



OPEN

Inclusion of double helix structural oligonucleotide (STexS) results in an enhance of SNP specificity in PCR

Jae Jong Kim^{1,4}, Hyoung-Min Park^{2,4}, A. Young Kyoung¹, In Kyung Park¹, Si-Kyu Lim¹✉ & Byoung Chul Park³✉

Genetic mutations such as single nucleotide polymorphisms (SNP) are known as one of the most common forms which related to various genetic disorders and cancers. Among of the methods developed for efficient detection of such SNP, polymerase chain reaction (PCR) methods are widely used worldwide for its cost and viable advantages. However, the technique to discriminate small amounts of SNP mixed in abundant normal DNA is incomplete due to intrinsic technical problems of PCR such as amplification occurring even in 3'mismatched cases because of high enzyme activity of DNA polymerases. To overcome the issue, specifically designed PCR platform, STexS (SNP typing with excellent specificity) using double stranded oligonucleotides was implemented as a means to emphasize the amplification of SNP templates by decreasing unwanted amplification of 3'mismatched DNA copies. In this study, the results indicate several EGFR mutations were easily detected specifically utilizing the STexS platform. Further trials show the novel method works effectively to discriminate mutations in not only general allele specific (AS)-PCRs, but also amplification refractory mutation system (ARMS)-PCR. The STexS platform will give aid in PCRs targeting potential SNPs or genetically mutated biomarkers in human clinical samples.

The investigation of genetic variations which involves in genetic disorders and cancers is essential for predicting cures, establishing treatment methods, and observing prognosis and relapse. Identifying these variations are not only critical for human welfare, but also for the selection and breeding of species and origin verification in agriculture¹.

Genetic variation can be caused by indigenous or environmental factors which living organisms are exposed throughout its lifespan. The most common form occurring in humans is the single nucleotide polymorphism (SNP) variant, which alters a single nucleotide within an individual. While the process of substitution of the nucleotides are constantly occurred among men, specific SNPs located in a critical genomic region tend to inflict differences in the susceptibility of various diseases such as sickle-cell anemia, cystic fibrosis, and Alzheimer's disease²⁻⁴. Studies regarding the discrimination of crucial SNPs are constant. The most common, efficient and viable method is utilizing the polymerase chain reaction (PCR). Nowadays the ability to estimate the nucleotide amount by PCR and quantitative real time PCR (qPCR) is widely used in public health, agriculture and environment preservation⁵. Although the advantage is certain, PCR also has shown limitations. The continuous effort to improve the overall mixture was mainly aimed to enhance reactions, diminish unnecessary primer dimers, reduce nonspecific binds due to mismatches, and block efficiency drops occurred by high G/C rates^{6,7}. While most of the issues were solved, primer template mismatches constantly tamper the overall specificity, resulting in unwanted false-positives. For example, "HotStart PCR" was implemented to block unwanted nonspecific amplifications caused by primer binding on unwanted target DNA (not alleles) at low temperature⁸, but this method still is not adequate for identifying a handful of mutant DNA mixed in abundant normal wild-type DNA (alleles). As a result, methods for diagnosing cancers and other disorders using tissue, blood, and saliva

¹GenoTech Corporation, 26-69, Gajeongbuk-ro, Yuseong-gu, Daejeon 34113, Republic of Korea. ²Certified Reference Material Lab, Korea Research Institute of Standards and Science, 267 Gajeong-ro, Yuseong-gu, Daejeon 43113, Republic of Korea. ³Disease Target Structure Research Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon 34141, Republic of Korea. ⁴These authors contributed equally: Jae Jong Kim and Hyoung-Min Park. ✉email: slim@genotech.co.kr; parkbc@kribb.re.kr

