Commentary

Identification of molecular subtypes and significantly mutated genes in gastric cancer using next-generation sequencing

Byungho Lim[1], **Seon-Young Kim**[1,2]

[1]Personalized Genomic Medicine Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon, Korea; [2]Department of Functional Genomics, University of Science and Technology, Daejeon, Korea

Correspondence to: Dr. Seon-Young Kim, PhD. Personalized Genomic Medicine Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), 125 Gwahak-ro, Yuseong-gu, Daejeon 305-806, Korea. Email: kimsy@kribb.re.kr.

Comment on: Li X, Wu WK, Xing R, *et al.* Distinct Subtypes of Gastric Cancer Defined by Molecular Characterization Include Novel Mutational Signatures with Prognostic Capability. Cancer Res 2016;76:1724-32.

Submitted May 20, 2016. Accepted for publication May 27, 2016.
doi: 10.21037/tcr.2016.05.27
View this article at: http://dx.doi.org/10.21037/tcr.2016.05.27

Gastric cancer is a complex disease that is caused by interactions among multiple genetic and extrinsic factors. The heterogeneous characteristics of gastric cancer make it difficult to select suitable treatment options for individual patients. According to a clinicopathologic perspective, gastric cancer is divided into two subtypes, intestinal or diffuse, which is also known as Lauren's classification. This classification helps to understand the pathogenesis of gastric cancer, but it is insufficient as a predictor of disease prognosis and drug treatment. Thus, more elaborate subgrouping of individual patients is required to develop a personalized therapeutic regimen.

In efforts to stratify gastric cancer patients, recent next-generation sequencing (NGS) studies have identified previously unrecognized molecular subtypes of gastric cancer. As a representative study, The Cancer Genome Atlas (TCGA) network classified gastric cancer into four subtypes, including Epstein-Barr virus (EBV)-positive, microsatellite instable (MSI), genomically stable (GS), and chromosomal instability (CIN) subtypes (1). This study used both sequencing and array-based approaches by investigating exome sequences, copy-number alterations, gene expression, DNA methylation, and protein activities in gastric cancer (2). As a result, this classification provides the most well-defined gastric cancer molecular subtypes to date. Nonetheless, more research is needed to determine significantly mutated genes, druggable targets, or prognosticators that represent each subtype.

In a recent report published in *Cancer Research*, Li *et al.* attempted to define novel molecular subtypes of gastric cancer using methodologies different from those of TCGA (3). In contrast to previous clustering strategies where transcriptome or methylome profiles were generally used, this study used only mutation data for clustering. First, a classification was conducted based on mutation loads by analyzing a total of 544 gastric cancer patients combined from five previous whole genome or whole exome studies. This unsupervised clustering stratified gastric cancer into two subtypes, referred to as regular-mutated (2.4 mutations/Mb; range from 0 to 8.3) and hyper-mutated (20.5 mutations/Mb; range from 9.6 to 200.2). The hyper-mutated subtype from Li *et al.* overrepresented MSI patients, but this subtype appears to be slightly different from TCGA's MSI subtype, as the hyper-mutated phenotype can be generated by defects of various genomic integrity maintenance mechanisms as well as mismatch repair, which is related to the MSI phenotype. Importantly, previous elegant studies from Rizvi *et al.* and Le *et al.* demonstrated that patients with high mutation loads showed sensitive responses to therapeutic blockade against immune checkpoints such as PD-1 (4,5). Thus, the MSI subtype from TCGA or the hyper-mutated subtype from Li *et al.* would clinically benefit from immune checkpoint blockade with PD-1 inhibitors (i.e., pembrolizumab) (5), suggesting the clinical relevance of the hyper-mutated subtype from Li *et al.*

The different mutation loads between subtypes indicate that tumors from different subtypes may go through different mutation processes during tumor evolution. In

fact, mutation signatures between the regular- and hyper-mutated subtypes were distinct. Notably, the regular-mutated subtypes acquired six times more mutations at TpCpA/T sequences than the hyper-mutated subtype (3). Given that TpCpA/T is a sequence motif for the DNA cytosine deaminase APOBEC3B (6), this result suggested that APOBEC3B may be involved in the mutation loads of the regular-mutated or microsatellite stable (MSS) subtypes. Moreover, the number of APOBEC signature mutations positively and significantly correlated with the *APOBEC3B* mRNA level. However, these data were interpreted with caution because the APOBEC signature of the hyper-mutated subtype can be attenuated by other predominant mutation processes, such as a deficiency in mismatch repair. Supporting this suspicion, the *APOBEC3B* mRNA levels were comparable between two subtypes. Therefore, the fact that the APOBEC signature more strongly contributes to the mutation processes of the regular-mutated subtype may be true, but we cannot say that the APOBEC signature is not important in the hyper-mutated subtype.

Li *et al.* further clustered the regular-mutated subtype into two groups, referred to as C1 and C2. This clustering was performed based on a binary mutation status matrix of significantly mutated genes that were found by three algorithms, MutSigCV, MutSigCL, and MutSigFN (7). The C1 cluster contained a significantly high proportion of TCGA's CIN subtype patients, whereas the C2 cluster was enriched with TCGA's GS subtype patients. Reflecting these patterns, *TP53* mutations, which are highly detected in the CIN subtype, were enriched in the C1 cluster, and *RHOA* and *CDH1* mutations, which are the representative mutations of the GS subtype and diffuse-type gastric cancer, were overrepresented in the C2 cluster. More importantly, the C2 cluster displayed poorer prognostic outcome than the C1 cluster, suggesting that the high proportion of GS subtype or diffuse-type GC within the C2 cluster could result in poor survival rates. However, multivariate analysis revealed that C1/2 clusters alone have prognostic value independent of Lauren's classification and the TNM staging. For easier C1/2 clustering, Li *et al.* defined eight genes (*TP53*, *ARID1A*, *CDH1*, *PIK3CA*, *XIRP2*, *APC*, *ERBB2*, and *RHOA*) as a classifier. Because the eight-gene classifier showed enough power to discriminate C1/2 clusters and acted as a significant independent prognostic marker, this method could be applied to gastric cancer patients for clinical uses.

