

SNP Genotyping Enhancements: From Traditional PCR to Double Helix Structural Oligonucleotides (STexS)

Hyoung Min Park*

Biometrology group, Korea Research Institute of Standards and Science, Republic of Korea

***Corresponding author:** Hyoung Min Park, Biometrology group, Korea Research Institute of Standards and Science, Republic of Korea



ARTICLE INFO

Received:  January 17, 2022

Published:  January 26, 2022

Citation: Hyoung Min Park. SNP Genotyping Enhancements: From Traditional PCR to Double Helix Structural Oligonucleotides (STexS). Biomed J Sci & Tech Res 41(2)-2022. BJSTR. MS.ID.006585.

ABSTRACT

Among the various genetic changes, single nucleotide polymorphisms (SNPs) are known to be the most frequently occurring. Detecting such variations at an early stage alongside abundant unaltered gene cohabitants was continuously practiced using methods related to the gene's molecular biology, biochemistry, and even pharmacodynamic character. Despite countless efforts to achieve accurate diagnosis, specificity and sensitivity of each method was not sufficiently increased, leading to misguided results. This review addresses the current status and limit of SNP genotyping and presents the novel platform of double helix structural oligonucleotide implemented PCR (STexS).

Keywords: Single Nucleotide Polymorphisms; Polymerase Chain Reaction; Oligonucleotides; SNP Genotyping; Molecular Biology

Introduction

As the Human Genome Project was accomplished, series of subtle variations of polymorphisms were reported within the population [1,2]. Among the various types, single nucleotide polymorphisms (SNPs) are known as the most abundant, with more than 9 million reports in the public databases [3,4]. In a molecular aspect, SNPs are nucleotide alterations at a specific region in the genome. While the term is applied to nucleotide replacements, SNPs can be broadened with biallelic variations such as insertions, deletions, and allele frequencies less than 1% [5,6]. SNPs generally occur around one per 1000 bp within the human genome [7]. However, the distribution is not even, as SNPs tend to cluster in noncoding regions more than several hundred folds compared to actual coding regions [8]. In hindsight, SNPs located in the noncoding region may be neglected. But various reports soon regarded both locales as an important region in disease developments and other genetic malfunctions [9,10]. As more SNPs were found to be an important overlap in certain conditions, it soon served as a crucial genetic or

pharmacogenetic marker for future practices. The reason for such importance lies on which specific region, coding or noncoding, the SNPs are found. Even though noncoding regions do not contribute to protein coding, regulatory sites with SNPs can affect transcription rates, which alters the overall production of protein encoding. This process later triggers malfunctions that may develop disorders and various cancers [11-13].

As important the SNPs were known for molecular, pharmacogenetic, and cancer targets, so too did the importance to accurately depict and sort out targeted SNPs from normal, far more abundant gene strands. A common approach for this identification is a patient-control study involving large SNP genotypes in patient groups and healthy control populations to compare genotypes differences for all phenotypes under study. And the relationship between specific genotypes and phenotypes characterizes disease-related sensitive genes and is used to find encrypted proteins for disease prevention and treatment. Another major application in

which SNP is used as an indicator is pharmacogenetic approaches [14-16]. Because the vast population size of the cohorts, discovering valuable SNPs that mark a specific disorder is rather trivial. However, the hurdle has been elevated as treatments shifted from prognostic to predictive medicine. While patients who already developed severely will inevitably have high portions of SNPs which indicates the disease, normal or soon to develop potential cohorts will have very low abundance which in some cases are not detected by previous methods [17,18]. Other reports indicate patients having false negative results in a certain detection test which results in failure to prevent further disease development or early treatments [19]. As technology developed, the method to efficiently detect SNPs also enhanced. The most common and efficiently known method is the utilization of Polymerase Chain Reaction (PCR).

