Clinical Applications and Limitations of Next-Generation Sequencing

Reva Kakkar Basho, MD, Agda Karina Eterovic, PhD, and FundaMeric-Bernstam, MD

Abstract

Growing interest in personalized cancer therapy has led to numerous advances in the field of cancer genomics. Next-generation sequencing (NGS) is one such development that has allowed for lower cost, higher-throughput genome sequencing. However, the vast number and types of genomic aberrations found in cancer means that interpretation of the data generated by NGS requires substantial analytical complexity. Here, we discuss the clinical applications of NGS and the obstacles that must be overcome prior to widespread use in clinical decision making.

Key words: Next-generation sequencing, review, genomics, cancer

Introduction

Personalized cancer therapy requires the use of molecular diagnostics to tailor treatments to individuals. At this time, only a few molecular biomarker-based therapies, such as erlotinib in *EGFR*-mutated lung cancer and vemurafenib in *BRAF*-mutated melanoma, have been widely accepted.[1,2] Next-generation sequencing (NGS) has the potential to revolutionize oncology through the classification of tumors and identification of biomarkers that can predict response to individualized therapy.

Until recently, the Sanger sequencing method was the most widely used sequencing method, and resulted in the only complete human genome sequence.[3] This technology relies on incorporation of chain-terminating dideoxynucleotides during DNA replication.[4] Fluorescently labeled terminators, capillary electrophoresis separation, and laser signal detection have improved the throughput of Sanger sequencing.[5] However, it remains labor-intensive, time-consuming, and expensive when done in large scale.[6] Therefore, the demand for faster, more accurate, and more cost-effective genomic information has led to the development of NGS methods.

NGS methods are high-throughput technologies with capabilities of sequencing large numbers of different DNA (massively parallel) sequences at once. NGS technologies monitor the sequential addition of nucleotides to immobilized DNA templates generated from target tissue.[7] Unfortunately, the increased throughput of NGS reactions comes at the cost of shorter sequences, as most sequencing platforms (Illumina, Roche, SoLiD) offer shorter read lengths (30–400 bp) than the conventional Sanger-based method.[8] These shorter sequences are then assembled into longer sequences such as complete genomes.

Common approaches to DNA sequencing include whole-genome sequencing, whole-exome sequencing, targeted exome sequencing, and "hotspot" sequencing. Whole-genome sequencing sequences the complete genome of a sample (ie, chromosomal DNA and mitochondrial DNA, which includes intronic and exonic regions). Whole-exome sequencing is a technique that sequences all of the protein-coding genes (ie, all exons in the genome). Targeted exome sequencing uses target-enrichment methods to capture genes of interest. This approach is becoming increasingly popular in oncology for assessing the full sequence of cancer-related gene panels. Targeted exome sequencing also facilitates sequencing at a greater depth, and thus the identification of subclonal mutations. Alternately, rather than sequencing the full sequence of selected genes, only selected regions of selected genes can be sequenced, focusing on cancer gene "hotspots"—regions with recurrent mutations. Although hotspot mutation testing facilitates large-scale sequencing of many samples, it does limit the knowledge that is acquired through sequencing because it limits the evaluation to small regions in selected genes. Consequently, it increases the possibility of omitting relevant mutations for which evaluation is not being conducted, thus limiting the clinical knowledge that is gained through NGS. Despite its drawbacks, it is becoming a widely accepted form of NGS.

In addition to nucleotide change detection (mutations and small insertions and deletions), NGS allows for DNA-copy number predictions. Further, NGS technology also can be applied to RNA in order to evaluate the transcriptome of a

Practical Application

- Provide a brief introduction to the methods of next-generation sequencing (NGS)
- Identify clinical applications of NGS including the identification of various molecular aberrations in different tumor types, the resultant design of molecular biomarker-driven clinical trials, and the potential to identify molecular aberrations that lead to disease progression and resistance
- Identify limitations of NGS, including the need for extensive analytic capabilities, the difficulties in identification of driver mutations, and the confounding factor of tumor heterogeneity
- Identify potential future applications of NGS

tumor. RNA-sequencing (RNA-seq) allows for the assessment of gene expression and transcriptional splice variant analysis in addition to detection of mutations. A typical NGS work flow from sample collection to the capture and sequencing of genes of interest and data analysis is illustrated in **Figure 1**.

