REVIEW ARTICLE

Evolution of DNA Sequencing

Hamid Nawaz Tipu[1] and Ambreen Shabbir[2]

ABSTRACT

Sanger and coworkers introduced DNA sequencing in 1970s for the first time. It principally relied on termination of growing nucleotide chain when a dideoxythymidine triphosphate (ddTTP) was inserted in it. Detection of terminated sequences was done radiographically on Polyacrylamide Gel Electrophoresis (PAGE). Improvements that have evolved over time in original Sanger sequencing include replacement of radiography with fluorescence, use of separate fluorescent markers for each nucleotide, use of capillary electrophoresis instead of polyacrylamide gel electrophoresis and then introduction of capillary array electrophoresis. However, this technique suffered from few inherent limitations like decreased sensitivity for low level mutant alleles, complexities in analyzing highly polymorphic regions like Major Histocompatibility Complex (MHC) and high DNA concentrations required. Several Next Generation Sequencing (NGS) technologies have been introduced by Roche, Illumina and other commercial manufacturers that tend to overcome Sanger sequencing limitations and have been reviewed. Introduction of NGS in clinical research and medical diagnostics is expected to change entire diagnostic approach. These include study of cancer variants, detection of minimal residual disease, exome sequencing, detection of Single Nucleotide Polymorphisms (SNPs) and their disease association, epigenetic regulation of gene expression and sequencing of microorganisms genome.

Key Words: *DNA sequencing. Sanger sequencing. Next generation sequencing.*

INTRODUCTION

DNA sequencing is the determination of precise order of nucleotides [Adenine (A), Guanine (G), Cytosine (C), Thymine (T)] in a DNA molecule/genome. It includes many methods that have evolved over a period of time and revolutionized the biological research, medical diagnostics, forensic sciences and biotechnology. The focus of this review will be the scientific basis of these methods, how these have evolved over past few decades, their impact and role in rapid advancement of research and where the future of sequencing might lead. For this purpose, an extensive literature search was performed primarily in Google, Google books, Google Scholar, PubMed, Medline and Science Direct databases using keywords "DNA sequencing techniques", "Sanger sequencing", "next generation sequencing" and "second generation sequencing techniques" from the fields of immunogenetics, molecular biology, molecular diagnostics and immunology. Maximum effort was made to include both original research articles and review articles published in last 6 years, i.e., till 2008; however, at many places due to historical nature of the subject, old references had to be consulted.

[1] *Department of Pathology, Combined Military Hospital, Khuzdar Cantt.*
[2] *Department of Oral Pathology, Institute of Dentistry, CMH Medical College, Lahore.*

Correspondence: Dr. Hamid Nawaz Tipu, Classified Pathologist and Consultant Immunologist, Combined Military Hospital, Khuzdar Cantt.
E-mail: hnt1779@yahoo.com

Received: March 13, 2014; Accepted: September 17, 2014.

Watson and Crick proposed the structure of DNA in their original paper; "this structure has two helical chains each coiled around the same axis."[1] It is now known that in this double helix, each strand is composed of a 2' deoxyribose sugar which is a pentose; its 1' carbon binds one of four nitrogenous bases while phosphate group at 5' carbon binds hydroxyl group at 3' carbon of next pentose in sequence (Figure 1). And then its usual binding of adenine with thymine and guanine with cytosine. Knowledge of this atomic level structure is critical in understanding various sequencing methods. Despite discovery of DNA structure, it was only until late 1970s that reliable techniques were developed to

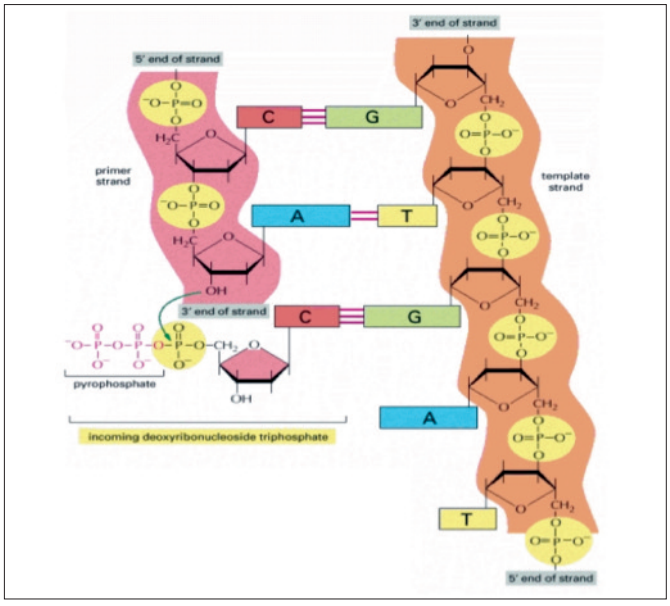

Figure 1: DNA biochemical structure. Source: Molecular Biology of the Cell. Garland Science 2002.