Hence, the overall format regarding the use of PCR has not been

replaced. The improvements were acquired by usually two ways; enhance the efficiency to estimate relative amplification rates or apply adjustment reagents to further boost the amplification rate of the targeted DNA template [20,21]. The method to elevate relative amplification rates was achieved through quantitative PCR (qPCR) [22]. Further enhancements to effectively sort out and specifically amplify targeted SNP templates were done by adding allele-specific primers, reagents which clamp unwanted DNA templates, or utilizing hydrolysis probes for melting curve analysis [23,24]. Many attempts to improve the detection rate of targeted SNPs bore fruit but was not complete. Because certain DNA strains would still attach the allele-specific primers and result in false positive misdiagnosis [25]. In order to achieve early detection of a target SNP while not binding with the normal DNA strand, novel approaches were needed. In this review, we address the methods of known SNP specific amplification methods and the recently introduced double helix structural oligonucleotide added PCR (STexS) [26].

Traditional SNP Genotyping Technologies

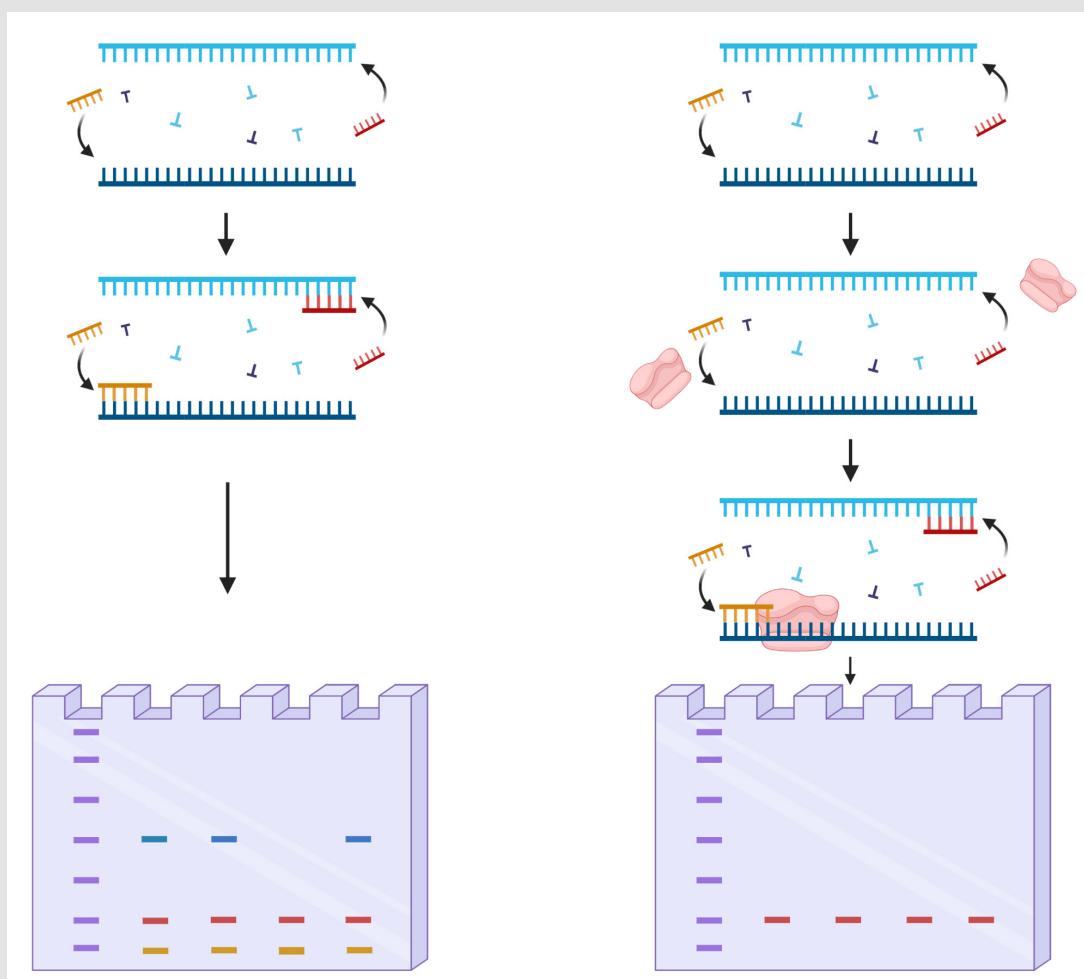


Figure 1: Basic schematics of the “HotStart” PCR.

