

CRITICAL REVIEWS IN ORAL BIOLOGY & MEDICINE

M. Jessri^{1,2} and C.S. Farah^{1,2*}

¹The University of Queensland, UQ Centre for Clinical Research, Herston, Qld 4029, Australia; and ²The University of Queensland, School of Dentistry, Brisbane, Qld 4000, Australia; *corresponding author, c.farah@uq.edu.au

J Dent Res 93(5):437-444, 2014

ABSTRACT

Advances in the management of patients with head and neck squamous cell carcinoma (HNSCC) have not significantly changed the prognosis of this tumor over the past five decades. Molecular heterogeneity of HNSCC and its association with HPV, in addition to the increase in the number of cancers arising in traditionally low-risk patients, are among some of the obstacles to the successful management of this group of tumors. Massively parallel sequencing, otherwise known as next-generation sequencing (NGS), is rapidly changing conventional patient management by providing detailed information about each patient's genome and transcriptome. Despite major advances in technology and a significant reduction in the cost of sequencing, NGS remains mainly limited to research facilities. In addition, there are only a few published studies that have utilized this technology in HNSCC. This paper aims to report briefly on current commercially available NGS platforms and discuss their clinical applications, ethical considerations, and utilization in personalized patient care, particularly as this relates to head and neck cancer.

KEY WORDS: head and neck neoplasms, squamous cell carcinoma of the head and neck, individualized medicine, DNA sequencing, high-throughput nucleotide sequencing, deep sequencing.

DOI: 10.1177/0022034514524783

Received September 25, 2013; Last revision January 14, 2014; Accepted January 29, 2014

© International & American Associations for Dental Research

Harnessing Massively Parallel Sequencing in Personalized Head and Neck Oncology

INTRODUCTION

Personalized medicine customizes health care based on an individual's unique environment, clinical profile, and genetics. Until recently, it was reasonable to assume that modification of oral care according to a patient's needs constituted personalized oral health. However, the oral sciences are being transformed by large quantities of data produced by high-throughput molecular profiling technologies. The next generation of clinicians will be expected to personalize their management according to the molecular profiles of their patients. At this stage, head and neck cancer (HNC) appears to be the primary beneficiary of this change. Although more than 90% of HNCs are head and neck squamous cell carcinoma (HNSCC), recent studies have revealed this tumor to have a very heterogeneous nature (Baxi *et al.*, 2012). These findings are mirrored in the differences in prognosis and recurrence rates of different subclasses of squamous cell carcinoma (Chung *et al.*, 2004).

HNSCC is considered an environmental tumor mostly attributable to tobacco use, alcohol consumption, and infection with human papilloma virus (HPV) (Hashibe *et al.*, 2007). There is a growing proportion of cases in younger low-risk patients with a poor prognosis and a distinctive clinical and histopathological pattern (Chung *et al.*, 2004; Dahlstrom *et al.*, 2008), which suggests a role for genetic factors contributing to carcinogenesis of these tumors (Liang *et al.*, 2008). Genetic and prognostic differences between HPV-positive and -negative HNSCC (Chung and Gillison, 2009), combined with the heterogeneous nature of this cancer, presents clinicians with an urgent need for a refined tumor classification system.

Massively parallel sequencing, more commonly known as next-generation sequencing (NGS), technologies have been applied to answer fundamental questions about the genetic pathogenesis of various diseases (Feldman *et al.*, 2011). The increasing number of research projects and clinical trials targeting therapeutic biomarkers based on specific genetic mutations and aiming to overcome drug resistance in cancer patients makes NGS a perfect means to achieve personalized care in dentistry.

There is currently limited literature concerning the application of NGS in overall patient management in dentistry, with an increasing body of work related to HNSCC. With the exponential expansion of NGS platforms, the reduction in sequencing costs, and the push for personalized oral care, we review current NGS technologies with special consideration given to clinical applications, limitations, and ethical considerations, particularly as they pertain to HNSCC.

TECHNOLOGIES FOR MASSIVELY PARALLEL SEQUENCING

Sanger sequencing was the first generation of sequencing methods, while the new technologies with their ability to sequence enormous volumes of data

