International Journal of Pharmacology

ISSN 1811-7775

science alert

ansinet
Asian Network for Scientific Information

OPEN ACCESS International Journal of Pharmacology

ISSN 1811-7775
DOI: 10.3923/ijp.2017.709.723

Review Article
Next-Generation Sequencing for Drug Designing and Development: An Omics Approach for Cancer Treatment

[1]Gautam Kumar, [2]Kamal Kishore Chaudhary, [1]Krishna Misra and [1]Anushree Tripathi

[1]Department of Applied Sciences, Computer Centre II, Indian Institute of Information Technology, 211015 Allahabad (IIIT-A), India
[2]Department of Biotechnology, Jaipur National University, Jaipur, 302017 Rajasthan, India

Abstract

Next Generation Sequencing (NGS) techniques retain the potential to open up avenue to identify novel therapeutic target guided drug to pursue and proving as an efficient alternative technology for denovo drug design. Its remarkable applications are established in relation to therapeutic diagnosis and treatment of multifactorial diseases including cancer. The NGS accurately predicts cellular reactions at genomic, transcriptomic or proteomic level. Its impact in novel anticancer drug designing relies on its therapeutic approaches, multiplexing of samples and high diagnostic sensitivity for genetic and epigenetic biomarkers. Development of new therapies and drug usually takes longer time and require recruiting considerable pools of patient. NGS cut down cost of drug development and time by using its unseen potential to identified specific therapeutic target and new pathophysiological pathways involvement in cancer, helped in designing of targeted drug and correct evaluation of its deleterious side effect. Furthermore, improved gene and disease interaction validations techniques changed entire cancer therapy methods. This is a critical review of multiple approaches of NGS in development of anti-cancer drug including biomarkers based diagnosis and recent trends of targeted capture technology based personalised medicine in cancer therapy. Additionally, emerging paradigm in disease diagnosis based on state-of-art third generation sequencing technologies have also discussed.

Key words: Next generation sequencing, therapeutic approach, biomarkers, personalised medicine, genome sequencing, cancer

Citation: Gautam Kumar, Kamal Kishore Chaudhary, Krishna Misra and Anushree Tripathi, 2017. Next-generation sequencing for drug designing and development: An omics approach for cancer treatment. Int. J. Pharmacol., 13: 709-723.

Corresponding Author: Gautam Kumar, Department of Applied Sciences, Computer Centre II, Indian Institute of Information Technology, 211015 Allahabad (IIIT-A), India Tel: +91 9889404804

Competing Interest: The authors have declared that no competing interest exists.

Data Availability: All relevant data are within the paper and its supporting information files.

INTRODUCTION

Next generation sequencing emerged as a promising approach for therapeutic diagnosis and treatment of complex pathophysiological conditions including cancer. At global level, cancer is considered as major health problem and now become leading cause of mortality. Cancer is a complex disease process accompanied by large numbers of correlated genetic alternations affecting deregulation of cell cycle[1]. According to World Health Organization (WHO) 8.8 million deaths were caused by cancer in 2015 along with a prediction of 12 million cancer death per year by 2030[2]. Breast, colorectal, lung and cervix are most frequently observed cancer in women whereas lung, prostate, colorectal, stomach and liver are main cancer types causing death in men[2]. In both developing as well as developed countries, lung and breast cancer are primary causes of deaths in males and females respectively[3-5]. Furthermore, liver and stomach cancer in males and cervical cancer in females were also significantly reduce survival rate of cancer patients as reported by National Cancer Institute-NIH[6].

In view of recent scenario in therapeutics, Next Generation Sequencing (NGS) technologies are recognised as a high-throughput sequencing and data analysis approach, essentially needed for therapeutic diagnosis and treatment of cancer. The NGS allows assessment of tumour genomic at individual gene or transcript level using massive parallel run and multiplexing of samples[7,8]. The longstanding challenges in clinical diagnosis and treatment of malignancies represent variability of treatment and gradual development of resistance to medication at sub-clones level[9]. The ability of NGS to work with comprehensive landscape from identification of genetic alternations, the main cause of multiple resistances makes it the right approach for designing the right drugs for cancer for right patients, with specific dose, at defined time interval.

History of sequencing methods: Advent of First Generation Sequencing (FGS) technology was started with the development of chemical method by Maxam-Gilbert[10] and dideoxy method by Sanger[11]. In the Maxam and Gilbert method, terminally labelled DNA fragments were cleaved chemically at adenine, guanine, cytosine or thymine and reaction products were separated by gel electrophoresis based on size of fragments to determine nucleotide sequence[12]. However, dideoxy method of Sanger determined DNA sequence through base specific termination of DNA synthesis using chain termination of 2′,3′-dideoxy and arabinonucleoside analogs[11,12]. Sanger method of DNA sequencing was adopted as primary sequencing technology among FGS for clinical as well as research[13]. The FGS was primarily based on radioactive or fluorescent materials and due to limited sequencing potential, gradual changes in applications of sequencing technologies in clinical medicines lead to dramatic change in sequencing scenario resulting in the development of Second Generation Sequencing (SGS) technology now referred as Next Generation Sequencing (NGS) technology[13].

NGS: A new vision for therapeutics: The significance of next generation sequencing technologies lies in the comparative analysis of clinical samples in a much faster and inexpensive way. The NGS methods are based on several techniques including, (1) Micro-chip based electrophoretic sequencing, (2) Sequencing by hybridization, (3) Real-time sequencing and (4) Cyclic-array sequencing[14]. The working mechanism is composed of certain basic steps that include genomic template preparation for downstream sequence analysis, generation of short sequence reads, alignment of reads on reference sequence, assembly of sequence from aligned reads[15]. Most NGS technologies utilize sequencing by synthetic[16] approaches and have been used by various commercial platforms that include Roche/454 life science, Illumina/Solexa, Applied Biosystem/SOLiD, Polonator and Helicos Biosciences and single molecule sequencing as tabulated in Table 1[17,18].

The NGS approach for designing targeted small-molecule cancer drugs capture the sense of excitement along with reducing the existing designing burden on researchers in cancer treatment. There have been many existing problems and challenges associated with traditional designing of drugs. These challenges represented as complexity of anti-cancer drug discovery process, low precision level of target identification, high cost of drug synthesis and clinical trials, limited knowledge of underlined molecular mechanism and lack of validated biomarkers for characterisation of tumour type[19]. Furthermore, the process of drug design and development (Fig. 1) has high level of failure rate at the stage of clinical trials. Another big challenge is to understand the way to overcome drug resistance that results relapse of tumour growth and considered as major hurdle in cancer therapy[20]. Significant progress has been marked by the development of anti-cancer drugs such as imatinib and erlotinib and leukemia patients who are not responding the imatinib, sequencing of patient genome profile used to recommend before changing the therapy that can be done by SGS technologies[21].

The NGS has wide range of applicability in the field of cancer research, specifically since cancers are mainly caused

