Next-Generation Sequencing of Lung Cancers

Lessons Learned and Future Directions

Siddhartha Devarakonda, MD[a], Ashiq Masood, MD[b], Ramaswamy Govindan, MD[c,*]

KEYWORDS

• Sequencing • Lung cancer • Treatment

KEY POINTS

- Lung cancer genomes from smokers have a high burden of mutations. This high mutational burden poses a challenge for the discovery of low-frequency driver alterations.
- Sequencing a large number of tumor samples and combining genomic data from multiple cancer types for analysis can yield enough statistical power to identify low-frequency driver alterations in cancer genomes — some of which may be targetable.
- Although different subtypes of lung cancers share certain genomic alterations, the majority of these alterations tend to be histology specific. It is possible that the heterogeneity in mutational processes underlying malignant transformation and differences in the cell of origin account for this observation.
- The clonal architecture of cancers is complex. The role of clonal heterogeneity as a prognostic and predictive biomarker is currently being investigated.
- Whole-exome and whole-genome sequencing data have the potential to guide immunotherapy and cancer vaccine development.

INTRODUCTION

Lung cancer continues to remain a serious global problem and one of the leading causes of cancer-related death worldwide.[1] The past decade has witnessed significant advances in next-generation sequencing technologies, which have made it

Equal contribution (S. Devarakonda and A. Masood).
[a] Division of Oncology, Washington University School of Medicine, 660 South Euclid Avenue, St Louis, MO 63110, USA; [b] Division of Oncology, Alvin J Siteman Cancer Center, Washington University School of Medicine, 660 South Euclid Avenue, Box 8056, St. Louis, MO 63110, USA; [c] Section of Medical Oncology, Division of Oncology, Alvin J Siteman Cancer Center, Washington University School of Medicine, 660 South Euclid Avenue, Box 8056, St. Louis, MO 63110, USA
* Corresponding author.
E-mail address: rgovindan@wustl.edu

Hematol Oncol Clin N Am 31 (2017) 1–12
http://dx.doi.org/10.1016/j.hoc.2016.08.008 **hemonc.theclinics.com**

possible to study cancer genomes in unprecedented detail and gain a better understanding of the alterations that underlie cancer development and progression.[2–5] Comprehensive genomic analyses of lung cancer have been reported by several groups and consortia, such as The Cancer Genome Atlas (TCGA).[4–7] The aim of this review is to highlight some of the important findings reported in these studies and discuss their clinical significance.

SOMATIC MUTATIONS IN LUNG CANCER

Acquisition of somatic point mutations is one of the common mechanisms by which normal cells undergo malignant transformation. Cancer cells continually accrue a variety of mutations and are exposed to stresses, such as hypoxia, treatment, and attacks by the host immune system. As a result, cancer cells with mutations that confer upon them a survival advantage (often referred to as driver mutations) are selected over time.[8] Because driver mutations increase the survival fitness of cancer cells, these mutations are likely to be over-represented and recurrent in cancer samples that are obtained from different patients, compared with other bystander or passenger mutations that do not offer a growth advantage to the cancer cell.[8] Studies often use sophisticated statistical algorithms to identify significantly mutated genes.[3,9,10] These algorithms take into account several factors that influence mutation rate, such as gene size, background mutation rate, DNA repair mechanisms (genes that are more actively transcribed into RNA have lower mutation burdens due to transcription-coupled repair), and replication timing (genes that are replicated later during cell division are more prone to mutations).[9,10] Although these statistical predictions by themselves do not imply a biological role for genes in cancer, they can be extremely useful in identifying gene alterations for further functional studies.

Lung cancer genomes have a high burden of mutations, with approximately 8 mutations/megabase (Mb) or million base-pairs compared with other cancer types, such as pediatric tumors or acute leukemias.[9] This high mutation burden in lung cancers is attributed to cigarette smoke exposure and abnormal activity of cell intrinsic mutagenic processes, such as APOBEC cytidine deaminase enzymes, and poses a challenge for identifying low-frequency driver alterations from passenger alterations.[11] Exposure to cigarette smoke also results in a characteristic mutation pattern in tumors. Genomes of tumors from smokers are enriched for transversions, where a pyrimidine (cytosine or thymine) is replaced by a purine (adenine or guanine) or vice versa.[4] In contrast, genomes of lung adenocarcinomas (LUADs) from never-smokers have an approximately 10-fold lower mutation burden (0.6 mutations/Mb) and are enriched for transitions, where a purine is replaced by a purine or a pyrimidine by a pyrimidine.[12]

Using statistical algorithms, studies have reported recurrent mutations in tumor suppressors, such as *TP53, STK11, NF1, RB1, PTEN*, and *CDKN2A*, and oncogenes, such as *KRAS, EGFR, MET*, and *PIK3CA* in lung cancer.[4,5,13] While some of these alterations are shared by different subtypes of lung cancer, others are histology specific. For instance, LUADs are characterized by mutations in genes, such as *EGFR, KRAS, BRAF, ERBB2*, and *MET*, that activate the receptor tyrosine kinase (RTK)/RAS/RAF signaling pathway.[4] Unlike LUADs, small cell lung cancers (SCLCs) rarely show mutations in the RTK/RAS/RAF signaling pathway.[13,14] These tumors are typically characterized by comutation of the tumor suppressors, *TP53* and *RB1*, and alterations in genes that regulate neuroendocrine differentiation.[13] Similarly, although LUAD and squamous cell lung cancer (SQLC) share mutations in genes, such as *PIK3CA,*