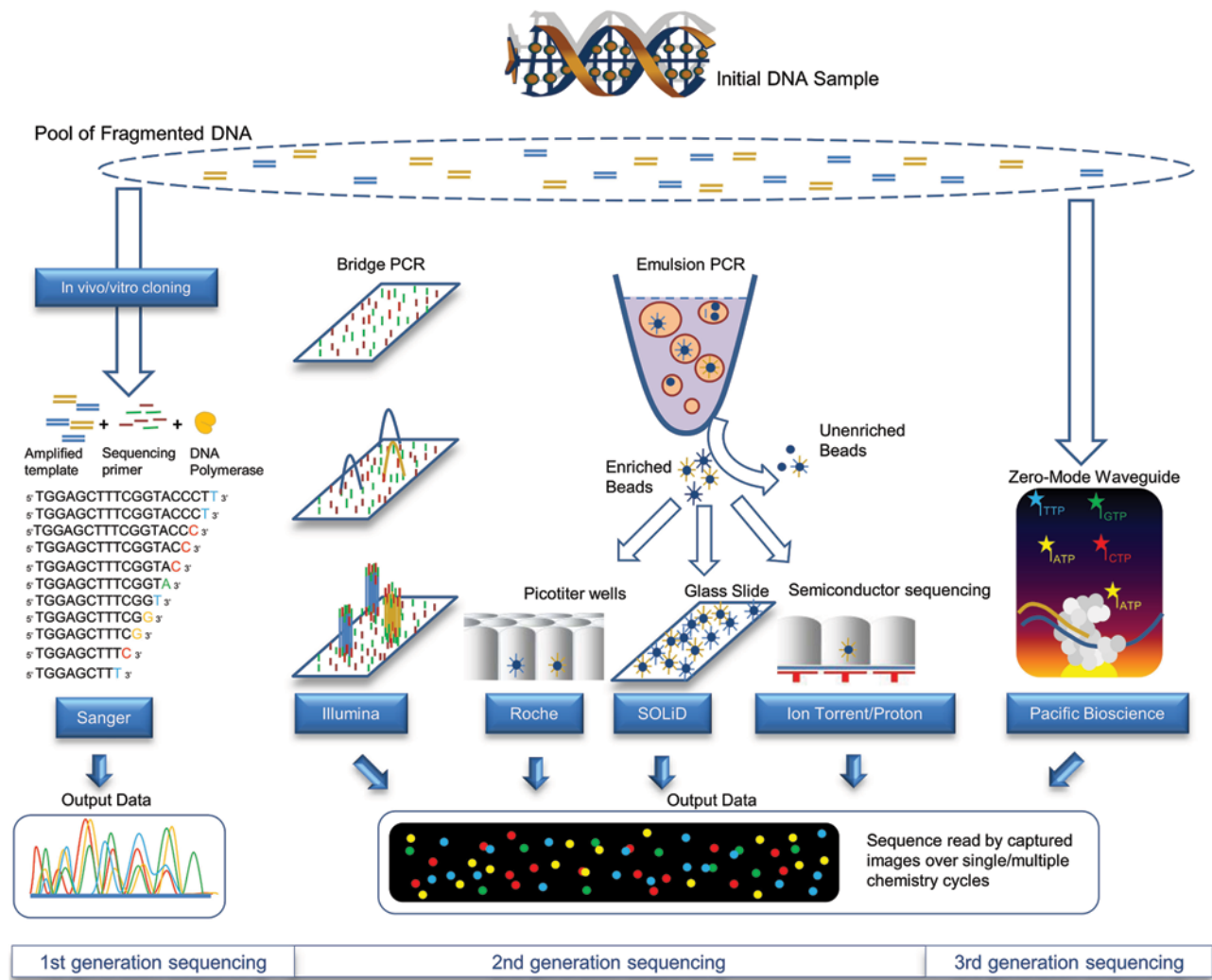


Figure 1. Technical characteristics including sample preparation, sequencing chemistries, and data output formats of different sequencing approaches.

cheaply are commonly considered “next-generation sequencing” techniques. NGS platforms have evolved to have an equal or lower error rate compared with the Sanger method (Glenn, 2011); however, due to differences in the nature of systematic errors, Sanger sequencing is still used to validate NGS data (Ulahannan *et al.*, 2013). Roche, Illumina®, and Life Technologies™ are three companies that currently offer NGS platforms. Technical characteristics of these platforms and their relevance and application in HNSCC have been discussed in greater detail elsewhere (Jessri and Farah, 2014); however, Fig. 1 briefly outlines the similarities and differences of these platforms.

Introduced in 2005, Roche454 (Qiagen, Venlo, Netherlands) was the first commercial NGS platform (Glenn, 2011). This technique is based on pyrosequencing and production of light as the by-product of nucleotide incorporation into the new strand of DNA. The amount of light measured is proportional to the number of nucleotides added (Ulahannan *et al.*, 2013).

Illumina® (Illumina® Inc., San Diego, CA, USA) platforms are the most commonly used and are based on reversible dye

terminator single-base synthesis (SBS) and immediate acquisition of digital images (Ross and Cronin, 2011). Briefly, both ends of each DNA fragment are ligated to specific adaptors and amplified by bridge PCR in a flow cell (Mardis, 2008). Nucleotides are incorporated in a step-wise manner, and the locations of fluorescently labeled nucleotide incorporations are recorded by scanning flow cells (Liu *et al.*, 2012).

Supported Oligonucleotide Ligation and Detection (SOLiD™, Life Technologies™, Carlsbad, CA, USA) relies on emulsion PCR and sequencing by ligation. Briefly, DNA fragments are deposited onto flow cells, and primers anneal to the adapter sequence on each fragment. Fluorescently labeled octamers are added to the DNA sequence, and di-base probes interrogate their first two bases, followed by fluorescence detection by CCD camera imaging (Mardis, 2008; Ross and Cronin, 2011). Recently, Life Technologies™ introduced the Ion Personal Genome Machine™ (PGM™), which relies on a hypersensitive semiconductor to detect hydrogen ions released during each cycle of polymerization (Rusk, 2011).

Table. Comparison of Current Commercially Available NGS Platforms in Terms of Price, Applications, and Error Rates

Company	Instrument	Instrument Price (cost per run)	Cost/Mb	Run Time	Primary Errors	Error Rate (%)*	Applications
Applied Biosystems	3730xl	\$95,000 (\$140)	\$2,400 [†]	20 min-3 hrs	Substitution	0.1-1 [†]	Validation, target sequencing
Roche	454 FLX Titanium	\$500,000 (\$7,000)	\$10 [†]	23 hrs	Indel	0.1 [†]	<i>De novo</i> whole genome, transcriptome, targeted re-sequencing, metagenomics
Illumina	HiSeq2000	\$690,000 (\$6,000)	\$0.07 [†]	11 days 27 hrs (rapid run)	Substitution	0.26 [‡]	<i>De novo</i> whole genome, transcriptome, targeted re-sequencing, metagenomics, mutation detection
	MiSeq (BT)	\$128,000 (\$1,040)	\$0.50 [‡]	4-27 hrs	Substitution	>0.8 [‡]	Transcriptome, targeted re-sequencing, metagenomics, mutation detection
Applied Biosystems	SOLiD 5500 XL	\$500,000 (\$15,000)	\$0.13 [†]	7 days to 4 wks	A-T bias	>0.01	Transcriptome, targeted re-sequencing, metagenomics, mutation detection, transcriptome counting
	Ion Torrent 318 chip	\$80,000 (\$939)	\$1 [‡]	2 hrs	Indel	1.71 [‡]	Transcriptome, target re-sequencing
	Ion Proton II	\$149,000 (\$1,000)	\$0.08 [§]	2 hrs	Indel	0.1 [§]	Transcriptome, target re-sequencing
PcBio	Pacific Bioscience	\$695,000 (\$300)	\$2 [‡]	2 hrs	GC deletions	12.86 [‡]	Targeted transcriptome, targeted re-sequencing, metagenomics

*Error rates are reported as percentage *per base* for single reads, and, due to differences in chemistry, error rates among platforms are not comparable.

[†]Liu *et al.*, 2012.