SNP genotyping generally regards generations of allele-specific templates commercially or custom-made to increase amplification of wanted SNP-containing genes. The continuous efforts to manufacture efficient primers lead to many discoveries of genes which associates with a specific disorder such as Myeloperoxidase, N-acetyltransferase 1, and N-acetyltransferase which affects cancer susceptibility in Lung cancer, Bladder cancer, and colon cancer [27]. Other discoveries include SNPs located in CYP3A4 leading to prostate cancer and Thiopurine S-methyltransferase which SNPs alter the pharmacokinetic effects of anticancer drugs [28]. Although the roles of genes with certain SNP types were found, the method to effectively detect such allele in drastically low abundant patients were consistently held back due to limitations. The major limitation was nonspecific binds of normal DNA strands leading to primer template mismatches, resulting in false-positive detections to personnel who has the potential to develop the disorder. The "HotStart" PCR was a method to overcome such nonspecific amplifications by increasing temperature during the annealing process (Figure 1) [29]. However, SNPs consisting less than 1% in a clinical sample required something more sophisticated to discriminate normal DNA. The recently introduced STexS PCR platform combines the concept of "HotStart" PCR and allele-specific primers and further restrains nonspecific binds using a known method called amplification refractory mutation system-PCR (ARMS-PCR) [30]. ARMS-PCR is meant to detect any mutation involved single nucleotide changes including insertions and deletions [31]. This method greatly improves the efficiency to sort mismatched primers, but ARMS-PCR alone declines the amplification rate drastically below the limit of detection tampering the overall SNP genotyping.

The Implementation of Oligonucleotides

Oligonucleotides has been used alongside the standard

PCR for various purposes including SNP detection [32,33]. Oligonucleotides are used as a probe to detect a specific sequence that are complementary to the oligonucleotides. Because it is mainly consisted of a relatively short 12~25 nucleotide, it is also referred as primers in a conventional PCR for detecting and multiplying wanted complementary gene targets. Oligonucleotides are majorly single stranded, thus when performing PCR, both forward and backward templates of a targeted template is required for maximum efficiency. However, in SNP detection, nonspecific bindings of mismatched gene strands lead to misleading results. While the process to completely block the large portion of normal genes to be amplified is a challenge, giving the targeted SNP strand enough cycles to amplify firsthand by delaying mismatched DNA copies will make enough difference to successfully perform SNP genotyping.

The STexS platform mentioned above uses a double helix forming oligonucleotide (discrimination boosting oligonucleotide, dbOligo) to fully utilize the concept of disrupting mismatched amplification. The annealing process is mainly contributed by the DNA polymerase (DNAP) that synthesizes matched or mismatched strands to a complete gene template, which in turn acts as an additional template for amplifying [34]. While intended matches of dbOligo does not boost or block the amplification rate compared to other commercially used primers, the kinetic energy of DNAP to attach and detach mismatched primers dwindles, resulting in a constant unstable state that ultimately delay the overall process of amplification (Figure 2). Combining with the "Hotstart" method, each dbOligo is deliberately designed to be at least 10 base pairs to adjust the annealing temperature to further block nonspecific bindings and primer dimers. Validation through previously reported SNPs contributing to several cancer types showed a reasonable contrast between matched and mismatched templates, further proving the importance of STexS method (Table 1).

