Original Article

To Investigate the Efficiency of Surveyor Nuclease in Error Correction during Gene Synthesis

***Zia-ur-Rehman Farooqi*[*]**
Andrew Curin**
***Imran Tipu*[**]**
***Philip J.R Day*[**]**
***Najam-us-Sahar Sadaf Zaidi*[*]**

**.Atta-ur-Rahman School of Applied Biosciences, National University of Sciences and Technology, Kashmir High ways Islamabad.*
*** Manchester Institute of Biotechnology, University of Manchester, UK*

Objective: To establish the efficacy of Surveyor Nuclease in gene synthesis process

Introduction: Synthetic biology provides tools to design "for the purpose" building of genome/gene and organisms. Approaches have adopted for the construction of different viral genomes like polio, etc. Error prone amplification of gene is the main hurdle in gene synthesis. The error may include insertion, deletion, substitution, etc. These error mainly arise from incorrect chemical synthesis of oligo,s or polymerase induced error. Therefore a sophisticated/efficient error correction method can give the better output of desired function of resulting protein.

Materials and Methods: In this study the efficiency of Surveyor enzyme (a mismatch repair enzyme) was analysed on 611 bp long hepatitis C virus (HCV) core gene. The synthesised core gene was cloned into pCDNA3.1^{+}, transformed in DH5α and after mini prep sent for the DNA sequencing.

Results: The result implicated that Surveyor enzyme showed maximum efficacy at 60min incubation. It can produce 60 per cent more error free gene as compared to Surveyor enzyme untreated genes.

Conclusion: This method can reduce the errors upto 60 %, when this method is utilised for gene synthesis

Key words: Gene Synthesis, Surveyor enzyme, Error correction

Address for Correspondence
Najam-us-Sahar Sadaf Zaidi,
Atta-ur-Rahman school of Applied Biosciences, National University of Sciences and Technology, Sector H-12, Islamabad,
E-mails: zaidi.sahar@gmail.com

Introduction

One of the latest applications of genetics, molecular, cell and computational biology knowledge has emerged as a new field of science called "synthetic biology". Synthetic biology covers the engineering of novel pathways, production of new biopharmaceuticals and even synthesis of different genomes.[1] *Invitro* synthesis of gene is the fundamental step to get the desired protein/phenotype in a particular organism. Currently genes are being synthesized in research labs by chemical synthesis of overlapping oligo,s covering whole gene. These sequences are than amplified enzymatically through PCR based amplification.But after successful amplification synthesis showed the common mutations like insertion, deletion, substitution of bases and gaps due to deletion of more than one nucleotide[2,3]. Current investigations showed that these mutations could be imparted during chemical synthesis of oligo's or polymerase induced mutations during thermo-cycling amplification. [3,4]

Naturaly, with *invivo* replication in prokaryotes (Bacteriophages &E.coli) showed the mutation rate ranging from 10^{-7} to 10^{-8}. Eukaryotes have a sophisticated DNA repair mechanism and polymerase proof reading ability but still they have about error rate of 10^{-6} [4]. Normal PCR in the presence of *Taq* Polymerase also showed the error rate about same as *invivo* replication of DNA in eukaryotes. These errors can be reduced from 10X to 100X by application of high fidelity polymerases like Phusion pol, Accu Prime, Q5 and etc [5]. Conversely, in polymerase amplification during gene synthesis reactions;typical high error rate of 10^{-1}–10^{-3}per kilo base-pairs (kbp) has been reported [4](Reviewed by Ma, Saaem and Tian, 2012).

