International Journal of Life science & Pharma Research

PYROSEQUENCING- PRINCIPLES AND APPLICATIONS

MD. FAKRUDDIN[1*], ABHIJIT CHOWDHURY[1], MD. NUR HOSSAIN[1], KHANJADA SHAHNEWAJ BIN MANNAN[2], REAZ MOHAMMAD MAZUMDAR[3]

[1]Institute of Food Science and Technology (IFST), Bangladesh Council of Scientific and Industrial Research (BCSIR), Dhaka, B angladesh.
[2]Center for Food & Waterborne Disease (CFWD), icddr'b, Dhaka, Bangladesh.
[3]BCSIR Laboratories Chittagong, Bangladesh Council of Scientific and Industrial Research (BCSIR), Bangladesh.

ABSTRACT

Pyrosequencing is the first alternative to the conventional Sanger method for *de novo* DNA sequencing. Pyrosequencing is a DNA sequencing technology based on the sequencing-by-synthesis principle. It employs a series of four enzymes to accurately detect nucleic acid sequences during the synthesis. Pyrosequencing has the potential advantages of accuracy, flexibility, parallel processing, and can be easily automated. Furthermore, the technique dispenses with the need for labeled primers, labeled nucleotides, and gel-electrophoresis. Pyrosequencing has opened up new possibilities for performing sequence-based DNA analysis. The method has been proven highly suitable for single nucleotide polymorphism analysis and sequencing of short stretches of DNA. Pyrosequencing has been successful for both confirmatory sequencing and *de novo* sequencing**.** By increasing the read length to higher scores and by shortening the sequence reaction time per base calling, pyrosequencing may take over many broad areas of DNA sequencing applications as the trend is directed to analysis of fewer amounts of specimens and large-scale settings, with higher throughput and lower cost. This article considers key features regarding different aspects of pyrosequencing technology, including the general principles, enzyme properties, sequencing modes, instrumentation, limitations, potential and future applications.

Key words: Pyrosequencing, DNA sequencing, Typing, Enzymatic.

1. INTRODUCTION

DNA sequencing is one of the most important platforms for the study of biological systems today. Sequence determination is most commonly performed using di-deoxy chain termination technology **(**Ronaghi, 2001**).** The chain termination sequencing method, also known as Sanger sequencing, was developed by Frederick Sanger and colleagues (Sanger *et al.*, 1977), has been the most widely used sequencing method since its advent in 1977 and still is in use after more than 29 years. Despite all the advantages, there are limitations in this method, which could be complemented with other techniques (Gharizadeh *et al.*, 2007). Recently, pyrosequencing has emerged as a new sequencing methodology **(**Ronaghi, 2001**).**

Many research groups around the world have made effort to develop alternative principles of DNA sequencing. Three methods that hold great promise are sequencing by hybridization (Drmanac *et al.*, 1989; Khrapko *et al.*, 1989), parallel signature sequencing based on ligation and cleavage (Brenner *et al.*, 2000), and pyrosequencing (Ronaghi *et al.*, 1996, Ronaghi *et al.*, 1998).

Pyrosequencing technology is a novel DNA sequencing technology, developed at the Royal Institute of Technology (KTH), and is the first alternative to the conventional Sanger method for *de novo* DNA sequencing.

This method relies on the luminometric detection of pyrophosphate that is released during primer-directed DNA polymerase catalyzed nucleotide incorporation. It is suited for DNA sequencing of up to one hundred bases and it offers a number of unique advantages (Gharizadeh *et al.*, 2003a**).** This technique is a widely applicable, alternative approach for the detailed characterization of nucleic acids. Pyrosequencing has potential advantages of accuracy, flexibility, parallel processing, and can be easily automated. Furthermore, the technique avoids the need for labeled primers, labeled nucleotides, and gel-electrophoresis. Pyrosequencing has been successful for both confirmatory sequencing and *de novo* sequencing (Ronaghi, 2001**).**

2. PYROSEQUENCING CHEMISTRY

Pyrosequencing technique is based on sequencing-by-synthesis principle (Hyman, 1988; Melamede, 1985) and on the detection of released pyrophosphate (PPi) during DNA synthesis (Ronaghi, 2001). It employs a series of four enzymes to accurately detect nucleic acid sequences during the synthesis. In Pyrosequencing (Nyren and Skarpnack, 2001) the sequencing primer is hybridized to a single-stranded DNA biotin-labeled template and mixed with the enzymes; DNA polymerase, ATP sulfurylase, luciferase and apyrase, and the substrates adenosine 5′ phosphosulfate (APS) and luciferin (Gharizadeh *et al.*, 2007).

Cycles of four deoxynucleotide triphosphates (dNTPs) are separately added to the reaction mixture iteratively. The cascade starts with a nucleic acid polymerization reaction in which inorganic PPi is released as a result of nucleotide incorporation by polymerase. Each nucleotide incorporation event is followed by release of inorganic pyrophosphate (PPi) in a quantity equimolar to the amount of incorporated nucleotide. Subsequently The released PPi is quantitatively converted to ATP by ATP sulfurylase in the presence of APS. The generated ATP drives the luciferase-mediated conversion of luciferin to oxyluciferin, producing visible light in amounts that are proportional to the amount of ATPs. The light in the luciferase-catalyzed reaction with a maximum of 560 nanometer wavelength is then detected by a photon detection device such as a charge coupled device (CCD) camera or photomultiplier. Apyrase is a nucleotide-degrading enzyme, which continuously degrades ATP and non-incorporated dNTPs in the reaction mixture. There is a certain time interval (usually 65 seconds) between each nucleotide dispensation to allow complete degradation. For this reason, dNTP addition is performed one at a time (Gharizadeh *et al.*, 2007). Because the added nucleotide is known, the sequence of the template can be determined (Ronaghi, 2001). A schematic representation of pyrosequencing is shown in figure 1.