sequence it. Chemical cleavage of DNA for sequencing was developed by Maxam and Gilbert in late 1970s.[2,3] Owing to complexities of procedure, extensive use of hazardous chemicals and autoradiography, and short read lengths (less than 500 bp); it is now obsolete[4] and is of historical interest only.

However, the method that began a new era of molecular biology was developed by Frederick Sanger and co-workers in 1977.[5] In his original article, Sanger describes the basic principle of his methods that lasts until recently, that when 2'3' dideoxythymidine triphosphate (ddTTP) is inserted into the growing oligonucleotide chain, it inhibits DNA polymerase to extend the chain any further, and carrying out the reaction with one 32P labeled ddTTP and other normal deoxyribose nucleotide triphosphates (dNTPs), a mixture of fragments with same 5' end but ddT residue at 3' end is obtained. By electrophoresing it on denaturing polyacrylamide gel, bands show distribution of ddT in DNA strand. Extrapolating the principle to remaining three nucleotides will generate bands of varying lengths from which sequence can be read.[5] The Sanger method still involved use of radioactive substances, which were later replaced as modifications in the original method continued.

Sanger di-deoxy chain termination sequencing method: During DNA replication in a cell (leaving other biochemical details and machinery aside), it extends in 5' to 3' direction only, and this has far reaching consequences in Sanger sequencing. Looking at the Figure 1 again, incoming nucleotide binds the previous one in chain by forming phosphodiester bond, releasing pyrophosphate making the reaction favorable energetically. The OH at 3' carbon of previously bound nucleotide binds phosphate at 5' carbon of incoming one. This extension cannot take place in reverse direction (3' to 5') as polymerase does not have this ability in human beings.[6] Now if there is no OH at 3' carbon of a nucleotide in growing chain, incoming nucleotide cannot be added to it due to lack of phosphodiester bond formation and growth of DNA strand stops here. This chemical modification of dNTPs into ddNTPs (removal of OH group from 3' carbon of pentose sugar) forms the basis of Sanger sequencing and was a novel approach then.

"Classical" Sanger sequencing relied on base specific termination of reaction when all four dNTPs plus one ddNTP (in limiting quantity) was added to the reaction. Consequently, four reactions needed to be set up, each containing one of four ddNTPs (ddATP, ddGTP, ddCTP, ddTTP). In the presence of polymerase and dNTPs, the extension of growing DNA chain terminates every time a corresponding ddNTP is added. This happens in all four reactions and results in cocktail of extension products where ddNTP has been added to 3' growing end, resulting in different sized DNA fragments. The other novelty of approach was labeling of extension products with radioactive phosphorus or sulfur isotopes incorporated into newly synthesized DNA strand through a labeled precursor (dNTP or sequencing primer).

Figure 2: Autoradiograph of acrylamide gel for DNA sequencing.[5]

As all the four reaction tubes contained products of varying sizes, next step was to separate them achieved with Polyacrylamide Gel Electrophoresis (PAGE) and subsequent radiography.[7] The sequence was deducted according to size of product on gel. Figure 2 shows autoradiograph from original work of Sanger used for DNA sequencing.[5] The hazards of radioactivity and its waste disposal issues prompted the scientists to replace it with other techniques. Chemiluminescence did not prove viable for DNA sequencing due to its indirect detection limitation. Fluorescence detection is simple, sensitive and easy to automate and have almost entirely replaced radioactive detection methods in DNA community.[8]

Improvements upon classical Sanger's method: Still the Sanger's method required labor intensive preparative protocol. Efforts to improve upon it continued. In 1986, Smith *et al*. showed that sequencing data could be collected directly to a computer, without autoradiography being involved. The sequencing primer was labeled with four different dyes and used in each of the four sequencing reactions. The reactions were combined and electrophoresed in a single polyacrylamide tube gel. DNA passed a detector near the bottom of gel and dyes distinguished by their color. Sequence deduced from the order in which four different dyes passed the detector.[9] This automation greatly increased throughput, reduced costs and enabled completion of large projects.[10]

Prober *et al*. developed an automated sequencing system based on four ddNTPs each carrying a different chemically tuned succinylfluorescein dye. The DNA fragments were resolved by PAGE in one sequencing lane and identified by a fluorescence detection system.[11] Primer labeling has since been substituted with color coding of terminating ddNTPs to allow all four synthesis reactions to be carried out in a single tube. Various commercial companies introduced automated DNA sequencers. Still however, the use of PAGE had its

limitations like preparation of gels which was labor intensive and tedious, use of toxic chemicals, problems of gel loading, thickness and electrophoresis related issues.