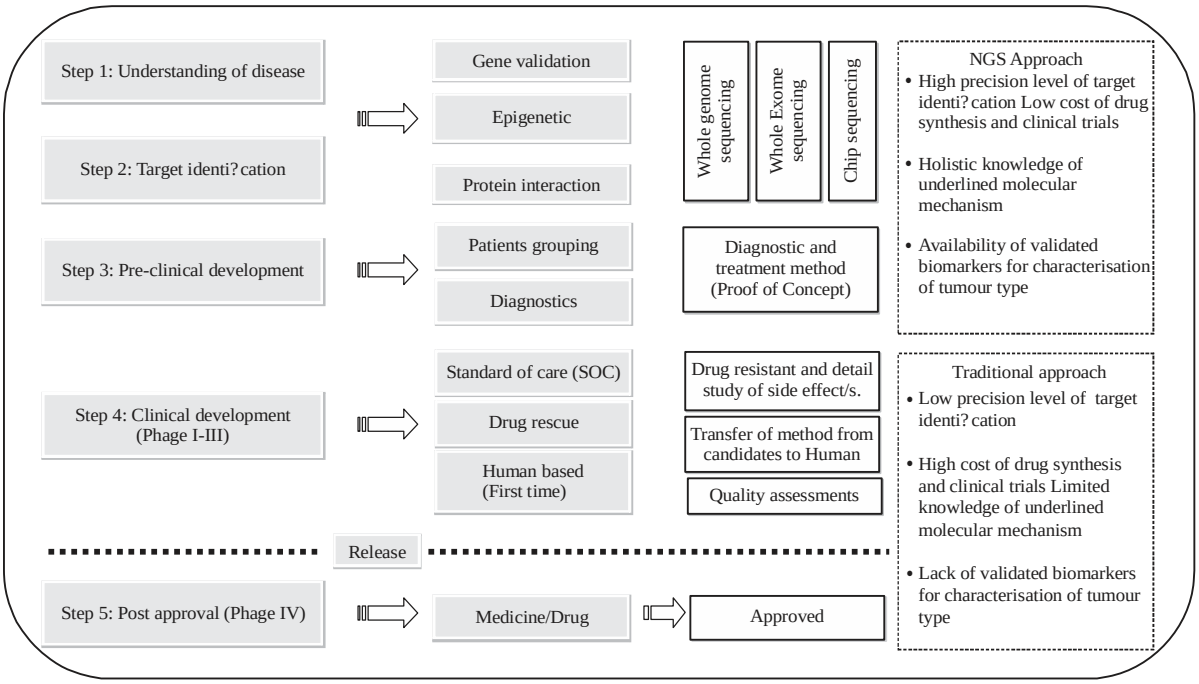

Fig. 1: Drug discovery process based on next generation sequencing

Table 1: Comparative overview of different NGS platforms

NGS platforms	Library preparation method	Sequencing technique	Run time	Read length	Base per run
Roche/454 life sciences	Emulsion PCR	Pyrosequencing	10 h	250 bp	500 Mb
Illumina/Solexa	Bridge PCR	Reversible termination	2.5 days	36 bp	18-35 Gb
Applied biosystem/SOLiD	Emulsion PCR	Ligation	6 days	35 bp	30-50 Gb
Helicos biosciences	Single molecule	Reversible terminator	8 days	37 gb	37Gb

by gene mutations. The NGS technology enables researcher for the identification of gene mutations, characterisation of tumour type, diagnosis of tumour progression by biomarker prediction. Finding of Ley *et al.*[22] reported whole-genome sequencing study on acute myeloid leukemic cells[22]. The impact of FLT3 gene mutation has been identified in acute myloid leukemic patients and solid tumour exomes of breast and colorectal cancer were first identified using NGS technology[23-25]. Multiple mutated genes were associated with malignancies depending on its location and type. In myeloma, BRAF, NRAS and KIT were observed as mutation causing genes. Such genes have targeted for reducing metastatic growth[26]. Four subtypes of breast cancer and their mutated genes have been discriminated based on exome analysis. The frequencies of TP53 and HER2 mutations were found to be highest in tumour[27]. Using NGS, breast cancer progression has been estimated by finding difference in read length of CAG repeats in terms of intra-tumour genetic heterogeneity of androgen receptor gene. It has been observed that shorter length of CAG repeats may have protective role against breast cancer[28]. The discovery of biomarkers like BRCA1, BRCA2, HER2, PR and ER are known to be extremely significant in molecular profiling, tumourigenicity and targeted drug designing[29]. The NGS plays potential role in finding mutation in heterogeneous population of cancer cells through biomarker detection[30].

Based on the analysis of genetic mutations, NGS technologies facilitate discovery of precision medicine in oncology. Broadly, there are three ways that confer NGS utility in cancer therapy. First, diagnosis of tumour type determined by mutations leads to genetic alternations. Second approach predicts targeted gene therapy against specific tumour type. Third strategy finds mutations that cause resistance to targeted therapy[31]. The NGS technologies integrate genomics, transcriptomic and epigenomic mutations in cancer biology as well as classify various types of cancers for early diagnosis and targeted therapy[32].

DEVELOPMENT OF NGS TECHNOLOGIES IN RELATION TO CANCER

NGS and pharmacogenomics: The NGS has wide spectra for drug development in connection to pharmacogenomics, deals with the study of association between genetic variation and

drug response for disease treatment[33]. It correlates drug efficacy and toxicity with genomic variation in drug targets that contributes in improving treatment response. In cancer, genetic variations of patient must be examined by considering both acquired (somatic) and germline (inherited) mutations due to their significance in drug efficacy. Analysis of somatic mutations plays peculiar role in enhancing treatment efficacy by defining genetic alterations during tumour development resulted in prediction of potential drug target such as mutations in TP53 and CYP19 are used to predict genetic constituent of breast cancer[34]. On the other hand, germline mutations find pharmacokinetic property of drug to understand treatment response to targeted therapy[35]. In oncology, there have been mammoth of challenging tasks, which increases the need of advanced NGS technologies and use of statistical analysis methods. The analysis of large amount of data generated from multiple samples of cancer patients and identification of rare genetic variants from those samples are major challenges in pharmacogenomics. The NGS technologies provide fast and robust approach to tackle such challenges[36]. Moreover, NGS determines molecular pathways associated with metastasis, finds polymorphisms in genes causing multidrug resistance and targets potential drugs against specific genetic mutation. Integration of genome-wide association study (GWAS) and NGS have great significance on survival rate of cancer patients with the advancement in cancer therapy[37].

NGS methods in cancer therapeutics: Early diagnosis of cancer development is now possible with discoveries of advanced NGS technologies. This leads to easier cancer genomic profiling and provides targeted therapy. For genomic profiling of cancer patients, formalin fixation and paraffin embedding are two commonly used pathological biopsy media[38]. The sensitivity and accuracy of profiling cancer genome is influenced by steps of genomic data generation protocol that includes pre-analytical methods (data collection, storage, extraction and manipulation), library preparation, sequencing and variant calling. Variations have observed in preparation issues of pre-analytical methods based on type of sample, selection of sequencing-based assays as illustrated in Table 2[39].

To understand small genetic alterations in cancer patients, whole genome sequencing, whole exome sequencing, targeted RNA panel, transcriptomie sequencing are used[30]. Whole Genome Sequencing (WGS) detects copy number variants with high resolution, regulatory regions like promoters and enhancers and determines intergenic regions[40]. This approach allows researchers to examine cancer genome for identification and categorization of novel mutations[41]. However, Whole Exome Sequencing (WES) is focused to cover 1-2% of entire genome[42,43]. The coverage of WES is up to 95% of exons much higher than WGS. The WES detects both somatic and germline mutations in cancer patients[43]. Transcriptome sequencing of cancer genome profiling has been carried out using mRNA expression analysis. In addition to gene expression, the analysis of DNA alterations makes the transcriptomic analysis method more effective[44]. The transcriptome sequencing has immense significance for non-coding RNA (i.e., miRNA, siRNA, piRNA and lncRNA) detection based biomarker development[44]. Transcriptome sequencing method assesses epigenome, proteome and metabolome in a broader way[44]. Targeted panel sequencing associated with precision oncology in which gene abnormalities have been identified in the panel of 20-500 genes build on amplicon based or hybridisation based techniques[45]. It has been used in detection of Single Nucleotide Variants (SNVs) and small insertions/deletions (Indels operation) for cancer therapeutics[46]. Targeted panel sequencing is quite useful in providing high depth and high exon coverage which are two critical factors for consistent variant calling. The depth of coverage defines repetition of a specific base in sequencing and alignment to a reference genome[45,46]. Apart from that, exon coverage also depicts spanning by atleast one sequencing read in terms of percentage[47,48]. The formalin-fixed paraffin-embedded tissue has been used in targeted panel sequencing method. Different aspects of NGS at epigenomic, transcriptomic and genomic level are depicted in Fig. 2 while comprehensive view of whole genome sequencing, whole exome sequencing, transcriptome analysis and targeted panel sequencing is shown in Table 3.