Identification of Cancer Genomics

In recent years, NGS has been used to characterize genomic alterations such as mutations, insertions/deletions, and copy number changes, and the frequency with which they occur in various tumor types. Efforts such as the International Cancer Genome Consortium (ICGC) and The Cancer Genome Atlas (TCGA) aim to catalog such genomic alterations across many tumor types.[9,10] However, the wealth of information that is generated through this process unveils potentially the largest hurdle of genomic medicine: How do we analyze the abundance of information that is generated to make informed decisions regarding therapy? Analysis of cancer genomes reveals that most tumors contain multiple alterations.[11-14] As a result, it is very important to distinguish the "driver" mutations that contribute to tumor development from the "passengers" that do not.[15]

Comparison of sequenced genomes to reference genomes allows for the identification of genome alterations that may be relevant in disease development and progression.[16] However, such comparison depends on the establishment of extensive and accurate reference genomes, which is a cumbersome task. Further, the complexity of genomic aberrations in cancer makes it difficult to rely on standard reference genomes.[8] Therefore, simplified methods of identifying driver mutations are required. Several theories exist for the potential identification of driver mutations. One such hypothesis is that mutations that occur with higher frequency are more likely to contribute to tumor development and growth.[17] Genome-wide association studies (GWAS) aim to compare the incidence of commonly known single nucleotide polymorphisms (SNPs) in genomes from patients with and without a specific disease. SNPs that occur at a higher frequency in the diseased population are identified as potentially causative. If a specific mutation is not found in high frequency, but the same molecular pathway contains frequent genomic alterations, those alterations may also be relevant.

Another theory is that alterations present in both germline and tumor tissue of the same patient are likely to be integral to tumor development. For example, some mutations in cancer-predisposition genes such as *BRCA1/2* clearly do contribute to the development and maintenance of cancer. This, however, requires that germline tissue be collected in each patient. Yet another theory is that sequencing DNA and RNA from the same sample will identify mutations that subsequently alter expression, and are thus significant. However, all of these methods only begin to narrow the spectrum of genomic alterations that may be clinically relevant. Chromosome-scale changes and epigenomic changes cannot be evaluated in this manner. Many studies are now focusing on the development of bioinformatic tools to aid in the identification of driver mutations.[18]