[‡]Quail *et al.*, 2012.

[§]Information based on the manufacturer's Web site.

Despite their differences, NGS platforms are potentially capable of identifying changes in gene-expression profiling or DNA-copy number with similar accuracy, as well as sequencing both ends of DNA molecules, commonly referred to as paired-end sequencing. An important consideration in the utilization of NGS in clinical practice is equipment entry costs and consumable expenses of platforms and their projected accuracy (Table). The accuracy of NGS platforms depends on adequate "coverage" and can be improved to reach more than 99.9% (Quail *et al.*, 2012). However, a 0.01% error rate corresponds to approximately 300,000 errors in sequencing the whole human genome, and systematic errors such as non-uniformity of sequence depth reported in short-read instruments cannot be overcome by higher coverage (Meldrum *et al.*, 2011).

Third-generation sequencing platforms offer the promise of sequencing single DNA molecules and are the latest achievement in NGS technologies. Pacific Biosciences Inc. commercialized their third-generation platform, PacBio RS II, in 2011. By combining single-molecule real-time sequencing (SMRT[®]) with highly sensitive fluorescence detection technology, this platform has 99% accuracy at 15-fold coverage (Eid *et al.*, 2009).

With the growing interest in the role of tumor microenvironment in cancer treatment, the quantification of gene expression in tumor cells and their surrounding tissues can provide important information about tumor growth, response/resistance to cancer therapy, and prognosis (Ignatiadis *et al.*, 2012). The application of NGS to cDNA libraries made by reverse transcription of RNA is termed RNA-seq. The coding RNA can be separated from the non-coding molecules by targeting the 3'

polyadenylated [poly(A)] tail; however, random inclusion of RNA has the advantage of higher sensitivity (Morin *et al.*, 2008). The sequence transcriptome is mapped to the reference level, and the number of mapped reads is considered to correlate with the abundance level and is a measure of gene/genome expression. Compared with cDNA gene expression microarrays, RNA-seq is more reliable in detecting exon/exon junctions; also, by accessing the reads that span exon/exon boundaries, sequencing provides valuable information on isoform diversity and abundance (Agarwal *et al.*, 2010). In analyzing RNA-seq data, one must not overlook the fact that, unlike gene-probe-based methods, expression signal in RNA-seq depends on the expression level of other transcripts and is influenced by the sequencing depth. To address this and other technical differences, several statistical algorithms have been developed to normalize RNA-seq data and detect differential expression (Rapaport *et al.*, 2013).

CLINICAL APPLICATIONS

Clinicians' approaches to classifying diseases and their diagnoses are rapidly changing due to the progress made in multigene analysis tools such as gene arrays, comparative genomic hybridization, and NGS. The most common application of molecular diagnosis is in the analysis of patients' genomic alterations to investigate susceptibility, risk, and target therapy. Based on the role they play in disease, genomic mutations can contribute to tumor development (driver mutations) or be a product of genomic instability of the tumor (passenger mutations) (Hanahan and Weinberg, 2000, 2011). The distinction between the two can

change through the course of tumor development (Desmedt *et al.*, 2012).

Prior to the application of NGS platforms in cancer research, array comparative genomic hybridization (a-CGH) and, more commonly, different types of polymerase chain-reaction (PCR) were applied to study cancer genome and its characteristics. Regardless of the promising picture portrayed by commercially available a-CGH and PCR kits, these techniques are mostly beneficial in the detection of known genomic alterations. Several other laboratory-based techniques – such as extensive tumor dissection (Maley *et al.*, 2006), isolation and analysis of individual nuclei (Park *et al.*, 2010), and examination of intra-tumor distribution of pre-identified markers (Maley *et al.*, 2006; Jovanovic *et al.*, 2008) – exist for the study of cancer cell sub-populations in any given tumor; nevertheless, despite standardized protocols for these techniques and the vast body of work in the literature, they are rapidly being superseded by NGS approaches.

The first two reported NGS studies of HNSCC were published simultaneously in 2011 (Agrawal *et al.*, 2011; Stransky *et al.*, 2011). Stransky *et al.* sequenced the whole exome of the tumors and blood of 74 HNSCC patients (Stransky *et al.*, 2011), while Agrawal *et al.* sequenced 18,000 protein-encoding genes in 32 discovery tumors and further investigated their findings in 88 additional HNSCC and matched non-neoplastic tissues (Agrawal *et al.*, 2011). The former found an average of 130 coding mutations *per* tumor, of which 321 mutations were queried by mass spectrometric genotyping and 89.7% validated (Stransky *et al.*, 2011). Both studies found the mutation rate of smoker patients/HPV-negative tumors to be higher than that of non-smoker/HPV-positive cases, respectively. Agrawal *et al.* also showed the mutation spectrum of all HNSCC patients to be similar to that of non-smoker patients with lung cancer (Agrawal *et al.*, 2011; Stransky *et al.*, 2011). Finally, both Agrawal *et al.* and Stransky *et al.* not only found mutations in genes already known to play a role in HNSCC [such as TP53 (50% of all cases), CDKN2A, PIK3CA, PTEN, and HRAS (Agrawal *et al.*, 2011; Stransky *et al.*, 2011)], but also highlighted a new gene, NOTCH1, previously not known to be involved in HNSCC. Although functional studies have shown NOTCH1, an important tumor suppressor gene, to be important in skin squamous cell carcinoma (Dotto, 2008), due to its large size (34 coding exons), it had not been identified by classic (Sanger) sequencing techniques.

Despite a competitive marketplace and extensive efforts for realization of the \$1,000 genome concept, the aforementioned two studies remain the most important attempts to utilize NGS platforms in head and neck oncology, while subsequent studies have focused on targeted sequencing. Mahjabeen *et al.* sequenced all 17 exons of XRCC1, a critical protein in base excision repair, in 300 HNC cases and 150 matched cancer-free controls (Mahjabeen *et al.*, 2012) and found two silent mutations in 45% and two missense mutations in 55% of cases, accounting for a total mutation frequency of 87%. Both silent mutations were distributed equally among males and females and smokers *vs.* non-smokers (Mahjabeen *et al.*, 2012). In another study, sequencing of all exons and adjacent introns of RAD51C revealed 5 distinct heterozygous sequence alterations in 5.8% of HNSCC cases (Scheckenbach *et al.*, 2014).