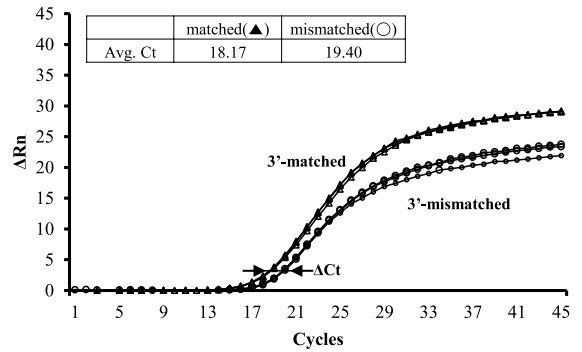
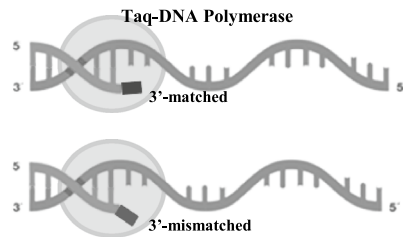
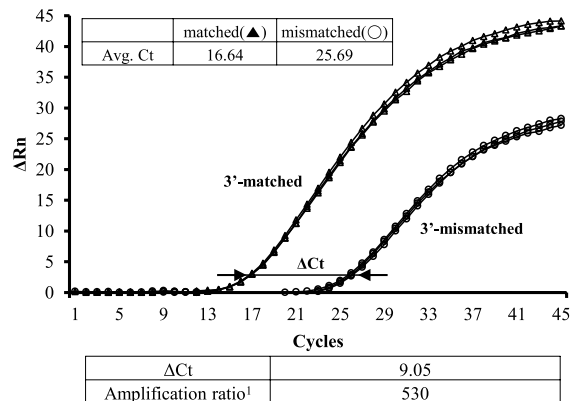
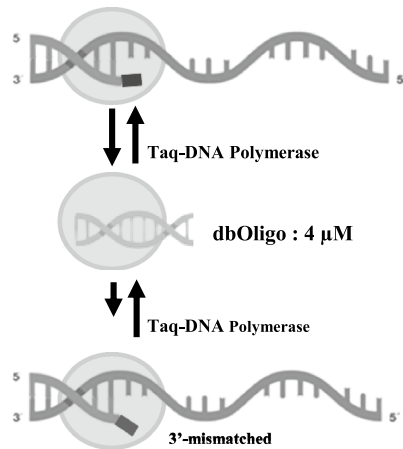
a. General allele specific PCR**b. STexS (SNP Typing-excellent specificity)**

Figure 1. Standard PCR model of DNA polymerase reaction utilizing dbOligo. **(a)** Schematics of general allele specific PCR (AS-PCR). **(b)** Schematics of the SNP Typing -excellent specificity (STexS) PCR. qPCRs were performed with 790-F58-1 for EGFR T790M, and 4 μM of AHP as dbOligo for STexS PCR was used.

are yet to be complete due to the fact that clinical samples are mainly consisted of normal DNA alongside a mere fraction of mutated DNA.

The context of discriminating a mutant within a clinical sample has to have high specificities which can detect 1% or less included mutated DNA alongside a high robustness for further clinical implementations. For these reasons, this study represents a refined method of PCR, STexS (SNP typing with excellent specificity) by adding a double helix forming oligonucleotide DNA (discrimination boosting oligonucleotides; “dbOligo”) to increase the discrimination between matched and mismatched PCR (Sup Fig. 1).

Results

Utilizing double stranded oligonucleotides in PCRs for discrimination boosting. The typical allele specific PCR (AS-PCR) proceeds even when a mismatch occurs in the 3' end of a primer. This is due to DNA polymerase reactions carrying on the next dNTP even when a mismatch happens, which after the next cycle converts to a normal match that continues amplifying. To overcome this issue, the annealing has to be repressed when the initial template detects a mismatch. The priority was to control the amount of DNA polymerase for adjusting reaction rates and additional melting temperature (T_m) calibrations above the lowest annealing temperature for discarding unwanted dimers.

To increase the amplification ratio [$2^{\Delta Ct}$; $\Delta Ct = Ct$ of 3' mismatched (normal DNA)— Ct of 3' matched (mutant DNA)] of qPCR, specifically designed dbOligo, a double helix forming oligonucleotide DNA, was added (Fig. 1). Optimization for the duplex forming ratio required each dbOligo to be designed at least 10 base pairs and adjusted to elevate the melting temperature higher than an average annealing temperature. Further optimization gave evidence in which among various factors such as T_m , length and GC contents, T_m represented to be much more crucial as GC contents and length factors become irrelevant once T_m is above the PCR extension temperature (Sup Figs. 2, 3, 4). In a 3' matched situation, a general PCR and dbOligo added PCR (STexS-PCR) did not show any big difference. Both Ct's were similar (3' matched in Fig. 1). This implies adding dbOligo within a PCR mixture does not influence the activity of the DNA polymerase. In a 3' mismatch situation, the K_{cat2d} of dbOligo-added PCR should still be same or slightly lower than the K_{cat2} of a general although highly lower than K_{cat1} and K_{cat1d} in a 3' match situation PCR ($K_{cat2} \approx K_{cat2d} \ll K_{cat1} \approx K_{cat1d}$). However, dbOligo-added PCR plot was highly decreased in a 3' mismatch situation (3' mismatched in Fig. 1 and Sup Fig. 2). The cause is considered to be the decrease of the association reaction (k_2) or the increase of the dissociation reaction (k_2) within the DNAP·P/ T_2 cluster (Sup Fig. 5c, d). In other words, the overall attach and detach rate of DNAP should have