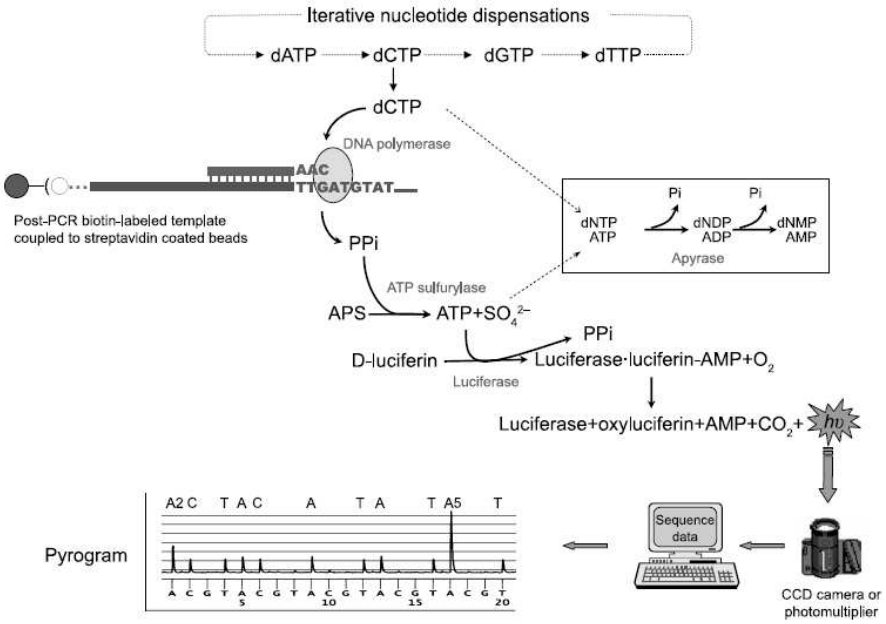

Figure 1: Schematic representation of pyrosequencing.

The generated light is observed as a peak signal in the pyrogram (corresponding to electropherogram in dideoxy sequencing) proportional to the number of nucleotides incorporated (a triple dGTP incorporation generates a triple higher peak) (Gharizadeh *et al.*, 2007). During this synthesis process, the DNA strand is extended by complementary nucleotides, and the DNA sequence is demonstrated by the pyrogram on a screen. The slope of the ascending curve in a pyrogram displays the activities of DNA polymerase and ATP sulfurylase, the height of the signal shows the activity of luciferase, and the slope of the descending curve demonstrates the nucleotide degradation (Gharizadeh *et al.*, 2003b). Base-callings are performed with integrated software, which has many features for related SNP and sequencing analysis (Gharizadeh *et al.*, 2007). The overall reaction from polymerization to light detection takes place within 3–4 sec at room temperature (Ronaghi, 2001). ATP sulfurylase converts PPi to ATP in approximately 1.5 seconds and the generation of light by luciferase takes place in less than 0.2 seconds (Nyren and Lundin, 1985).

Standard pyrosequencing uses the Klenow fragment of *Escherichia coli* DNA Pol I (Benkovic and Cameron, 1995). The ATP sulfurylase used in pyrosequencing is a recombinant version from the yeast *Saccharomyces cerevisiae* (Karamohamed *et al.*, 1999) and the luciferase is from the American firefly *Photinus pyralis* (Ronaghi, 2001). The apyrase is from *Solanum tuberosum (*Pimpernel variety) (Espinosa *et al.*, 2003; Nourizad *et al.*, 2003).

3. PYROSEQUENCING TECHNOLOGY

Removal or degradation of unincorporated or excess dNTPs was a crucial factor for sequencing-by-synthesis in order to be applied for DNA sequencing. In pyrosequencing nucleotide removal is performed in two different ways: (*i*) the solidphase pyrosequencing, which utilizes a three-coupled enzymatic procedure with washing steps and (*ii*) the liquid-phase pyrosequencing technique, which employs a cascade of four enzymes with no washing steps (Gharizadeh *et al.*, 2003a).

(1) Solid phase pyrosequencing

Solid-phase method is based on a combination of the sequencing-by-synthesis technique and a solid-phase technique. The four nucleotides are dispensed sequentially in the reaction system and a washing step removes the unincorporated nucleotides after each addition (Gharizadeh *et al.*, 2003b). There are different immobilization techniques that can be used in the solid-phase approach. In one approach, the biotin-labeled DNA template with annealed primer is immobilized to streptavidin coated magnetic beads (Ronaghi *et al.*, 1996). The immobilized, primed single-stranded DNA is incubated with three enzymes: DNA polymerase, ATP sulfurylase and luciferase. After each nucleotide addition to the reaction mixture, the DNA template is immobilized by a magnet system and the unincorporated nucleotides are removed by a washing step. Loss of DNA templates in the washing procedure, repetitive addition of enzymes, unstable baseline fluctuations and automation difficulties are drawbacks of this approach (Gharizadeh *et al.*, 2003a).

(2) Liquid phase pyrosequencing

Liquid-phase pyrosequencing employs a cascade of four enzymes and the DNA sequencing is monitored in real-time. The sequencing reaction is initiated by annealing a sequencing primer to a single-stranded DNA template.

Breakthrough in pyrosequencing by the liquid-phase approach came about by the introduction of a nucleotide-degrading enzyme, called apyrase (Nyren, 1994; Ronaghi *et al.*, 1998; Nyren and Skarpnack, 2001). The implementation of this enzyme in the pyrosequencing system excluded the use of solid-phase separation, and consequently, eliminated extra steps such as washes and repetitive enzyme additions. Apyrase shows high catalytic activity and low amounts of this enzyme in the reaction system efficiently degrade the unincorporated nucleoside triphosphates to nucleoside diphosphates and subsequently to nucleoside monophosphates. In addition, apyrase stabilizes the baseline with no fluctuations in the sequencing procedure as the same enzymes catalyze the reaction continuously (Gharizadeh *et al.*, 2003b).