Capillary electrophoresis was introduced by Jorgensen in early 1980s as an alternative to gel electrophoresis.[12,13] These high purity fused silica capillaries hold a sieving medium that allows DNA fragments separation based on their size. A laser detector near the end of capillaries detects fluorescent signals emitted by incorporated labeled ddNTPs. Although capillaries reduce Joule heat to negligible levels and allow use of very high electric fields for rapid DNA fragment separation, it is the ease of automation that has resulted in its induction in sequencing instruments, in place of slab gel sequencers.[14] A single capillary instrument did not provide much of advantage over slab gel instrument as latter could run multiple samples, this led to the introduction of Capillary Array Electrophoresis (CAE).[15] Since it relied on a single laser scanning across multiple capillaries, in a 96 capillary instrument, each capillary was scanned 1% of the time and 99% DNA was lost without detection. Instruments were developed that could detect signals from all capillaries simultaneously by introduction of sheath flow cuvette. DNA sequencing fragments run as discrete streams from each capillary within sheath fluid. A laser beam focused into cuvette skims beneath capillary tips thus scanning all capillaries simultaneously. These CAE instruments have increased the sequencing throughput several folds than slab gel instruments.[14]

Next generation sequencing technologies: Sanger sequencing dominated the industry for about two decades and led to many accomplishments, including the completion of finished grade human genome sequence.[16] Although it was sufficient for majority of clinical applications, the level of sensitivity afforded could be insufficient for some clinically relevant low level mutant alleles, as the length of DNA that could be reliably sequenced was limited due to limited power of discrimination between fragment sizes during electrophoresis. Secondly, the analysis of highly polymorphic regions such as Major Histocompatibility Complex (MHC) which is the most polymorphic gene known,[17] could generate complex data due to multiple heterozygous positions in the sequence. Thirdly, DNA must be present in high concentration before sequencing. This showed a need for new and improved technologies for sequencing large number of human genomes as Sanger platform was not readily scalable to achieve a throughput capable of analyzing complex and diploid genomes, and that too at low cost. Since Sanger method was the “first generation” method, newer methods introduced were called the “Next Generation Sequencing” (NGS) methods. These methods have not yet been fully integrated into clinical diagnostics and are yet limited to research purposes only; besides, each of these methods require a separate review, so here we will discuss only the brief principle of these methods individually.

1. **Roche/454 life sciences:** In 2005, the first NGS platform was introduced (454 genome sequencer)[18] that utilizes pyrosequencing technology. The DNA fragments are ligated with flanking adapter sequences which are then used to immobilize these library fragments to complementary oligonucleotides on the surface of capture agarose beads, such that each bead is associated with a single fragment. These fragment-bead complexes are each isolated into oil water micelles containing Polymerase Chain Reaction (PCR) reactants and thermal cycling (called emulsion PCR) of micelles produces about one million copies of each DNA fragment on each bead. Then emulsions are broken with solvent and beads containing amplified product are enriched and annealed with sequencing primer. These beads are then arrayed onto a silica picotiter plate. Each of the several hundred thousand wells holds a single bead. The actual pyrosequencing step occurs when the addition of a dNTP during extension step liberates pyrophosphate. Pyrophosphate is converted to ATP through the action of sulfurylase. ATP is subsequently used to convert luciferin to oxyluciferin by luciferase. The light generated is measured by camera as each nucleotide solution is introduced in a step wise fashion, with an imaging step after each nucleotide incorporation step.[19,20]

dNTP incorporation → Inorganic pyrophosphate (PPi)