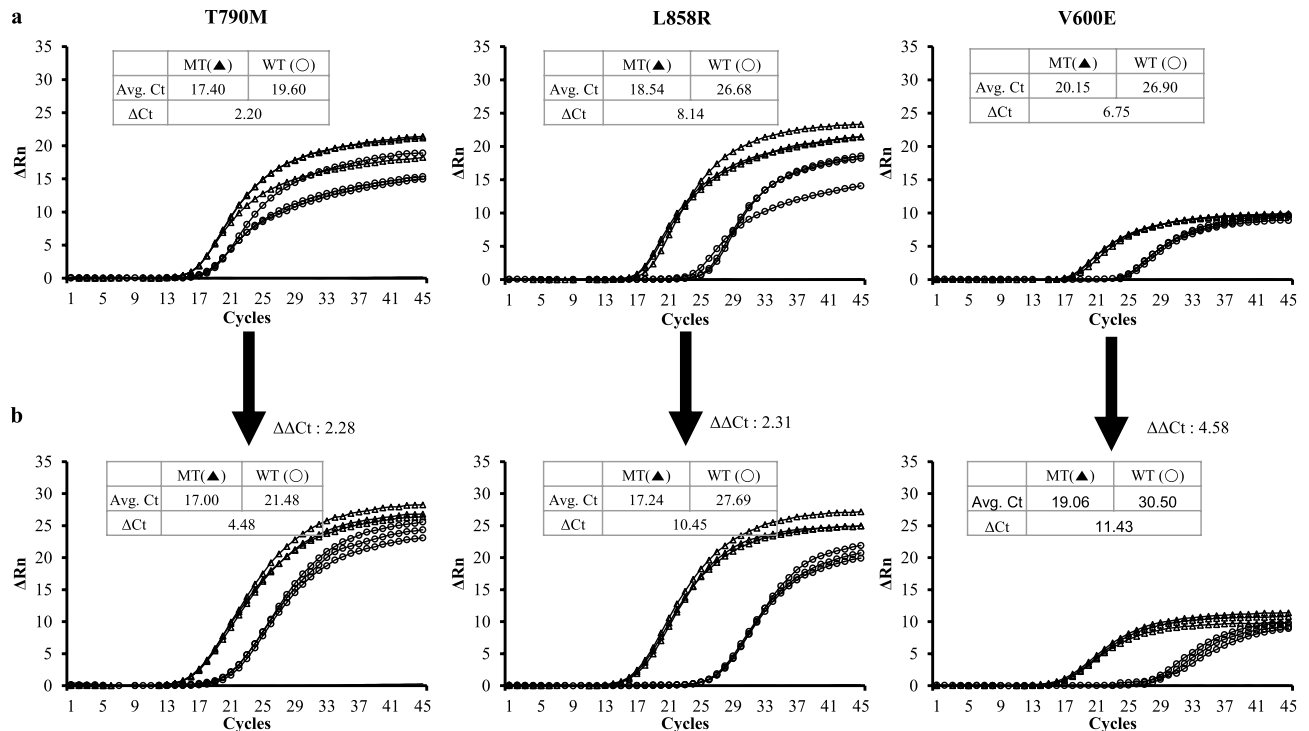


Figure 2. Efficiency test of standard AS-PCR and STExS utilized PCR. qPCR plots of general (a) and STExS PCR (b). Each Primer was designed to target three types of variation of the gene EGFR; T790M, L858R and BRAF V600E. Each general PCR/STExS PCR was performed to evaluate whether the primer can effectively differentiate matched against mismatched templates. qPCRs were performed with 790-F58-1 for EGFR T790M, 790-F58-1 for EGFR T790M 858-F25 for EGFR L858R and 1799-F11 for BRAF V600E, and 1 μ M of AHP3 as dbOligo for STExS PCR was used.