4. AUTOMATION OF PYROSEQUENCING

The availability of an automated system for liquid-phase pyrosequencing (PSQ 96 system,

http://www.pyrosequencing. com) has allowed the technique to be adapted for high-throughput analyses (Ronaghi, 2001). Pyrosequencing is now being applied in microfluidic format commercially by 454 Life Sciences Corporation (Branford, CT, USA). The microfluidic pyroseqeuncing has been integrated with emulsion PCR and DNA sequencing assembly software. The DNA-sequence-signal peaks are presented in flowgrams. Sequence reads of up to 100 bases are generated, allowing massive parallel fashion sequencing in picolitre volumes. The new platform has the capacity to sequence up to 300,000 samples and generate up to 20-40 million bases at an accuracy of 99% per 4 hours sequencing run. In brief, the entire genomic DNA is sheared, and the fragments are ligated to adapters. The fragments are bound to the beads that favour one fragment per bead by limiting dilution approach. The beads are captured in the droplets of a PCR-reaction-mixture-in oil (Margulies *et al.*, 2005). The DNA fragments are clonally amplified and the emulsions are broken after complete amplification. After single-strand treatments, the beads are deposited into the picotitre wells (etched in a fibre-optic slide). Each well has the capacity for a single bead. Smaller beads carrying immobilized enzymes for pyrosequencing reaction are added to the wells and the reagents flow cyclically through the wells. The generated photons are detected by a CCD camera and presented in flowgrams by integrated software (Gharizadeh *et al.*, 2007).

5. APPLICATIONS OF PYROSEQUENCING

Pyrosequencing has opened up new possibilities for performing sequence-based DNA analysis (Ronaghi, 2001). Pyrosequencing is well suited for *de novo* sequencing and resequencing (Ronaghi, 2001). Currently, pyrosequencing method is broadly being used in many applications such as Single Nucleotide Polymorphism (SNP) genotyping (Ahmadian *et al.*, 2000a; Nordstrom *et al.*, 2000; Milan *et al.*, 2000), identification of bacteria (Gharizadeh *et al.*, 2003a; Grahn *et al.*, 2003; Jonasson *et al.*, 2002), fungal (Gharizadeh *et al.*, 2005; Trama *et al.*, 2005a; Trama *et al.*, 2005b) and viral typing (Gharizadeh *et al.*, 2001; Adelson *et al.*, 2005; Elahi *et al.*, 2003; Gharizadeh *et al.*, 2003a; Gharizadeh *et al.*, 2005). Moreover, the method has demonstrated the ability to determine difficult secondary structures (Ronaghi *et al.*, 1999) and perform mutation detection (Ahmadian *et al.*, 2000b; Garrcia *et al.*, 2000), DNA methylation analysis (Neve *et al.*, 2002; Rickert *et al.*, 2002; Uhlmann *et al.*, 2002), multiplex sequencing (Gharizadeh *et al.*, 2003b; Gharizadeh *et al.*, 2006), tag sequencing of cDNA library (Nordstrom *et al.*, 2001) and clone checking (Nourizad *et al.*, 2003). Another highly significant application is whole genome sequencing (Margulies *et al.*, 2005).

Table: 1: A Collection of Microbial Applications Using Pyrosequencing (Diggle and Clarke, 2004)

Organism	Application	Target	Reference
Eubacteria	Profiling and identification	In stomachs of Mongolian gerbils with or without *Helicobacter pylori*	Sun *et al.*, 2003
Mixed bacteria	Identification	In DNA-contaminated PCR amplifications of 16S DNA variable VR1 and VR3 regions	Grahn *et al.*, 2003
General bacteria	Classification, identification and subtyping	Analysis of 16S rDNA fragments	Jonasson *et al.*, 2002
Lactobacilli	Identification	DNA VRs within colonies collected from normal vaginal fluid	Tarnberg *et al.*, 2002
Helicobacter pylori	Profiling,	NudA protein	Lundin *et al.*, 2003
	Identification and subtyping	16S DNA VR1 and VR3	Monstein *et al.*, 2001; Sun *et al.*, 2003
Listeria	Grouping	Single nucleotide polymorphisms	Unnerstad *et*

monocytogenes		(SNP) in the inlB gene	*al.*, 2001
Papillomavirus	Typing	Human papillomavirus (HPV)	Gharizadeh *et al.*, 2001
Human immunodeficiency virus (HIV)	Monitoring	Resistance to HIV type 1 protease inhibitors (PI)	O'Meara *et al.*, 2001
Viruses	Quantitative	Estimation of viral fitness	Lahser *et al.*, 2003

Some of the potential applications of pyrosequencing have been described here.

5.1 Genotyping of Single-Nucleotide Polymorphisms

For analysis of single-nucleotide polymorphisms (SNPs) by pyrosequencing, the 3'-end of a primer is designed to hybridize one or a few bases before the polymorphic position. In a single tube, all the different variations can be determined as the region is sequenced. A striking feature of pyrogram readouts for SNP analysis is the clear distinction between the various genotypes; each allele combination (homozygous or heterozygous) will give a specific pattern compared to the two other variants (Ahmadian *et al.*, 2000a; Alderborn *et al.*, 2000; Ekstrom *et al.*, 2000; Nordstrom *et al.*, 2000). This feature makes typing extremely accurate and easy. Relative standard deviation values for the ratio between key peaks of the respective SNPs and reference counterparts are ≤0.1 (Alderborn *et al.*, 2000). Simple manual comparison of predicted SNP patterns and the raw data obtained can score an SNP, especially as no editing is needed. Because specific patterns can be readily achieved for individual SNPs, it is also be possible to automatically score the allelic status by pattern recognition software **(**Ronaghi, 2001**)**.