PPi + adenosine 5' phosphosulfate (APS) + sulfurylase → ATP

ATP + Luciferase → Light

2. Illumina genome analyzer: This system differs from 454 in that clonal amplification takes place *in situ* on the surface of flow cell, rather than in a separate emulsion PCR reaction. DNA fragments are first ligated to oligonucleotide adaptors that bind anchor nucleotides covalently linked to the surface of flow cell. The template DNA molecules are clonally amplified by bridge PCR. In this, DNA molecules can form a bridge with adjacent anchor oligonucleotide. This results in generation of several million individual clusters containing over one thousand copies of clonally amplified DNA molecules on the surface of the flow cell. Clusters are then denatured and a sequencing primer hybridized to the strand. During each sequencing cycle, clusters are exposed to DNA polymerase and a mixture of four nucleotides, each labeled with a unique fluorescent label. Here Illumina uses reversible terminator chemistry; the nucleotides are modified at 3' end with a cleavable terminator moiety. At the end of each cycle, the fluorescent signal is measured for each cluster and then both fluorescent label and 3' terminator moiety are removed allowing another cycle of nucleotide addition.[19,20]

3. Applied biosystems SOLiD sequencer: This approach is similar to 454 in that emulsion PCR is used to generate a clonally amplified, adapter attached DNA molecule bound to a bead, bead attached to surface of a glass slide in flow cell. A sequencing primer is attached to this DNA template. For sequencing, 8 mer labelled oligonucleotide probes of a total 16 possible dinucleotide combinations are used. The first two nucleotides represent combination while rest six are degenerate. DNA ligase is used to attach the probe to sequencing primer and fluorescence recorded. Probe is cleaved, seven cycles performed, newly synthesized strand denatured and new sequencing primer annealed to template DNA. The new primer is one nucleotide less than initial sequencing primer (n-1). The SOLiD instrument performs seven cycles of ligation, each from total of five different sequencing primers, thus resulting in read length of 35 bases.[19,20]

4. Complete genomics: Complete genomics is a California based life sciences company that is different from other DNA sequencing technologies that it provides only an outsourced model that provides its customers with finished variant reports, by combining proprietary genome sequencing technology with its data management software. Template DNA is cleaved by restriction endonucleases and DNA fragments ligated with adapter sequences to create DNA nanoballs. These nanoballs are packed very tightly together on a silicon chip, now called DNA nanoball array. Then it uses probe ligation sequencing chemistry similar to SOLiD sequencing.[21,22]

5. Helicos: It is unique from other commercially available sequencing platforms in that it does not require amplified DNA templates for sequencing. Adenosines attached to 3' end of template DNA allow annealing to poly T-anchor oligonucleotides covalently attached to flow cell surface. Initial adenosine fluorescence is cleaved and template exposed to polymerase and one of the four fluorescently labeled nucleotides. After each round, signal is measured by detection system.[19]

6. Pacific biosciences single molecule real time (SMRT) sequencing: It is relatively new technology that aims at determining sequence in real time as DNA polymerase synthesizes DNA from a template strand. The Zero Mode Waveguide (ZMW) design reduces the observation volume thus reducing stray fluorescent molecules that enter the detection layer.[23] The reaction occurs on a plate containing thousands of nanometer sized wells. Polymerase molecules are bound to well and optical system measures fluorescence emitted from bottom of the well. Wells are exposed to fluorescently labeled nucleotides which emit fluorescence when incorporated, within the detection volume of optical system. Fluorescence moiety moves out of detection volume and polymerase continues to next base incorporation.[19]

NGS and its clinical applications: The development of high throughput sequencing technologies has enabled researchers to widely increase research domain into microbes genome sequencing, finding genetic variants by targeting specific genomic regions, understanding human gene expression variations, characterizing the transcriptomes by RNA sequencing and profiling of various proteins and epigenetic markers.[24]

Cancer is not a single disease entity but arises through a multistage mechanism in which genetic alterations lead to changes in gene sequence, structure or copy number.[25] Therefore, identification of genetic events that result in cancer development will improve our understanding of tumors and lead to discovery of approaches for diagnosis, prognosis, prevention and treatment. The advances in NGS technologies have enabled researchers to sequence a large number of entire cancer genomes and thus characterize and study cancer genomes at genomic, transcriptomic and epigenetic levels. Several groups have performed comprehensive analysis of a variety of cancer genomes including acute myelogenous leukemia,[26,27] lung cancer[28] and melanoma.[29] These studies demonstrate enormous power of NGS in detecting DNA damage and mutations that underlie cancers. Aberrant mRNA expression is the hallmark of cancers, which if we recall central dogma of molecular biology, directly reflects deranged cellular processes. High throughput sequencing approaches have also been adopted in transcriptomes analysis, through generation of cDNA from RNA and then sequencing that.[30] Detection of circulating tumor cells, Minimal Residual Diseases (MRD), serves as a prognostic marker in several solid organ tumors.[31] qPCR is usually accurate clinical assay to detect MRD where nucleic acid target is similar for majority of patients. However, certain cancers may exhibit heterogeneous molecular defects for which each patient may require tailor made unique primers for detection. High throughput sequencing can improve MRD detection as each patient's genomic alterations are specifically characterized.