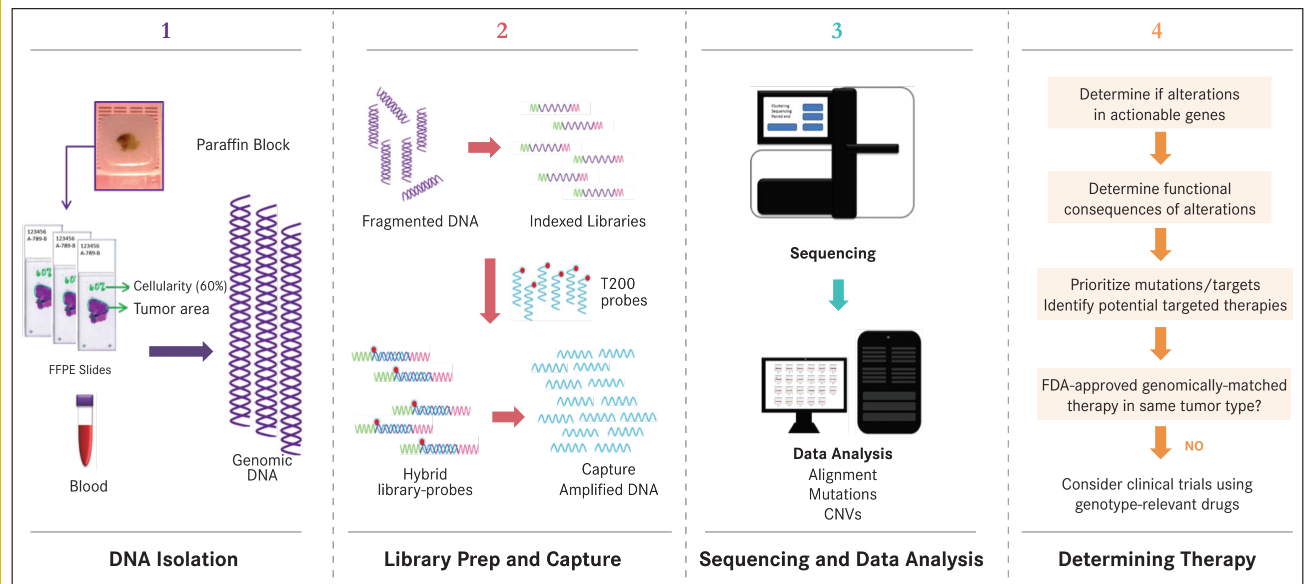
Clinical Decision Support

Once driver mutations have been identified in a tumor, the next step is to assess whether those mutations are "actionable." Actionable alterations affect the function of a cancer-related gene and can be targeted with approved or investigational therapies. Assessing functionality is a difficult task and requires predictive knowledge of genome alterations. Often, early-phase studies are used to assess the role various mutations based on rates of response to targeted therapies. However, enrollment in such studies requires that physicians be aware of genome alterations and potential trials for each patient.

A survey of 160 physicians at a tertiary-care National Cancer Institute (NCI)-designated comprehensive cancer center revealed that a considerable percentage of physicians have low confidence in their genomic knowledge.[19] As a result, many institutions have instituted tumor boards to increase awareness of and access to appropriately targeted therapies.[20] Similarly, the American Society of Clinical Oncology has monthly presentations that explore current treatment strategies and novel therapeutics in various tumor types to increase knowledge of newer targeted therapies. Other trials such as NCI Molecular Analysis for Therapy Choice Program (NCI-MATCH) has streamlined the decision making by designing algorithms and creating rules to designate alterations as actionable, and to prioritize targets if more than one target is identified. In this signal-seeking trial, 3000 patients will undergo tumor NGS to match genomic alterations to smaller histology-agnostic phase 2 trials of Food and Drug Administration (FDA)-approved agents (in other diseases) and investigational therapeutics (**Figure 2**).

If a response signal is seen in early-phase trials, the clinical relevance and therapeutic implications of actionable mutations can be assessed through thoughtful biomarker-driven research. Hypothesis-driven preclinical studies and clinical trials to assess targeted therapies in various tumor types can be designed. Such trials allow for the recruitment of selected patients into clinical

FIGURE 1. Overview of a Potential Next-Generation Sequencing Work Flow



(1) Slides are cut from tumor samples embedded in paraffin blocks. For each tumor sample, hematoxylin and eosin stains are performed and cellularity is assessed. Matching peripheral blood is also collected for each patient. Genomic DNA is isolated from both the formalin-fixed, paraffin-embedded tissue and blood. Alternately, frozen fresh tissue samples as well as other normal DNA sources such as saliva, buccal swab, or normal tissue can be used. (2) DNA is fragmented and libraries are made by ligating indexed adaptors (Indexed Libraries) that allow for sample pooling. Hybridization with probes is performed; the captured DNA is washed and amplified and proceeded to DNA sequencing. (3) Captured DNA is sequenced; after sequencing, samples are demultiplexed, or separated, and the raw data is submitted to data analysis for mutations and copy number variations identification. (4) The genomic alterations are reviewed and alterations in actionable genes are identified. Functional impact of alterations in actionable genes is assessed and therapeutic implications of known and predicted functional alterations are determined.

