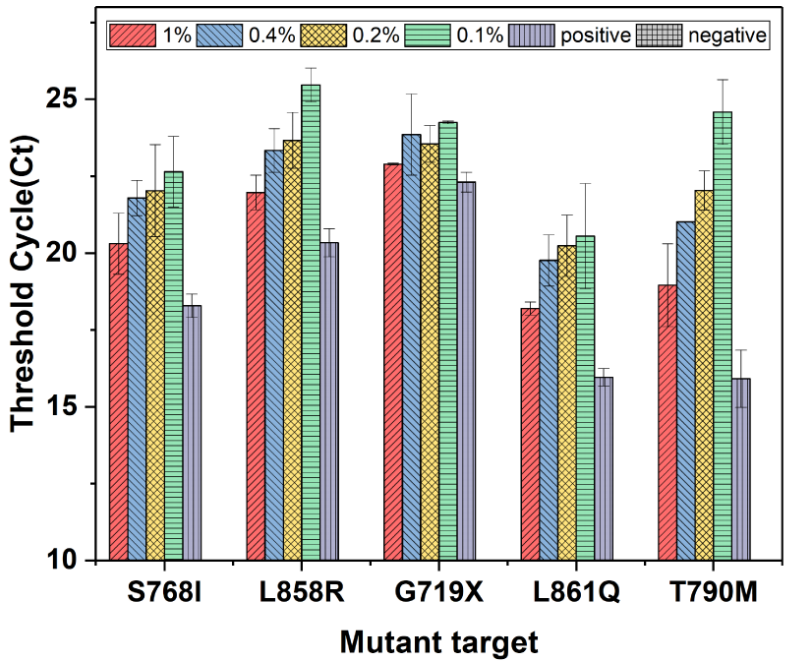


Posted on Authorea 17 May 2021 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.162128311.16782690/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.



Posted on Authorea 17 May 2021 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.162128311.16782690/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

A



B

Sample Tag Number	Commercial kit	LNA clamping system (Ct)	Correlation
1	G719X/S768I	G719X（13.64/13.61） S768I（13.44/13.54）	match
2	G719X	G719X（17.11/16.82）	match
3	G719X	G719X（18.36/18.87）	match
4	S768I/L858R	S768I（18.15/18.29） L858R（18.33/18.60）	match
5	S768I	S768I（20.54/22.92）	match
6	L861Q	L861Q（17.05/16.99）	match
7	L861Q	L861Q（17.94/17.99）	match
8	L861Q	L861Q（13.93/13.93）	match
9	L861Q	L861Q（14.08/14.30）	match
10	L858R	L858R（17.32/17.60）	match
11	L858R	L858R（17.96/17.73）	match
12	L858R/T790M	L858R（19.97/19.70） T790M（18.23/19.38）	match
13	G719X	G719X（18.42/17.89）	match
14	G719X	G719X（17.90/17.00）	match
15	L861Q	L861Q（15.31/15.15）	match
16	negative	negative	match
17	negative	negative	match
18	L858R	L858R（14.85/15.21）	match
19	L858R	L858R（17.54/17.92）	match
20	L858R	L858R（22.30/21.42）	match
21	G719X	G719X（16.48/16.92）	match
22	negative	negative	match
23	negative	negative	match
24	negative	negative	match