As pyrosequencing signals are very quantitative, it is possible to use this strategy for the studies of allelic frequency in large population. This system allows >5000 samples to be analyzed in 8 h. Furthermore, pyrosequencing enables determination of the phase of SNPs when they are in the vicinity of each other allowing the detection of haplotypes (Ahmadian *et al.*, 2000b).

5.2 Resequencing

Pyrosequencing is currently the fastest method for sequencing a PCR product. Because pyrosequencing generates an accurate quantification of the mutated nucleotides, the resequencing of PCR-amplified disease associated genes for mutation scanning will be one of the interesting applications. Using this technique for resequencing results in longer read length than de novo sequencing because nucleotide delivery can be specified according to the order of the sequence **(**Ronaghi, 2001**)**. Programmed dispensing generates a signal for each addition in a pyrogram, therefore variation in the pattern indicates the appearance of a mutation. This strategy has been used for resequencing of the p53 tumor suppressor gene, where mutations were successfully determined and quantified (Garrcia *et al.*, 2000).

5.3 Tag Sequencing

The sequence order of nucleotides determines the nature of the DNA. Theoretically, eight or nine nucleotides in a row should define a unique sequence for every gene in the human genome. However, it has been found that, to uniquely identify a gene of a complex organism such as human, a longer sequence of DNA is needed. In a pilot study, 98% of genes in a human cDNA library could be uniquely identified by sequencing a length of 30 nucleotides. Pyrosequencing was used to sequence this length for gene identification and the results were in complete agreement with longer sequence data obtained by Sanger DNA sequencing.

Pyrosequencing offers high-throughput analysis of cDNA libraries because 96 samples can be analyzed in less than one hour. Like Sanger DNA sequencing, pyrosequencing also has the advantage of library screening as the original cDNA clone is directly available for further analysis (Ronaghi, 2001).

5.4 Analysis of Difficult Secondary Structures

Hairpin structures are common features in genomic materials and have been proposed to have regulatory functions in gene transcription and replication. However, analyzing these sequences by

conventional DNA sequencing usually gives rise to DNA sequence ambiguities seen as "run-off" or compressions. These problems have been associated with gel electrophoresis (Ronaghi, 2001). Pyrosequencing was successfully applied to decipher the sequence of such regions (Ronaghi *et al.*, 1999).

5.5 Microbial and viral typing

Near instantaneous detection of pathogens from clinical materials is important for diagnosis, treatment and prophylactic measurements. The Microbial and viral threats have been altered in the course of history as a substantial number of pathogens have been eradicated or controlled. New pathogens are emerging and some are developing resistance. Accurate and specific typing of microbial and viral pathogens is of utmost importance in clinical diagnosis (Gharizadeh *et al.*, 2003a). Many microorganisms usually lack adequate morphological details for easy identification. Furthermore, the development of appropriate therapies/ vaccines requires implementation of sufficient parameters for microbial and viral detection, and a reliable and robust genotyping method is necessary for accurate follow-up during clinical trials and serveillance of treatment (Van Doorn *et al.*, 2001). Pyrosequencing technique, due to its many advantages, has been turned out to be an useful implement for microbial and viral typing. It is currently being applied for rapid typing of a large number of bacteria, yeasts, and viruses (Ronaghi, 2001).

5.6 Fungal identification

Pyrosequencing appears to be a good diagnostic tool for detection and identification of fungal pathogens. It can be utilized in typing of fungi clinically isolated from immune-compromised patients suffering from proven invasive fungal infections. For typing, the DNA has to be amplified by general consensus primers (Einsele *et al.*, 1997) having complementary to a highly conserved region within the 18S rRNA gene allowing the amplification of a broad range of fungal species. The amplified DNA fragments are then sequenced up to 40 bases to identify the samples. All the samples can be accurately identified by the pyrosequencing technique. The reproducibility of the technique has been confirmed by sequencing the amplicons several times with identical sequence data each time for every sample. The sequence data obtained by pyrosequencing technique suggested that 18 to 32 bases are sufficient for identification of *C. albicans, C. glabrata, C. krusei, C. parapsilosis, C. tropicalis* and *Aspergillus* spp. (Gharizadeh *et al.*, 2003b).

5.7 Bacterial identification

DNA markers used for typing normally contain both conserved and variable regions. A DNA primer complementary to the conserved or semiconserved region is usually employed to sequence the variable region. In bacteria, 16s rRNA gene is commonly used to identify different species and strains. By analyzing a sequence between 20–100 nucleotides on 16S rRNA gene, it is possible to group different bacteria taxonomically and, in many cases to get information about strains. Further development of bacterial identification by pyrosequencing could be brought about in order for the method to be suitable for regions other than 16s rRNA of different bacteria (Gharizadeh *et al.*, 2003a).

6. ADVANTAGES OF PYROSEQUENCING

Pyrosequencing has emerged as an alternative method of sequencing. Although it has read-length limitations compared with di-deoxy sequencing, it is a fast method with real-time read-out that is highly suitable for sequencing short stretches of DNA (Gharizadeh *et al.*, 2007).

Pyrosequencing employs co-operativity of several enzymes to monitor DNA synthesis. Parameters such as stability, fidelity, specificity, sensitivity, K_M, and *k*cat are mandatory for the optimal performance of the enzymes used in the sequencing reaction. The kinetics of the enzymes can be studied in realtime (Gharizadeh *et al.*, 2003b).

