



RESEARCH ARTICLE



Received on: 29/04/2014

Accepted on: 09/06/2014

Published on: 15/06/2014

Corresponding Author:

Smruti Prava Das

Department of Chemistry, Ravenshaw University, Cuttack - 753003, Odisha, India

Email: dassmrutiprava@yahoo.in



QR Code for Mobile users

Conflict of Interest: None Declared !

DOI: 10.15272/ajbps.v4i32.495

DNA sequencing by polymer synthesis with variable ratio of deoxynucleotide triphosphate and fluorescent dideoxynucleotide triphosphate

Devinder Singh Negi¹, Pankaj Shrivastava², Smruti Prava Das*¹,¹ Department of Chemistry, Ravenshaw University, Cuttack - 753003, Odisha, India² DNA Fingerprinting Unit, State Forensic Science Laboratory, Sagar - 470001, Madhya Pradesh, India

Abstract

The chemical mechanism of double stranded DNA polymer synthesis is a complex process. The nucleotides are joined together by phosphodiester bond between 5' C of deoxynucleotide triphosphate and O at 3' C of another deoxynucleotide triphosphate with the release of pyrophosphate and hydrogen ion. The synthesis reaction is performed on the template strand based on the complimentary addition of nucleotide attached by hydrogen bonds. The nucleotide adenine A pairs with thymine T and cytosine C pairs with guanine G by double and triple hydrogen bonds respectively. The deoxynucleotide triphosphate dNTP are the synthesizing molecules and fluorescent dideoxynucleotide triphosphate ddNTP are the terminator molecules in the polymer synthesis. The fluorescent ddNTP lacking O at 3' C, terminate the further addition of nucleotide and fluorescence of attached dye is captured by laser light excitation and absorption on the detector. The fluorescence detection read the nucleotide in the synthesis cycle. The DNA sequencing is the reading of nucleotide in the polymer synthesis with the dNTP and fluorescent ddNTP. We have investigated a chemical mechanism of DNA polymer synthesis with variable ratio of dNTP and ddNTP to elucidate the best design of DNA sequencing for robust, error free long read length to scan the genome. The short tandem repeat sequence STR and single nucleotide polymorphism with the reference sequence SNP are molecular markers reside in the size range of 100 to 1000 nucleotides sequence of genome. The variation in the molecular markers is analyzed for human genetic identification and diversity.

Keywords: DNA, Polymer, dNTP, ddNTP, PCR, Sequencing, Genome

Cite this article as:

Devinder Singh Negi, Pankaj Shrivastava, Smruti Prava Das. DNA sequencing by polymer synthesis with variable ratio of deoxynucleotide triphosphate and fluorescent dideoxynucleotide triphosphate. Asian Journal of Biomedical and Pharmaceutical Sciences; 04 (32); 2014. 32-38.

Introduction:

Deoxyribo Nucleic Acid DNA a polymer of nucleotides is a self-synthesizing hereditary molecule, forms the genetic anatomy of all living organisms. The sequencing is a chemical method to read the nitrogenous base of each nucleotide in the polymer. The four bases in the DNA molecule are adenine A, thymine T, cytosine C and guanine G. Nucleotides in the DNA molecule are joined by phosphodiester bond in condensation reaction that releases pyrophosphate, containing two phosphate groups and hydrogen ion. DNA is a double stranded molecule. The synthesis reaction is performed by complimentary addition of nucleotide to the template strand. The addition of nucleotide is based on the hydrogen bond pairing of A with T and C with G. The phosphodiester bond is formed between 5' C of deoxynucleotide triphosphate and O at 3' C of another deoxynucleotide triphosphate. The addition of fluorescent labelled dideoxynucleotide triphosphate ddNTP, lacking in O at 3' C in place of deoxynucleotide triphosphate dNTP makes it unable to form phosphodiester bond for the next nucleotide addition and terminate the chain. The concurrent and competitive, complimentary addition of any of dNTP and fluorescent dye labelled ddNTP i.e. JOE-ddATP, FAM-ddCTP, TAMRA-ddGTP, ROX-ddTTP on template DNA makes the synthesis of strand by dNTP and termination by ddNTP. The addition of each ddNTP is sensed by fluorescence excitation by laser light and emission capture by detector. The nucleotide sequence is read in the cycle of synthesis and termination. We have investigated a chemical mechanism of dNTP and ddNTP with variable ratio in the reaction mix to read the base of each nucleotide with maximum possible read length. The tandem repeat sequences of nucleotides and substitution of nucleotide with the reference control sequences form the short tandem repeat STR and single nucleotide polymorphism SNP molecular markers respectively. The STR and SNP markers regions are generally identified by sequencing in the range of 100 to 1000 nucleotides of the genome. The sequence variations of the STR and SNP molecular markers are associated with physical phenotypes and infer the human genetic identification and diversity.