Mutation data combined from 544 gastric cancer patients enabled detection of previously unrecognized significantly mutated genes with a high statistical power. In the regular-mutated subtype, 31 significantly mutated genes were identified. In addition to well-known cancer driver genes such as *TP53*, *ARID1A*, *CDH1*, *PIK3CA*, *APC*, *RHOA*, *SMAD4*, *ERBB4*, *KRAS*, *ERBB2*, and *CTNNB1*, Li *et al.* found unreported significantly mutated genes, including *XIPR2*, *NBEA*, *COL14A1*, *AKAP6*, *CNBD1*, and *ITGAV*. *NBEA* is located on chromosome 13 and is frequently deleted in multiple myeloma (8,9). *COL14A1* is down-regulated by promoter hyper-methylation in renal cell carcinoma (10). *AKAP6* and *CNBD1* are mutated in esophageal adenocarcinoma and gastroenteropancreatic neuroendocrine tumors, respectively (11,12). The functional roles of *XIPR2*, compared to those of other genes, have not been elucidated in the cancer field. Because *XIPR2* is one of the eight genes used in the classifier for C1/2 clustering, preferential functional validation is required to demonstrate the activity and role of *XIPR2* in gastric cancer.

Although this study revealed novel aspects of gastric cancer subtypes, there are several limitations. One limitation of this study is the lack of functional validation of the novel significantly mutated genes that were identified. Another limitation is that the statistical power to detect less frequent mutations (below 2%) is still not high enough with the current 544 samples. Interestingly, Lim *et al.* recently analyzed a total of 629 gastric cancer patients and identified several significantly mutated genes (i.e., *DHFR*, *GHSR*, *GLI3*, *GRM8*, *KIF2B*, and *PREX2*) (2) not reported in the Li *et al.* analysis. With more samples, more significantly mutated genes are expected to be discovered. As another limitation, it is unclear whether the subtyping from Li *et al.* is superior to the TCGA's gastric cancer subtypes. For instance, the C1 cluster from Li *et al.* contained a significantly enriched number of *ARID1A* and *PIK3CA* mutations, which were included in TCGA's EBV subtype, suggesting that the C1 cluster may be a mixture of TCGA's EBV subtype and other subtypes. Even four subtypes from TCGA may oversimplify the complexity of gastric cancer. Therefore, further efforts to categorize each molecular subtype from TCGA and Li *et al.* may be required to facilitate precision medicine of gastric cancer patients.

Acknowledgments

Funding: This work was supported by the grants funded by the Ministry of Science, ICT, and Future Planning (NRF-2012M3A9D1054670 and NRF-2014M3C9A3068554) and KRIBB Research Initiative.

Footnote

Provenance and Peer Review: This article was commissioned and reviewed by the Section Editor Zhen-Yu Lin, MD (Cancer center, Union hospital, Huazhong University of Science and Technology Wuhan, China).

Conflicts of Interest: Both authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/tcr.2016.05.27). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Cancer Genome Atlas Research Network. Comprehensive molecular characterization of gastric adenocarcinoma. Nature 2014;513:202-9.
2. Lim B, Kim JH, Kim M, et al. Genomic and epigenomic heterogeneity in molecular subtypes of gastric cancer. World J Gastroenterol 2016;22:1190-201.
3. Li X, Wu WK, Xing R, et al. Distinct Subtypes of Gastric Cancer Defined by Molecular Characterization Include Novel Mutational Signatures with Prognostic Capability. Cancer Res 2016;76:1724-32.
4. Rizvi NA, Hellmann MD, Snyder A, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. Science 2015;348:124-8.
5. Le DT, Uram JN, Wang H, et al. PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. N Engl J Med 2015;372:2509-20.
6. Burns MB, Lackey L, Carpenter MA, et al. APOBEC3B is an enzymatic source of mutation in breast cancer. Nature 2013;494:366-70.
7. Lawrence MS, Stojanov P, Polak P, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature 2013;499:214-8.
8. Fonseca R, Oken MM, Harrington D, et al. Deletions of chromosome 13 in multiple myeloma identified by interphase FISH usually denote large deletions of the q arm or monosomy. Leukemia 2001;15:981-6.
9. O'Neal J, Gao F, Hassan A, et al. Neurobeachin (NBEA) is a target of recurrent interstitial deletions at 13q13 in patients with MGUS and multiple myeloma. Exp Hematol 2009;37:234-44.
10. Morris MR, Ricketts C, Gentle D, et al. Identification of candidate tumour suppressor genes frequently methylated in renal cell carcinoma. Oncogene 2010;29:2104-17.
11. Dulak AM, Stojanov P, Peng S, et al. Exome and whole-genome sequencing of esophageal adenocarcinoma identifies recurrent driver events and mutational complexity. Nat Genet 2013;45:478-86.
12. Lim B, Kim C, Kim JH, et al. Genetic alterations and their clinical implications in gastric cancer peritoneal carcinomatosis revealed by whole-exome sequencing of malignant ascites. Oncotarget 2016;7:8055-66.

Cite this article as: Lim B, Kim SY. Identification of molecular subtypes and significantly mutated genes in gastric cancer using next-generation sequencing. Transl Cancer Res 2016;5(S1):S81-S83. doi: 10.21037/tcr.2016.05.27