The protein coding region of all genes (Exome) account for 85% of DNA mutations that affect human diseases, despite the fact that they constitute very little (1 - 2%) of entire human genome. Mutations in regulatory sequences might not have been studied as much. Using NGS for selectively sequencing protein coding regions reduces sequencing cost.[32]

Although the relationship between DNA variation and diseases has long been a major focus of genetic research, the identification of specific genetic loci of multifactorial and multigenic diseases and their complex interplay poses a daunting task.[33,34] High density nucleotide arrays used in International HapMap Project for Single Nucleotide Polymorphism (SNP) detection

have been the major methodology but are limited by density of array.[35] A comprehensive database of sequence variants discovered using high throughput technologies has been compiled as a part of 1000 genomes project that will improve predictive power of genome wide association studies and build upon our understanding of complex disease trait loci.[36]

Epigenetic regulation of gene expression, the most understood mechanism of DNA methylation, is aberrant in a number of diseases especially cancer. The advent of pharmacological agents that can demethylate and thus reactivate repressed genes, have resurrected the interest in quantification of methylation status. NGS can not only describe genome wide methylation patterns but can also help in selecting patients for demethylation therapies and monitor their response.[37,38]

Culturing microorganisms is a laborious and time consuming technique. Microbiology is increasingly relying on molecular biology for detection of microbes.[39] Quantitative PCR (qPCR) of microbial nucleic acid requires prior knowledge/suspicion of organism, while sequencing offers an unbiased approach.

CONCLUSION

The technical ability to analyze thousands of genes by high throughput sequencing has far outpaced our skill to interpret the data in a clinically meaningful manner. There are numerous issues that currently impede introduction of high throughput sequencing in routine clinical diagnostics like its accuracy, genotype-phenotype correlation, clinical utility and ethical considerations. Such technical considerations are difficult to understand by clinicians with no laboratory experience, as for them, the laboratory must provide current and accurate interpretation of sequence variations. However, as geneticists accumulate sequence data from large number of individuals across different ethnic backgrounds and health issues, the ability to characterize a variation will continue to improve.