Materials and Methods

DNA was extracted from the biological samples blood by standard phenol chloroform extraction method¹. 100ng of template DNA was subjected to PCR amplification reaction with forward and reverse flanking sequence oligonucleotide primers. The target region of DNA was amplified for generating multiple copies by polymerase chain reaction PCR. The PCR

reaction mix contained 1mM dNTP, 10mM Tris-HCl, 50mM KCl, 2.5mM MgCl₂, 2.5 U polymerase enzyme at pH 8.3. The PCR cycle conditions were 95 °C for 11 min, 35 cycles of 95 °C for 30sec, 59 °C for 30 sec, 72 °C for 45 sec and 72 °C for 15 min. The PCR reaction of 35 cycles generated 10⁹ copies of target DNA. The quantification of DNA was performed on agarose gel electrophoresis. The amplified DNA samples were cleaned with 1 U mix of exonuclease shrimp alkaline phosphatase at 37 °C for 15min and thereafter 80 °C for 30min to inactivate the enzyme. The 100 ng of template DNA samples were taken for sequencing reaction PCR amplification in reaction mix of 0.1 to 1mM dNTP, 1 to 0.1mM ddNTP, 10mM Tris-HCl, 50mM KCl, 2.5mM MgCl₂, 2.5 U polymerase enzyme at pH 8.3. The ratio of dNTP and ddNTP concentration was changed in different sets of reaction. The PCR cycle conditions were 95 °C for 11 min, 35 cycles of 95 °C for 30sec, 59 °C for 30 sec, 72 °C for 30 sec and 72 °C for 15 min. The PCR products were cleaned by ethanol sodium acetate precipitation. The sequencing reaction results were analysed in capillary electrophoresis Applied Biosystems 3500 genetic analyzer sequencer, and 9700 PCR system & chemical reagents were used as per manufacturer instructions for experiments. The ratio of dNTP and ddNTP was designed on the basis of chemical mechanism of their addition to the synthesizing polymer.

Results

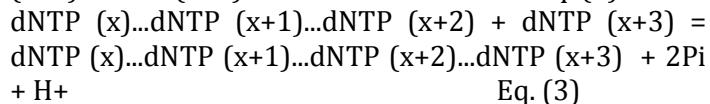
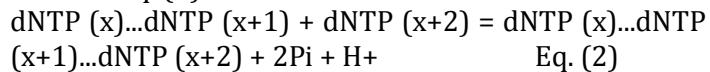
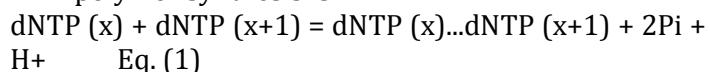
The chemical reaction for synthesis of DNA strand by complimentary nucleotide addition with dNTP and fluorescent ddNTP in variable ratio was analysed. A chemical model was prepared for the DNA polymer chain elongation and reading of nucleotide by addition of fluorescent ddNTP on the copies of target template DNA strand. The DNA polymer chain has 5' forward and 3' carbon reverse ends. The forward reaction is carried on the reverse template and vice versa. A chemical mechanism of DNA polymer synthesis with the induction of synthesizer and terminator molecules in a variable ratio in the synthesis cycle was designed. The polymer is synthesized with the deoxynucleotide triphosphate dNTP. The terminator of the reaction is fluorescent dye labelled dideoxynucleotide triphosphate ddNTP, lacking in polymer connecting oxygen for next addition of nucleotide²⁻⁶. The additions of fluorescent ddNTP in the synthesis reaction were detected in the recorder by laser excitation and capture of emission wavelength. The absorption and emission spectra of fluorescent dye attached to ddNTP are shown in table 1.