been altered by addition of dbOligo ($K_{-2}/K_2 < < K_{-2d}/K_{2d}$). Due to these reasons, dbOligo-added PCR easily repress 3' mismatched amplification without altering 3' matched PCR.

qPCR validation of the STExS dbOligo. To further address the significance of dbOligo added PCR, the term will be named as SNP Typing-excellent specificity (STExS) PCR. For further validation of STExS, several trials of PCR were performed targeting several variations of the EGFR gene. The forward primers were designed to match or mismatch 1 base pair of the EGFR c.2369 C > T (p.T790M); EGFR c.2573T > G (p.L858R) and BRAF c.1799 rc. A > T (p.V600E) (Sup Table 1). In a general AS-PCR, the discrepancy of wild-type and mutant qPCR versions were not sufficient to address apparent difference. When STExS was implemented, the contrast between matched and mismatched were escalated with increased $\Delta\Delta$ Ct values ($\Delta\Delta$ Ct = Δ Ct of not STExS PCR – Δ Ct of STExS PCR) between 2.20 to a maximum of 4.48 (Fig. 2). Repeated trials of each method consistently represented STExS PCR repressing non-significant templates (Sup Fig. 6). The increased discrimination of EGFR T790M and wild-type was enough to doubtlessly consider the matched template of the mutated variate was differentially sorted with STExS PCR.

Amplification refractory mutation system-polymerase chain (ARMS-PCR) is known as a simple and economical method for genotyping SNPs⁹. To demonstrate whether the STExS also contributes in an ARMS-PCR, several primers were designed to target an identical gene region EGFR T790M with (an) additional 1–2 mismatch bases introduced within each primer. As a result, all normal ARMS-PCRs increased the discrimination between wild-type and T790M. When dbOligo was added in each ARMS-PCR, the discrimination was further increased showing $\Delta\Delta$ Ct values of 2.39, 3.97 and 3.24 (Fig. 3). Interestingly, the results indicate the effectiveness of STExS for discrimination when adopted in not only AS PCR but also ARMS PCR to detect a single mutation.

Validation of STExS PCR with various DNA polymerases resulted as an effective way to differentiate SNPs. Recent studies suggest the use of modified DNAPs to enhance the differentiating efficiency of 3' mismatches¹⁰. To test whether adding modified DNAPs further increase the overall amplification gap between matched and mismatched templates, wild type Taq or mutant Taq (R536K) was added in each standard and STExS PCRs. The mutated Taq polymerase itself increased the discrimination compared to the WT Taq polymerase (Fig. 4a and b). When dbOligo was added to PCR, each $\Delta\Delta$ Ct values of wild type Taq and mutant Taq (R536K) were 6.76 and 8.38 (Fig. 4c and d). The result implies that STExS PCRs perform in a synergistic manner when modified DNAP is implemented. Although the performance of the combined STExS and R536K mutated Taq DNAP expressed an impressive outcome, not all SNP targeting genes will be implemented to sufficient discrimination.