Table 2: Common preclinical and sequencing assays in cancer genome profiling

Sample type	Sequencing test
Formalin-fixed, paraffin-embedded	Whole genome sequencing
Fresh frozen tissue of bulk cells	Whole exome sequencing
Single cells	Large gene panel
Liquid biopsy	Small gene panel
	Hotspot panel
	Transcriptome
	Targeted RNA panel

Other methods of NGS are de-novo sequencing, non-coding RNA sequencing and epigenomics sequencing. De-novo sequencing deals with alignment of reads to generate sequencing of complete genome specifically when reference genome is unavailable[49]. In non-coding RNA sequencing, the regulation of differential gene expressions has been analysed as silencers or repressors[50]. Epigenomic approach includes methylation sequencing, ChIP sequencing

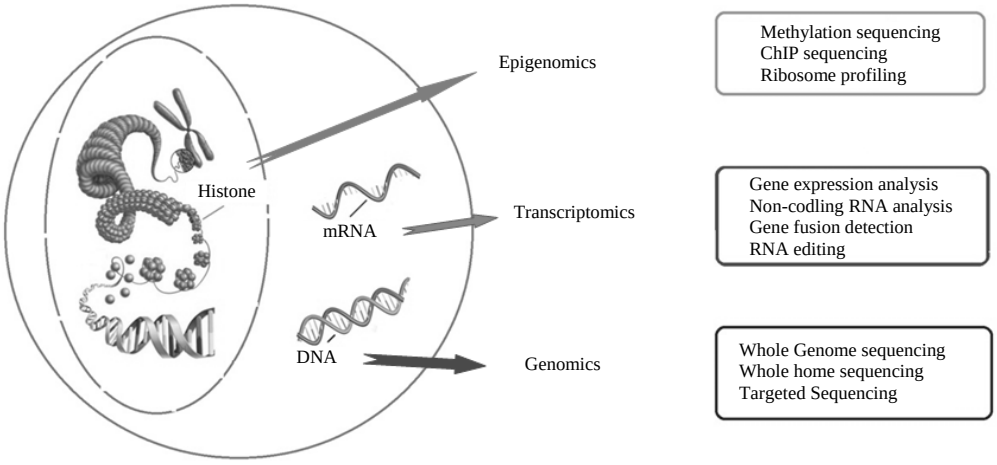

Fig. 2: Different approaches of Next generation sequencing technologies

Table 3: Comparative aspects of common NGS sequencing techniques

Sequencing assay	Target regions	Benefits	Drawback
Whole genome sequencing	Genes, Exons and Non-coding regions	• Comparative genome analysis • Highest resolution of genetic aberrations • Low cost • SNVs detection	• Expensive • Lack of specificity
Whole exome sequencing	Genes, Exons	• Prediction of CNVs • Detection of novel variants • Low cost	• Low CNVs resolution • Time consuming
Transcriptome sequencing	mRNAs	• Deep assessment of epigenome, proteome and metabolome	• Requirement of fresh tissue with genetic variability • Difficult in standardisation
Targeted panel sequencing	Fusion genes	• High depth coverage • Easy in implementation	• Limited detection of complex genetic alterations

CNV: Copy number variation, SNVs: Single nucleotide variants

and ribosome profiling[51,52]. The study of cytosine methylation profile in the region of hetrochromatin and promoters, suggests deeper insight into the regulation of gene expression patterns[53]. Epigenomic sequencing allows identification of methylation up to single nucleotide level and DNA fragmentation up to 100-150 bp followed by construction of standard libraries for NGS analysis[53]. Chromatin immune-precipitation (ChIP) sequencing enables diagnosis of any diseased state through the study of protein-DNA or protein-RNA interaction[54]. Ribosome profiling mainly focus on active mRNA fragments captured by ribosome during translation processes provides overall activity of cell at specific time scale and facilitates identification of active forms of proteins that modulate cellular processes[55].

APPROACHES USED BY NGS TECHNOLOGIES IN TUMOUR IDENTIFICATION

From the past two decades, next generation sequencing technologies emerged as a high-throughput technology to analyze biological information relatively at very low cost and provided novel platform for biological research. The NGS allowing the researchers to perform almost any type of analysis and identified potential therapeutic target at genomics, transcriptomics or proteomics level by considering genetic or epigenetic factors associate with disease.

The urgent need of sequencing for understanding holistic nature of complex disease leads to the foundation of Human Genome Project (HGP) primarily based on Sanger method of sequencing. However, extremely high cost of genome sequencing was remained as a major barrier to further implement it in the area of clinical and personalised medicines[56,57]. The introduction of next generation sequencing technology in the 2000s dramatically drop down the cost of genome sequencing by almost 50000 folds, as it was roughly $300 million estimated for generating the first initial draft of human genome sequence under Human Genome Project[58]. Figure 3 clearly shown fortunes of NGS techniques and downfall of cost over year. At present, one can get complete sequence of cell after expending approximately at $1000 or below[59].

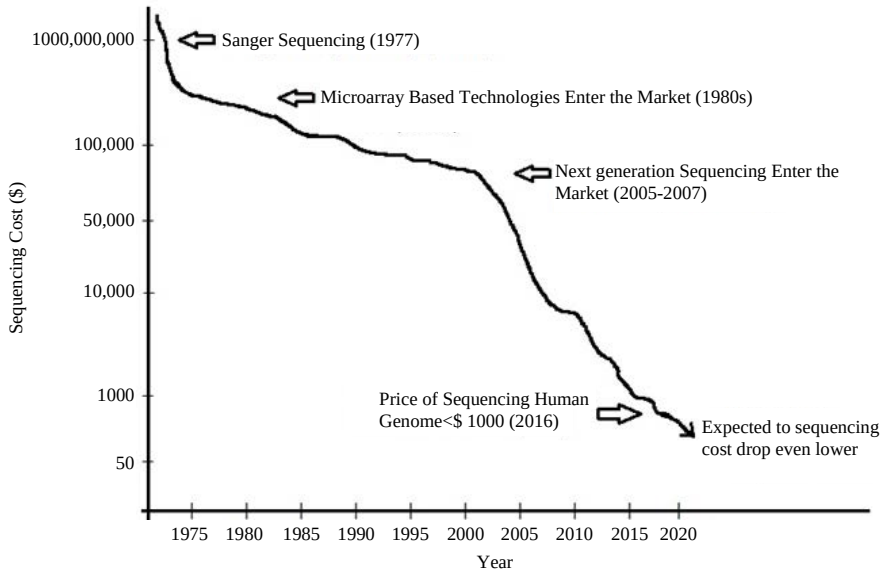

Fig. 3: Comparative fortune of NGS techniques and downfall of cost

In last one decade, the next generation sequencing has undergone several technological up-gradation and has evolved as a reliable platform for the current era of genomics. The first automated genome sequencing machine (AB370) was launched by Applied Biosystem in 1987[60]. AB370 was able to detect 96 bases simultaneously with 500k bases in a day considering maximum read length capacity of 600 bases as compared with the current AB3730 machine that can detect up to 2.88M bases per day[61].

Basic workflow of NGS technologies: Therapeutic applications recruited NGS guided approaches mostly based on sample type and diagnostic questions to be addressed for a specific disease in defined pathophysiological condition. NGS approaches can be grouped into 'DNA-seq', 'RNA-seq', 'ChIP-seq' and 'methyl-seq' analysis based on the sample type[62,63]. However, the diagnostic approach and overall treatment protocols may vary among different techniques. The present review highlights the base methodology (Fig. 4) of sequencing in reference to one of the most reliable Illumina MiSeq and Ion Torrent PGM machine[64].

Step 1: Sample/library preparation: Genomic materials isolated from diseased tissue fragmented either enzymatically or mechanically up to required fragment size, inputted to sequencer. However, sample amplification may be preferred for 4-10 cycles using PCR in most of the cases and based on the initial amount of genomic material[65]. The genomic information produced in form of 'reads', stored using .SRA, .FASTA or .FASTAQ file formats.

Step 2: Quality check: Performed to remove any bad quality reads, with quality score less than standard cut-off defined by scale of Phred quality score[66]. Phred quality score (Q-score) used to measure the base calling accuracy of sequencers and indicates the incorporation of incorrect bases. Q score, Q_{score} is measured as negative log of base calling error probabilities, P_{bcp} (Eq. 1)[67]:

$$Q_{score} = -10 \log_{10} \times P_{bcp} \quad (1)$$

Equation 1 explained probability of incorrect base insertion 1 in 1000 run of sequencers. Phred quality score for base is 30 represents 99.9% base accuracy[66]. Table 4 list base calling error probabilities, P_{bcp} with Phred score, Q_{score}.