Table 1:

S. No.	Fluorescent Dye	Absorption Wavelength (nm)	Emission Wavelength (nm)
1.	JOE	520	548
2.	FAM	494	518
3.	TAMRA	555	580
4.	ROX	575	602

JOE: 4',5'-dichloro-2',7'-dimethoxy-5(6)-carboxyfluorescein, FAM: 6-carboxyfluorescein amidite TAMRA: Carboxytetramethylrhodamine ROX: 6-carboxyl-x-rhodamine. All four dyes excited at 488 nm wavelength. The dyes are attached to the nucleotide as JOE-ddATP, FAM-ddCTP, TAMRA-ddGTP, ROX-ddTTP.

The concentration of dNTP and ddNTP added in the DNA polymer reaction was designed to enable the synthesis cycle with dNTP and ddNTP to run for maximum unit length, so that each nucleotide is read in DNA sequencing to scan the genome. The equations of DNA polymer synthesis is:

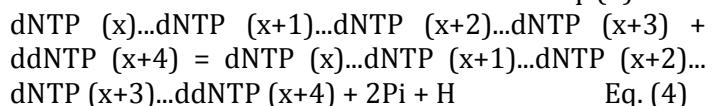
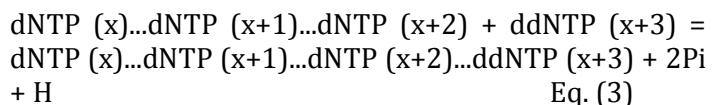
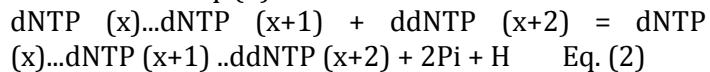
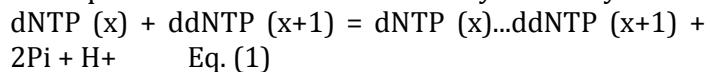


The equation (1) steps to (2), (3) and then so on continues to (n):



The synthesis reaction was performed in a thermal cycler and around 10^9 copies of DNA polymer were generated in the synthesis cycle. The n is saturation point of reaction and is the length of synthesis cycle. The nucleotide N in the reaction is either of nitrogenous base adenine A, cytosine C, thymine T and guanine G. The reaction is catalysed by *Thermus aquaticus* Taq DNA polymerase enzyme. The synthesis reaction is carried out on the template DNA strand of nucleotides, where A complements with T, C complements with G and vice versa. Hence the addition of nucleotide in the form of dNTP is dependent on template strand nucleotides. The unit x is nucleotide on the complimentary strand and so on addition goes in the polymer synthesis.

The equations of termination on the synthesis cycle:



The equation (1), (2), (3), (4) goes on till the last synthesized nucleotide attached with ddNTP and reaction terminated with no further addition of dNTP or ddNTP. The fluorescent dye attached to the ddNTP is excited by laser and emission wavelength is captured by detector. The termination reaction occurs on the synthesized units of polymer. The synthesis reaction is set A and termination reaction is set B. Then to complete the reading of nucleotides in the synthesis, set B will be set A+1. The ddNTP has to search the nucleotide on the complimentary strand for addition and also to compete with dNTP in the copies of polymer strand for addition. This makes it a complex set of chemical reactions where the quantity of dNTP and ddNTP, with their competitive addition in the growing polymer chain has to be balanced. The complex reaction carried out in thermal cycler to generate billion of copies of DNA polymer strands. The addition of nucleotide is dependent on the complimentary nucleotide sequence where the synthesis is performed. The termination gives the reading of nucleotide is also dependent on the copies of polymer.