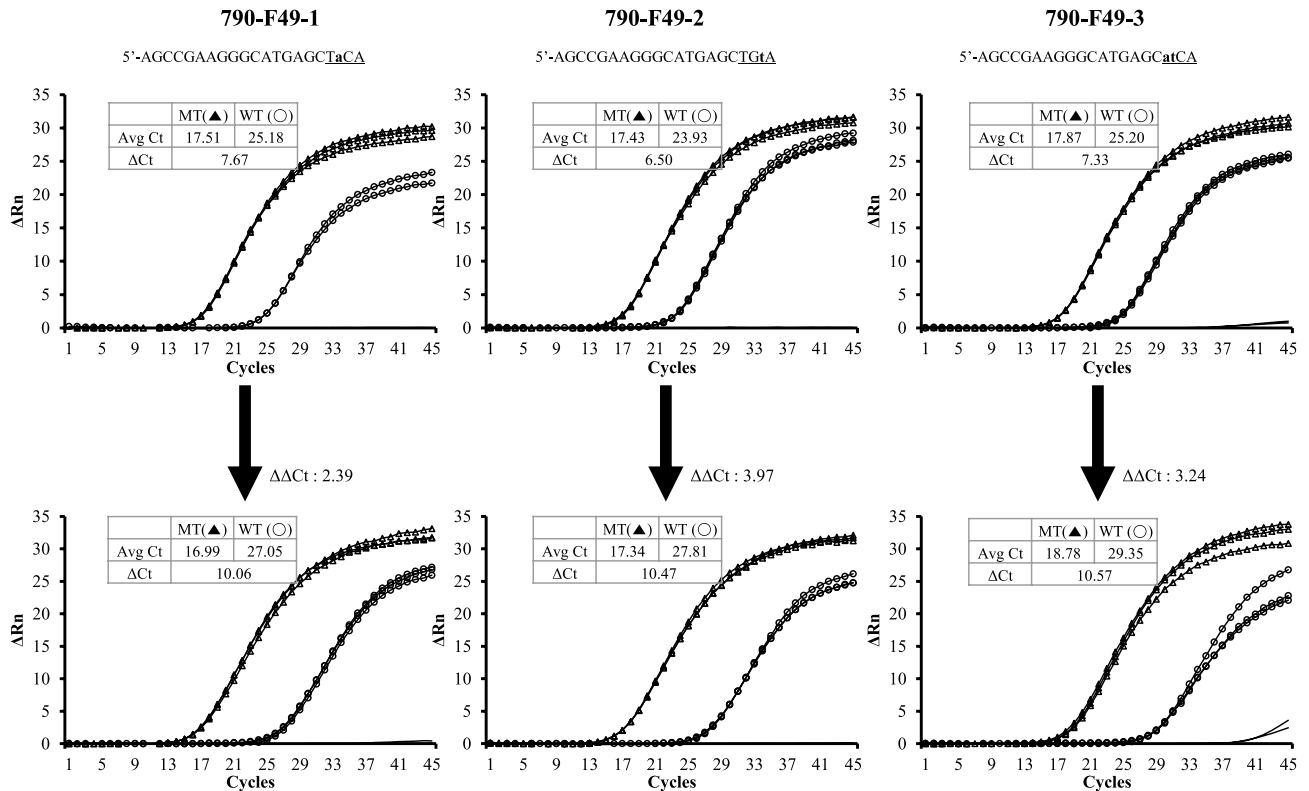


Figure 3. Efficiency of ARMS-PCR using standard and STExS utilized methods. Each ARMS primer is designed to target the EGFR gene variation T790M. Additional one or two mismatch base(s) were inserted to the primers and were mixed with the WT primers for quantification. Both general ARMS-PCR and STExS PCRs were identically performed. qPCRs were performed with 790-F49-1, 790-F49-2 and 790-F49-3 for EGFR T790M, and 1 μ M of AHP3 as dbOligo for STExS PCR was used.

STExS PCR was further tested using 5X PCR mixtures that adjusted primer set and dbOligo amount (4 μ M) for the high discrimination to evaluate the significance in SNP typing using the considerable amount (1×10^5 copies) of samples. As a result, in a conventional 5X Taq PCR, the difference between matched and mismatched templates was slim. Surprisingly, the STExS PCR expressed drastic contrasts between matched and mismatched, repressing the mismatched template well below the detection value (Fig. 5). Alongside conventionally used PCR mixtures, plasmid copied DNA were tested with commercially used human gDNA to detect possible differences between cell-line made DNA and real samples. Repetitive trials targeting 60 types of SNP resulted in no difference between commercially used human gDNA and plasmid copied DNA (Sup Fig. 7). Overall results show the STExS PCR platform to be a successful method for amplifying mutant DNAs without unwanted duplicated mismatches.

Discussion

It has been reported that the kinetics of forming the complex (DNAP·P/T) consisting of DNA polymerase (DNAP), primer(P) and template(T) is highly different according to 3' matched or mismatched¹¹. Within the DNA annealing process, 3' matched complexes (DNAP·P/T₁) are reported to be 100–1000 times higher of K_{cat}/K_m than 3' mismatched complexes (DNAP·P/T₂). This difference is made by a drastic decrease of K_{cat} (around 10–600 times) and a slight increase of K_m (maximum 3 times) caused by 3' mismatched. According to the kinetics of PCR in the report mentioned above, polymerization is continued without DNA polymerase detach due to high K_{cat1} of 3' matched complex (DNAP·P/T₁), but DNA polymerase will repeatedly be detached and attached in the event of discontinuous polymerization due to the low K_{cat2} of 3' mismatched complex (DNAP·P/T₂) (Sup Fig. 1c and d). The specifically designed dbOligo focuses on the detachment of DNAP with mismatched templates to further emphasize the amplification ratio of targeted SNPs.

