Published in final edited form as:

Hepatology. 2017 July ; 66(1): 280–282. doi:10.1002/hep.29220.

Regulome Networks and Mutational Landscape in Liver Cancer: An Informative Path to Precision Medicine

Song Yi^{1,†} and Nidhi Sahni^{1,2,†}

¹Department of Systems Biology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

²Graduate Program in Structural and Computational Biology and Molecular Biophysics, Baylor College of Medicine, Houston, TX 77030, USA

The power of systems-level analysis of molecular profiles and perturbation patterns in human disease such as liver cancer is increasingly appreciated. A recent article by Fujimoto et al. (1) (Fujimoto A, Furuta M, Totoki Y, Tsunoda T, Kato M, Shiraishi Y, Tanaka H, et al. Whole-genome mutational landscape and characterization of noncoding and structural mutations in liver cancer. *Nat Genet* 2016;48:500-509) reported a comprehensive study on whole-genome mutational signatures associated with liver carcinogenesis and frequently mutated coding and noncoding regions. These results are invaluable in identifying potential cancer driver mutation candidates, and provide possible biomarkers for targeted precision medicine.

Liver cancer is one of the most frequent cancers leading to death worldwide. By the end of this decade a large number of human genomic sequences derived from patients stricken by liver cancer will be available. The first whole-genome sequencing of a hepatocellular carcinoma genome revealed more than 11,000 somatic substitutions (2). To date, ~1000 hepatocellular carcinoma genomes have been sequenced, which provide a mutational landscape of liver cancer. Most common mutations, including *TERT* promoter mutations, occur in the non-coding genome (3), affecting gene expression regulation and signaling pathways in cancer. Current gene-targeted liver cancer therapies are primarily hampered by the extreme genetic heterogeneity observed across cancer patient populations. One key challenge is to distinguish causal cancer mutations from non-pathogenic variants, and the functional impact of such driver mutations (4, 5). In liver cancer, while a handful of non-coding variants have been functionally characterized, the mechanistic impact of most variants remains elusive.

Fujimoto et al. undertook a systematic approach to analyze different types of mutations in liver cancer, using whole-genome sequencing and RNA-seq of 300 Japanese liver cancer patients. The authors first distinguished histological types (hepatocellular carcinoma ‘HCC’ versus intrahepatic cholangiocarcinoma ‘ICC’) of samples by a principal-component

[†]Correspondence should be addressed to Song Yi (Syi2@mdanderson.org) and Nidhi Sahni (Nsahni@mdanderson.org).

Potential conflict of interest

The authors declare no conflict of interest.

analysis by fractions of different trinucleotide mutational contexts (5' and 3' to each mutation site). Non-negative matrix factorization (NMF) was applied to identify 7 mutational signatures in Japanese HCC cohort. These signatures were validated based on high similarity with COSMIC (Catalog of Somatic Mutations in Cancer) signatures. However, it was found that some liver cancer mutational signatures identified in previous studies for Chinese and French HCC cases were missing from this study. The authors concluded that these signatures might be population- or sample-specific, but this was limited and did not include a detailed study of the molecular profiling or genomic aberrations of these samples. 9 groups of tumors were identified by hierarchical clustering based on the enrichment of specific mutational signatures. Both mutational signatures and tumor clusters were significantly associated with clinical features.

Mutational landscapes in liver cancer were revealed by examining multiple types of genomic aberrations, including structural variations (STVs), virus integration, coding and noncoding mutations (Fig. 1). To search for virus integration events in liver cancer, the authors aligned whole-genome sequencing reads to viral genomes, and found the number of HBV integration sites was positively correlated with the number of STVs. The microbiome analysis identified AAV and AAV integrations in cancer and non-cancer liver tissues. HCV is basically not integrated into human genome as its viral feature. To identify driver genes in liver cancer, it was found that 25 protein-coding genes, including *TP53*, *CTNNB1*, *ARID2*, *ARID1A*, *RB1*, *AXIN1*, *RPS6KA3* and *HNF4A*, had significantly recurrent mutations (primarily loss-of-function mutations), and exhibited the functional enrichment of 'chromatin regulator' and 'regulation of cell cycle'. Moreover, 6 lincRNA genes, including *NEAT1* and *MALAT1*, 10 promoters, including those of *TERT*, *WDR74* and *MED16*, and 9 UTRs, including those of *BCL6* and *AFF4*, were identified as significantly mutated noncoding regions.

While the study proposes an intriguing genome-wide alteration map that may underlie tumor heterogeneity in liver cancer, there are some shortcomings. The authors primarily investigated the correlation between mutational signatures and molecular profiles or clinical parameters, but the exact functional impact of each mutation was largely unexplored. A significant association between mutations and replication timing was identified. While point mutations were more prevalent in late-replicating regions, STVs showed distinct patterns: CNAs and chromosomal translocations were frequent in early-replicating regions, but deletions were frequent in late-replicating regions. Although some identified driver genes and promoters were validated by knockdown and reporter assays respectively, the mechanistic insights of mutations remained unclear. Furthermore, the authors reported a strong association between RNA-seq gene expression and CNAs, inversions and chromosomal translocations. The expression levels of nearby oncogenes (*CCND1* and *MET*) were increased, whereas those of tumor-suppressor genes (*CDKN2A* and *APC*) were suppressed, indicating that genomic mutations can promote liver cancer development by altering gene expression.