As summarized in Fig. 2, applications of NGS in dentistry are not necessarily confined to cancer patients. With advances in our knowledge of oral diseases, the link between genetics and oral, dental, and craniofacial disorders is becoming increasingly clear. However, a careful patient-selection algorithm based on individual needs, along with the associated costs and perceived benefits of NGS for each case, should be considered. In addition, NGS results must be analyzed appropriately utilizing standardized criteria and data analysis pipelines before clinical decisions are made. In an attempt to translate the results of NGS studies, Mroz and Rocco (2012) introduced mutant-allele tumor heterogeneity (MATH) as a measure of the distribution of mutant-allele fractions among loci. According to those authors, in 74 HNSCC cases from the study by Stransky *et al.* (2011), MATH ranged between 19 and 55 and was reported to be lower in HPV-positive tumors. It was also noted that the group of tumors having the highest intra-tumor heterogeneity showed the worst outcome (Mroz and Rocco, 2012).

NEXT-GENERATION SEQUENCING IN PERSONALIZED MEDICINE

The ultimate purpose of research aimed at understanding the genetic mechanisms contributing to development and progression of disease is prevention and the improvement of therapeutic outcomes for patients. Currently, NGS is being used to study genomic alterations in different diseases, and although frequently mutated genes have been identified, Mardis notes that most of these are mutated in only a small fraction of tumors (Mardis, 2012). This finding has brought forth the idea of personalized care on a patient-by-patient basis to manage the physiological effects of each genomic alteration (Jones *et al.*, 2010).

Recent randomized trials have shown that most available targeted therapy drugs bear low or no anti-tumor activity, due to intrinsic (primary) and extrinsic (secondary) drug resistance, and their effects are limited to a few months. Long-term follow-up has shown either low or no therapeutic effects for most of these drugs (Ziogas *et al.*, 2012). To overcome these challenges, biomarker-based cancer therapy that aims at the discovery and identification of new molecular targets and new therapies is being extensively studied (Ross *et al.*, 2010).

Until recently, patients' mutation status was determined by different techniques that were based mostly on PCR. In an example of the utilization of the amplification refractory mutation system (ARMS) in targeted therapy, Mok *et al.* (2009) observed that non-smoking Asian patients with epidermal growth factor receptor (EGFR) mutation have a better progression-free survival rate when their pulmonary adenocarcinoma was treated with gefitinib (Mok *et al.*, 2009). In a similar clinical study on non-resectable stages IIIC and IV melanoma cases with BRAF V600E mutation detected by real-time PCR, the progression-free survival rate of patients treated with vemurafenib was significantly higher than that of those treated with different cytotoxic drugs (Chapman *et al.*, 2011). These two studies are good examples of the importance of personalizing treatment according to patients' molecular profile. However, despite its convenience and time efficiency, PCR lacks sensitivity and has limited ability in finding unknown mutations.

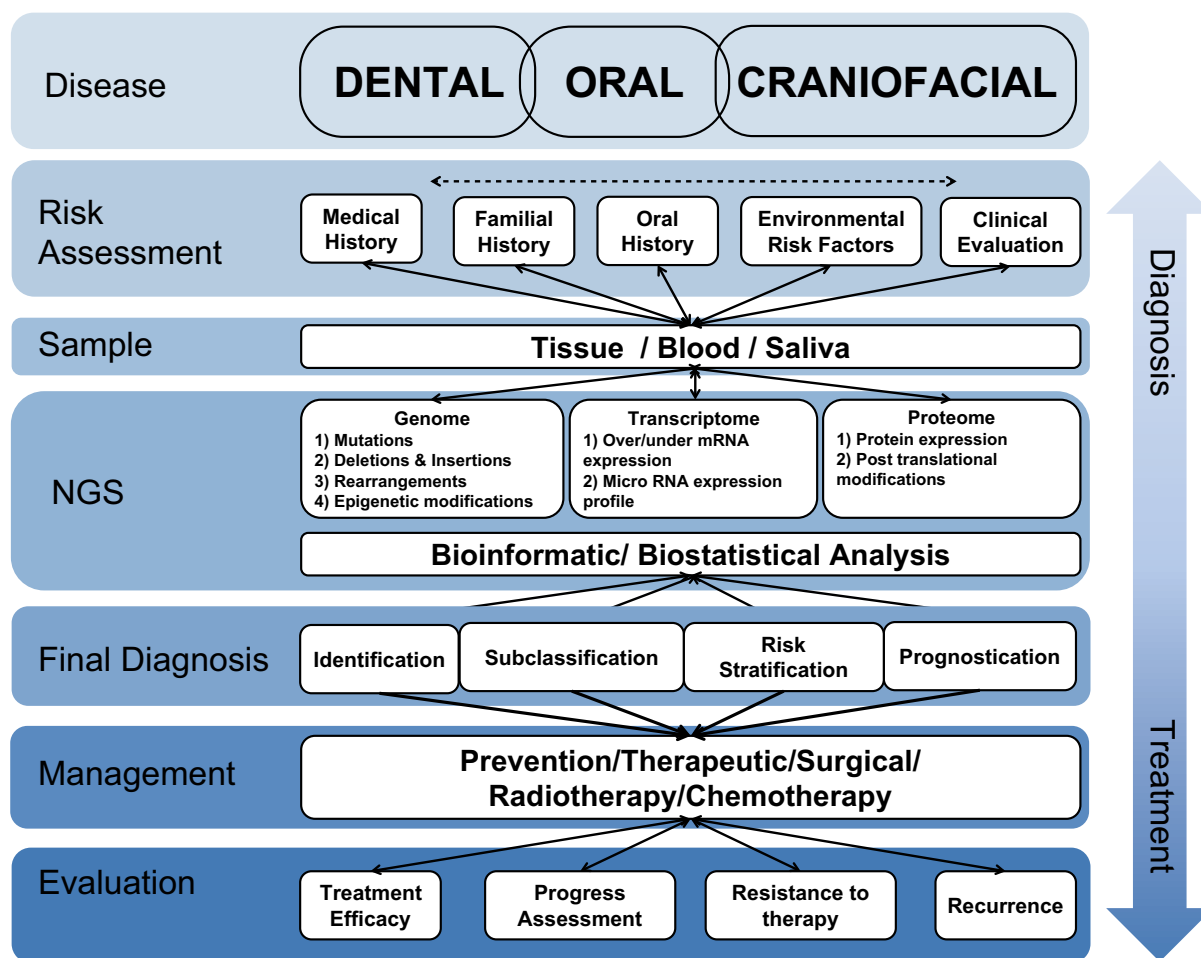


Figure 2. Application of next-generation sequencing in personalized risk assessment, prevention, and treatment of patients with dental, oral, and craniofacial diseases.