Unlike Sanger sequencing, which lays a reading gap of roughly 20-30 bases from the sequencing primer, pyrosequencing can generate sequence signals immediately downstream of the primer. As sequencing starts with the first base next to the annealed primer, making primer design becomes more flexible in this method. Sample and single-strand DNA preparation process is also relatively rapid (about 15 min), while sample preparation takes approximately 4 hours for Sanger sequencing (60 min for PCR cleanup, 3-4 hours for cyclic amplification and 15 minutes for dye

cleanup). The reagent costs are considerably lower for sequencing short stretches of DNA compared to currently available methods (Gharizadeh *et al.*, 2007).

The pyrosequencing technology has many unique advantages over other DNA sequencing technologies. One advantage is that the order of nucleotide dispensation can be easily programmed and alterations in the pyrogram pattern reveal mutations, deletions and insertions. Moreover, this technique is carried out in real-time, as nucleotide incorporations and base callings can be observed continuously for each sample. In addition, the Pyrosequencing method can be automated for large-scale screenings (Gharizadeh *et al.*, 2003a).

Table 2: **Pros and Cons of pyrosequencing.**

Pros	Cons
1. Expedient (1day)	1. Expensive
2. High thoroughput (>200,000)	2. Short reads for phylogenetic inference (next generation)
3. Sequence length can be as high as >200bp	3. Bioinformatics cumbersome
4. Unlimited sample number	4. Relatively high error rate (0.0098)
5. Simple incorporation into PCR protocols	5. Long single dNTP strings unreliable (8bp linearity)
6. Simple frequency data	6. Long fusion primers may bring bias
7. It is a fast method with real-time read-out that is highly suitable for sequencing short stretches of DNA.	
8. It can generate sequence signals immediately downstream of the primer.	
9. Sample preparation and single-strand DNA process is also relatively rapid	
10. The reagent costs are considerably lower for sequencing short stretches of DNA compared to current available methods.	

7. CHALLENGES WITH PYROSEQUENCING

1. Earlier pyrosequencing was earlier limited to sequencing of short stretches of DNA, due to the inhibition of apyrase. The natural dATP was a substrate for luciferase, resulting in false sequence signals. dATP was substituted by dATP-α-S (Ronaghi *et al.*, 1996). The dATP- α -S consisted of two isomers, Sp and Rp. The Rp isomer was not incorporated in the DNA template as it was not a substrate for DNA polymerase, and its presence in the sequencing reaction simply inhibits apyrase activity. By introducing the dATP- α -S Sp isomer, substantial longer reads were achieved. This improvement had a major impact on pyrosequencing read length and allowed sequencing of up to one hundred bases (Gharizadeh *et al.*, 2002) and opened up avenues for numerous applications (Gharizadeh *et al.*, 2007).

2. Homopolymer Ts (more than 3-4) are a challenge in Pyrosequencing. Homopolymer string (mainly homopolymeric T) regions can influence synchronized extension and synthesis of the DNA strand causing non-uniform sequence peak heights, affecting the read-length and possibly causing sequence errors. Studies have shown that the incorporation of dATP-α-S in T homopolymeric regions results in uneven sequence signals and reduced sequence quality directly downstream of such homopolymers (Gharizadeh *et al.*, 2002), (Eriksson *et al.*, 2004). As mentioned earlier, during the development of pyrosequencing technique, the natural dATP was replaced by dATP- α -S to increase the signal-to-noise ratio. Inefficient dATP- α -S incorporation by the exonuclease-deficient Klenow DNA polymerase causes the template extensions to go out of the phase making sequence peak signals asynchronous and ambiguous that could be demonstrated in the pyrograms as uniformly reduced sequence signal

peak-height. Therefore, efforts were made to compensate this in the reading of those sections. By employing Sequenase (Gharizadeh *et al.*, 2004), an exonuclease deficient T7 DNA polymerase, the poly-T homopolymer string reads were significantly improved by generation of significantly more synchronized sequence and uniform signal peaks after homopolymeric T regions (Gharizadeh *et al.*, 2007).

3. An important factor in pyrosequencing is primer design for PCR and sequencing. Sequencing primers should be checked for self-looping, primer-dimer (primer-primer hybridizations) and cross-hybridization (when more than one sequencing primer is used). Single-stranded DNA binding protein (SSB) (Ronaghi, 2000) is inhibitory to primer-template complications in pyrosequencing. Furthermore, it is highly recommended that sequencing primers be designed before PCR primers in order to biotin-label the suitable DNA strand. This facilitates avoiding possible difficult homopolymers or regions that can interfere with efficient sequencing and base-callings (Gharizadeh *et al.*, 2007).

4. An inherent problem with the described method is *de novo* sequencing of polymorphic regions in heterozygous DNA material. In most cases, it will be possible to detect the polymorphism. If the polymorphism is a result of substitution, it is possible to obtain a synchronized extension after the substituted nucleotide. If the polymorphism is due to deletion or insertion of the same kind as the adjacent nucleotide on the template, the sequence after the polymorphism will be synchronized (Ronaghi, 2001).

However, if the polymorphism is because a deletion or insertion of another type, the sequencing reaction can become out of phase, making the interpretation of the subsequent sequence difficult. If the polymorphism is known, it is always possible to use programmed nucleotide delivery to keep the extension of different alleles synchronized after the polymorphic region. It is also possible to use a bidirectional approach (Ronaghi *et al.*, 1999) whereby the complementary strand is sequenced to decipher the sequence flanking the polymorphism (Ronaghi, 2001).