Targeting mutated DNAs within a gene can be beneficial in various circumstances. The importance of SNPs as prognostic or predictive cancer biomarkers are implemented throughout various aspects such as SNP targeting cancer drugs, precision medicine and biomarker derived anticancer therapy¹². As the development of genomic sequencing continues to improve, more and more mutations within the human gene will be discovered to better predict, diagnose and treat various disorders. Alongside with the increasingly found mutated targets, superior methods and better solvents will be required to detect the elusive SNPs within the abundant normal DNA. To further provide a better platform to detect and easily amplify, we introduce a novel platform regarding the use of oligo-nucleotide DNA. The designed STExS method performed well above the normal AS-PCR to differentiate mutated genes alongside repression of normal genes. The STExS platform also excels in specific PCR methods such as ARMS-PCR and provide a superior method for genotyping SNPs of very small amounts. It is well known that

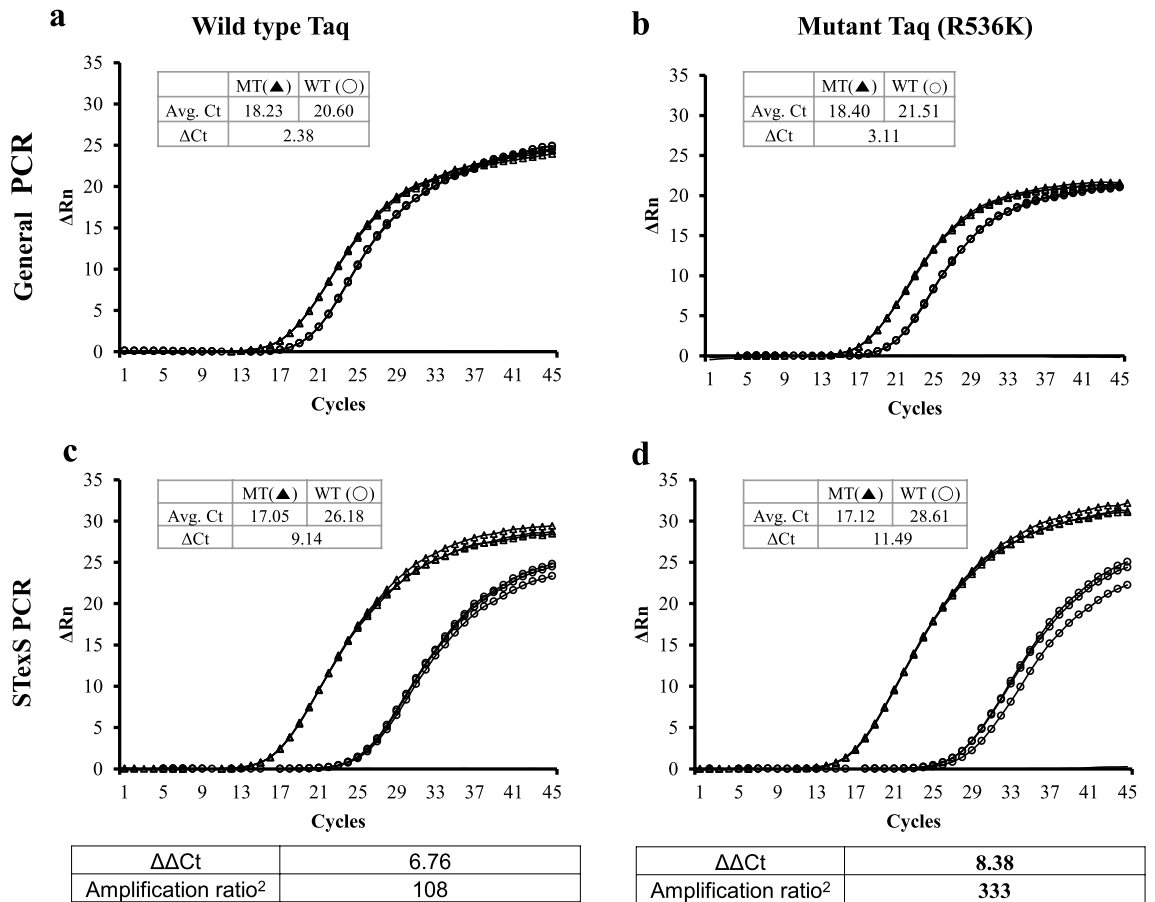


Figure 4. Efficiency of STeXS PCR with additional modified DNA polymerase. Each qPCR trial was performed using wild type Taq and mutant Taq (R536K). Each amplification ratio is quantified by $\Delta\Delta Ct$ of wild type (a, c) and mutant (b, d) Taq. qPCRs were performed with 790-F58-1 for EGFR T790M, and 4 μM of AHP3 as dbOligo for STeXS PCR was used. The amplification ratio² was calculated by using the $\Delta\Delta Ct$ values of each wild type Taq and mutant Taq to compare between general (a, b) and STeXS (c, d) PCR.