Step 3: Mapping/assembly of reads: This is one of the crucial steps of NGS technologies marked by either finding of overlapping zone among all reads of DNA-seq or RNA-seq data followed by construction of contigs or mapping of reads against known reference genome. Overall performance has been measured by calculating length (maximum, average, combined) and N50[68]. N50 best represented as set of largest

Table 4: Quality check score for sequencer: Base calling error probabilities, with Phred score, Q_{score}

Probability of incorrect base call	Q_{score}	P_{bcp} (%)
1 in 10	10	90.00
1 in 1,000	30	99.90
1 in 10,000	40	99.99
1 in 100,000	50	99.999

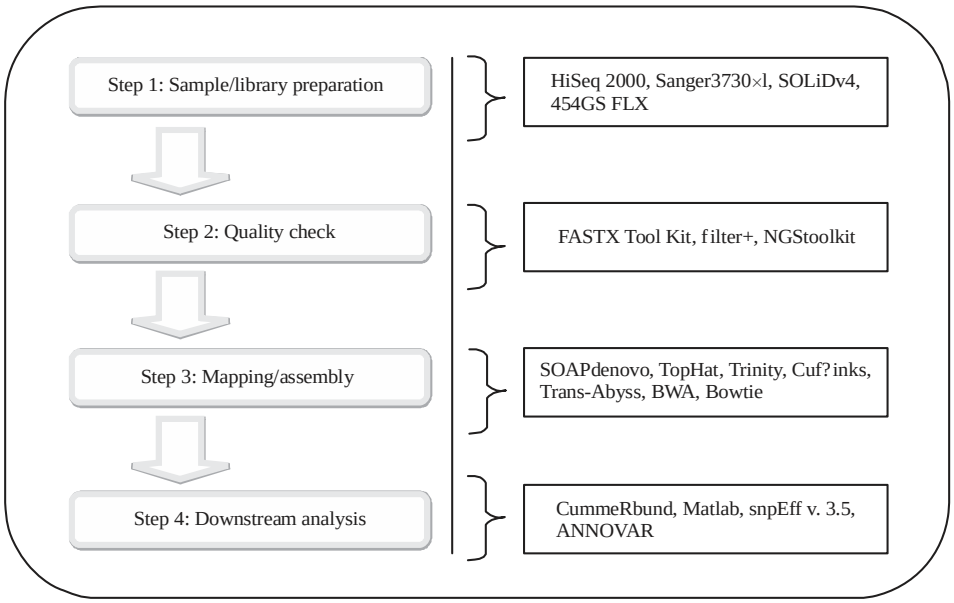

Fig. 4: Underline steps of next generation sequencing technology with associated software

contigs whose sum of length is more than 50% length of the assembly[69]. However, some of the researchers mentioned that assembly accuracy is difficult to measure[69].

Step 4: Downstream analysis: Overall flavour of NGS technologies are to find out pattern of novel gene expression from cell type in particular therapeutic condition. DNA-seq mostly used to find out homozygous and heterozygous single nucleotide polymorphism (SNPs) and mutations, locus of insertion and deletion (InDels), structural variants, and, copy number variations (CNVs)[59,70,71]. While, RNA-seq used for finding splicing variant regions, differential expression of genes, gene regulatory networks, signaling pathways and networks[70,72]. Furthermore, several standard platforms developed for downstream analysis includes SAMtools[73], GATK[74] and cummeRbund[75] which have been specially designed for cuffdiff output for RNA-seq data.

In cancer studies, most of tumours are resulted due to disturbance of genetic state of cell. This disturbance may be of genetic or epigenetic in nature. The variance between tumour and normal cell lines well demonstrated with high accuracy and sensitivity by using GATK, SAMtools, mpileup and Isaac variant caller[76].

NGS AND TARGET CAPTURE TECHNOLOGY

Next generation sequencing technology utilized several approaches for parallel sequencing of high-throughput data but most of them contribute in 'Research Use Only' (RUO)[77]. Application of NGS in therapeutic diagnosis needs consistent accuracy and high performance as per the guidelines of state agencies regulating the clinical laboratory[78]. Several potential biomarkers have been identified at genomic or transcriptomic level by researchers for clinical diagnosis of cancer (Table 5) and further validated through molecular diagnosis, based on the applications of NGS technologies.

Traditional methods of sequencing required considerable amount of DNA or RNA for clinical identification of therapeutic target. However, NGS technologies need comparatively small amount of genomic material for screening of several genes simultaneously[70]. Sequencing of known list of biomarkers for clinical screening provides alternative and efficient approach for routine analysis of samples of therapeutic importance. Several commercial gene panels or kits are available, designed to screen list of genes of particular therapeutic interest. The content of NGS panel is an important consideration for screening list of genes to be targeted and sequenced[79].

Target Capture Technologies (TCT) has been designed to take advantage of known biomarkers and target for small potential set of therapeutic genes[80]. The TCT mostly depends on two factors, first is sample type and second is quality and quantity of DNA or RNA[81]. Table 6 represents the list of important target capture technologies, underlying principle and amount of nucleic acid obligatory for analysis.

NGS: DRUG DESIGNING TO CANCER THERAPY

In the last five decades, the approvals of new drugs by competing authorities have almost remains constant while the cost of clinical diagnosis to treatment of diseases, has

Table 5: Globally identified selected biomarkers for tumour as per the guidelines of 'The American Society of Clinical Oncology' (ASCO)

Biomarkers	Tumour Type	Diagnostic material
ALK gene	Non-small cell lung cancer and anaplastic large cell lymphoma	Tumour
Alpha-fetoprotein (AFP)	Liver cancer and germ cell tumours	Blood
Beta-human chorionic gonadotropin (Beta-hCG)	Choriocarcinoma and germ cell tumours	Urine or blood
BRCA1 and *BRCA2*	Ovarian cancer	Blood
BCR-ABL fusion gene	Chronic myeloid leukemia, acute lymphoblastic leukemia, and acute myelogenous leukemia	Blood and/or bone marrow
C-kit/CD117	Gastrointestinal stromal tumour and mucosal melanoma	Tumour
CA-125	Ovarian cancer	Blood
CD20	Non-Hodgkin lymphoma	Blood
Cytokeratin fragment 21-1	Lung cancer	Blood
Estrogen receptor (ER)/progesterone receptor (PR)	Breast cancer	Tumour
HE4	Ovarian cancer	Blood
KRAS	Colorectal cancer and non-small cell lung cancer	Tumour
Prostate-specific antigen (PSA)	Prostate cancer	Blood
Nuclear matrix protein 22	Bladder cancer	Urine
Lactate dehydrogenase	Germ cell tumours, lymphoma, leukemia, melanoma, and neuroblastoma	Blood

Source: https://www.cancer.gov/, NIH: National Cancer Institute

Table 6: Target capture technologies, area of enrichment application and amount of nucleic acid required for analysis

Target capture technologies	Underlying method	Amount of DNA required
TruSeq (Illumina)	DNA probe-based	1 µg to 250 ng*
AmpliSeq (Life Technologies)	PCR-based	10 ng
Haloplex (Agilent Technologies)	DNA fragmentation followed by probe-based target enrichment	200 ng
GeneRead (Qiagen)	PCR-based	40 ng
SureSelect (Agilent Technologies)	Hybridization and capture using cRNA-baits	200 ng – 3 µg

*Depending on the version of kit-TruSeq Amplicon (recent version v1.5), illumina, North America

Table 7: List of approved drugs in relation to cancer therapy

Approved drug	Therapeutic area	Therapeutic target
Arsenic Trioxide	Oncology	PML/RARα
Brentuximab Vedotin	Oncology	CD30
Busulfan	Oncology	Ph Chromosome
Capecitabine	Oncology	DPD
Cetuximab (1)	Oncology	EGFR
Cetuximab (2)	Oncology	KRAS
Cisplatin	Oncology	TPMT
Dasatinib	Oncology	Ph Chromosome
Denileukin Diftitox	Oncology	CD25
Erlotinib	Oncology	EGFR
Crizotinib	Oncology	ALK
Everolimus	Oncology	Her2/neu
Exemestane	Oncology	ER &/PgR receptor
Fulvestrant	Oncology	ER receptor
Imatinib (1)	Oncology	C-Kit
Imatinib (2)	Oncology	Ph Chromosome
Imatinib (3)	Oncology	PDGFR
Imatinib (4)	Oncology	FIP1L1-PDGFRα
Irinotecan	Oncology	UGT1A1
Lapatinib	Oncology	Her2/neu
Mercaptopurine	Oncology	TPMT
Letrozole	Oncology	ER &/PgR receptor
Nilotinib (1)	Oncology	Ph Chromosome
Nilotinib (2)	Oncology	UGT1A1
Panitumumab (1)	Oncology	EGFR
Panitumumab (2)	Oncology	KRAS
Pertuzumab	Oncology	Her2/neu
Rasburicase	Oncology	G6PD
Tamoxifen	Oncology	ER receptor
Thioguanine	Oncology	TPMT
Tositumomab	Oncology	CD20 antigen
Trastuzumab	Oncology	Her2/neu
Vemurafenib	Oncology	BRAF

Source: http://www.fda.gov

increased to three times measuring in the yearly scale of 1990[82]. The major cause of this bottleneck can be grouped into technological and administrative barriers. Technological barrier represented as fall back of developmental efficiency of new drug and limitation of production models while administrative barriers have been adversely affective the overall process including strict regulatory and experimental framework, continuously rising cost of scientific query and complexities of patent process and its effect after expirations[83].