NGS has also been used in the personalized treatment of non-small-cell lung cancer with epidermal growth factor receptor (EGFR) deletions (Guo *et al.*, 2011). The best example for the application of NGS in individualizing treatment is systematic pharmacogenomic analysis of a whole genome from a Malay male volunteer, which resulted in a map containing a list of genes involved in the metabolism and transformation of drugs/therapeutic agents, and information about the predicted deleterious variants and their potential effect on transport, metabolism, and drug end targets (Salleh *et al.*, 2013). Despite sporadic impressive case reports, the application of NGS in clinical oncology is progressing slowly; however, with improvements in NGS technology and data analysis pipelines, the preparation of a comprehensive molecular profile prior to designing personalized therapy in a clinically acceptable timeframe is edging ever so closely to completion.

To test the feasibility of applying targeted sequencing to clinical practice, Han *et al.* tested alterations of 183 genes in tumor and matched normal samples of 60 colorectal adenocarcinoma patients (Han *et al.*, 2013). They found their results to be of clinical use and consistent with Sanger sequencing data. These findings are suggestive of the need for implementation of

personalized care based on evaluation of somatic mutations in cancer patients. To test the practicality of NGS in clinical situations, Roychowdhury *et al.* sequenced the whole genome and transcriptome of tumors and targeted whole-exome of tumors and normal DNA in two patients with advanced or refractory cancer (one metastatic colorectal cancer and one metastatic melanoma) (Roychowdhury *et al.*, 2011). A multidisciplinary board deliberated on the clinical consequences of genomic alterations, and within 24 days of biopsy, personalized treatments were designed (Roychowdhury *et al.*, 2011). Although most of the findings in this pilot study were not clinically significant, the approach proved NGS to be clinically relevant and practical in terms of cost and time.

The identification of driver mutations in cancers with a high mutation rate and a heterogeneous/more complex nature is challenging. As illustrated in Fig. 3, based on the purpose and nature of the study, different portions of genetic material should be targeted by NGS platforms. While NGS seems very likely to become part of routine patient management, overall costs, clinical utility and relevance, timeline requirements, and bioinformatic analysis are among the unresolved issues that need to be debated and discussed further before this becomes a clinical reality.

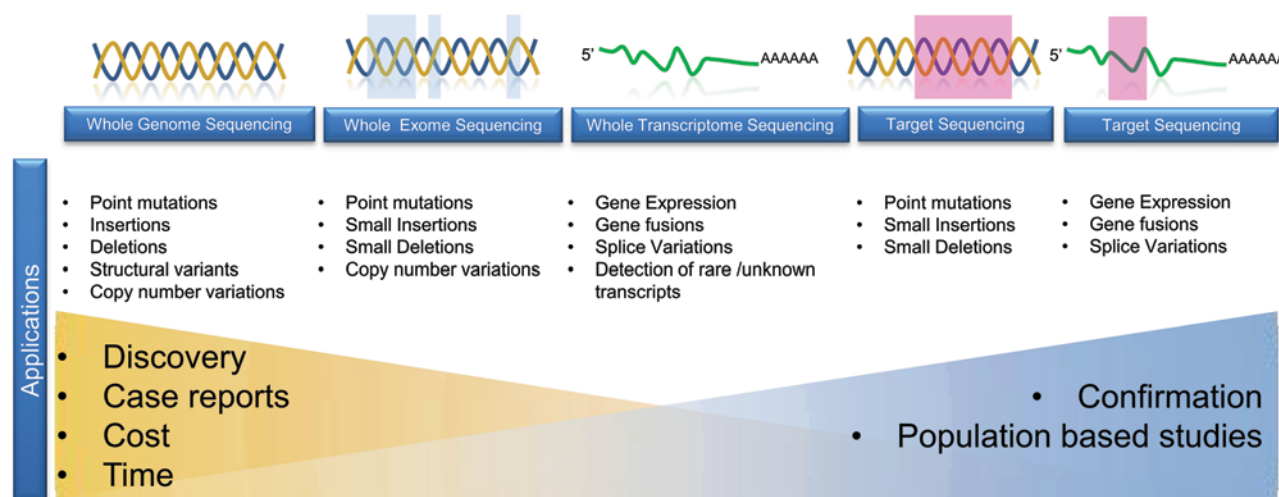


Figure 3. Potential applications of NGS strategies in research and clinical practice.

ETHICAL CONSIDERATIONS

There are a trio of factors to be considered when it comes to NGS ethical issues: (a) the techniques and facilities that lead to production of large amounts of data and the resultant information held by scientists; and (b) the fact that research groups tend to share their results with each other. These first two factors seem reasonably well founded and thought out with no possible challenge until one realizes that these genomic methods are applied in a social context where (c) storage and unauthorized use of data are becoming major concerns. In this context, the fundamental ethical requirements of patient consent for access to their data, treatment, and research should be re-evaluated. With the progress in mathematical and bioinformatics techniques, the identification of patients becomes easier through linking points and scores. Moreover, more data are being shared through informal means or deposition in cloud-based databases, and this compromises participant privacy (Zonta *et al.*, 2012). Whole-genome sequencing data are normally protected by complex and sometimes controversial national laws, and it is difficult for ethics review committees to protect participants when the regulations are nationally based. The national governance of data also fails to address international collaborations, which are increasing in number (Kaye *et al.*, 2010).