Modified from Chen et al. Clin Chem.2015.[33]

trials to enhance the assessment of those targeted therapies. Ultimately, the goal is to implement randomized clinical trials to assess molecularly targeted therapy in a biomarker-selected or biomarker-stratified fashion. The Adjuvant Lung Cancer Enrichment Marker Identification and Sequencing Trials (ALCHEMIST) is an example of such a trial in which patients with early-stage adenocarcinoma of the lung are screened for *EGFR* and *ALK* mutations, and subsequently randomized into trials of relevant targeted therapy if mutations are found. With NGS technology, a high throughput of patients can undergo testing to assess their eligibility for clinical trials within a clinically reasonable timeframe.

Genomic Evolution and Intertumor and Intratumor Heterogeneity

Further complicating the implementation of genomic medicine is the fact that driver mutations can evolve during the course of cancer. As tumors are treated or as they grow, a variety of acquired genomic alterations may emerge. For example, melanoma treated with BRAF or MEK inhibitors has been shown to acquire *BRAF* amplifications and downstream alterations that lead to reactivation of the MAP kinase pathway.[21-23] Similarly, increased signaling via the phosphatidylinositol 3-kinase/Akt pathway may contribute to trastuzumab resistance in HER2-positive breast cancer.[24] Thus, the dynamic nature of cancer requires that genomic information be applicable in real time in order for clinical use. As a result, archived tissue from biopsies may not be relevant for therapy selection at the time of progression.

In addition to genomic evolution, tumors may also develop intertumor and intratumor heterogeneity. Intertumor heterogeneity refers to differences in alterations of tumors at different sites, while intratumor heterogeneity refers to differences in alterations within a tumor. Both intertumor and intratumor heterogeneity can further complicate the determination of relevant mutations because it means that tissue for NGS has to be obtained from relevant sites as well as at a relevant time point

in the treatment course. This can result in repeated biopsies. Additionally, metastatic sites such as bone and brain can be difficult to test. However, comparison of primary tumors with matched metastases has shown relatively high concordance in their mutational profiles, suggesting that additional biopsies may not always be necessary.[25,26]

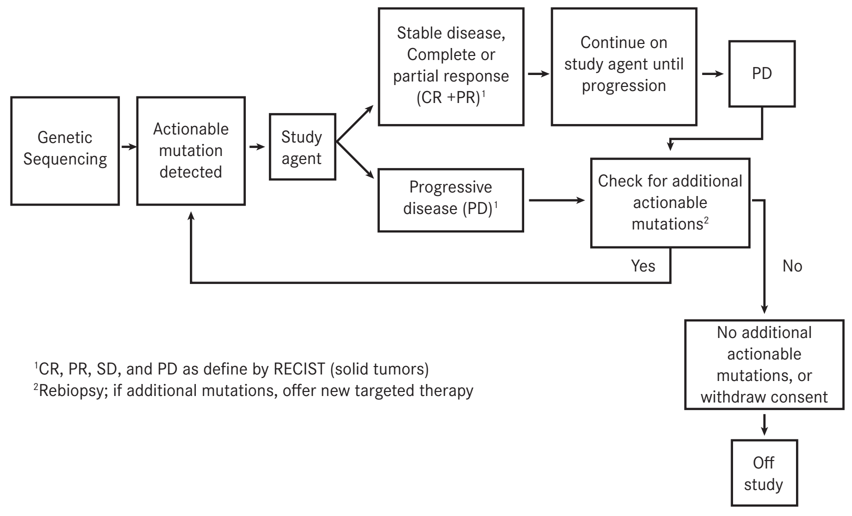
Although genomic evaluation makes it difficult to identify relevant aberrations, recognizing genomic evolution is a powerful tool to better understand the progression of cancer. Genomic analysis of cancer at different stages, from precancerous lesions to localized tumors to metastatic disease, can identify genetic events that drive tumor growth. For example, genomic studies that analyze genomic alterations in breast ductal carcinoma in situ (DCIS) can help to design a predictive model for lesions that are likely to progress to carcinoma versus those that are not.[27] Similarly, NGS-based analysis of drug-resistant cells can help identify mechanisms of resistance. For instance, sequencing tumors from patients with estrogen receptor (ER)–positive breast cancer that recurred or progressed after treatment with antiestrogen therapy revealed mutations in the *ESR1* gene; these mutations were constitutively active.[28] Interestingly, *ESR1* gene mutations were not seen in TCGA analysis, which included primary tumors only.[11] Together, these studies suggest that activating mutations in the *ESR1* gene are an acquired mechanism of resistance to antiestrogen therapy. Similarly, RNA sequencing of tamoxifen-sensitive and -resistant breast cancer cells revealed gene expression changes implicating a series of resistance mechanisms that could be grouped in ER functions, cell cycle regulation, transcription/translation, and mitochondrial dysfunction.[29]