To overcome all barriers in the process of drug development and designing of new models for cancer therapy, both biotechnology and pharmaceutical based interdisciplinary companies, have started utilizing the concept of comparatively more efficient NGS guided genome sequencing based approaches to overcome the technological and administrative bottlenecks[84]. Researchers has been utilizing the concept of sequencing based approach and remarkably several new discovered drugs have been approved for further clinical diagnosis and therapeutic uses[84]. Recently, more than 100 drugs of pharmaceutical importance in distant therapeutic area have been listed by US Food and Drug Administration (FDA) (http://www.fda.gov). Table 7 show the advancement in cancer therapy and development of anticancer drugs.

NGS guided approach has been provided promising platform for pharmaceutical industries but still the success rate to launched new drug has remained constant and

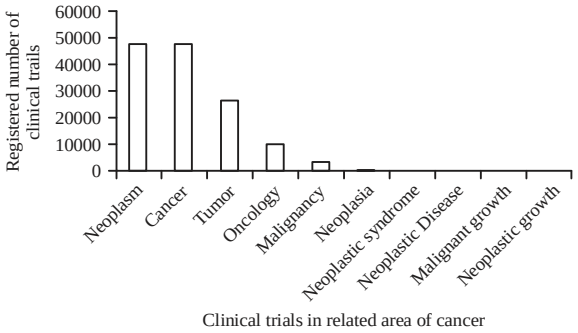

Fig. 5: Number of clinical trials registered for cancer according to Food and Drug Administration Amendments Act-2007

Source: https://clinicaltrials.gov/ct2/results/details?term=cancer

surprisingly, it has been observed that only one out of ten drugs qualified the preclinical testing criteria[85,86]. In the year of 2010, Ginsburg *et al*. noted that 45 different drugs failed in phase III of clinical diagnosis and testing stage, causing the loss of huge investment against average recorded cost of about $100 million per drug in phase III[87].

CHANGING PARADIGM: NGS GUIDED PERSONALISED ONCOGENIC TREATMENT

Over last two decades, there has been many fold increase in application of NGS technologies in diagnosis of potential genes alterations, splicing sites and epigenetic guided alternations considering multifactorial nature of disease[88]. These technologies are primarily used to identify the causing factors of a specific therapeutic condition that leads to the discovery of new diagnostic techniques for designing of novel drugs. Total 60046 hits were observed (Fig. 5) against searching keyword 'cancer' over the database of ClinicalTrials which was launched to implement 'Food and Drug Administration Amendments Act' of 2007 (FDAAA) clearly highlights the use of NGS in research. Large number of comprehensive cancer projects have been launched globally such as The Cancer Genome Atlas (TCGA)[89] (https://cancergenome.nih.gov/), apply genome analysis techniques for diagnosis of molecular basis of cancer along with comprehensive identification of co-related changes of cancer sub-types[88], International Cancer Genome Consortium (ICGC)[90] (http://icgc.org/) working on 50 different type of cancer of clinical and social importance and Cancer Genome Project (CGP)[91] (http://www.sanger.ac.uk/genetics/CGP/) involves in diagnosis of novel set of genes in cancer development.

Researchers have discovered plethora of diagnostic techniques for diagnosis of novel and rare mutations associated with a particular therapeutic condition. The NGS has been successfully recruited in most of the diagnostic purposes, including common cancer like lung cancer, prostate cancer and breast cancer[48].

To date, number of oncogenic biomarkers (Table 5) have been identified and characterised in application of clinical practices as per the clinical practice guideline issued by 'The American Society of Clinical Oncology' (ASCO)[92]. Current interest of researchers hasmoved in the area of personalised diagnosis and treatment of cancer. Next generation sequencing revolutionized an era of genome sequencing and medical science provided, comparatively much faster tools for screening of biomarkers, sequencing of whole genome or targeted gene of interest, reliable identification of genetic interaction, which were unidentified with traditional cytogenetic techniques[74].

Recent studies have clearly witnessed the efficient transformation of therapeutic workflows of NGS technologies into comparatively very high accuracy using WGS, WES or Epigenetic based analysis with broader clinical impact in personalised medicine. Walsh *et al*.[93] demonstrated the clinical diagnostic capability of NGS to target germ line mutations from patient suffering from primary level of carcinoma[93]. In other experiment, Holbrook *et al*. (year) identified the genes signature (AURKA, EGFR, FGFR2, KRAS, NET and PIK3CA) in gastric cancer based on Illumina sequencing technology and successfully used this technique on personalised treatment[94]. Now it is possible to have comprehensive disease history of patient for personalised treatment of therapeutic regimens based on sequencing of normal and tumour genomes. However, in spite of tremendous utility in cancer therapeutics, NGS based technologies have number of challenges that need to be addressed for further clinical investigation and its implementation for personalised medicines and treatment of cancer.

CHALLENGES OF NGS AND FUTURE TECHNOLOGY FOR MALIGNANCY

Next generation technologies revolutionised diagnosis processes of complex diseases including cancer and established Omics approach for understanding of disease complexities and further taking holistic approach of treatment. Cancer research based on NGS has witnessed exceptionally fast, more reliable approaches within the accepted guidelines of clinical as well as molecular diagnosis

of malignancies. Recent achievements of NGS platform can be best represented in diagnosis and identified therapeutic targets for breast cancer[95-97], ovarian cancer[98], lung cancer[99], melanoma[100], colorectal cancer[97] and head and neck cancer[101]. Although, NGS personified plethora of information for diagnosis and treatment of malignancy, still there have been many computational and clinical challenges remains that need to be addressed.

Computational barrier can be best described in the use of statistical approaches to identify set of differentially expressed genes, estimate overall accuracy of mapping and measure alignment of reads, derived from especially repetitive segments of DNA. Current algorithms of NGS technologies are difficult to explain for differential identification of genes, new isoforms or rare splicing junctions[102]. One of our research works in connection to identify potential set of probes guided subset of genes considering high throughput screening of microarray replicates for breast cancer explained the advantages and disadvantage of statistical approaches[103]. The work was considered three methods, fisher discriminate ratio[104], vector norm[105] and t-test (paired)[106] used to revalidate the potential set of gene. Furthermore, there is a need of high performance computational architecture to stores, analyses and interprets the NGS data[107]. Additionally, Clinical challenges are mainly concerned with design of novel anti-cancer drugs and applications of personalised medicines utilising concept of targeted therapy. However, sometimes biomarkers based therapy selection process made to be very difficult, if tumour spreading resulted into sub-clones and every sub-clone identified by separate genetic biomarkers[108].

The NGS based outcome are generally difficult to explain particularly in connection to cancer research. To overcome the heterogeneous complexity of inter-regulated tumour biomarkers, Third Generation Sequencing (TGS) technologies appears to lead the entire clinical diagnosis and treatment processes. The most significant aspect of TGS lies in the capability of interpreting even a single altered nucleotide from small genomic sequence instead of working with considerable amount of patients data[65]. Nanopore guided sequencing techniques emerged as promising candidate among all TGS techniques and its first version (MinION sequencer) has already been distributed among selected group of researchers for diagnostic and testing purposes in 2014[109]. Single Molecule Real Time (SMRT) Sequencing is one of the most recognised TGS techniques introduced in 2009 and its main advantage includes production of long-reads over the other sequencing techniques[110]. Researchers have been trying to address this problem by introducing the concept of 'Hybrid technologies' by combining short and long-read sequencing technology that may work as milestone in cancer research[110.] Single Molecule DNA Sequencing (SMDS) and Direct RNA Sequencing (DRS), is the first commercial available third generation sequencers allows up to 50 parallel run with capacity of almost 30×10^6 reads in each run[111,112]. Third generation sequencing technology hitherto catalysed the genome assemblies and sequencing speed many folds, still there have been many issues that need improvement particularly in specificity of statistical approaches, accuracy in output error rate and efficiency of result interpretation .

CONCLUSION

The advancement in sequencing technology has exceedingly supported the therapeutic diagnosis and treatment of human diseases. Omic approach of Next Generation Sequencing (NGS) provides much faster, cost effective and comparatively more realistic way to decode life mysteries and improve life qualities. Precise analysis of genomic data has opened fascinating opportunities in medical sciences including designing of novel drugs, identification of therapeutic targets, use of targeted capture technologies to speed up overall disease screening processes and reliable use of personalised medicine. Interesting to mention that, NGS shifted the bottleneck of diagnosis process from sequencing of large-reads of genomic or transcriptomic sequence to computational management of clinical data.