5. Another inherent problem is the difficulty in determining the number of incorporated nucleotides in homopolymeric regions, due to the nonlinear light response following incorporation of more than 5–6 identical nucleotides. The polymerization efficiency over homopolymeric regions has been investigated and the results indicate that it is possible to incorporate ≤10 identical adjacent nucleotides in the presence of apyrase (Ronaghi, 2000). However, to elucidate the correct number of incorporated nucleotides, it may be necessary to use specific software algorithms that integrate the signals. For resequencing, it is possible to add the nucleotide twice for a homopolymeric region to ensure complete polymerization (Ronaghi, 2001**).**

8. FUTURE OF PYROSEQUENCING

Genome sequencing in general provides bulk of information that can be used in different areas of biology (Ronaghi, 2001). Future applications require more robust and efficient DNA sequencing techniques for sequence determination. The Pyrosequencing method has already shown evidence of high accuracy in DNA sequencing and analysis of polymorphic DNA fragments in many clinical and research settings. It is a relatively straightforward and user-friendly method possessing unique methodological characteristics and this technique is currently being used in multidisciplinary fields in academic, clinical and industrial settings. By increasing the read length to higher scores and by shortening the sequence reaction time per base calling, pyrosequencing may take over many broad areas of DNA sequencing applications as the trend is directed to analysis of fewer amounts of specimens and large-scale settings, with higher throughput and lower cost (Gharizadeh *et al.*, 2007).

The massively parallel pyrosequencing of emulsified PCR-based templates holds great promise to revolutionize high-throughput sequencing. However, there is concern over the potentially high degree of error (Huse *et al.*, 2007).

Pyrosequencing has shown excellent accuracy in analysis of polymorphic DNA fragments. This technology has also been used for quantification of allelic frequency in populations. While the variations are characterized, correlation of variation to phenotype can be performed. Pyrosequencing will have a large impact in that area because a large number of samples can be

pooled in one pyrosequencing reaction. A high throughput version of this technology can potentially be used for resequencing of genomes. Pyrosequencing technology is relatively new and there lies ample room for versatile developments in both chemistry and instrumentation. This technology is already time- and cost-competitive as compared to the most conventional sequencing methods. Work is underway to further improve the chemistry, to measure the sequencing efficiency at elevated temperatures, and to run the reaction in miniaturized formats (Ronaghi, 2001).

By removal of inhibitory factors and improvement of chemistry it is feasible to ensure enhanced sequence quality and read-length, which are likely to open new avenues for many new applications relating to *de novo* sequencing as well as re-sequencing, mutation detection, and microbial and viral typing.

Multiple sequencing primer method has also contributed to typing of samples containing a multitude of types/ species and unspecific amplified products in clinical settings, eliminating the need for stringent PCR reactions, nested PCRs and gene cloning. This strategy has many other potential applications as automation is being more and more integrated into clinical settings and the cost of DNA sequencing is dropping.

9. CONCLUDING REMARKS

The pyrosequencing method has emerged as a versatile DNA sequencing technology suitable for numerous applications in the field of modern biology. It is the first alternative to the conventional Sanger di-deoxy method for *de novo* DNA sequencing. Pyrosequencing was earlier restricted to sequencing and analysis of SNPs and short stretches of DNA. However, it has now gone through phenomenal developments and improvements both in chemistry and instrumentation. More importantly, there are still opportunities for further breakthroughs in this area as only very few research groups have focused on this method. In future, pyrosequencing technology is expected to achieve longer read length, to reduce the sequencing time frame, to decrease the sample quantity and to make further improvements in automation.

10. ACKNOWLEDGEMENT

The author is thankful to Zakir Hossain, Research Fellow, Bangladesh Council of Scientific and Industrial Research (BCSIR) for his kind efforts to review this article critically.

11. REFERENCES

1. Adelson M.E., Feola M., Trama J., Tilton R.C. and Mordechai E. Simultaneous detection of herpes simplex virus types 1 and 2 by real-time PCR and pyrosequencing. *J. Clin. Virol.* 2005: **33**: 25-34.
2. Ahmadian A., Gharizadeh B., Gustafsson A.C., Sterky F., Nyren P., Uhlen M. and Lundeberg J. Single-nucleotide polymorphism analysis by pyrosequencing. *Anal. Biochem*. 2000a: **280**: 103–110.
3. Ahmadian A., Lundeberg J., Nyren P., Uhlen M. and Ronaghi M. Analysis of the p53 tumor suppressor gene by pyrosequencing. *Biotechniques,* 2000b: **28**: 140–147.
4. Alderborn A., Kristofferson A. and Hammerling U. Determination of single nucleotide polymorphisms by real-time pyrophosphate DNA sequencing. *Genome Res*. 2000: **10**: 1249–1258.Bains W. and Smith G.C. A novel method for nucleic acid sequence determination. J. *Theoretical Biol.* 1988: **135**: 303-307.
5. Brenner S., Williams S.R., Vermaas E.H., Storck T. and Moon K. *et al. In vitro* cloning of complex mixtures of DNA on microbeads: Physical separation of differentially expressed cDNAs. *Proc. Natl. Acad. Sci.* 2000: **97**: 1665-1670
6. Benkovic S.J. and Cameron C.E. Kinetic analysis of nucleotide incorporation and misincorporation by Klenow fragment of *Escherichia coli* DNA polymerase I. *Methods Enzymol.* 1995: **262**: 257-269.
7. Drmanac R., Labat I., Brukner I. and Crkvenjakov R. Sequencing of megabase plus DNA by hybridization: Theory of the method. *Genomics*. 1989: **4**: 114–128.
8. Diggle M.A. and Clarke S.C. Pyrosequencing ™: Sequence typing at the speed of light. *Mol. Biotechnol.* 2004: **28**: 129-138.