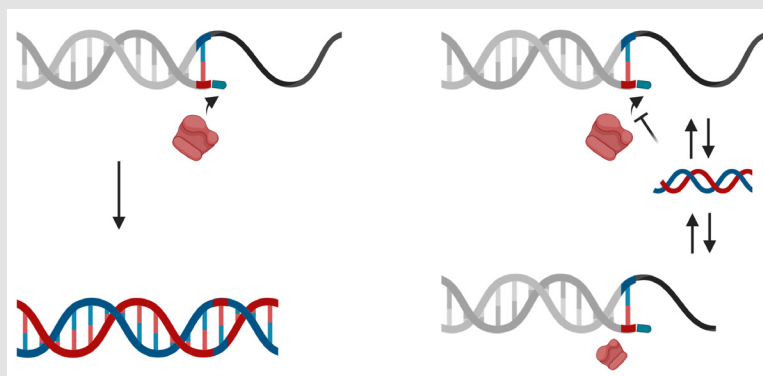


Figure 2: Difference of mismatched amplification between conventional PCR and STexS PCR.

Table 1: SNP targets of different cancer and amplification difference between matched and mismatched DNA.

Targets /purpose	Associated cancer	Sequence Type	ΔCt between mismatched
EGFR c.2369 C>T (p.T790M)	Nonsmall cell Lung cancer	790-F58-1	10.06
	Nonsmall cell Lung cancer	790-F49-2	10.47
	Nonsmall cell Lung cancer	790-F49-3	10.57
EGFR c.2573 T>G (p.L858R)	Lung Adenocarcinoma	858-F25	10.45
BRAF c.1799 rc. A>T (p.V600E)	papillary thyroid cancer	1799-F11	11.43

Conclusion and Future Directions

The ability to sort out even one mismatch between two different DNA strands is a fundamental technology not only limited to PCR techniques, but to other molecular extents. Fully discriminating matched targets from unwanted mismatched strands will also lead to improved specificity and sensitivity. Current diagnostics with various arrays and detection kits are challenged not just with its rate of success, but also constant incidents of false negatives misleading practitioners to judge poorly [35,36]. The importance of not neglecting cases of false negatives drastically increased, as contagious viral infections tend to spread among patients who cannot be scanned with conventional methods [37-39]. The recent spread of SARS-CoV-2 influenced generations of detection methods to effectively tackle and contain infections from further progression. However, due to the rapid mutation of the virus's nature, the ongoing struggle seems to go ever on [40]. The STexS will definitely serve as a suitable bedrock for future detections arrays. Not just with currently known lethal mutated variants, the STexS platform can successfully detect hundreds of variants regarding both precedented and novel, hazardous and non-lethal but with potential to develop. Future applications for the STexS will be focused not only with SNPs related to early cancer detection, but also in precise diagnosis of the ongoing COVID-19. Further, the technology will improve molecular biology and medical science and reshape the current health care industry.