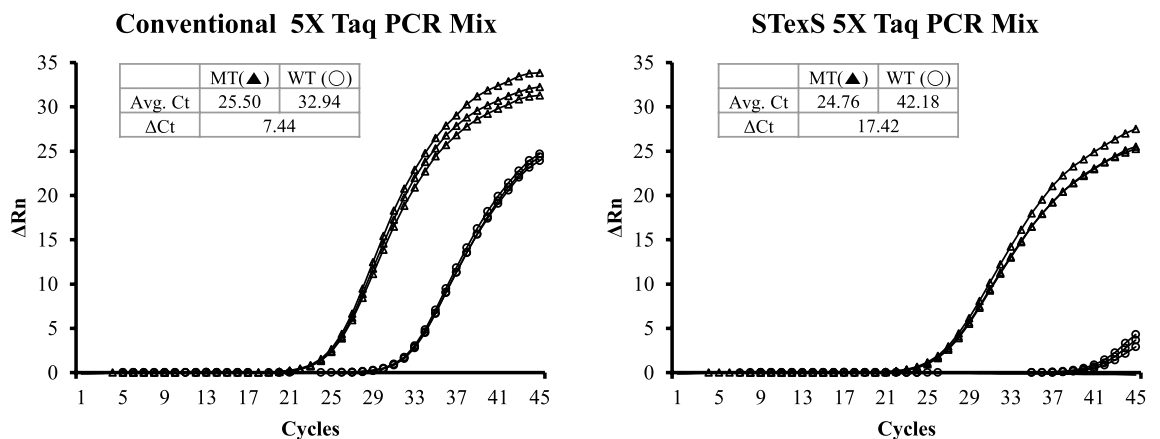


Figure 5. Amplification rates of conventional Taq PCR and STeXS implemented Taq PCR. Each qPCR was performed using the conventional 5X Taq PCR mixture. The amplification ratio was calculated using the $\Delta\Delta Ct$ value obtained by subtracting each Ct of the WT template and mutated template. qPCRs were performed with 790-F49-2 for EGFR T790M and 4 μM of AHP3 as dbOligo for STeXS PCR was used.

conventional allele-specific PCR lacks in specificity when it comes to SNP detection. The STeXS platform shows clear evidence of suppressed amplification in non-specific DNA. This further leads to a decline in false negative rates and will reveal patients which were previously not detected due to undistinguishable amplification rates between mutant and normal DNA. For example, the EGFR gene is well known to be mutated as various forms in non-small cell lung cancer¹⁵. As it is best to predict such genetic disorders, capturing such mutations in the

early stage is essential to lung cancer treatment. The STexS PCR method easily separates the various mutations of EGFR compared to the standard PCR, offering a stable and viable method to detect lung cancer in clinical samples such as blood and lung tissue.

The contents of PCR continue to improve to induce higher activities of polymerases, obtain better amplified copies and reduce unwanted debris such as primer dimers. To better discriminate target SNP from normal DNA, research in polymerase mutation was introduced for the alternative solution. While the mutated DNAP itself was effective at amplifying SNP templates, the synergistic effect of the combined modified DNAP and STexS increased the discrimination ratio² more than 1000 times ($\Delta\Delta Ct = 9.98$) the standard PCR (Fig. 5).

As shown above, the STexS PCR will serve to be an effective method for future SNP detection and give further aid in developing potential biomarkers.

Materials and methods

dbOligo design and use. Each dbOligo was designed for efficient forming of double stranded duplex to have at least 10 bases pairing in a single strand. AHP3 is one of the designed dbOligo with the highest performance listed (Sup Table 1). Designed dbOligo were added 1–80 pmol of dOligo per 20 ul of PCR mixtures for the expected 0.05–4 uM of final.

