

## DNA Diagnosis in Oncology

Milica Kontic<sup>1\*</sup>, Marko Kontic<sup>2</sup>, Miodrag Ognjanovic<sup>3</sup>, Dragana Jovanovic<sup>1</sup> and Simona Ognjanovic<sup>4</sup>

<sup>1</sup>Clinic for Pulmonology, Clinical Center of Serbia, Belgrade, Serbia

<sup>2</sup>Military Medical Academy, Belgrade, Serbia

<sup>3</sup>Department of Environmental Health and Safety, University of Minnesota, Minnesota, USA

<sup>4</sup>Mayo Graduate School, Mayo Clinic, Rochester, Minnesota, USA

### Abstract

It is well established that cancers are caused by the accumulation of genomic and epigenomic alterations. However, molecular changes occurring in the early stages of cancer development or in precursor lesions remain to be poorly understood. The employment of molecular characterization of tumors prior to therapy has opened the door for personalized therapy of individual patients. Such genetic alterations that are utilized to predict response to therapy are considered predictive biomarkers. The development of high throughput technologies lead to the establishment of a relatively new field of cancer genomics. These genome-wide approaches are increasingly more important in cancer diagnostic, prognosis, and treatment. Here we briefly discuss cancer transcriptome, cancer genome and epigenome analyses approaches and their clinical utility. Gene expression patterns in tumors (cancer transcriptome) are analyzed by expression microarrays who can serve as a diagnostic tool which facilitates distinguishing different cancer subtypes. Another high throughput technology that may revolutionize personalized cancer therapy, lead to the development of new genetic diagnostic tests and biomarkers is next generation sequencing (NGS). NGS has an ability to fully sequence large number of genes in a single test and simultaneously detect deletions, insertions, copy number alterations, translocations and exome-wide base substitutions in cancer-related genes. NGS methods analyses DNA methylation, detection of modified histones, mapping of transcription factor occupancy, and epigenetic regulators. DNA diagnosis often trumps clinical diagnosis in oncology. This will cause the shift toward testing minute amounts of DNA and specimens including needle biopsies, circulating tumor cells, and circulating cell-free DNA are likely to become clinically relevant. Thus, as NGS enters clinical testing it will impact clinical decisions and cancer outcomes. Finally, improved understanding of cancer biology will lead to further development of targeted therapies and the use of combinations of targeted approaches simultaneously, increasing treatment effectiveness and supporting highly individualized patient care.

**Keywords:** DNA; Diagnosis; Oncology; NGS; Microarrays; Cancer transcriptome

It is well established that cancers are caused by the accumulation of genomic and epigenomic alterations trigger by environmental factors leading to inappropriate activation or inactivation of specific genes and resulting in neoplastic transformation. However, molecular changes occurring in the early stages of cancer development or in precursor lesions remain to be poorly understood. Numerous sequencing efforts are directed towards improved understanding of changes in genome, epigenome, and deregulated pathways that precede and promote tumor development and to distinguish between functionally important changes ("driver mutations") and functionally non-important changes ("passenger mutations"). These insights will not only improve the understanding of cancer biology, but will also affect diagnosis, prognosis and therapy.

### Genetic Diagnosis in Oncology: Clinical Applications Involving a Single or a Limited Number of Genes

#### Prognostic biomarkers

Initially, diagnostic efforts in oncology were focused on germline detection of few well-defined driver mutations, which underly familial syndromes that increase the risk of cancer development. Alterations in these specific genes are considered prognostic biomarkers because they facilitate the assesment of cancer risk and determine prognosis. In the US, it is a requirement that any genetic testing, which results will be used to guide patient care, should be performed at a Clinical Laboratory Improvement Amendments (CLIA)-certified laboratory. This ensures the accuracy, reliability and timeliness of patient test results regardless of where the test was performed. The National Institutes of

Health (NIH) developed an outstanding resource, Genetest website, which provides information on genetic testing and its use in diagnosis in order to promote the appropriate use of genetic services [1]. This includes the information on CLIA laboratories worldwide that perform a genetic test of interest [1]. For example, germline mutations in *TP53* gene underlie Li-Fraumeni syndrome (LFS). LFS is characterized by early onset of a wide variety of childhood- and adult-onset cancers, including soft-tissue and bone sarcomas, breast cancer, adrenal cortical carcinoma, brain tumors, and leukemia [2] and very high lifetime cumulative cancer risk [3]. There are at least fourteen CLIA-certified laboratories in the US, which offer *TP53* testing, and over thirty additional laboratories worldwide. The actual number of CLIA laboratories performing specific gene tests, including *TP53* testing, is likely to be much higher, because reporting to Genetest database is voluntary and many CLIA-certified laboratories may choose not to be listed in this database. The majority of these laboratories rely on direct, single gene sequencing (Sanger sequencing remains the gold standard). While this method offers complete gene sequence analysis at high