In genomic studies, the potential of infringement on privacy is far more important than the potential of physical harm to a patient. Patients should be adequately “informed” about the possibility of data-sharing before consenting to participate in genomic sequencing, either for research purposes or as part of their medical care. Providing patients with the results of the study in which they participated is both non-controversial and a welcome practice (Kaye *et al.*, 2010). In contrast, the feedback of individual sequencing studies, especially whole genome sequencing, remains an extensively debated and controversial issue (Knoppers *et al.*, 2006). Currently the common belief is that the same standards that apply to MRI and CT studies should be applied to whole genome sequencing, implying that when there is a serious finding with significant, clear, and verifiable

clinical relevance, in case it is treatable, the moral obligation of the research team is to inform the participant (Wolf *et al.*, 2012).

CONCERNS AND CONCLUSIONS

The drive from both the clinical and research spheres to improve and upgrade existing technologies has resulted in the production of a new generation of NGS platforms with higher accuracy, time-efficiency, and cost-effectiveness. Projects such as The Cancer Genome Atlas (<http://cancergenome.nih.gov>) and the International Cancer Genome Consortium (<http://icgc.org>) attempt to comprehensively profile different types of cancers by studying different genomic alterations through NGS platforms.

NGS provides the option of producing more than one billion base-pairs in a 4-day run; nevertheless, to interpret the produced data, one needs sophisticated bioinformatics methods. Analysis and translation of such data into the clinic are further complicated by the heterogeneous nature of most cancers. Finally, regardless of how exciting the large body of produced data and potential discoveries may sound, studies in this field are limited by cost, complexity of the data output, and scarcity of samples with high-integrity DNA. To partly resolve these issues, the third generation of NGS platforms focuses on sequencing a single molecule of DNA which, despite its promise, poses an important question of bias in sample selection (Ulahannan *et al.*, 2013). Among other limiting factors in NGS studies are the quantity and quality of DNA. Most archival tumor samples are preserved as formalin-fixed paraffin-embedded (FFPE) tissue; hence, the quality of the starting material for these studies can be compromised. Potential for increased protein carryover, and high amounts of degraded and single-stranded DNA are among the limitations of using FFPE samples (Schäfer *et al.*, 2009). Despite these limitations, a few studies have successfully used FFPE samples to study mutations, copy number variation, and determination of germline variations (Kerick *et al.*, 2011; Schweiger *et al.*, 2011; Menon *et al.*, 2012).

With more clinically relevant data from NGS studies, it is expected that NGS will expand from a cancer/mutation-focused

technique to include translocation, gene copy number alterations, and epigenetic changes of the genome. With advances in technology and our understanding of intra-tumor heterogeneity, the starting sample and the germline control may also change. In addition, the new-generation oral health practitioners are embracing the concepts of evidence-based and minimal intervention medicine that make personalized treatment of oral diseases even more practical (Farah *et al.*, 2013).

However, between the production of NGS data and the application of these findings to day-to-day clinical needs, there are a few major hurdles to overcome. First, detection of a genetic alteration does not necessarily make it a driver mutation. One step closer to the clinical application, an identified mutation is neither necessarily preventable nor definitely treatable. Silencing mutations in tumor suppressor genes are very complex to reverse, and in many instances their inhibitors are of no use (Brakenhoff, 2011). Most cancer-related pathways in cells are very critical, and inhibition or alteration of these could be lethal for cells; however, the presence of compensating pathways makes synthetic lethality a successful strategy (Fong *et al.*, 2009). Finally, as Garcia *et al.* have discussed, the side-effects of targeted therapies, such as drug resistance and toxicity, should not be overlooked (Garcia *et al.*, 2013).

While significant progress has been made in the identification of mutations underlying the development of HNSCC, the first forays into whole-genome and exome sequencing of HNSCC have proven that we are only beginning to understand the complex landscape of this group of malignancies. Heterogeneity of HNSCC, combined with racial and geographical differences in the human genome (Rosenberg *et al.*, 2002), necessitates extreme attention to detail in the design of NGS studies. Continuing to map samples to a “normal pool” or a common reference genome should be re-evaluated. Similarly, in light of the advantages and disadvantages of NGS and the potential of bioinformatic models to improve the data, attention should be paid to family-based studies in non-consanguineous cases (Hou *et al.*, 2013; Kamphans *et al.*, 2013).

NGS is a powerful tool with high potential, and its application is a great asset to researchers; however, further optimization and detailed studies are required to shed light on the molecular pathogenesis of cancer to assist patients in everyday clinical scenarios in a personalized and meaningful manner. The extended time between the studies by Stransky (Stransky *et al.*, 2011) and Agrawal (Agrawal *et al.*, 2011) and subsequent publications in the field, along with differences in the patient cohorts studied, are symptomatic of obstacles faced by scientists and clinicians in this area. Planning and utilization of NGS techniques require not only physical resources but also clinicians with an interest in the field, in addition to technical and bioinformatic knowledge. Furthermore, advancement in this area will be dependent on research groups willing to collaborate and share resources.

With competitive increases in the capacity of platforms, and decreases in the cost and volume of starting material, NGS offers great potential for the unraveling of oral diseases. NGS has the potential to transform the reactive and treatment-based nature of oral health care, to actively predict the risk, and aim to prevent them. Truly personalized dentistry shares its challenges

with NGS; both need cost reduction and solutions to technical shortcomings, in addition to connecting the generated data to meaningful clinical findings. Genetic profiling of patients with oral diseases is still in its infancy, and oral health practitioners will need to broaden their understanding of the molecular basis of disease before this becomes useful in routine clinical practice. The next generation of oral health clinicians should acknowledge that the term ‘head and neck cancer’ does not adequately reflect the complexity of this group of tumors, in a manner similar to that espoused by Ellis and Perou in relation to breast cancer (Ellis and Perou, 2013). As they have correctly noted, the application of all available technologies does not make our patient management personalized, but rather helps us reach a “real diagnosis” and an “appropriate etiology-matched treatment” which should become standard care in the near future (Ellis and Perou, 2013).