Allele specific and ARMS PCR. PCR trials were performed using a wild-type and mutated template, forward and reverse primers and dual labeled hydrolysis probes for signal detection (Sup Table 1). Primers for the targeted EGFR gene were designed to match the mutated three variations EGFR T790M, L858R and BRAF V600E alongside mismatching one base pair of each normal genes. All templates for the targets were synthesized artificially and cloned into pUC19 and their copy numbers were measured and calculated with nano drop spectrometry.

Enzymes for each PCR were 2 units of Taq DNA polymerase (GenoTech) to suit the expected 0.05–0.08 uM. PCR buffer were consisted with 10 mM Tris (pH 9.0), 1.5 mM MgCl₂, 60 mM KCl, 10 mM (NH₄)₂SO₄ mixed to a total of 20 ul. Primers were used 20 pmols and probes were used 10 pmols in 20 ul reactions. Specific cycles were set to the standard ABI 7500 qPCR system, starting from 95°C to 55°C for 45 cycles. All trials were tested as a triplicate.

In qPCR, $\Delta Ct = Ct$ of 3'-mismatched – Ct of 3'-matched, $\Delta\Delta Ct = \Delta Ct$ of STexS – ΔCt of no STexS, amplification ratio¹ = $2^{\Delta Ct}$ and amplification ratio² = $2^{\Delta\Delta Ct}$ were calculated to explain for the discrimination ability.

Received: 23 June 2021; Accepted: 13 September 2021

Published online: 27 September 2021

References

1. Ndjioudjop, M. N. *et al.* Development of species diagnostic SNP markers for quality control genotyping in four rice (*Oryza L.*) species. *Mol. Breed.* **38**(11), 131 (2018).
2. Ingram, V. M. A specific chemical difference between the globins of normal human and sickle-cell anaemia haemoglobin. *Nature* **178**(4537), 792–794 (1956).
3. Hamosh, A. *et al.* Cystic fibrosis patients bearing both the common missense mutation Gly--Asp at codon 551 and the delta F508 mutation are clinically indistinguishable from delta F508 homozygotes, except for decreased risk of meconium ileus. *Am. J. Hum. Genet.* **51**(2), 245–250 (1992).
4. Wolf, A. B. *et al.* APOE and neuroenergetics: an emerging paradigm in Alzheimer's disease. *Neurobiol. Aging* **34**(4), 1007–1017 (2013).
5. Valones, M. A. *et al.* Principles and applications of polymerase chain reaction in medical diagnostic fields: A review. *Braz. J. Microbiol.* **40**(1), 1–11 (2009).
6. Bai, Y. *et al.* Enhancement of PCR sensitivity and yield using thiol-modified primers. *Sci. Rep.* **8**(1), 14858 (2018).
7. Zhong, Y. *et al.* Enhancing the specificity of polymerase chain reaction by graphene oxide through surface modification: Zwitterionic polymer is superior to other polymers with different charges. *Int. J. Nanomed.* **11**, 5989–6002 (2016).
8. Kellogg, D. E. *et al.* TaqStart antibody: "Hot start" PCR facilitated by a neutralizing monoclonal antibody directed against Taq DNA polymerase. *Biotechniques* **16**(6), 1134–1137 (1994).
9. Medrano, R. F. & de Oliveira, C. A. Guidelines for the tetra-primer ARMS-PCR technique development. *Mol. Biotechnol.* **56**(7), 599–608 (2014).
10. Drum, M. *et al.* Variants of a *Thermus aquaticus* DNA polymerase with increased selectivity for applications in allele- and methylation-specific amplification. *PLoS ONE* **9**(5), e96640 (2014).
11. Rejali, N. A., Moric, E. & Wittwer, C. T. The effect of single mismatches on primer extension. *Clin. Chem.* **64**(5), 801–809 (2018).
12. Kamel, H. F. M. & Al-Amodi, H. Exploitation of gene expression and cancer biomarkers in paving the path to era of personalized medicine. *Genomics Proteomics Bioinformatics* **15**(4), 220–235 (2017).
13. da Cunha Santos, G., Shepherd, F. A. & Tsao, M. S. EGFR mutations and lung cancer. *Annu. Rev. Pathol.* **6**, 49–69 (2011).

Author contributions

J.J.K. designed experiments and prepared Figs. 3, 4 and 5, H.M. Park wrote the main manuscript text, A.Y.K., I.K.P. prepared Figs. 1, 2, 3, 4 and 5, S.L. and B.C.P. conduct all this experiments and edit the main manuscript and supplementary figures. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-98610-8>.

Correspondence and requests for materials should be addressed to S.-K.L. or B.C.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021