**\*Corresponding author:** Milica Kontic MD, PhD, Clinic for Pulmonology, Clinical Center of Serbia, Visegradska 26, Belgrade 11000, Serbia, Tel: +381 11 361 5556, +381 63 55 49 63; E-mail: milicakontic@yahoo.com

**Received** January 24, 2013; **Accepted** March 20, 2013; **Published** March 23, 2013

**Citation:** Kontic M, Kontic M, Ognjanovic M, Jovanovic D, Ognjanovic S (2013) DNA Diagnosis in Oncology. J Mol Biomark Diagn S5: 001. doi:10.4172/2155-9929.S5-001

**Copyright:** © 2013 Kontic M, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

sensitivity, the disadvantages include that it cannot detect deletions, translocations or copy number changes and it is fairly time-consuming [4]. Some of the CLIA labs therefore also use other methods, such as quantitative PCR, long-range PCR, multiplex ligation-dependent probe amplification (MLPA), and chromosomal microarray (CMA) to analyze the gene/chromosome segment of interest in order to detect duplications or deletions. Examples of cancer susceptibility syndromes and their underlying genes, for which CLIA testing is available in the US, are shown in table 1, which summarizes the information extracted from Genetest database [1].

### Predictive biomarkers

Cancer genome sequences can influence the selection of anticancer therapy. The employment of molecular characterization of tumors prior to therapy has opened the door for personalized therapy of individual patients. Such genetic alterations that are utilized to predict response to therapy are considered predictive biomarkers. These biomarkers can predict either resistance or sensitivity to specific therapy. Examples of genetic alterations that influence therapy selection in oncology practice are shown in table 2 [5-12]. Interestingly, sometimes different genetic alterations within the same gene (EGFR) can predict different outcomes: while the majority of EGFR activating mutations confer sensitivity to first-line EGFR tyrosine kinase inhibitors (gefitinib and erlotinib), rare EGFR mutations have been reported that confer resistance to such therapy [7].

### Whole Genome Approaches and their Utility in Oncology Practice

In contrast to the above described single-gene examinations characteristic of cancer genetics, the development of high throughput technologies lead to the establishment of a relatively new field of cancer genomics (also known as oncogenomics). These genome-wide approaches are increasingly more important in cancer diagnosis, prognosis, and treatment. Below we briefly discuss cancer transcriptome, cancer genome and epigenome analyses approaches and their clinical utility.

### Cancer transcriptome

Gene expression patterns in tumors (cancer transcriptome) are analyzed by expression microarrays. These microarray analyses of gene expression pattern in tumor cells can serve as a diagnostic tool which facilitates distinguishing different cancer subtypes, as shown by Alizadeh and colleagues [13] for diffuse large B-cell lymphoma. Based on the gene expression patterns the authors were able to distinguish two DLBCL subtypes with significantly different prognosis (overall survival) [13]. This is an example of how molecular classification of tumors identifies distinct clinical-pathological entities, which were not known before. Similarly, gene expression pattern of BRAF-wild type colorectal cancer tumors identified a subgroup of these tumors associated with poor prognosis, which expression pattern resembled

Cancer predisposing syndrome	Malignancy	Implicated gene(s)	Minimal number of CLIA labs offering genetic testing
Gorlin syndrome	BCC**	PTCH1*	1
Bloom syndrome	Broad range of leukemias, lymphomas and carcinomas	RECQL3	21
Hereditary breast/ovarian cancer syndrome	Breast and ovarian cancer	BRCA1	7
		BRCA2	6
Lynch syndrome	CRC	EPCAM	9
		MLH1	17
		MLH3	1
		MSH3	10
		MSH6	14
		PMS1	1
		PMS3	12
Familial Adenomatous Polyposis (FAP)	CRC	APC	16
Li-Fraumeni syndrome (LFS)	A wide range of tumors: sarcomas, breast cancer, brain tumors, adrenal cortical carcinomas, leukemia	TP53	14
Xeroderma pigmentosum	Skin cancer including BCC, melanoma, SCC	ERCC3	1
		POLH	1
		XPA	2
		XPC	2

\*Only those genes for which CLIA lab testing is available are shown here.