9. Espinosa V., Kettlun A.M., Zanocco A., Cardemil E. and Valenzuela M.A. Differences in nucleotide-binding site of isoapyrases deduced from tryptophan fluorescence. *Phytochemistry*. 2003: **63**: 7-14.
10. Eriksson J., Gharizadeh B., Nourizad N. and Nyren P. 7-Deaza-2'-deoxyadenosine-5'-triphosphate as an alternative nucleotide for the pyrosequencing technology. *Nucleosides Nucleotides Nucleic Acids*. 2004: **23**: 1583-1594.
11. Ekstrom B., Alderborn A. and Hammerling U. Pyrosequencing for SNPs. *Progress Biomed. Optics*. 2000: **1**: 134–139.
12. Einsele H., Hebart H., Roller G., Loffler J. and Rothenhofer I. *et al.* Detection and identification of fungal pathogens in blood by using molecular probes. *J. Clin. Microbiol.* 1997: **35**: 1353-1360.
13. Elahi E., Pourmand N., Chaung R., Rofoogaran A. and Boisver J. *et al.* Determination of hepatitis C virus genotype by pyrosequencing. *J. Virol. Methods*. 2003: **109**: 171-176.
14. Garrcia C.A., Ahmadian A., Gharizadeh B., Lundeberg J., Ronaghi M. and Nyren P. Mutation detection by pyrosequencing: Sequencing of exons 5-8 of the p53 tumour suppressor gene. *Gene*. 2000: **253**: 249-257.
15. Gharizadeh B., Ghaderi M. and Nyren P. Pyrosequencing technology for short DNA sequencing and whole genome sequencing. *Technology*. 2007: **47**: 129-132.
16. Gharizadeh B., Nordstrom T., Ahmadian A., Ronaghi M. and Nyren P. Long-read pyrosequencing using pure 2'-deoxyadenosine-5'-O'-(1-thiotriphosphate) Sp-isomer. *Anal. Biochem.* 2002: **301:** *82*-90.
17. Gharizadeh B., Eriksson J., Nourizad N., Nordstrom T. and Nyren P. Improvement in pyrosequencing technology by employing sequenase polymerase. *Anal. Biochem*. 2004: **330**: 272-280.
18. Gharizadeh B., Ohlin A., Molling P., Backman A., Amini B., Olcen P. and Nyren P. Multiple group-specific sequencing primers for reliable and rapid DNA sequencing. *Mol. Cell Probes*. 2003a: **17**: 203-210.
19. Gharizadeh B., Ghaderi M., Donnelly D., Amini B., Wallin K.L. and Nyren P. Multiple-primer DNA sequencing method. *Electrophoresis*. 2003b: **24**: 1145-1151.
20. Gharizadeh B., Kaller M., Nyren P., Andersson A., Uhlen M., Lundeberg J. and Ahmadian A. Viral and microbial genotyping by a combination of multiplex competitive hybridization and specific extension followed by hybridization to generic tag arrays. *Nucl. Acids Res.* 2003a: **31**: e146.
21. Gharizadeh B., Oggionni M., Zheng B., Akom E. and Pourmand N. *et al.* Type-specific multiple sequencing primers: A novel strategy for reliable and rapid genotyping of human papillomaviruses by pyrosequencing technology. *J. Mol. Diagn.* 2005: **7**: 198-205.
22. Gharizadeh B., Zheng B., Akhras M., Ghaderi M. and Jejelowo O. *et al.*, Sentinel-base DNA genotyping using multiple sequencing primers for high-risk human papillomaviruses. *Mol. Cell Probes*. 2006: **20**: 230-238.
23. Gharizadeh B., Kalantari M., Garcia C.A., Johansson B. and Nyren P. Typing of human papillomavirus by pyrosequencing. *Lab. Invest.* 2001: **81**: 673-679.
24. Grahn N., Olofsson M., Ellnebo-Svedlund K., Monstein H.J. and Jonasson J. Identification of mixed bacterial DNA contamination in broadrange PCR amplification of 16S rDNA V1 and V3 variable regions by pyrosequencing of cloned amplicons. *FEMS Microbiol. Lett.* 2003: **219**: 87–91
25. Hyman E.D. A new method of sequencing DNA. *Anal. Biochem.* 1988: **174**: 423-436.
26. Huse S.M., Huber J.A., Morrison H.G., Sogin M.L. and Welch D.M. Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biol.* 2007: **8**: R143.
27. Jonasson J., Olofsson M. and Monstein H.J. Classification, identification and subtyping of bacteria based on pyrosequencing and signature matching of 16S rDNA fragments. *APMIS*. 2002: **110**: 263–272.
28. Karamohamed S., Nilsson J., Nourizad K., Ronaghi M., Pettersson B. and Nyren P. Production, purification and luminometric analysis of recombinant *Saccharomyces cerevisiae* MET3 adenosine triphosphate sulfurylase expressed in *Escherichia coli*. *Protein Exp. Purif.* 1999: **15**: 381-388.
29. Khrapko K.R., Lysov Y.P., Khorlyn A.A., Shick V.V., Florentiev V.L. and Mirzabekov A.D. An oligonucleotide hybridization approach to DNA sequencing. *FEBS Lett.* 1989: **256**: 118–122.