References

- Collins FS, LD Brooks, A Chakravarti (1998) A DNA polymorphism discovery resource for research on human genetic variation. *Genome Res* 8(12): 1229-1231.
- Collins FS, M Morgan, A Patrinos (2003) The Human Genome Project: lessons from large-scale biology. *Science* 300(5617): 286-290.
- (2005) International HapMap C A haplotype map of the human genome. *Nature* 437(7063): 1299-1320.
- Rocha D, Ivo Gut, Alec J Jeffreys, Pui-Yan Kwok, Anthony J Brookes, et al. (2006) Seventh international meeting on single nucleotide polymorphism and complex genome analysis: Ever bigger scans and an increasingly variable genome. *Hum Genet* 119(4): 451-456.
- Vignal A, Denis Milan, Magali San Cristobal, André Eggen (2002) A review on SNP and other types of molecular markers and their use in animal genetics. *Genet Sel Evol* 34(3): 275-305.
- Kim S, A Misra (2009) SNP genotyping: Technologies and biomedical applications. *Annu Rev Biomed Eng* 9: 289-320.
- Wang DG, JB Fan, CJ Siao, A Berno, P Young, et al. (1998) Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science* 280(5366): 1077-1082.
- Nickerson DA, SL Taylor, K M Weiss, A G Clark, R G Hutchinson, et al. (1998) DNA sequence diversity in a 9.7-kb region of the human lipoprotein lipase gene. *Nat Genet* 19(3): 233-240.
- Ingram VM (1956) A specific chemical difference between the globins of normal human and sickle-cell anaemia haemoglobin. *Nature* 178(4537): 792-794.
- Hamosh A, TM King, BJ Rosenstein, M Corey, H Levison, et al. (1992) 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.
- Jin HJ, Segun Jung, Auditi R DebRoy, Ramana V Davuluri (2016) Identification and validation of regulatory SNPs that modulate transcription factor chromatin binding and gene expression in prostate cancer. *Oncotarget* 7(34): 54616-54626.
- Mullany LE, Roger K Wolff, Jennifer S Herrick, Matthew F Buas, Martha L Slattery, et al. (2015) SNP Regulation of microRNA Expression and Subsequent Colon Cancer Risk. *PLoS One* 10(12): e0143894.
- Huo Y, Shiwu Li, Jiewei Liu, Xiaoyan Li, Xiong Jian Luo et al. (2019) Functional genomics reveal gene regulatory mechanisms underlying schizophrenia risk. *Nat Commun* 10(1): 670.
- Guzey C, O Spigset (2004) Genotyping as a tool to predict adverse drug reactions. *Curr Top Med Chem* 4(13): 1411-1421.
- Pirmohamed M, BK Park (2001) Genetic susceptibility to adverse drug reactions. *Trends Pharmacol Sci* 22(6): 298-305.
- Marsh S, P Kwok, HL McLeod (2002) SNP databases and pharmacogenetics: Great start, but a long way to go. *Hum Mutat* 20(3): 174-179.
- Dvornyk V, Ji Rong Long, Dong Hai Xiong, Peng Yuan Liu, Lan Juan Zhao, et al. (2004) Current limitations of SNP data from the public domain for studies of complex disorders: A test for ten candidate genes for obesity and osteoporosis. *BMC Genet* 5: 4.
- Schwarz G, Stefan Bäumler, Annette Block, Friedrich G Felsenstein, Gerhard Wenzel, et al. (2004) Determination of detection and quantification limits for SNP allele frequency estimation in DNA pools using real time PCR. *Nucleic Acids Res* 32(3): e24.
- Lejman M, Monika Włodarczyk, Borys Styka, Agata Pastorczak, Joanna Zawitkowska, et al. (2020) Advantages and Limitations of SNP Array in the Molecular Characterization of Pediatric T-Cell Acute Lymphoblastic Leukemia. *Front Oncol* 10: 1184.