ACKNOWLEDGMENTS

The authors are currently undertaking next-generation sequencing of oral cancer and oral potentially malignant lesions utilizing SOLiD™ and Ion™ technologies, funded by grants held by author Camile S. Farah from the Queensland Government Smart Futures Co-Investment Fund and Cancer Australia, in collaboration with Life Technologies and Agilent Technologies. The authors declare no potential conflicts of interest with respect to the authorship and/or publication of this article.

REFERENCES

- Agarwal A, Koppstein D, Rozowsky J, Sboner A, Habegger L, Hillier LW, *et al.* (2010). Comparison and calibration of transcriptome data from RNA-Seq and tiling arrays. *BMC Genomics* 11:383.
- Agrawal N, Frederick MJ, Pickering CR, Bettgowda C, Chang K, Li RJ, *et al.* (2011). Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. *Science* 333(6046):1154-1157.
- Baxi S, Fury M, Ganly I, Rao S, Pfister DG (2012). Ten years of progress in head and neck cancers. *J Natl Compr Canc Netw* 10:806-810.
- Brakenhoff RH (2011). Another NOTCH for cancer. *Science* 333:1102-1103.
- Chapman PB, Hauschild A, Robert C, Haanen JB, Ascierto P, Larkin J, *et al.* (2011). Improved survival with vemurafenib in melanoma with BRAF V600E mutation. *N Engl J Med* 364:2507-2516.
- Chung CH, Gillison ML (2009). Human papillomavirus in head and neck cancer: its role in pathogenesis and clinical implications. *Clin Cancer Res* 15:6758-6762.
- Chung CH, Parker JS, Karaca G, Wu J, Funkhouser WK, Moore D, *et al.* (2004). Molecular classification of head and neck squamous cell carcinomas using patterns of gene expression. *Cancer Cell* 5:489-500.
- Dahlstrom KR, Little JA, Zafereo ME, Lung M, Wei Q, Sturgis EM (2008). Squamous cell carcinoma of the head and neck in never smoker-never drinkers: a descriptive epidemiologic study. *Head Neck* 30:75-84.
- Desmedt C, Voet T, Sotiriou C, Campbell PJ (2012). Next-generation sequencing in breast cancer: first take home messages. *Curr Opin Oncol* 24:597-604.
- Dotto GP (2008). Notch tumor suppressor function. *Oncogene* 27:5115-5123.
- Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, *et al.* (2009). Real-time DNA sequencing from single polymerase molecules. *Science* 323:133-138.
- Ellis MJ, Perou CM (2013). The genomic landscape of breast cancer as a therapeutic roadmap. *Cancer Discov* 3:27-34.
- Farah CS, Bhatia N, John K, Lee BW (2013). Minimum intervention dentistry in oral medicine. *Aust Dent J* 58(Suppl 1):85-94.