Future Applications and Directions

Several additional applications of NGS are under development. One potential future application of NGS is the evaluation of circulating tumor cells or free-plasma DNA to detect early relapse or residual cancer.[20] Once tumor-specific genome alterations have been identified by NGS, PCR assays could be used to detect circulating tumor cells or free-plasma DNA harboring the same alterations. Disease status, drug responsiveness, and relapse could be serially assessed. The monitoring strategy would, however, require that the mutation being tested be present in all tumor cells and remain present throughout the course of disease. As discussed previously, due to genomic evolution and tumor heterogeneity, such mutations are difficult to identify. Optimally, mutations used for monitoring would be truncal mutations—mutations in the "trunk-branch" model of heterogeneity, and thus representing ubiquitous driver mutations present in every tumor subclone and region.[31] However, serial monitoring could also identify new alterations that occur under the selection pressure of treatment, which could give insights into mechanisms of acquired resistance.

Another potential application of NGS is to improve the diagnosis of cancer. Poor tissue sampling and processing can often make a histological diagnosis difficult. Additionally, mixed tumor phenotypes can sometimes make it difficult to determine the origin of the tumor. However, NGS-based analysis of tissue can be performed on small amounts of viable tissue and is accurate when sufficient information regarding causative mutations is known. An evaluation of 143 benign and malignant thyroid nodules revealed that genotyping of fine-needle aspiration (FNA) samples of the nodules using a broad NGS panel provided high sensitivity and specificity in the diagnosis of these samples.[32] Such diagnoses would require clinical validation prior to widespread use. Furthermore, as the genomics of different tumors become apparent, NGS can be

FIGURE 2. Overview of NCI-MATCH Trial Design



A schematic of the NCI-MATCH trial design. Patient tumors undergo genetic sequencing and are assessed for actionable mutations. If found, patients are enrolled in smaller phase 2 trials of approved or investigational therapeutics until progression of disease. At the time of disease progression, patient tumors are again assessed for other actionable mutations. If found, patients are again enrolled in another phase 2 trial. If no further actionable mutations are found, patients are taken off study.

Figure adapted from Abrams et al. ASCO *Educational Book*. 2014.[34]

used to identify different molecular subtypes, which is already becoming commonplace with sarcoma fusion proteins.

Finally, NGS can identify molecular aberrations that render tumors exquisitely sensitive to certain therapies, resulting in exceptional responses. Such extraordinary outcomes can improve our understanding of molecular features that can predict response to certain drugs. For this purpose, the NCI has undertaken the Exceptional Responders Initiative, through which tumors of exceptional responders will undergo DNA and RNA sequencing to define genetic alterations that might have resulted in such responses.