To accomplish accurate gene synthesis of required target, it is necessary to remove these errors. So far different methods like enzymatic mismatch cleavage, mismatch-binding proteins, synthesis of higher quality oligonucleotides, site-directed mutagenesis and pyro-sequencing has been employed to remove the errors. In

this study we aimed to employ the error correction method based on nucleotide bases mismatch repairmechanism. In this method Surveyor Nuclease was used. This enzyme excises the DNA in two ways; mismatch-specific endonuclease and 3'–5' exonuclease activity. These excise products were reassembled by overlap extension polymerase chain reaction (OE-PCR) into intact gene. [4]

Materials and Methods

Gene synthesis: The core gene of hepatitis C virus (HCV) accession No. AAC-42195 was selectedfrom gene bank data base by random selection.Briefly, The core gene was synthesised by polymerase cycling assembly (PCA).For cloning pcDNA3.1 vector was linearized by *Nhe1* restriction enzyme (New England Bio, UK). Both, linearized vector and HCV-core were digested with T5 exonuclease to generate the overhanging fragments. Later one these were ligated by *Taq* ligase (New England Bio, UK) and phusion polymerase (New England Bio, UK). The ligated vector (pcDNA-HCV-core) was transformed in E.coliDH5-α for amplification. About twenty positive colonies were picked randomly and re-inoculated in LB media, plasmid isolated and were sent for the sequencing. After overnight incubation plasmids were isolated and confirmed by colony PCR (data not shown) [3,7,8].

Error correction: Normally, Surveyor enzyme cannot recognise the mutation in newly synthesize product. To visualise the product it is necessary that mutation should appear as a clear mismatch. This can be accomplished by simple denaturation. Briefly, product was denatured at 95 ^{0}C for two mins and then temperature was ramped from 95^{0}C - 58^{0}C at the rate of 2^{0}C /sec, then a hold at 25^{0}C for 10 mins. A mixture of new hybridised gene (5ul), Surveyor endonuclease (2ul) (Transgenomic, USA)), enhancer (1ul) and 1-HF buffer (New England Bio, UK) was incubated at 48^{0}C for 2hrs. To stop the endonuclease activity samples were purified by PCR purification kit (Qiagen, UK).[3,4] The excised products were subjected to OE-PCR to complete the amplification of full gene. Final amplified product was subjected to ligation in pcDNA3.1$^+$, transformation in DH5α andscreened by the procedure adopted earlier.

Selected plasmids were sent for the Sanger's Sequencing (Uni. Of Manchester Sequencing Facility) and Sequencing data was visualised by Bio edit tool. The sequencing data was analysed by using CLC sequence viewer (CLC Bio & Qaigen).