- Feldman AL, Dogan A, Smith DI, Law ME, Ansell SM, Johnson SH, *et al.* (2011). Discovery of recurrent t(6;7)(p25.3;q32.3) translocations in ALK-negative anaplastic large cell lymphomas by massively parallel genomic sequencing. *Blood* 117:915-919.
- Fong PC, Boss DS, Yap TA, Tutt A, Wu P, Mergui-Roelvink M, *et al.* (2009). Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers. *N Engl J Med* 361:123-134.
- Garcia I, Kuska R, Somerman MJ (2013). Expanding the foundation for personalized medicine: implications and challenges for dentistry. *J Dent Res* 92(7 Suppl):3S-10S.
- Glenn TC (2011). Field guide to next-generation DNA sequencers. *Mol Ecol Resour* 11:759-769.
- Guo G, Gui Y, Gao S, Tang A, Hu X, Huang Y, *et al.* (2011). Frequent mutations of genes encoding ubiquitin-mediated proteolysis pathway components in clear cell renal cell carcinoma. *Nat Genet* 44:17-19.
- Han SW, Kim HP, Shin JY, Jeong EG, Lee WC, Lee KH, *et al.* (2013). Targeted sequencing of cancer-related genes in colorectal cancer using next-generation sequencing. *PLoS ONE* 8:e64271.
- Hanahan D, Weinberg RA (2000). The hallmarks of cancer. *Cell* 100:57-70.
- Hanahan D, Weinberg RA (2011). Hallmarks of cancer: the next generation. *Cell* 144:646-674.
- Hashibe M, Brennan P, Benhamou S, Castellsague X, Chen C, Curado MP, *et al.* (2007). Alcohol drinking in never users of tobacco, cigarette smoking in never drinkers, and the risk of head and neck cancer: pooled analysis in the international head and neck cancer epidemiology consortium. *J Natl Cancer Inst* 99:777-789.
- Hou L, Faraci G, Chen DTW, Kassem L, Schulze TG, Shugart YY, *et al.* (2013). Amish revisited: next-generation sequencing studies of psychiatric disorders among the Plain People. *Trends Genet* 29:412-418.
- Ignatiadis M, Singhal SK, Desmedt C, Haibe-Kains B, Criscitiello C, Andre F, *et al.* (2012). Gene modules and response to neoadjuvant chemotherapy in breast cancer subtypes: a pooled analysis. *J Clin Oncol* 30:1996-2004.
- Jessri M, Farah CS (2014). Next generation sequencing and its application in deciphering head and neck cancer. *Oral Oncol* [Epub ahead of print 1/15/2014] (in press).
- Jones SJ, Laskin J, Li YY, Griffith OL, An J, Bilenky M, *et al.* (2010). Evolution of an adenocarcinoma in response to selection by targeted kinase inhibitors. *Genome Biol* 11:R82.
- Jovanovic L, Delahunt B, McIver B, Eberhardt NL, Grebe SK (2008). Most multifocal papillary thyroid carcinomas acquire genetic and morphotype diversity through subclonal evolution following the intra-glandular spread of the initial neoplastic clone. *J Pathol* 215:145-154.
- Kamphans T, Sabri P, Zhu N, Heinrich V, Mundlos S, Robinson PN, *et al.* (2013). Filtering for compound heterozygous sequence variants in non-consanguineous pedigrees. *PLoS ONE* 8:e70151.
- Kaye J, Boddington P, De Vries J, Hawkins N, Melham K (2010). Ethical implications of the use of whole genome methods in medical research. *Eur J Hum Genet* 18:398-403.
- Kerick M, Isau M, Timmermann B, Sultmann H, Herwig R, Krobisch S, *et al.* (2011). Targeted high throughput sequencing in clinical cancer settings: formaldehyde fixed-paraffin embedded (FFPE) tumor tissues, input amount and tumor heterogeneity. *BMC Med Genomics* 4:68.
- Knoppers BM, Joly Y, Simard J, Durocher F (2006). The emergence of an ethical duty to disclose genetic research results: international perspectives. *Eur J Hum Genet* 14:1170-1178.
- Liang XH, Lewis J, Foote R, Smith D, Kademani D (2008). Prevalence and significance of human papillomavirus in oral tongue cancer: The Mayo Clinic experience. *J Oral Maxillofac Surg* 66:1875-1880.
- Liu L, Li Y, Li S, Hu N, He Y, Pong R, *et al.* (2012). Comparison of next-generation sequencing systems. *J Biomed Biotechnol* 2012:251364.
- Mahjabeen I, Baig RM, Masood N, Sabir M, Inayat U, Malik FA, *et al.* (2012). Genetic variations in XRCC1 gene in sporadic head and neck cancer (HNC) patients. *Pathol Oncol Res* 19:183-188.
- Maley CC, Galipeau PC, Finley JC, Wongsurawat VJ, Li X, Sanchez CA, *et al.* (2006). Genetic clonal diversity predicts progression to esophageal adenocarcinoma. *Nat Genet* 38:468-473.
- Mardis ER (2008). Next-generation DNA sequencing methods. *Annu Rev Genomics Hum Genet* 9:387-402.
- Mardis ER (2012). Applying next-generation sequencing to pancreatic cancer treatment. *Nat Rev Gastroenterol Hepatol* 9:477-486.
- Meldrum C, Doyle MA, Tothill RW (2011). Next-generation sequencing for cancer diagnostics: a practical perspective. *Clin Biochem Rev* 32:177-195.
- Menon R, Deng M, Boehm D, Braun M, Fend F, Boehm D, *et al.* (2012). Exome enrichment and SOLiD sequencing of formalin fixed paraffin embedded (FFPE) prostate cancer tissue. *Int J Mol Sci* 13:8933-8942.
- Mok TS, Wu YL, Thongprasert S, Yang CH, Chu DT, Saijo N, *et al.* (2009). Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N Engl J Med* 361:947-957.
- Morin RD, Bainbridge M, Fejes A, Hirst M, Krzywinski M, Pugh TJ, *et al.* (2008). Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing. *BioTechniques* 45:81-94.
- Mroz EA, Rocco JW (2012). MATH, a novel measure of intratumor genetic heterogeneity, is high in poor-outcome classes of head and neck squamous cell carcinoma. *Oral Oncol* 49:211-215.
- Park SY, Gönen M, Kim HJ, Michor F, Polyak K (2010). Cellular and genetic diversity in the progression of in situ human breast carcinomas to an invasive phenotype. *J Clin Invest* 120:636-644.
- Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, *et al.* (2012). A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics* 13:341.
- Rapaport F, Khanin R, Liang Y, Pirun M, Krek A, Zumbo P, *et al.* (2013). Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. *Genome Biol* 14:R95.
- Rosenberg NA, Pritchard JK, Weber JL, Cann HM, Kidd KK, Zhivotovsky LA, *et al.* (2002). Genetic structure of human populations. *Science* 298:2381-2385.
- Ross JS, Cronin M (2011). Whole cancer genome sequencing by next-generation methods. *Am J Clin Pathol* 136:527-539.
- Ross JS, Torres-Mora J, Wagle N, Jennings TA, Jones DM (2010). Biomarker-based prediction of response to therapy for colorectal cancer: current perspective. *Am J Clin Pathol* 134:478-490.
- Roychowdhury S, Iyer MK, Robinson DR, Lonigro RJ, Wu YM, Cao X, *et al.* (2011). Personalized oncology through integrative high-throughput sequencing: a pilot study. *Sci Transl Med* 3:111ra121.
- Rusk N (2011). Torrents of sequence. *Nat Methods* 8:44.
- Salleh MZ, Teh LK, Lee LS, Ismet RI, Patowary A, Joshi K, *et al.* (2013). Systematic pharmacogenomics analysis of a Malay whole genome: proof of concept for personalized medicine. *PLoS ONE* 8:e71554.
- Schäfer SC, Pfür M, Yerly S, Fandel TM, Jichlinski P, Lehr HA (2009). Cryopreservation of prostate cancer tissue during routine processing of fresh unfixed prostatectomy specimen: demonstration and validation of a new technique. *Prostate* 69:191-197.
- Scheckenbach K, Baldus SE, Balz V, Freund M, Pakropa P, Sproll C, *et al.* (2014). RAD51C – a new human cancer susceptibility gene for sporadic squamous cell carcinoma of the head and neck (HNSCC). *Oral Oncol* 50:196-199.
- Schweiger MR, Kerick M, Timmermann B, Isau M (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.
- Stransky N, Egloff AM, Tward AD, Kostic AD, Cibulskis K, Sivachenko A, *et al.* (2011). The mutational landscape of head and neck squamous cell carcinoma. *Science* 333:1157-1160.
- Ulahannan D, Kovac MB, Mulholland PJ, Cazier JB, Tomlinson I (2013). Technical and implementation issues in using next-generation sequencing of cancers in clinical practice. *Br J Cancer* 109:827-835.
- Wolf SM, Crock BN, Van Ness B, Lawrenz F, Kahn JP, Beskow LM, *et al.* (2012). Managing incidental findings and research results in genomic research involving biobanks and archived data sets. *Genet Med* 14:361-384.
- Ziogas DE, Katsios CS, Tzaphlidou M, Roukos DH (2012). Targeted therapy: overcoming drug resistance with clinical cancer genome. *Expert Rev Anticancer Ther* 12:861-864.
- Zonta MA, Monteiro J, Santos G Jr, Pignatari AC (2012). Oral infection by the human papilloma virus in women with cervical lesions at a prison in São Paulo, Brazil. *Braz J Otorhinolaryngol* 78:66-72.