\*\*Acronyms: BCC=basal cell carcinoma, CRC=colorectal cancer, SCC=squamous cell carcinoma

**Table 1:** Examples of cancer predisposing syndromes with known underlying genetic abnormality for which testing is available in the US.

Genetic aberration	Malignancy	Therapy	Aberration conveys
KRAS mutation	CRC	Cetuximab/ Panitumumab [5]	Resistance
BRAF mutation	CRC	Cetuximab/ Panitumumab [5]	Resistance
EGFR mutation	NSCLC	Erlotinib/Gefitinib [7]	Sensitivity
EML4_ALK translocation	NSCLC	Crizotinib [8]	Sensitivity
BRAF mutation	melanoma	Vemurafenib [9]	Sensitivity
cKIT mutation or amplification	melanoma	Imatinib/Sumitinib [10]	
BCR_ABL translocation	CML	Imatinib/Dasatinib [11]	Sensitivity
HER2 amplification	breast cancer	Transtuzumab [12]	Sensitivity

**Table 2:** Examples of biomarkers predicting response to therapy.

tumors carrying BRAF mutation [6]. Thus, tumor genetic signature revealed common biology that was not captured by BRAF mutation status alone and provided a novel classification tool with prognostic feature that may guide therapeutic strategies [6]. Such insights not only improve diagnosis, but also support the development of personalized cancer therapy.

### Cancer genome sequences

Another high throughput technology that may revolutionize personalized cancer therapy, lead to the development of new genetic diagnostic tests, and identification of biomarkers, is next generation sequencing (NGS). The availability of the complete human genome sequence was a prerequisite for NGS to become one of the main tools for exploration of cancer genomes. NGS enables rapid, routine, and relatively inexpensive interrogation of entire cancer genomes, transcriptomes and interactomes. NGS encompasses a number of different methodologies which have emerged since 2005. This high-throughput technology is anticipated to lead to the completion of catalogue of driver mutations and increase the efficiency of detection of somatic cancer genome alterations, including nucleotide substitutions, small insertions and deletions, copy number alterations, and chromosomal rearrangements [14]. NGS led to full appreciation of the immense heterogeneity of cancer genomes when the number of somatically-acquired genetic changes are concerned. For example - some cancer genomes carry more than 100,000 point mutations whereas others have fewer than 1000. One of the suggested explanations for such diversity include previous heavy mutagenic exposures or the presence of DNA repair defects [15]. It is also anticipated that NGS will facilitate exploration of the functional roles of cancer genes, such as transcriptomes analyzed by microarrays (as described above) or interactomes, which involve analysis of DNA-protein interactions [16].

NGS holds a number of advantages over traditional methods, such as ability to fully sequence large number of genes in a single test and simultaneously detect deletions, insertions, copy number alterations, translocations and exome-wide base substitutions in cancer-related genes [14]. Advances in NGS technology are anticipated to lead to further cost reduction, increased turnaround time, and to improvements required to analyze smaller specimens, such as circulating tumor cells or circulating cell-free DNA in plasma [17]. Improvements in technology will also be required using NSG for sequencing diagnostic tumor tissue, due to small amount of heterogenous, degraded, and fixative-affected DNA extracted from Formalin-Fixed, Paraffin-Embedded (FFPE) tissue. Different methods of sequence enrichment prior to NGS application have been tested (e.g., PCR-based or hybridization capture enrichment) [18,19], but further validations are needed before the cancer diagnostics use of NGS on DNA extracted from FFPE tissues.

Analyses of NGS data represent another significant challenge. The sheer amount of data and the scale of analyses, terminology, and the need for computer language expertise all contribute to the NGS data analyses challenges. In research setting, bioinformatics experts are often engaged to develop novel platforms for data analyses, as commercial software packages offered by the producers of NGS equipment may not be as advanced as open-source tools developed by large genome sequencing centers [20]. The difficulty of NGS data analyses, together with the time it takes to analyze whole genome data make it impractical for clinical utilization. Namely, the decision on treatment would ideally be made within a few days after diagnosis was established and waiting much longer, as is currently needed to obtain a complete patient's cancer genome information, limits the clinical utility of NGS [20].