This systematic analysis provided a practical approach to identifying candidate cancer genes emerging from genome-wide somatic mutation sequencing projects. Based on mutation frequency in patients, new genes associated with liver cancer were found, including certain

nuclear receptor signaling genes, DNA-repair genes, chromatin regulators, and liver metabolism genes. This study attempted to find all driver cancer genomic changes but with the limitation that a long catalog of just the genome alterations found in liver cancer, although necessary, would not be sufficient to accurately reveal their functional consequences. Although the authors speculated that these genomic mutations might affect protein binding affinity, they did not explore further to distinguish driver from passenger mutations.

It is critical to determine the ‘regulome’ network mechanisms that underlie genetic heterogeneity and potentiate functional diversification across different liver cancer subtypes in patients. In cellular regulome networks, multiple types of molecular interactions are present to carry out biological function, including but not limited to protein-protein, protein-DNA, and protein-RNA interactions (4, 6). Recent works indicate that gaining a systems-level understanding of human cancer requires strategies to integrate computational and experimental technology to prioritize cancer variants. Such strategies will further help understand precise control of gene regulation inside the cell. In liver cancer, signal transduction transfers information from microenvironmental cues to intracellular effectors via dynamic molecular interactions. Genes and regulatory elements interact with each other in molecular networks (7). Biological systems and cellular networks underlie most genotype-phenotype relationships (6).

A systematic survey of ~1,000 human genetic disorders revealed that an unexpectedly large fraction (>50%) of missense disease mutations perturb protein interactions, about half of which exhibited specific network perturbations (5). This was in stark contrast to common variants found in healthy people that showed significant enrichment for wild-type regulome network. It becomes more appreciated that different mutations of the same gene often lead to several distinct phenotypes (allelic heterogeneity). Network perturbation model suggests that mutations result in distinct phenotypic outcomes through alternative interaction perturbations. Another study indicates that most individuals have unique repertoires of TF binding activities, which contributes to phenotypic variation (8). For instance, a ‘T to C’ mutation in the CYBB promoter causes loss of binding for two TFs, IRF2 and IRF5, resulting in a reduction in expression in chronic granulomatous disease. In liver cancer, the vast majority of disease-associated mutations are in non-coding regions. However, few non-coding cancer-associated mutations have been studied as of date. TFs often bind to non-coding regulatory DNAs to mediate transcription, therefore protein-DNA assays such as ChIP-seq, yeast one-hybrid (Y1H) and protein-binding microarray (PBM) can be used to assess the function of non-coding mutations (9). Moreover, *in vivo* assay of non-coding mutational effect is needed to gain better insights into carcinogenesis. Recent advances in the CRISPR technology enables genome-wide screens to evaluate functional consequences of non-coding regions *in vivo* (9). Systematic experimental and analytical platforms will be insightful in identifying disease-causing mutations, and uncovering patient mutation-specific disease mechanisms in liver cancer.

References

1. Fujimoto A, Furuta M, Totoki Y, Tsunoda T, Kato M, Shiraishi Y, Tanaka H, et al. Whole-genome mutational landscape and characterization of noncoding and structural mutations in liver cancer. *Nat Genet* 2016;48:500–509. [PubMed: 27064257]
2. Totoki Y, Tatsuno K, Yamamoto S, Arai Y, Hosoda F, Ishikawa S, Tsutsumi S, et al. High-resolution characterization of a hepatocellular carcinoma genome. *Nat Genet* 2011;43:464–469. [PubMed: 21499249]
3. Khurana E, Fu Y, Chakravarty D, Demichelis F, Rubin MA, Gerstein M. Role of non-coding sequence variants in cancer. *Nat Rev Genet* 2016;17:93–108. [PubMed: 26781813]
4. Sahni N, Yi S, Zhong Q, Jailkhani N, Charlotiaux B, Cusick ME, Vidal M. Edgotype: a fundamental link between genotype and phenotype. *Curr Opin Genet Dev* 2013;23:649–657. [PubMed: 24287335]
5. Sahni N, Yi S, Taipale M, Fuxman Bass JI, Coulombe-Huntington J, Yang F, Peng J, et al. Widespread macromolecular interaction perturbations in human genetic disorders. *Cell* 2015;161:647–660. [PubMed: 25910212]
6. Vidal M, Cusick ME, Barabasi AL. Interactome networks and human disease. *Cell* 2011;144:986–998. [PubMed: 21414488]
7. Fuxman Bass JI, Sahni N, Shrestha S, Garcia-Gonzalez A, Mori A, Bhat N, Yi S, et al. Human gene-centered transcription factor networks for enhancers and disease variants. *Cell* 2015;161:661–673. [PubMed: 25910213]
8. Barrera LA, Vedenko A, Kurland JV, Rogers JM, Gisselbrecht SS, Rossin EJ, Woodard J, et al. Survey of variation in human transcription factors reveals prevalent DNA binding changes. *Science* 2016;351:1450–1454. [PubMed: 27013732]
9. Yi S, Lin S, Li Y, Zhao W, Mills GB, Sahni N. Functional variomics and network perturbation: connecting genotype to phenotype in cancer. *Nat Rev Genet* 2017;27 Mar 2017 online.

Figure 1. Complex genetic heterogeneity and regulome network rewiring in liver cancer. Model of pleiotropic clinical outcomes resulting from distinct regulome network perturbation patterns induced by distinct cancer mutations. STVs, structural variations. Rectangles, coding genes. Stars, mutations. Coloured circles, different macromolecules. Edges, interactions between molecules.