REFERENCES

1. Watson JD, Crick FH. Molecular structures of nucleic acids. *Nature* 1953; **71**:737.
2. Gaastra W. Chemical cleavage (Maxam and Gilbert) method for DNA sequence determination. *Methods Mol Biol* 1985; **2**:333-41.
3. Maxam AM, Gilbert W. A new method for sequencing DNA. *Proc Natl Acad Sci USA* 1977; **74**:560-4.
4. Schackelford RE. Maxam-Gilbert DNA sequencing. Chapter DNA sequencing [Internet]. 2011. Available from: http://www.pathologyoutlines.com/topic/molecularpathdnaseqmaxam.html
5. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain terminating inhibitors. *Proc Natl Acad Sci USA* 1977; **74**: 5463-7.
6. Albert B, Johnson A, Lewis J, editors. Molecular biology of the cell. 4th ed. New York: *Garland Science*; 2002.
7. Janitz M. Next generation genome sequencing: towards personalized medicine. Philadelphia: *Wiley;* 2008.
8. Sensen CW, editor. Essentials of genomics and bioinformatics. Philadelphia: *John Wiley & Sons;* 2002.
9. Smith LM, Sanders JZ, Kaiser RJ, Hughes P, Dodd C, Connell CR, *et al.* Fluorescence detection in automated DNA sequence analysis. *Nature* 1986; **321**:674-9.
10. Slatko BE, Kieleczawa J, Ju J, Gardner AF, Hendrickson CL, Ausubel FM. "First generation" automated DNA sequencing technology. *Curr Protoc Mol Biol* 2011; Chapter 7: Unit 7.2.
11. Prober JM, Trainor GL, Dam RJ, Hobbs FW, Robertson CW, Zagursky RJ, *et al.* A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science* 1987; **238**:336-41.
12. Jorgenson JW, Lukacs KD. Zone electrophoresis in open-tubular glass capillaries. *Anal Chem* 1981; **53**:1298-302.
13. Karger BL, Guttman A. DNA sequencing by capillary electrophoresis. *Electrophoresis* 2009; **30**:196-202.
14. Dovichi NJ, Zhang J. How capillary electrophoresis sequenced the human genome. *Angew Chem Int Ed* 2000; **39**:4463-8.
15. Zagursky RJ, McCormick RM. DNA sequencing sparations in capillary gels on a modified commercial DNA sequencing instrument. *Biotechniques* 1990; **9**:74-9.
16. International human genome sequencing consortium. Finishing the euchromatic sequence of the human genome. *Nature* 2004; **431**:931-45.
17. Tipu HN, Ahmed TA, Ahmed S. Validity of human leukocyte antigen class II typing in-house assay in comparison with commercial kit. *Pak J Pathol* 2008; **19**:121-5.
18. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, *et al.* Genome sequencing in microfabricated high density picoliter reactors. *Nature* 2005; **437**:376-80.
19. Anderson MW, Schrijver I. Next generation DNA sequencing and the future of genomic medicine. *Genes* 2010; **1**:38-69.
20. Mardis ER. Next generation DNA sequencing methods. *Annu Rev Genomics Hum Genet* 2008; **9**:387-402.
21. Drmanac R, Sparks AB, Callow MJ, Halpern AL, Burns NL, Kermani BG, *et al.* Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science* 2010; **327**:78-81.
22. Complete genomics analysis platform [Internet]. [cited 2013, Dec 31] Available from: http://www.completegenomics.com/technology/
23. Metzker ML. Sequencing technologies: the next generation. *Nature Rev* 2010; **11**:31-46.
24. Su Z, Ning B, Fang H, Hong H, Perkins R, Tong W, *et al.* Next generation sequencing and its applications in molecular diagnostics. *Expert Rev Mol Diagn* 2011; **11**:333-43.
25. Bell DW. Our changing view of the genomic landscape of cancer. *J Pathol* 2010; **220**:231-43.
26. Ley TJ, Mardis ER, Ding L, Fulton B, McLellan MD, Chen K, *et al.* DNA sequencing of a cytogenetically normal acute myelogenous leukemia genome. *Nature* 2008; **456**:66-72.
27. Mardis ER, Ding L, Dooling DJ, Larson DE, McLellan MD,

ChenK, *et al.* Recurring mutations found by sequencing an acute myelogenous leukemia genome. *N Engl J Med* 2009; **361**:1058-66.

28. Pleasance ED, Stephens PJ, O'Meara S, McBride DJ, Meynert A, Jones D, *et al.* A small cell lung cancer genome with complex signatures of tobacco exposure. *Nature* 2010; **463**: 184-90.
29. Pleasance ED, Cheetham RK, Stephens PJ, McBride DJ, Humphray SJ, Greenman CD, *et al.* A comprehensive catalogue of somatic mutations from a human cancer genome. *Nature* 2010; **463**:191-6.
30. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 2009; **10**:57-63.
31. Allan AL, Keeney M. Circulating tumour cell analysis: technical and statistical considerations for application to the clinic. *J Oncol* 2010; **426**:218.
32. Choi M, Scholl UI, Ji WZ, Liu T, Tikhonova IR, Zumbo P, *et al.* Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc Natl Acad Sci USA* 2009; **106**: 19096-101.
33. Manolio TA, Collins FS. The HapMap and the genome-wide association studies in diagnosis and therapy. *Annu Rev Med* 2009; **60**:443-56.
34. Tipu HN, Ahmed TA, Bashir MM. Human leukocyte antigen class II susceptibility conferring alleles among non-insulin dependent diabetes mellitus patients. *J Coll Physicians Surg Pak* 2011; **21**:26-9.
35. La Framboise T. Single nucleotide polymorphism arrays: a decade of biological, computational and technological advances. *Nucleic Acids Res* 2009; **37**:4181-93.
36. 1000 Genomes: a deep catalogue of human genetic variation [Internet]. [cited 2014 Jan 9]. Available from http://www.1000genomes.org/
37. Hegi ME, Diserens AC, Gorlia T, Hamou MF, de Tribolet N, Weller M, *et al.* MGMT gene silencing and benefit from temozolomide in glioblastoma. *N Engl J Med* 2005; **352**:997-1003.
38. Piekarz RL, Bates SE. Epigenetic modifiers: basic understanding and clinical development. *Clin Cancer Res* 2009; **15**: 3918-26.
39. Muldrew KL. Molecular diagnostics of infectious diseases. *Curr Opin Pediatr* 2009; **21**:102-11.