The most common applications of NGS include whole genome sequencing and whole exome sequencing. Whole genome sequencing incorporates analysis of more than 20,000 protein-coding genes as well as functional elements in intronic and intergenic DNA, while whole-exome sequencing is used to capture the 1 to 2 percent of the human genome that is protein coding and contains the majority of disease-causing mutations [15]. Thus, it is not surprising that currently much of oncology research is still focused on mutations in exons or exon-intron boundaries.

NGS can also be used to detect structural variation, including copy number alterations, which are frequent in cancer. Identification of structural variation, such as rearrangements including translocations, tandem duplications, inversions, is more challenging compared to single nucleotide variations detection, and the results obtained vary widely depending on the method used for structural variation detection [21]. Never the less, copy number variation analyses have reached their clinical application and an example is shown in table 2 (detection of HER2 amplifications in breast cancer). Whether copy number of EGFR gene predicts sensitivity to tyrosine kinase inhibitor (TKI), erlotinib has sparked contradicting reports [22]. Namely, EGFR mutations frequently demonstrate high EGFR copy number, posing a question whether the copy number alone has true predictive value [23]. This controversy has been resolved only recently, when the investigators of Iressa Pan-Asia Study (IPAS) showed that the progression free survival (PFS) benefit associated with erlotinib therapy was confined only to those patients with high copy numbers who also carried an EGFR mutation [22]. This led to the recommendation that copy number alone should not be used clinically as a selection tool of frontline treatment of NSCLC with EGFR TKIs [23].

### Epigenetics and epigenomics

NGS methods have been developed for analyses DNA methylation, detection of modified histones, mapping of transcription factor occupancy, and epigenetic regulators [20]. Analysis of DNA methylation on the level of whole genome is the most frequently used method and employs bisulphite sequencing [24]. However, single or multiple gene interrogation of DNA methylation is currently more common approach and a number of reports investigated clinical utility of epigenetic profiles as predictors of cancer risk, most notably non-small cell lung cancer (NSCLC) risk [25]. However, to our knowledge, no multi-gene or genome-wide methylation assays have yet been adopted in oncology practice.

### Conclusions and Future Directions

DNA diagnosis often trumps clinical diagnosis in oncology, as illustrated by a number of examples throughout this manuscript. Relying on DNA expression patterns or mutational signatures of cancers can lead to the identification of subgroups of patients who will fair poorly and/or not respond to specific therapies, or those with better prognosis, sensitive to therapy, when these factors could not be inferred from the clinical features. High throughput assays (microarrays, NGS) will continue to make big strides towards their clinical application and replacement of standard single-gene analyses (Sanger sequencing, PCR), even for analyses of cancer predisposition syndromes, where Sanger sequencing is still the gold standard. While NGS has a much higher throughput and significantly lower cost per analyzed base compared to Sanger sequencing, the challenges include data management and interpretation dependence on sophisticated bioinformatics analyses programs. Research efforts towards improvement of NGS sensitivity and specificity for detecting genetic abnormalities in FFPE specimens,

circulating tumor DNA or circulating tumor cells DNA isolated from plasma, are anticipated to soon lead to validation of NGS used for these challenging specimens and enhance their clinical utility. This will cause the shift toward testing minute amounts of DNA and specimens including needle biopsies, circulating tumor cells, and circulating cell-free DNA are likely to become clinically relevant. Likewise, current focus on mutations and gene copy number variations will expand to include translocations, and assessment of epigenetic changes. The NGS-based assays most useful to clinicians are those that detect genetic abnormalities that can inform treatment decisions and/or be used to direct patients to ongoing clinical trials. Thus, as NGS enters clinical testing it will impact clinical decisions and cancer outcomes. Finally, improved understanding of cancer biology will lead to further development of targeted therapies and the use of combinations of targeted approaches simultaneously, increasing treatment effectiveness and supporting highly individualized patient care.