Conclusion

Next-generation sequencing has opened a broad new area of research with the potential to revolutionize personalized cancer medicine. However, further development of this field requires real-time knowledge of genome alterations that can be used in clinical decision making. This requires a robust data infrastructure, continuous improvement in sequencing technology, development of analytical tools, and ongoing biomarker-driven preclinical and clinical trials. Ultimately, however, NGS data have the potential to guide clinicians in tailoring treatment to dynamic genomic changes in individual tumors.

Affiliations: Reva Kakkar Basho, MD, is from the Department of Cancer Medicine; Agda Karina Eterovic, PhD, is from the Department of Systems Biology; Funda Meric-Bernstam, MD, is from the Departments of Investigational Cancer Therapeutics, Surgical Oncology, and the Sheikh Khalifa Bin Zayed Al Nahyan Institute for Personalized Cancer Therapy at The University of Texas MD Anderson Cancer Center, Houston, TX.

Disclosures: Drs Basho and Eterovic have no relevant conflicts of interest to disclose. Dr Meric-Bernstam has served as a consultant or on a paid advisory board for Novartis and Roche; has received grants from Novartis, AstraZeneca, Taiho Pharmaceutical, Genentech, Calithera Biosciences, Debiopharm, and Bayer; has received honoraria from Genentech, Roche Diagnostics, and Sysmex Corporation; and has served on an advisory committee or review panel for Genentech, Novartis, and Roche.

Funding support: This work was supported in part by the Sheikh Khalifa Bin Zayed Al Nahyan Institute for Personalized Cancer Therapy and the MD Anderson Cancer Center Support grant (P30 CA016672).

Address correspondence to: Funda Meric-Bernstam, MD, UT MD Anderson Cancer Center, 1400 Holcombe Boulevard, Unit 455, Houston, TX 77030. Phone: 713-792-3737; fax: 713-563-0566; email: fmeric@mdanderson.org.

REFERENCES

1. Rosell R, Carcereny E, Gervais R, et al. Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced EGFR mutation-positive non-small-cell lung cancer (EURTAC): a multicentre, open-label, randomised phase 3 trial. *Lancet Oncol*. 2012;13(3):239-246.

2. Chapman PB, Hauschild A, Robert C, et al. Improved survival with vemurafenib in melanoma with BRAF V600E mutation. *N Engl J Med*. 2011;364(26):2507-2516.

3. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature*. 2004;431(7011):931-945.

4. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA*. 1977;74(12):5463-5467.

5. Rizzo JM, Buck MJ. Key principles and clinical applications of "next-generation" DNA sequencing. *Cancer Prev Res*. 2012;5(7):887-900.

6. Metzker ML. Emerging technologies in DNA sequencing. *Genome Res*. 2005;15(12):1767-1776.

7. Metzker ML. Sequencing technologies—the next generation. *Nat Rev Genet*. 2010;11(1):31-46.

8. Ulahannan D, Kovac MB, Mulholland PJ, et al. Technical and implementation issues in using next-generation sequencing of cancers in clinical practice.*Br J Cancer.* 2013;109(4):827-835.

9. International Cancer Genome Consortium; Hudson TJ, Anderson W, Artez A, et al. International network of cancer genome projects. *Nature*. 2010;464(7291):993-998.

10. Cancer Genome Atlas Research Network; Weinstein JN, Collisson EA, Mills GB, et al. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet*. 2013;45(10):1113-1120.

11. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature*. 2012;490(7418):61-70.

12. Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature*. 2012;487(7407):330-337.

13. Cancer Genome Atlas Network. Integrated genomic analyses of ovarian carcinoma. *Nature*. 2011;474(7353):609-615.

14. Shah SP, Roth A, Goya R, et al. The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature*. 2012;486(7403):395-399.

15. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144(5):646-674.

16. Nielsen R, Paul JS, Albrechtsen A, Song YS. Genotype and SNP calling from next-generation sequencing data. *Nat Rev Genet*. 2011;12(6):443-451.