The structures of nucleotide, nucleotide pairing by hydrogen bond, deoxynucleotide triphosphate dNTP and dideoxynucleotide triphosphate ddNTP, lacking in O at 3' C are shown in figure 1. The structure and mechanism of polymer formation is described in figure 2.

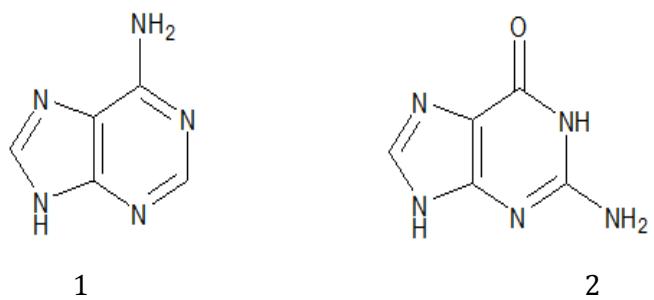
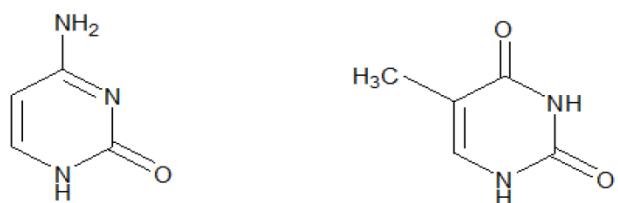
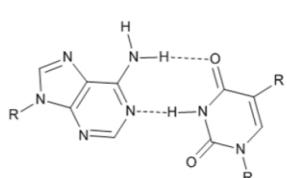


Figure 1: (1) Adenine (2) Guanine

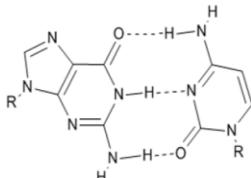


3

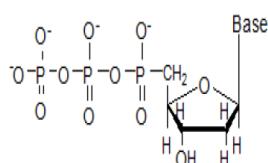
4



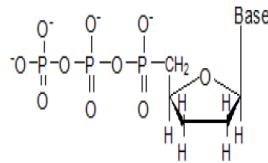
5



6



7



8

Fig 2(continued.....)

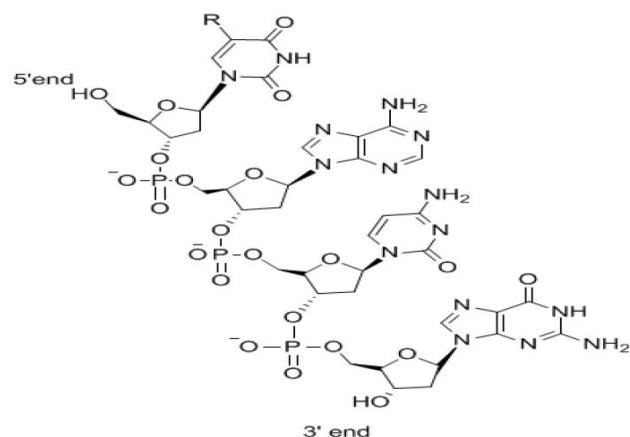
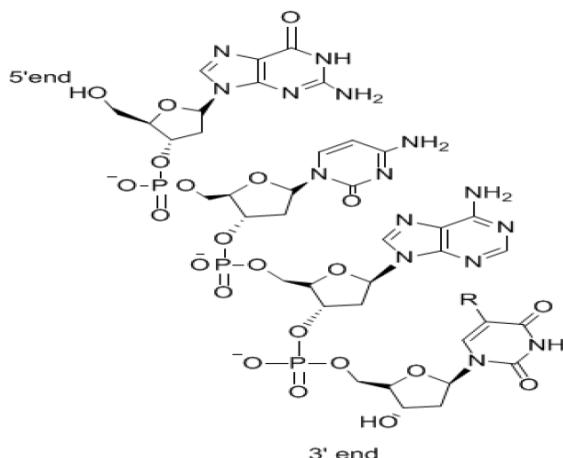
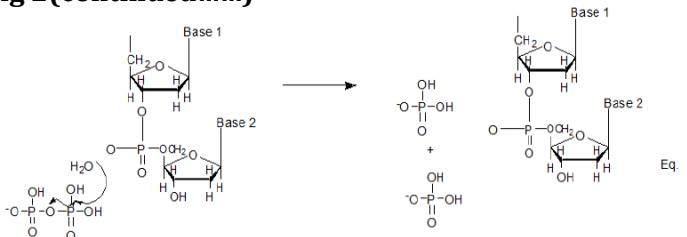