## References

1. <http://www.ncbi.nlm.nih.gov/sites/GeneTests/lab?db=genetests> Last accessed on January 22, 2013.
2. Palmero EI, Achatz MI, Ashton-Prolla P, Olivier M, Hainaut P (2010) Tumor protein 53 mutations and inherited cancer: beyond Li-Fraumeni syndrome. *Curr Opin Oncol* 22: 64-69.
3. Mai PL, Malkin D, Garber JE, Schiffman JD, Weitzel JN, et al. (2012) Li-Fraumeni syndrome: report of a clinical research workshop and creation of a research consortium. *Cancer Genet* 205: 479-487.
4. Cronin M, Ross JS (2011) Comprehensive next-generation cancer genome sequencing in the era of targeted therapy and personalized oncology. *Biomark Med* 5: 293-305.
5. Monzon FA, Ogino S, Hammond ME, Halling KC, Bloom KJ, et al. (2009) The role of KRAS mutation testing in the management of patients with metastatic colorectal cancer. *Arch Pathol Lab Med* 133: 1600-1606.
6. Popovici V, Budinska E, Tejpar S, Weinrich S, Estrella H, et al. (2012) Identification of a poor-prognosis BRAF-mutant-like population of patients with colon cancer. *J Clin Oncol* 30: 1288-1295.
7. Dy GK (2012) EGFR tyrosine kinase inhibitors in lung cancer: clinical implications of mutations and treatment resistance. *International journal of targeted therapies in cancer* 50-54.
8. Rodig SJ, Shapiro GI (2010) Crizotinib, a small-molecule dual inhibitor of the c-Met and ALK receptor tyrosine kinases. *Curr Opin Investig Drugs* 11: 1477-1490.
9. Chapman PB, Hauschild A, Robert C, Haanen JB, Ascierto P, et al. (2011) Improved survival with vemurafenib in melanoma with BRAF V600E mutation. *N Engl J Med* 364: 2507-2516.
10. Guo J, Si L, Kong Y, Flaherty KT, Xu X, et al. (2011) Phase II, open-label, single-arm trial of imatinib mesylate in patients with metastatic melanoma harboring c-Kit mutation or amplification. *J Clin Oncol* 29: 2904-2909.
11. Copland M, Hamilton A, Elrick LJ, Baird JW, Allan EK, et al. (2006) Dasatinib (BMS-354825) targets an earlier progenitor population than imatinib in primary CML but does not eliminate the quiescent fraction. *Blood* 107: 4532-4539.
12. Roukos DH (2011) Trastuzumab and beyond: sequencing cancer genomes and predicting molecular networks. *Pharmacogenomics* 11: 81-92.
13. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, et al. (2000) Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 403: 503-511.
14. Ross JS, Cronin M (2011) Whole cancer genome sequencing by next-generation methods. *Am J Clin Pathol* 136: 527-539.
15. Stratton MR, Campbell PJ, Futreal PA (2009) The cancer genome. *Nature* 458: 719-724.
16. Roukos DH (2010) Novel clinico-genome network modeling for revolutionizing genotype-phenotype-based personalized cancer care. *Expert Rev Mol Diagn* 10: 33-48.
17. Kontic M, Ognjanovic M, Jovanovic D, Ognjanovic S (2012) Clinical utility of circulating tumor cells in non-small-cell lung carcinoma: are we there yet? *J Thorac Dis* 4: 450-452.
18. Dressman D, Yan H, Traverso G, Kinzler KW, Vogelstein B (2003) Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations. *Proc Natl Acad Sci U S A* 100: 8817-8822.
19. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, et al. (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437: 376-380.
20. Meldrum C, Doyle MA, Tothill RW (2011) Next-generation sequencing for cancer diagnostics: a practical perspective. *Clin Biochem Rev* 32: 177-195.
21. Alkan C, Coe BP, Eichler EE (2011) Genome structural variation discovery and genotyping. *Nat Rev Genet* 12: 363-376.
22. Mok TS, Wu YL, Thongprasert S, Yang CH, Chu DT, et al (2009) Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N Engl J Med* 361: 947-957.
23. Shepherd FA (2011) Molecular selection trumps clinical selection. *J Clin Oncol* 29: 2843-2844.
24. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, et al. (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* 462: 315-322.
25. Kontic M, Stojic J, Jovanovic D, Bunjevacki V, Ognjanovic S, et al. (2012) Aberrant promoter methylation of CDH13 and MGMT genes is associated with clinicopathologic characteristics of primary non-small-cell lung carcinoma. *Clin Lung Cancer* 13: 297-303.

This article was originally published in a special issue, **DNA Diagnosis** handled by Editor(s). Dr. George J. Netto, Johns Hopkins University, USA; Dr. Yang Yongliang, Harvard University, USA; Dr. Shahrokh Francois Shariat, Cornell University, USA