17. Sjoblom T, Jones S, Wood LD, et al. The consensus coding sequences of human breast and colorectal cancers. *Science*. 2006;314(5797):268-274.

18. Merid SK, Goranskaya D, Alexeyenko A. Distinguishing between driver and passenger mutations in individual cancer genomes by network enrichment analysis. *BMC Bioinformatics*.

2014;15:308.
19. Gray SW, Hicks-Courant K, Cronin A, et al. Physicians' attitudes about multiplex tumor genomic testing. *J Clin Oncol*. 2014;32(13):1317-1323.
20. Schwaederle M, Parker BA, Schwab RB, et al. Molecular tumor board: the University of California-San Diego Moores Cancer Center experience. *Oncologist*. 2014;19(6):631-636.
21. Wagle N, Van Allen EM, Treacy DJ, et al. MAP kinase pathway alterations in BRAF-mutant melanoma patients with acquired resistance to combined RAF/MEK inhibition. *Cancer Discov*. 2014;4(1):61-68.
22. Mao Y, Chen H, Liang H, et al. CanDrA: cancer-specific driver missense mutation annotation with optimized features. *PLoS One*. 2013;8(10):e77945.
23. Van Allen EM, Wagle N, Sucker A, et al. The genetic landscape of clinical resistance to RAF inhibition in metastatic melanoma. *Cancer Discov*. 2014;4(1):94-109.
24. Nahta R, Yu D, Hung MC, et al. Mechanisms of disease: understanding resistance to HER2-targeted therapy in human breast cancer. *Nat Clin Pract Oncol*. 2006;3(5):269-280.
25. Meric-Bernstam F, Frampton GM, Ferrer-Lozano J, et al. Concordance of genomic alterations between primary and recurrent breast cancer. *Mol Cancer Ther*. 2014;13(5):1382-1389.
26. Meric-Bernstam F, Frampton GM, Ferrer-Lozano J, et al. Concordance of genomic alterations between primary and recurrent breast cancer. *Mol Cancer Ther*. 2004;13(5):1382-1389.
27. Kaur H, Mao S, Shah S, et al. Next-generation sequencing: A powerful tool for the discovery of molecular markers in breast ductal carcinoma in situ. *Expert Rev Mol Diagn*. 2013;13(2):151-165.
28. Segal CV, Dowsett M. Estrogen receptor mutations in breast cancer—new focus on an old target. *Clin Cancer Res*. 2014;20(7):1724-1726.
29. Huber-Keener KJ, Liu X, Wang Z, et al. Differential gene expression in tamoxifen-resistant breast cancer cells revealed by a new analytical model of RNA-Seq data. *PLoS One*. 2012;7(7):e41333.
30. McBride DJ, Orpana AK, Sotiriou C, et al. Use of cancer-specific genomic rearrangements to quantify disease burden in plasma from patients with solid tumors. *Genes Chromosom Cancer*. 2010;49(11):1062-1069.
31. Yap TA, Garlinger M, Futreal PA, et al. Intratumor heterogeneity: seeing the wood for the trees. *Sci Transl Med*. 2012;4(127):127ps10.
32. Nikiforov YE, Carty SE, Chiosea SI, et al. Highly accurate diagnosis of cancer in thyroid nodules with follicular neoplasm/suspicious for a follicular neoplasm cytology by ThyroSeq v2 next-generation sequencing assay. *Cancer*. 2014;120(23):3627-3634.
33. Chen K, Meric-Bernstam F, Zhao H, et al. Clinical actionability enhanced through deep targeted sequencing of solid tumors [published online January 27, 2015]. *Clin Chem*. 2015.pii:clinchem.2014.231100.
34. Abrams J, Conley B, Mooney M, et al. National Cancer Institute's Precision Medicine Initiatives for the New National Clinical Trials Network. *ASCO Educational Book*. 2014;71-76.