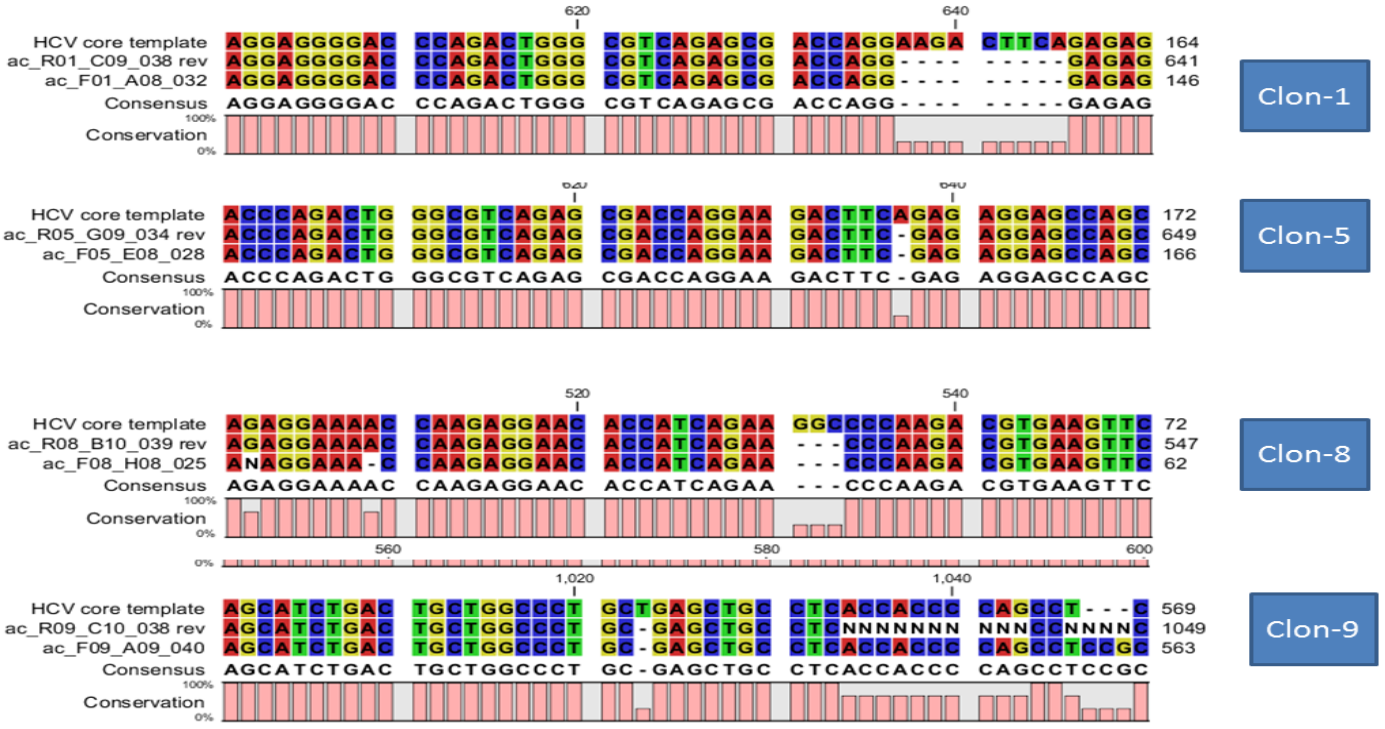

Figure1: Representative clones are showing multiple mutational events. The comparisons of the sequencing before and after the EC are shown in table-I.

Results

Total 10 random clones were selected for the analyses of mutational events in each pre EC and post EC by DNA sequencing. The data was analysed by CLC sequencing viewer.In pre EC sample, data show multiple insertion, deletion and gaps in all sequencing data except one. A global alignment was performed between reference sequence (accession No. AAC-42195), forward strand (FWD) core and sequencing of reverse strand (RVS) core. CLC output of few clones is being presented in figure 1. Only one clone showed 100 % match with the reference sequence.

The EC Surveyor assay contains three step, firstdenaturation, it facilitates the hetero-duplexes formation of DNA and this expose the errors on newly synthesised gene.[9] Second step digestion (main action of Surveyor enzyme), and third OE-PCR (to reconstruct the gene). To optimise the EC interval the digestion was performed for 60 mins (Lane-1 to Lane-3) and 20 min (lane-4 to lane-6) with reference to control (Lane-7 to Lane-9). Figure-2 showed the better results of EC when these amplified products (core gene) were treated for 60 mins. Therefore this time was chosen for all EC experiments.

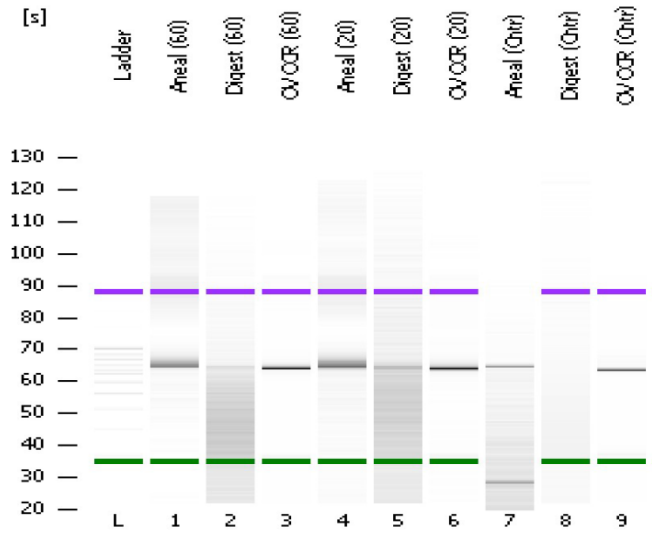

Figure 2: Elecrophrograph of EC experiment performed on core gene of 611 pb. Lanes show denaturation, digestion and OE-PCR respectively. Lane 1-3 represent the incubation time of 60 mins, Lane 4-6 represents incubation time of 20 mins, Lane 7-9 is the positive control for this experiment.