Figure 1 continued: (3) Cytosine (4) Thymine (5) Adenine-Thymine base pair (6) Cytosine-Guanine base pair (7) deoxynucleotide triphosphate (8) dideoxynucleotide triphosphate

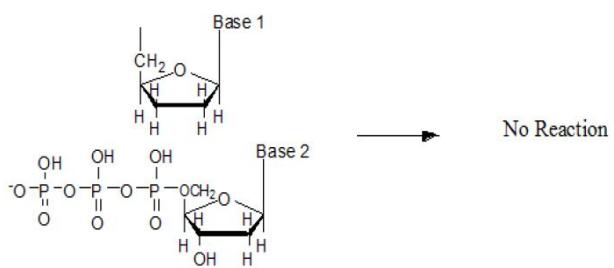
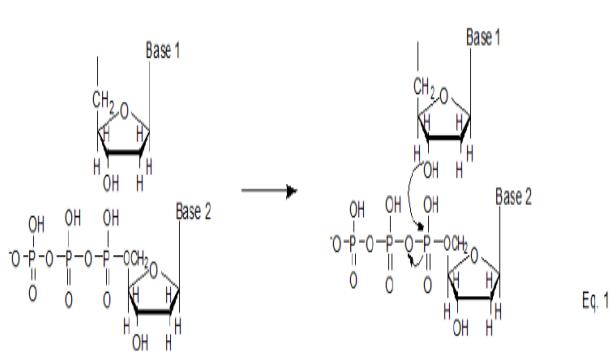


Figure 2: The DNA polymer synthesis with dNTP. The sequence of GCAT and TACG nucleotides. The ddNTP terminate the synthesis chain. (continued)

Figure 2: The DNA polymer synthesis with dNTP. The sequence of GCAT and TACG nucleotides. The ddNTP terminate the synthesis chain. (continued)

The ratio of dNTP to ddNTP was measured in variable ratio to analyse the efficiency of synthesis reaction and termination on the synthesis cycle to read the nucleotide of the polymer. The reading of nucleotide deciphers the sequence of the target region of genome for design of molecular markers. The results of effects

of ratio of dNTP to ddNTP in mM concentration on DNA sequencing read length is shown in table 2. The concentration of dNTP was kept from 0.1 to 1 mM and concentration of ddNTP from 1 to 0.1mM. The lower values of dNTP/ddNTP ratio synthesized short length of polymer with reading of nucleotide of the chain. The high ratio synthesized longer length of chain with reading of adding nucleotide. The ratio zero, terminates all the synthesis and reading of attached ddNTP is recorded as single base extension. The infinite ratio, no terminating ddNTP, has only synthesis without reading of adding nucleotide. The concentration of dNTP and ddNTP were in inverse proportion to the synthesis reaction as shown in figure 3.