Applications of NGS to deal with inter-tumour and intra-tumour heterogeneity turned up into an idea that each tumour is unique and showing distinct mutation profiles even derived within a single tumour. Additionally, clinical method to identify malignancies (biopsy) only covers about 55% of tumour mutation profiles. Therefore, treatment of disease was atomistic in nature. NGS technology has provided more robust and sensitive platform followed by diagnosis of specific tumour type and taking holistic approach of treatment.

In the last decade, pathway based analysis of genetic variations in human cancers has been centre of attraction. Now-a-days, NGS based applications are integrated with Genome-Wide Association Study (GWAS) to detect any systematic changes and co-regulated heterogeneity in genome expression profile of tumour. Apart from addressing the issues of tumour heterogeneity and pathway based analysis, target based treatment has been continuously emerging as reported by several clinical trials. It is now, believed that rapid expanding of NGS based clinical trials and identified biomarkers leading to further advancement in drug efficacy.

In recent years, technological advancements of next generation sequencing have revolutionized whole medical

sciences and clinical approaches dealing with human cancers. Short read-alignment processes have been shifting towards large-read mapping and subsequently effective library preparation using SMRT and Nanopore based third generation sequencing techniques have further improved the disease classification and early diagnosis processes. It is hoped that in the next decade, the muti-targeted and ultra deep sequencing techniques shall become more effective and way of diagnosis process and treatments will be improved further specifically with respect to cancers.

SIGNIFICANCE STATEMENT

Next Generation Sequencing (NGS) technologies providing a new platform for efficient identification of cancer biomarkers guided therapeutic targets, improved and reliable approaches for designing of novel anticancer drugs and time effective application of cancer therapy. Several NGS based omics approaches have been recruited for successful identification of cancer biomarkers and designing of personalized medicines even if, patients stop responding in conventional therapy. The NGS technologies have number of potential significances over traditional approach including remarkable changes from morphological to genetical identification of tumor type, targeted based drug designing and comparatively real and holistic view of overall treatment process.

ACKNOWLEDGMENTS

All the authors greatly acknowledge the logistic support of their respective Universities and Institutes.

REFERENCES

1. Hindorff, L.A., E.M. Gillanders and T.A. Manolio, 2011. Genetic architecture of cancer and other complex diseases: Lessons learned and future directions. Carcinogenesis, 32: 945-954.
2. WHO., 2017. Cancer. World Health Organization (WHO), Fact Sheet, February 2017. http://www.who.int/mediacentre/factsheets/fs297/en/
3. WHO., 2013. Global cancer burden rises to 14.1 million new cases in 2012: Marked increase in breast cancers must be addressed. World Health Organization (WHO), The International Agency for Research on Cancer, December 12, 2013.
4. Reis-Filho, J.S., B. Weigelt, D. Fumagalli and C. Sotiriou, 2010. Molecular profiling: Moving away from tumor philately. Sci. Trans. Med., Vol. 2. 10.1126/scitranslmed.3001329.
5. Colombo, P.E., F. Milanezi, B. Weigelt and J.S. Reis-Filho, 2011. Microarrays in the 2010s: The contribution of microarray-based gene expression profiling to breast cancer classification, prognostication and prediction. Breast Cancer Res., Vol. 13. 10.1186/bcr2890.
6. Torre, L.A., F. Bray, R.L. Siegel, J. Ferlay, J. Lortet-Tieulent and A. Jemal, 2015. Global cancer statistics, 2012. CA: Cancer J. Clin., 65: 87-108.
7. Liu, L., Y. Li, S. Li, N. Hu and Y. He *et al*., 2012. Comparison of next-generation sequencing systems. J. Biomed. Biotechnol., 10.1155/2012/251364.
8. Lighten, J., C. van Oosterhout and P. Bentzen, 2014. Critical review of NGS analyses for *de novo* genotyping multigene families. Mol. Ecol., 23: 3957-3972.
9. Voelkerding, K.V., S.A. Dames and J.D. Durtschi, 2009. Next-generation sequencing: From basic research to diagnostics. Clin. Chem., 55: 641-658.
10. Ansorge, W.J., 2009. Next-generation DNA sequencing techniques. New Biotechnol., 25: 195-203.
11. Strausberg, R.L., S. Levy and Y.H. Rogers, 2008. Emerging DNA sequencing technologies for human genomic medicine. Drug Discov. Today, 13: 569-577.
12. Goodman, M.F. and L.J. Reha-Krantz, 1996. Chain-terminating nucleotides for DNA sequencing methods. U.S. Patent 5547859A. https://www.google.com/patents/US5547859.
13. Schuster, S.C., 2008. Next-generation sequencing transforms today's biology. Nat. Methods, 5: 16-18.
14. Pavlopoulos, G.A., A. Oulas, E. Iacucci, A. Sifrim and Y. Moreau *et al*., 2013. Unraveling genomic variation from next generation sequencing data. BioData Mining, Vol. 6. 10.1186/1756-0381-6-13.
15. Dewey, F.E., S. Pan, M.T. Wheeler, S.R. Quake and E.A. Ashley, 2012. DNA sequencing: Clinical applications of new DNA sequencing technologies. Circulation, 125: 931-944.
16. Reuter, J.A., D.V. Spacek and M.P. Snyder, 2015. High-throughput sequencing technologies. Mol. Cell, 58: 586-597.
17. Behjati, S. and P.S. Tarpey, 2013. What is next generation sequencing? Arch. Dis. Childhood-Educ. Practice Edn., 98: 236-238.
18. Shendure, J. and H. Ji, 2008. Next-generation DNA sequencing. Nat. Biotechnol., 26: 1135-1145.
19. Hoelder, S., P.A. Clarke and P. Workman, 2012. Discovery of small molecule cancer drugs: Successes, challenges and opportunities. Mol. Oncol., 6: 155-176.
20. Yadav, N.K., P. Shukla, A. Omer, S. Pareek and R.K. Singh, 2014. Next generation sequencing: Potential and application in drug discovery. Sci. World J., Vol. 2014. 10.1155/2014/802437
21. Cagle, P.T. and L.R. Chirieac, 2012. Advances in treatment of lung cancer with targeted therapy. Arch. Pathol. Lab. Med., 136: 504-509.