20. Garibyan L, N Avashia (2013) Polymerase chain reaction. *J Invest Dermatol* 133(3): 1-4.
21. Ishmael FT, C Stellato, (2008) Principles and applications of polymerase chain reaction: basic science for the practicing physician. *Ann Allergy Asthma Immunol* 101(4): 437-443.
22. Abbott MA, BJ Poiesz, BC Byrne, S Kwok, JJ Sninsky, et al. (1988) Enzymatic gene amplification: qualitative and quantitative methods for detecting proviral DNA amplified *in vitro*. *J Infect Dis* 158(6): 1158-1169.
23. Gaudet M, Anna Giulia Fara, Isacco Beritognolo, Maurizio Sabatti (2009) Allele-specific PCR in SNP genotyping. *Methods Mol Biol* 578: 415-424.
24. Matsuda K (2017) PCR-Based Detection Methods for Single-Nucleotide Polymorphism or Mutation: Real-Time PCR and Its Substantial Contribution Toward Technological Refinement. *Adv Clin Chem* 80: 45-72.
25. Blais J, Sébastien B Lavoie, Sylvie Giroux, Johanne Bussi eres, Carmen Lindsay, et al. (2015) Risk of Misdiagnosis Due to Allele Dropout and False-Positive PCR Artifacts in Molecular Diagnostics: Analysis of 30,769 Genotypes. *J Mol Diagn* 17(5): 505-514.
26. Kim JJ, Hyoung Min Park, A Young Kyoung, In Kyung Park, Si Kyu Lim, et al. (2021) Inclusion of double helix structural oligonucleotide (STexS) results in an enhance of SNP specificity in PCR. *Sci Rep* 11(1): 19098.
27. Erichsen HC, SJ Chanock (2004) SNPs in cancer research and treatment. *Br J Cancer* 90(4): 747-751.
28. Kroemer HK, M Eichelbaum (1995) "It's the genes, stupid". Molecular bases and clinical consequences of genetic cytochrome P450 2D6 polymorphism. *Life Sci* 56(26): 2285-2298.
29. Kellogg, DE, I Rybalkin, S Chen, N Mukhamedova, T Vlasik, et al. (1994) TaqStart Antibody: "Hot start" PCR facilitated by a neutralizing monoclonal antibody directed against Taq DNA polymerase. *Biotechniques* 16(6): 1134-1137.
30. Medrano RE, CA de Oliveira (2014) Guidelines for the tetra-primer ARMS-PCR technique development. *Mol Biotechnol* 56(7): 599-608.
31. Little S (2001) Amplification-refractory mutation system (ARMS) analysis of point mutations. *Curr Protoc Hum Genet Chapter 9: Unit 9.8.*
32. Vincent J, H Gurling, G Melmer (1994) Oligonucleotides as short as 7-mers can be used for PCR amplification. *DNA Cell Biol* 13(1): 75-82.
33. Guo Z, Mark S Gatterman, Lee Hood, John A Hansen, Effie W Petersdorf, et al. (2002) Oligonucleotide arrays for high-throughput SNPs detection in the MHC class I genes: HLA-B as a model system. *Genome Res* 12(3): 447-457.
34. Ishino S, Y Ishino (2014) DNA polymerases as useful reagents for biotechnology - the history of developmental research in the field. *Front Microbiol* 5: 465.
35. Braunstein GD, Lori Schwartz, Pamela Hymel, Jonathan Fielding (2021) False Positive Results With SARS-CoV-2 RT-PCR Tests and How to Evaluate a RT-PCR-Positive Test for the Possibility of a False Positive Result. *J Occup Environ Med* 63(3): e159-e162.
36. Kanji JN, Nathan Zelyas, Clayton MacDonald, Kanti Pabbaraju, Muhammad Naeem Khan, et al. (2021) False negative rate of COVID-19 PCR testing: a discordant testing analysis. *Virology* 18(1): 13.
37. Jarvis KF, JB Kelley (2021) Temporal dynamics of viral load and false negative rate influence the levels of testing necessary to combat COVID-19 spread. *Sci Rep* 11(1): 9221.
38. Bhattacharyya R, Ritoban Kundu, Ritwik Bhaduri, Debashree Ray, Lauren J Beesley, et al. (2021) Incorporating false negative tests in epidemiological models for SARS-CoV-2 transmission and reconciling with seroprevalence estimates. *Sci Rep* 11(1): 9748.
39. Alizad Rahvar AR, Safar Vafadar, Mehdi Totonchi, Mehdi Sadeghi (2021) False Negative Mitigation in Group Testing for COVID-19 Screening. *Front Med (Lausanne)* 8: 661277.
40. Justo Arevalo S, Daniela Zapata Sifuentes, C esar J Huallpa, Gianfranco Landa Bianchi, Adriana Castillo Ch avez, et al. (2021) Dynamics of SARS-CoV-2 mutations reveals regional-specificity and similar trends of N501 and high-frequency mutation N501Y in different levels of control measures. *Sci Rep* 11(1): 17755.

ISSN: 2574-1241

DOI: 10.26717/BJSTR.2022.41.006585

Hyoung Min Park. Biomed J Sci & Tech Res



This work is licensed under Creative Commons Attribution 4.0 License

Submission Link: <https://biomedres.us/submit-manuscript.php>



Assets of Publishing with us

- Global archiving of articles
- Immediate, unrestricted online access
- Rigorous Peer Review Process
- Authors Retain Copyrights
- Unique DOI for all articles

<https://biomedres.us/>