S. No.	Deoxynucleotide triphosphate (dNTP) mM	Dideoxynucleotide triphosphate (ddNTP) mM	Ratio	Read Length
1.	0	1	0	1
2.	0.1	1	0.1	10
3.	0.2	0.9	0.2	20
4.	0.3	0.8	0.3	30
5.	0.4	0.7	0.5	40
6.	0.5	0.6	0.8	80
7.	0.6	0.5	1.2	120
8.	0.7	0.4	1.7	170
9.	0.8	0.3	2.6	260
10.	0.9	0.2	4.5	450
11.	0.95	0.15	6.3	630
12.	0.98	0.13	7.5	750
13.	0.99	0.12	8.2	820
14.	1	0.1	10	1000
15.	1	0	1	Nil

Table 2: The ratio of dNTP and ddNTP followed a direct proportion to the read length of synthesizing DNA polymer

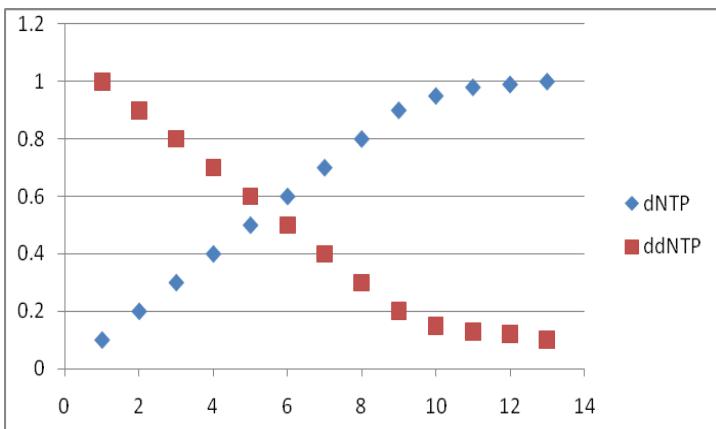


Figure 3: The concentration of dNTP and ddNTP were in ascending and descending order respectively and in inverse proportion in the synthesis reaction.

The DNA sequencing electropherograms showing the read of nucleotide bases are shown in supplementary figure 1. The dNTP/ddNTP ratio of 0.1, 0.2, 0.3, 0.5, 0.8, 1.2, 1.7, 2.6, 4.5, 6.3, 7.5, 8.2 and 10 provide the

corresponding 10, 20, 30, 40, 80, 120, 170, 260, 450, 630, 750, 820 and 1000 nucleotide sequence. The ratio 10 of synthesizer dNTP and terminator ddNTP in the chemical reaction of polymer synthesis is optimum for best long read length and ratio follow a direct proportion to read length.

Discussion and Conclusion

The DNA sequencing is performed by reading the addition of nucleotide complimentary to the template strand. The DNA template strand is the native strand of the target genome. The strand consists of polymer of four nucleotides adenine A, thymine T, cytosine C and guanine G. The adenine and guanine are purine compounds. The cytosine and thymine are pyrimidine compounds. The A pairs with T by two hydrogen bonds and C pairs with G by three hydrogen bonds. This pairing of nucleotide forms a unit of base pair. The pairing of nucleotide makes DNA a double strand molecule. The reading of nucleotide in the sequence is initiated by unwinding of two strands by heat labile hydrogen bond cleavage. The two strands becomes the template strand and any one can be selected for reading, making the other strand complimentary of the read one. The addition of nucleotide is based on addition of complimentary nucleotide to the template strand as A pairs with T and C pairs with G. The 5' C of incoming nucleotide couple with oxygen atom of 3' C and linked by phosphodiester bond releasing pyrophosphate and hydrogen ion. The different fluorescent dye labelling to the nucleotides read the addition signal. The reading is made exclusive for single nucleotide addition and any further addition is terminated. That is accomplished by making O at 3' C unavailable for the 5' C of incoming nucleotide. The most widely used chemical process of sequencing by synthesis is based on the fluorescent tagged nucleotide⁷⁻¹⁰. The sequencing with ddNTP addition is done collectively with dNTP. The dNTP addition elongates the chain and ddNTP addition terminates the reaction. Therefore the chemical binding ratio of dNTP/ ddNTP is measured to the length of sequence read. The target site of genome is picked by the oligonucleotide sequence complimentary to the conserved flanking region. The synthesis of complimentary nucleotide sequence is performed on the template strand of DNA. The read length variation of 100 to 1000 base pairs of DNA molecule is scanned for characterization of molecular markers. DNA sequencing is the revolutionary chemical method to read the each nucleotide of the genome. The sequencing process initiated with the unwinding of two complimentary strands of DNA on breaking of hydrogen bonds connecting the purines and pyrimidines at 95°C. The array of a designed 6pmole oligonucleotide of 20-25 bases called primer