22. Ley, T.J., E.R. Mardis, L. Ding, B. Fulton and M.D. McLellan *et al*., 2008. DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. Nature, 456: 66-72.
23. Ley, T.J., P.J. Minx, M.J. Walter, R.E. Ries and H. Sun *et al*., 2003. A pilot study of high-throughput, sequence-based mutational profiling of primary human acute myeloid leukemia cell genomes. Proc. Nat. Acad. Sci. USA., 100: 14275-14280.
24. Mardis, E.R., L. Ding, D.J. Dooling, D.E. Larson and M.D. McLellan *et al*., 2009. Recurring mutations found by sequencing an acute myeloid leukemia genome. New Engl. J. Med., 361: 1058-1066.
25. Shen, T., S.H. Pajaro-Van de Stadt, N.C. Yeat and J.C.H. Lin, 2015. Clinical applications of next generation sequencing in cancer: From panels, to exomes, to genomes. Frontiers Genet., Vol. 6. 10.3389/fgene.2015.00215
26. Kunz, M., M. Dannemann and J. Kelso, 2013. High-throughput sequencing of the melanoma genome. Exp. Dermatol., 22: 10-17.
27. Koboldt, D.C., R.S. Fulton, M.D. McLellan, H. Schmidt and J. Kalicki-Veizer *et al*., 2012. Comprehensive molecular portraits of human breast tumours. Nature, 490: 61-70.
28. Gottlieb, B., C. Alvarado, C. Wang, B. Gharizadeh and F. Babrzadeh *et al*., 2013. Making sense of intratumor genetic heterogeneity: Altered frequency of androgen receptor CAG repeat length variants in breast cancer tissues. Hum. Mutation, 34: 610-618.
29. Desmedt, C., T. Voet, C. Sotiriou and P.J. Campbell, 2012. Next-generation sequencing in breast cancer: First take home messages. Curr. Opin. Oncol., 24: 597-604.
30. Grumbt, B., S.H. Eck, T. Hinrichsen and K. Hirv, 2013. Diagnostic applications of next generation sequencing in immunogenetics and molecular oncology. Trans. Med. Hemother., 40: 196-206.
31. Gagan, J. and E.M. Van Allen, 2015. Next-generation sequencing to guide cancer therapy. Genome Med., Vol. 7. 10.1186/s13073-015-0203-x.
32. LeBlanc, V.G. and M.A. Marra, 2015. Next-generation sequencing approaches in cancer: Where have they brought us and where will they take us? Cancers, 7: 1925-1958.
33. March, R., 2000. Pharmacogenomics: The genomics of drug response. Yeast, 17: 16-21.
34. Ayoub, N., C. Lucas and A. Kaddoumi, 2011. Genomics and pharmacogenomics of breast cancer: Current knowledge and trends. Asian Pac. J. Cancer Prev., 12: 1127-1140.
35. Filipski, K.K., L.E. Mechanic, R. Long and A.N. Freedman, 2014. Pharmacogenomics in oncology care. Frontiers Genet., Vol. 5. 10.3389/fgene.2014.00073.
36. Wheeler, H.E., M.L. Maitland, M.E. Dolan, N.J. Cox and M.J. Ratain, 2013. Cancer pharmacogenomics: Strategies and challenges. Nat. Rev. Genet., 14: 23-34.
37. Lu, D.Y., T.R. Lu, B. Xu and J. Ding, 2015. Pharmacogenetics of cancer therapy: Breakthroughs from beyond? Future Sci. OA, Vol. 1. 10.4155/fso.15.80
38. Carrick, D.M., M.G. Mehaffey, M.C. Sachs, S. Altekruse and C. Camalier *et al*., 2015. Robustness of next generation sequencing on older formalin-fixed paraffin-embedded tissue. PloS One, Vol. 10. 10.1371/journal.pone.0127353.
39. Lennon, N.J., V.A. Adalsteinsson and S.B. Gabriel, 2016. Technological considerations for genome-guided diagnosis and management of cancer. Genome Med., Vol. 8. 10.1186/s13073-016-0370-4
40. Ding, L., T.J. Ley, D.E. Larson, C.A. Miller and D.C. Koboldt *et al*., 2012. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. Nature, 481: 506-510.
41. Puente, X.S., M. Pinyol, V. Quesada, L. Conde and G.R. Ordonez *et al*., 2011. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. Nature, 475: 101-105.
42. Bamshad, M.J., S.B. Ng, A.W. Bigham, H.K. Tabor, M.J. Emond, D.A. Nickerson and J. Shendure, 2011. Exome sequencing as a tool for Mendelian disease gene discovery. Nat. Rev. Genet., 12: 745-755.
43. Choi, M., U.I. Scholl, W. Ji, T. Liu and I.R. Tikhonova *et al*., 2009. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. Proc. Nat. Acad. Sci. USA., 106: 19096-19101.
44. Martin, J.A. and Z. Wang, 2011. Next-generation transcriptome assembly. Nat. Rev. Genet., 12: 671-682.
45. Davey, J.W., P.A. Hohenlohe, P.D. Etter, J.Q. Boone, J.M. Catchen and M.L. Blaxter, 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nat. Rev. Genet., 12: 499-510.
46. Harismendy, O., P.C. Ng, R.L. Strausberg, X. Wang and T.B. Stockwell *et al*., 2009. Evaluation of next generation sequencing platforms for population targeted sequencing studies. Genome Biol., Vol. 10. 10.1186/gb-2009-10-3-r32.
47. Metzker, M.L., 2010. Sequencing technologies-the next generation. Nat. Rev. Genet., 11: 31-46.
48. Horak, P., S. Frohling and H. Glimm, 2016. Integrating next-generation sequencing into clinical oncology: Strategies, promises and pitfalls. ESMO Open, Vol. 1. 10.1136/esmoopen-2016-000094.
49. Luo, R., B. Liu, Y. Xie, Z. Li and W. Huang *et al*., 2012. SOAPdenovo2: An empirically improved memory-efficient short-read *de novo* assembler. Gigascience, Vol. 1. 10.1186/2047-217X-1-18.
50. Mercer, T.R., M.E. Dinger and J.S. Mattick, 2009. Long non-coding RNAs: Insights into functions. Nat. Rev. Genet., 10: 155-159.

51. Ku, C.S., N. Naidoo, M. Wu and R. Soong, 2011. Studying the epigenome using next generation sequencing. J. Med. Genet. 10.1136/jmedgenet-2011-100242
52. Zhang, Y. and A. Jeltsch, 2010. The application of next generation sequencing in DNA methylation analysis. Genes, 1: 85-101.
53. Pedersen, J.S., E. Valen, A.M.V. Velazquez, B.J. Parker and M. Rasmussen *et al*., 2014. Genome-wide nucleosome map and cytosine methylation levels of an ancient human genome. Genome Res., 24: 454-466.
54. Fanelli, M., S. Amatori, I. Barozzi and S. Minucci, 2011. Chromatin immunoprecipitation and high-throughput sequencing from paraffin-embedded pathology tissue. Nat. Protocols, 6: 1905-1919.
55. Ingolia, N.T., G.A. Brar, S. Rouskin, A.M. McGeachy and J.S. Weissman, 2012. The ribosome profiling strategy for monitoring translation *in vivo* by deep sequencing of ribosome-protected mRNA fragments. Nat. Protocols, 7: 1534-1550.
56. Ware, J.S., A.M. Roberts and S.A. Cook, 2011. Next generation sequencing for clinical diagnostics and personalised medicine: Implications for the next generation cardiologist. Heart. 10.1136/heartjnl-2011-300742.
57. Simon, R. and S. Roychowdhury, 2013. Implementing personalized cancer genomics in clinical trials. Nat. Rev. Drug Discov., 12: 358-369.
58. Collins, F.S., A. Patrinos, E. Jordan, A. Chakravarti, R. Gesteland and L. Walters, 1998. New goals for the U.S. human genome project: 1998-2003. Science, 282: 682-689.
59. Van Dijk, E.L., H. Auger, Y. Jaszczyszyn and C. Thermes, 2014. Ten years of next-generation sequencing technology. Trends Genet., 30: 418-426.
60. Git, A., H. Dvinge, M. Salmon-Divon, M. Osborne and C. Kutter *et al*., 2010. Systematic comparison of microarray profiling, real-time PCR and next-generation sequencing technologies for measuring differential microRNA expression. RNA., 16: 991-1006.
61. Bentley, D.R., 2006. Whole-genome re-sequencing. Curr. Opin. Genet. Dev., 16: 545-552.
62. Nagalakshmi, U., Z. Wang, K. Waern, C. Shou, D. Raha, M. Gerstein and M. Snyder, 2008. The transcriptional landscape of the yeast genome defined by RNA sequencing. Science, 320: 1344-1349.
63. Wang, Z., M. Gerstein and M. Snyder, 2009. RNA-Seq: A revolutionary tool for transcriptomics. Nat. Rev. Genet., 10: 57-63.
64. Quail, M.A., M. Smith, P. Coupland, T.D. Otto and S.R. Harris *et al*., 2012. A tale of three next generation sequencing platforms: Comparison of ion torrent, pacific biosciences and Illumina MiSeq sequencers. BMC Genomics, Vol. 13. 10.1186/1471-2164-13-341.
65. Schweiger, M.R., M. Kerick, B. Timmermann and M. Isau, 2011. The power of NGS technologies to delineate the genome organization in cancer: From mutations to structural variations and epigenetic alterations. Cancer Metastasis Rev., 30: 199-210.
66. Wan, R., V.N. Anh and K. Asai, 2012. Transformations for the compression of FASTQ quality scores of next-generation sequencing data. Bioinformatics, 28: 628-635.
67. Illumina, 2011. Quality scores for next-generation sequencing. Technical Note: Sequencing. Illumina Inc., USA. https://www.illumina.com/documents/products/technotes/technote_Q-Scores.pdf
68. Del Fabbro, C., S. Scalabrin, M. Morgante and F.M. Giorgi, 2013. An extensive evaluation of read trimming effects on Illumina NGS data analysis. PLoS One, Vol. 8. 10.1371/journal.pone.0085024.
69. Miller, J.R., S. Koren and G. Sutton, 2010. Assembly algorithms for next-generation sequencing data. Genomics, 95: 315-327.
70. Werner, T., 2010. Next generation sequencing in functional genomics. Briefings Bioinform., 11: 499-511.
71. Zhao, M., Q. Wang, Q. Wang, P. Jia and Z. Zhao, 2013. Computational tools for Copy Number Variation (CNV) detection using next-generation sequencing data: Features and perspectives. BMC Bioinform., Vol. 14. 10.1186/1471-2105-14-S11-S1.
72. Trapnell, C., A. Roberts, L. Goff, G. Pertea and D. Kim *et al*., 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat. Protocols, 7: 562-578.
73. Li, H., B. Handsaker, A. Wysoker, T. Fennell and J. Ruan *et al*., 2009. The sequence alignment/map format and SAMtools. Bioinformatics, 25: 2078-2079.
74. Bahassi, E.M. and P.J. Stambrook, 2014. Next-generation sequencing technologies: breaking the sound barrier of human genetics. Mutagenesis, 29: 303-310.
75. Rapaport, F., R. Khanin, Y. Liang, M. Pirun and A. Krek *et al*., 2013. Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. Genome Biol., Vol. 14. 10.1186/gb-2013-14-9-r95
76. Bao, R., L. Huang, J. Andrade, W. Tan, W.A. Kibbe, H. Jiang and G. Feng, 2014. Review of current methods, applications and data management for the bioinformatics analysis of whole exome sequencing. Cancer Inform., 13: 67-82.
77. Gullapalli, R.R., M. Lyons-Weiler, P. Petrosko, R. Dhir, M.J. Becich and W.A. LaFramboise, 2012. Clinical integration of next-generation sequencing technology. Clin. Lab. Med., 32: 585-599.
78. Gargis, A.S., L. Kalman, M.W. Berry, D.P. Bick and D.P. Dimmock *et al*., 2012. Assuring the quality of next-generation sequencing in clinical laboratory practice. Nat. Biotechnol., 30: 1033-1036.