The sequencing data showed the multiple mutational events and high error rate in clones without Surveyor enzyme treatment, (Table-1). The study results indicated the marked reduction in the mutations after treatment with Surveyor enzyme.

Table-1: Showing the mutational events during gene synthesis of HCV core

Colony	without error correction	with error correction
clon-1	**159-160 (D)**	**467 G(D)**
Clon-2	**388 G (I)**	**482 G(D)**
clon-3	**295 A>G (S)**	100 % match
Clon-4	**118-121 4bp (D)**	**126 C(D), 413 C(D)**
Clon-5	**159 A (D),331 G (D)**	100 % match
Clon-6	**477-479 3bp (D)**	100 % match
Clon-7	**464-465 2bp (D)**	**229 G (D)**
Clon-8	**53-55 3bp (D)**	100 % match
clon-9	**542 T (D)**	100 % match
clon10	100 % match	100 % match

** I(insertion), D (deletion), S(substitution)*

Discussion

In synthetic biology procedures corrected gene synthesis is being used to enhance the efficiency of genes and for production of transgenic organisms. Therefore multiple methods has been devised to remove the oligos errors and polymerase induced errors. Currently, DNA repair mechanisms are natural choice of error corrections in synthetic constructs. Studies have applied multiple methods like enzymatic mismatch cleavage, mismatch-binding proteins, synthesis of higher quality oligonucleotides, site-directed mutagenesis and pyro-sequencing[6] The error removal methods have been described in all three phases of gene synthesis *i.e*. pre-synthesis errors, synthetic error and post synthetic error.These all procedures are laborious time consuming and require special set up. Moreover, in comparison to other enzyme base methods like MutS, Surveyor method is an improved technology with 16 fold better results.Consistent to the Surveyor enzyme most recent polymerase having 3'-5' proof reading abilities are also being used to reduce the error rate in gene synthesis. [9]

The main improvement employed in this study is the elimination of pre-synthesis oligo purification step. It also has been considered that this method is good for largergenes. Studies also reported that interestingly smaller genes have high frequency error rates as compared to longer chain.[4] The presence of nine error positive clones could be consistent with these findings.

On error correction (EC) by using Surveyor method (Table-1) and sequencing data show the 60 % improvement in the sequence.Consistent with the studies in which the reduction in error frequency was investigated with reference to incubation time, 58 % reduction in error frequency was observed when the incubation time was 60 mins instead of 20mins.[4] In spite of its high efficiency novel polymerase enzyme with proof reading abilities can reduce the gene synthesis time and errors.[5]

References

1. Jain, A., Bhatia, P. & Chugh, A. Microbial synthetic biology for human therapeutics. *Syst. Synth. Biol.* 6, 9–22 (2012).
2. Sharifian, H. Errors induced during PCR amplification. (2010).
3. Currin, a., Swainston, N., Day, P. J. & Kell, D. B. SpeedyGenes: an improved gene synthesis method for the efficient production of error-corrected, synthetic protein libraries for directed evolution. *Protein Eng. Des. Sel.* 27, gzu029– (2014).
4. Saaem, I., Ma, S., Quan, J. & Tian, J. Error correction of microchip synthesized genes using Surveyor nuclease. *Nucleic Acids Res.* 40, e23 (2012).
5. McInerney, P., Adams, P. & Hadi, M. Z. Error Rate Comparison during Polymerase Chain Reaction by DNA Polymerase. *Mol. Biol. Int.* 2014, 1–8 (2014).
6. Ma, S., Saaem, I. & Tian, J. Error correction in gene synthesis technology. *Trends Biotechnol.* 30, 147–154 (2012).
7. Gibson, D. G., Smith, H. O., Iii, C. A. H., Venter, J. C. & Merryman, C. Chemical synthesis of the mouse mitochondrial genome. 7, 11–15 (2010).
8. Gibson, D. G. Programming biological operating systems : genome design , assembly and activation. 11, (2014).
9. Kosuri, S. & Church, G. M. Large-scale de novo DNA synthesis: technologies and applications. *Nat. Methods* 11, 499–507 (2014).