complimentary to starting forward and end reverse conserved regions of target site is allowed to bind to the 30ng template strand at annealing temperature of 55 to 60°C, depending on the melting point of primer and thereafter the gap between the two ends was filled by adding free 1mM dNTP in a mix of 10mM Tris-HCl, 50mM KCl, 2.5mM MgCl₂, 2.5 U polymerase enzyme, pH 8.3, in extension step at 72°C. This process was performed in 20μl polymerase chain reaction PCR volume. The reaction cycle went to 35 to amplify the target site DNA sequence to 10⁹ copies. The amplified region was subjected to chemical method of reading the nucleotide by running the two concurrent process of synthesis by adding the deoxynucleotide triphosphate dNTP and termination by dideoxynucleotide triphosphate ddNTP. The ddNTP has no oxygen at 3°C, essential for formation of phosphodiester bond between two nucleotides. The mix of dNTP and ddNTP with 5pmole of forward primer and other reaction components, and similarly with reverse primer were taken for sequencing PCR. In this PCR the synthesis and termination run concurrently. The rate of synthesis should be balanced with the rate of termination. The adding dNTP are non fluorescent and terminating ddNTP are labelled with fluorescent dye attached to nitrogenous base of nucleotide. The adding of complimentary fluorescent dye labelled ddNTP to the pre-attached dNTP, do not allow any further addition of dNTP or ddNTP and the synthesis terminated. The fluorescence is read by laser excitation and detection of emission. The attached fluorescent ddNTP in this process, after completion of reaction passed through laser excitation. The emitted fluorescence is detected in the genetic analyzer, on passing the sequencing reaction products through the optical capillaries filled with the poly acrylamide gel to record the addition of ddNTP. The reaction products moved in electric field from negative to positive end of capillary. The sequencing of genomic regions is dependent on the addition reaction of ddNTP. The addition terminates the following addition of complimentary nucleotide to the template strand. In this cycle the complimentary nucleotide of the template DNA strand is read and sequencing process was completed. The ratio of dNTP to ddNTP is the limiting factor of the sequencing reaction. The ratio is balanced and measured for short and long read of the genome. The ratio also determines the starting and saturation point of the reaction. We have designed a chemical system of dNTP and ddNTP mix for providing robust, accurate nucleotide read of DNA sequences, so that desired length of target genome can be sequenced. The ratio of dNTP to ddNTP expressed a direct proportion in the form of 0.1, 0.2, 0.3, 0.5, 0.8, 1.2, 1.7, 2.6, 4.5, 6.3, 7.5, 8.2 and 10 values correspond to 10, 20, 30, 40, 80,