79. Mauer, C.B., S.M. Pirzadeh-Miller, L.D. Robinson and D.M. Euhus, 2013. The integration of next-generation sequencing panels in the clinical cancer genetics practice: An institutional experience. Genet. Med., 16: 407-412.
80. Lin, X., W. Tang, S. Ahmad, J. Lu, C.C. Colby, J. Zhu and Q. Yu, 2012. Applications of targeted gene capture and next-generation sequencing technologies in studies of human deafness and other genetic disabilities. Hearing Res., 288: 67-76.
81. Summerer, D., 2009. Enabling technologies of genomic-scale sequence enrichment for targeted high-throughput sequencing. Genomics, 94: 363-368.
82. Munos, B., 2009. Lessons from 60 years of pharmaceutical innovation. Nat. Rev. Drug Discov., 8: 959-968.
83. Begley, C.G. and L.M. Ellis, 2012. Drug development: Raise standards for preclinical cancer research. Nature, 483: 531-533.
84. Morgan, S., P. Grootendorst, J. Lexchin, C. Cunningham and D. Greyson, 2011. The cost of drug development: A systematic review. Health Policy, 100: 4-17.
85. DiMasi, J.A., 2001. Risks in new drug development: Approval success rates for investigational drugs. Clin. Pharmacol. Ther., 69: 297-307.
86. Hait, W.N., 2010. Anticancer drug development: The grand challenges. Nat. Rev. Drug Discov., 9: 253-254.
87. Ginsburg, G.S. and J.J. McCarthy, 2001. Personalized medicine: revolutionizing drug discovery and patient care. Trends Biotechnol., 19: 491-496.
88. Tomczak, K., P. Czerwinska and M. Wiznerowicz, 2015. The Cancer Genome Atlas (TCGA): An immeasurable source of knowledge. Contemp. Oncol. (Pozn), 19: A68-A77.
89. Weinstein, J.N., E.A. Collisson, G.B. Mills, K.R.M. Shaw and B.A. Ozenberger *et al*., 2013. The cancer genome atlas pan-cancer analysis project. Nat. Genet., 45: 1113-1120.
90. Zhang, J., J. Baran, A. Cros, J.M. Guberman and S. Haider *et al*., 2011. International cancer genome consortium data portal-a one-stop shop for cancer genomics data. Database. 10.1093/database/bar026.
91. Forbes, S.A., N. Bindal, S. Bamford, C. Cole and C.Y. Kok *et al*., 2011. COSMIC: Mining complete cancer genomes in the catalogue of somatic mutations in cancer. Nucl. Acids Res., 39: D945-D950.
92. Robson, M.E., C.D. Storm, J. Weitzel, D.S. Wollins and K. Offit, 2010. American society of clinical oncology policy statement update: Genetic and genomic testing for cancer susceptibility. J. Clin. Oncol., 28: 893-901.
93. Walsh, D., L. Rybicki, K.A. Nelson and S. Donnelly, 2002. Symptoms and prognosis in advanced cancer. Supportive Care Cancer, 10: 385-388.
94. Holbrook, J.D., J.S. Parker, K.T. Gallagher, W.S. Halsey and A.M. Hughes *et al*., 2011. Deep sequencing of gastric carcinoma reveals somatic mutations relevant to personalized medicine. J. Trans. Med., Vol. 9. 10.1186/1479-5876-9-119.
95. Rakha, E.A., J.S. Reis-Filho, F. Baehner, D.J. Dabbs and T. Decker *et al*., 2010. Breast cancer prognostic classification in the molecular era: The role of histological grade. Breast Cancer Res., Vol. 12. 10.1186/bcr2607.
96. Rothe, F., J.F. Laes, D. Lambrechts, D. Smeets and D. Vincent *et al*., 2014. Plasma circulating tumor DNA as an alternative to metastatic biopsies for mutational analysis in breast cancer. Ann. Oncol., 25: 1959-1965.
97. Wood, L.D., D.W. Parsons, S. Jones, J. Lin and T. Sjoblom *et al*., 2007. The genomic landscapes of human breast and colorectal cancers. Science, 318: 1108-1113.
98. Gates, M.A., B.A. Rosner, J.L. Hecht and S.S. Tworoger, 2010. Risk factors for epithelial ovarian cancer by histologic subtype. Am. J. Epidemiol., 171: 45-53.
99. Vignot, S., G.M. Frampton, J.C. Soria, R. Yelensky and F. Commo *et al*., 2013. Next-generation sequencing reveals high concordance of recurrent somatic alterations between primary tumor and metastases from patients with non-small-cell lung cancer. J. Clin. Oncol., 31: 2167-2172.
100. Xia, J., P. Jia, K.E. Hutchinson, K.B. Dahlman and D. Johnson *et al*., 2014. A meta-analysis of somatic mutations from next generation sequencing of 241 melanomas: A road map for the study of genes with potential clinical relevance. Mol. Cancer Ther., 13: 1918-1928.
101. Stransky, N., A.M. Egloff, A.D. Tward, A.D. Kostic and K. Cibulskis *et al*., 2011. The mutational landscape of head and neck squamous cell carcinoma. Science, 333: 1157-1160.
102. Treangen, T.J. and S.L. Salzberg, 2012. Repetitive DNA and next-generation sequencing: Computational challenges and solutions. Nat. Rev. Genet., 13: 36-46.
103. Kumar, G., T. Lahiri and R. Kumar, 2016. Statistical discrimination of breast cancer microarray data. Proceedings of the International Conference on Bioinformatics and Systems Biology, March 4-6, 2016, Allahabad, pp: 1-4.
104. Sugiyama, M., T. Ide, S. Nakajima and J. Sese, 2010. Semi-supervised local Fisher discriminant analysis for dimensionality reduction. Mach. Learning, 78: 35-61.
105. Horn, R.A. and C.R. Johnson, 1990. Norms for Vectors and Matrices. In: Matrix Analysis, Horn, R.A. and C.R. Johnson (Eds.). Cambridge University Press, Cambridge, England.
106. Snedecor, G.W. and W.G. Cochran, 1989. Statistical Methods. 8th Edn., Iowa State University Press, Ames, Iowa, ISBN-13: 9780813815619.
107. Shokralla, S., J.L. Spall, J.F. Gibson and M. Hajibabaei, 2012. Next generation sequencing technologies for environmental DNA research. Mol. Ecol., 21: 1794-1805.
108. Cronin, M. and J.S. Ross, 2011. Comprehensive next-generation cancer genome sequencing in the era of targeted therapy and personalized oncology. Biomarkers Med., 5: 293-305.

109. Morozova, O. and M.A. Marra, 2008. Applications of next-generation sequencing technologies in functional genomics. Genomics, 92: 255-264.
110. Mamanova, L., A.J. Coffey, C.E. Scott, I. Kozarewa and E.H. Turner *et al*., 2010. Target-enrichment strategies for next-generation sequencing. Nat. Methods, 7: 111-118.
111. Gupta, P.K., 2008. Single-molecule DNA sequencing technologies for future genomics research. Trends Biotechnol., 26: 602-611.
112. Schadt, E.E., S. Turner and A. Kasarskis, 2010. A window into third-generation sequencing. Hum. Mol. Genet., 19: R227-R240.