30. Lundin A., Nilsson C., Gerhard M., Andersson D.I., Krabbe M. and Engstrand L. The NudA protein in the gastric pathogen Helicobacter pylori is an ubiquitous and constitutively expressed dinucleoside polyphosphate hydrolase. *J. Biol. Chem.* 2003: **278**: 12574–12578.
31. Lahser F.C., Wright-Minogue J., Skelton A. and Malcolm B.A. Quantitative estimation of viral fitness using pyrosequencing. *Biotechniques.* 2003: **34**: 26–28.
32. Melamede R.J. Automatable process for sequencing nucleotide. 1985: U.S. Patent 4863849. http://www.freepatentsonline.com/4863849.html.
33. Margulies M., Egholm M., Altman W.E., Attiya S. and Bader J.S. *et al.* Genome sequencing in microfabricated high-density picolitre reactors. *Nature.* 2005: **437**: 376-380.
34. Milan D., Jeon J.T., Looft C., Amarger V. and Robic A. *et al.* A mutation in PRKAG3 associated with excess glycogen content in pig skeletal muscle. *Science.* 2000: **288**: 1248-1251.
35. Monstein H.J., Nikpour-Badr S. and Jonasson J. Rapid molecular identification and subtyping of *Helicobacter pylori* by pyrosequencing of the 16S rDNA variable V1 and V3 regions. *FEMS Microbiol. Lett.* 2001: **199**: 103–107.
36. Nyren P. Apyrase immobilized on paramagnetic beads used to improve detection limits in bioluminometric ATP monitoring. *J. Bioluminescence Chemiluminescence.* 1994: **9**: 29-34.
37. Nyren P. and Skarpnack S.E. Method of sequencing DNA based on the detection of the release of pyrophosphate and enzymatic nucleotide degradation. 2001: U.S. Patent 6258568. http://ip.com/patent/US6258568.
38. Nourizad N., Ehn M., Gharizadeh B., Hober S. and Nyren P. Methylotrophic yeast *Pichia pastoris* as a host for production of ATP-diphosphohydrolase (apyrase) from potato tubers (*Solanum tuberosum*). *Protein Exp. Purif.* 2003: **27**: 229-237.
39. Nyren P. and Lundin A. Enzymatic method for continuous monitoring of inorganic pyrophosphate synthesis. *Anal. Biochem.* 1985: **151**: 504-509.
40. Nordstrom T., Ronaghi M., Forsberg L., de Faire U., Morgenstern R. and Nyren P. Direct analysis of single-nucleotide polymorphism on double-stranded DNA by pyrosequencing. *Biotechnol. Applied Biochem.* 2000: **31**: 107-112.
41. Nordstrom T., Gharizadeh B., Pourmand N., Nyren P. and Ronaghi M. Method enabling fast partial sequencing of cDNA clones. *Anal. Biochem.* 2001: **292**: 266-271.
42. Nordstrom T., Ronaghi M., Forsberg L., de Faire U., Morgenstern R. and Nyren P. Direct analysis of single-nucleotide polymorphism on double-stranded DNA by pyrosequencing. *Biotechnol. Appl. Biochem.* 2000: **31**: 107–112.
43. Neve B., Froguel P., Corset L., Vaillant E., Vatin V. and Boutin P. Rapid SNP allele frequency determination in genomic DNA pools by pyrosequencing. *Biotechniques.* 2002: **32**: 1138-1142.
44. Nourizad N., Gharizadeh B. and Nyren P. Method for clone checking. *Electrophoresis.* 2003: **24**: 1712-1715.
45. O'Meara D., Wilbe K., Leitner T., Hejdeman B., Albert J. and Lundeberg J. Monitoring resistance to human immunodeficiency virus type 1 protease inhibitors by pyrosequencing. *J. Clin. Microbiol.* 2001: **39**: 464–473.
46. Ronaghi M. Pyrosequencing sheds light on DNA sequencing. *Genome Research.* 2001: **11**: 3-11 doi:10.1101/gr.150601
47. Ronaghi M., Uhlen M. and Nyren P. A sequencing method based on real-time pyrophosphate. *Science.* 1998: **281**: 363- 365.
48. Ronaghi M. Improved performance of pyrosequencing using single-stranded DNA-binding protein. *Anal. Biochem.* 2000: **286**: 282-288
49. Ronaghi M., Nygren M., Lundeberg J. and Nyren P. Analyses of secondary structures in DNA by pyrosequencing. *Anal. Biochem.* 1999: **267**: 65-71.
50. Ronaghi M., Karamohamed S., Pettersson B., Uhlen M. and Nyren P. Real-time DNA sequencing using detection of pyrophosphate release. *Anal. Biochem.* 1996: **242**: 84–89.
51. Rickert A.M., Premstaller A., Gebhardt C. and Oefner P.J. Genotyping of Snps in a polyploid genome by pyrosequencing. *Biotechniques.* 2002: **32**: 592-600.

52. Sanger F., Nicklen S. and Coulson A.R. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA*. 1977: **74**: 5463-5467.
53. Sun Y.Q., Monstein H.J., Nilsson L.E., Petersson F. and Borch K. Profiling and identification of Eubacteria in the stomach of *Mongolian gerbils* with and without *Helicobacter pylori* infection. *Helicobacter*. 2003: **8**: 149–157.
54. Trama J.P., Mordechai E. and Adelson M.E. Detection of *Aspergillus fumigatus* and a mutation that confers reduced susceptibility to itraconazole and posaconazole by real-time PCR and pyrosequencing. *J. Clin. Microbiol.* 2005a: **43**: 906 - 908.
55. Trama J.P., Mordechai E. and Adelson M.E. Detection and identification of *Candida* species associated with *Candida Vaginitis* by real-time PCR and pyro sequencing. *Mol Cell Probes*. 2005b: **19(2)**: 145-152.
56. Tarnberg M., Jakobsson T., Jonasson J. and Forsum U. Identification of randomly selected colonies of lactobacilli from normal vaginal fluid by pyrosequencing of the 16S rDNA variable VR1 and VR3 regions. *APMIS*. 2002: **110**: 801–810.
57. Unnerstad H., Ericsson H., Alderborn A., Tham W., Danielsson-Tham M.L. and Mattsson J.G. Pyrosequencing as a method for grouping of *Listeria monocytogenes* strains on the basis of single-nucleotide polymorphisms in the inlB gene. *Applied Environ. Microbiol.* 2001: **67**: 5339-5342.
58. Uhlmann K., Brinckmann A., Toliat M.R., Ritter H. and N'rnberg P. Evaluation of a potential epigenetic biomarker by quantitative methyl-single nucleotide polymorphism analysis. *Electrophoresis*. 2002: **23**: 4072-4079.
59. Van Doorn L.J., Kleter B. and Quint W.G. Molecular detection and genotyping of human papillomavirus. *Expert Rev. Mol. Diagn.* 2001: **1**: 394-402.