120, 170, 260, 450, 630, 750, 820 and 1000 read length nucleotide sequence. The ratio value 10 is best for DNA sequencing by synthesis of polymer alongwith the reading of nucleotide sequence. We further propose to explore and analyse chemical methods of DNA sequencing by capture of the products of the synthesis reaction, pyrophosphate and hydrogen ion to read the addition of nucleotide. The repeated nucleotides sequences, substitution of nucleotide, insertion & deletion of nucleotide and chemical modification to the nucleotide comprises the molecular markers. The DNA sequencing method by synthesis with the induction of terminator in the reaction has revolutionized the chemical mechanism of reading the nitrogenous bases of the genome². The nitrogenous bases of the nucleotide unit in the DNA polymer code all the genetic information of an individual. The molecular markers are scattered throughout the genome of an organism. The molecular markers encode the diversity of human in their diverse arena of physical, chemical and biological conditions.¹¹⁻¹³. The human groups are settled in different habitat and developed phenotypes to survive in their niche. This makes the human genetically diverse as per their physical and biological territory of existence. The molecular markers encode the information exclusive for an individual and these markers are applied in the human genetic identification and diversity.

Acknowledgement

We are thankful to Department of Chemistry, Ravenshaw University, Cuttack, India and Forensic Science Laboratory, Sagar, India for providing all support to carry out the work.

References

1. Sambrook J, Fritsch EF, Maniatis T. Molecular cloning. In: A laboratory manual, 2nd (edn.) Cold Spring Harbor Laboratory, Cold Spring Harbor, NY; 1989.
2. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci. 1977;74:5463-7.
3. Franc LTC, Carrilho E, Kist TBL. A review of DNA sequencing techniques. Reviews of Biophysics. 2002;35:169-200.
4. Fan J, Ranu RS, Smith C, Ruan C, Fuller CW. DNA Sequencing with [α -³³P]-Labeled ddNTP Terminators: A New Approach to DNA Sequencing with Thermo Sequenase™ DNA Polymerase. BioTechniques. 1996;21:1132-7.
5. Jia Guo J, Lin Yu L, Nicholas J, Turro NJ, Ju J. An Integrated System for DNA Sequencing by Synthesis Using Novel Nucleotide Analogues. Acc Chem Res. 2010;20: 551-63.
6. Kircher M, Kelso J. High-throughput DNA sequencing – concepts and limitations. Bioessays. 2010;32:524-36.
7. Guo J, Xu N, Li Z, Zhang S, Wu J, Kim DH, Marma MS, Meng Q, Cao H, Li X, Shi S, Yu L, Kalachikov S, Russo JJ, Turro NJ, Ju J. Four-color DNA sequencing with 3'-O-modified nucleotide reversible terminators and chemically cleavable fluorescent dideoxynucleotides. Proc Natl Acad Sci. 2008;105:9145-50.

8. Seo TS, Bai X, Ruparel H, Li Z, Turro NJ, Ju J. Photocleavable fluorescent nucleotides for DNA sequencing on a chip constructed by site-specific coupling chemistry. Proc. Natl Acad Sci. 2004;101:5488-93.
9. Ruparel H, Bi L, Li Z, Bai X, Kim DH, Turro NJ, Ju J. Design and synthesis of a 3-O-allyl photocleavable fluorescent nucleotide as a reversible terminator for DNA sequencing by synthesis. Proc. Natl Acad Sci. 2005;102:5932-37.
10. Ju J, Kim DH, Bi L, Meng Q, Bai X, Li Z, Li X, Marma MS, Shi S, Wu J, Edwards JR, Romu A, Turro NJ. Four-color DNA sequencing by synthesis using cleavable fluorescent nucleotide reversible terminators. Proc. Natl Acad Sci. 2006;103:19635-40.
11. Negi DS, Alam M, Bhavani SA, Nagaraju J. Multi-step microsatellite mutation in maternally transmitted locus D13S317: A case of maternal allele mismatch in the child. Int J of Legal Med. 2006;120:286-92.
12. Shrivastava P, Neetu M, Sharma NC, Trivedi VB, Negi DS, Verma MK. Autosomal STR genotyping analysis of juvenile delinquents of Madhya Pradesh: A pilot Study. Adv Bio Tech. 2013;13:20-4.
13. Shendure J, Mitra RD, Varma C, Church GM. Advanced sequencing technologies: methods and goals. Nat Rev Genet. 2004; 5:335-44.