



Published in final edited form as:

*Ann Surg Oncol.* 2013 November ; 20(12): 3707–3708. doi:10.1245/s10434-013-3071-x.

## Prognostic gene signatures for hepatocellular carcinoma – what are we measuring?

**Kenneth K. Tanabe, MD** and

Division of Surgical Oncology, Massachusetts General Hospital, Boston, MA 02114,  
ktanabe@partners.org

**Yujin Hoshida, MD, PhD**

Liver Cancer Program, Tisch Cancer Institute, Division of Liver Diseases, Department of  
Medicine, Icahn School of Medicine at Mount Sinai, New York, USA, yujin.hoshida@mssm.edu

In this issue of *Annals of Surgical Oncology*, Lim and colleagues report on prognostic gene-expression signatures from tumor and non-tumor liver tissues from hepatocellular carcinoma (HCC) patients who underwent surgical resection at a single institution in Korea (1). The signatures were developed from a set of patients with relatively early-stage diseases, and then evaluated in an independent validation cohort of patients with more advanced-stage diseases. The authors report that their gene signature derived from the tumor tissues and not their gene signature from the non-tumor tissue accurately predicted disease-free survival in the validation cohort. Their findings are consistent with previously published data suggesting that tumor-derived molecular information is more prognostic in advanced-stage HCC, whereas non-tumor liver-derived molecular information is more prognostic in earlier-stage HCC (2–4).

Prognostic gene signatures using genome-wide gene-expression microarrays have been explored in many types of cancer with an eye toward improving prognostic models, guiding treatment recommendations and identifying molecular pathways to target for therapy. In the case of HCC, dozens of prognostic gene signatures have been reported (5). HCC differs from most cancers in that there are two biological processes to simultaneously consider. HCC commonly arises within a liver compromised by cirrhosis, thereby requiring effective management of two disease processes. Thus, death from disease progression may arise from distinctly different causes – recurrence and/or dissemination of the HCC, de novo new HCC arising in the cirrhotic liver, and complications of progressive cirrhosis including hepatic decompensation. Patients with advanced-stage HCC are more likely to die due to progression of the primary HCC tumor, whereas patients with earlier-stage diseases are more likely to die due to additional primary HCC tumors (e.g. clonally distinct from the first HCC) or complication of cirrhosis (2, 3). Nonetheless, in many cases the precise causes are difficult to accurately determine. Accordingly, development of prognostic gene signatures for HCC requires careful consideration and delineation of the patient population as well as careful consideration of specific endpoints for training the signature.

Of note, Lim and colleagues chose disease-free survival for the end-point on which to train their prognostic gene signatures. This endpoint combines all three processes noted above – recurrence and/or dissemination of HCC, de novo new HCC arising in the cirrhotic liver,

and complications of progressive cirrhosis. Thus, rather than capturing a distinct biological process, this signature likely captures a blend of biological processes. And because the nature of the blend is highly dependent on the patient population, it is difficult to determine whether this signature can be accurately applied to other patient populations. In this sense, the use of more specific clinical outcomes, e.g., time to recurrence and time to HCC-specific death, to train the signature would provide for a more straightforward interpretation of the results as well as determination of their applicability to other patient populations.

Rapidly growing public genomics databases such as NCBI's Gene Expression Omnibus ([www.ncbi.nlm.nih.gov/geo](http://www.ncbi.nlm.nih.gov/geo)) have enabled *in silico* cross-cohort validation of prognostic gene signatures (6). The authors are encouraged to deposit their genomic dataset to such databases as systematic evaluation of previously reported gene signatures on the authors' dataset will help determine whether there is any unique advantage of their signature (6). In addition, validation of the authors' gene signature in other publicly reported patient cohorts will provide information on the potential clinical utility of the signature, given that the signature was derived from a single-assay platform and single-institution patient cohort.

The methodology of signature gene selection is another important component to ensure robust predictive/prognostic performance of a signature. In this study Lim and colleagues used a "parsimonious" method to choose genes independently associated with disease-free survival. With this approach, one would expect that each gene in the signature represents activation status of a distinct molecular pathway involved in the disease. This is theoretically an effective strategy to capture maximum information with minimum number of measurements. However, this could be a limitation in analyzing real-world clinical specimens, in which gene expression measurements are less reproducible at single gene level, thereby increasing the likelihood of missing in the signature information from key pathways. This issue is particularly relevant given the recent trend towards use of partially degraded RNA isolated from archived formalin-fixed tissues for expression profiling (7, 8). One possible solution to the problem of missing important genes in the signature may be to include multiple genes to more robustly capture the status of each molecular pathway, even when the signals from some genes are missed due to sample quality issue, etc. (9).

One of the major challenges in genome-wide gene-expression microarrays has been experimental variation that results from amplification of cDNA by polymerase chain reaction or similar techniques. Recent development of highly sensitive signal detection technologies has enabled digital counting of transcripts without amplification, which may have a potential to overcome this limitation (10, 11). When this progress in technology is combined with the research community's movement towards public sharing of data sets for cross-cohort validation, the future for prognostic gene signatures in all cancers – including HCC – is bright. These gene signatures are expected to provide a better understanding of the prognosis of patients with cirrhosis and HCC. More importantly, it is a reasonable expectation that these signatures will divulge important clues on the molecular pathways that are critically important for these diseases (both cirrhosis and HCC) and for which effective targeting will lead to effective therapies.

## References

1. Lim HY, Sohn I, Deng S, Lee J, Jung SH, Mao M, Xu J, et al. Prediction of disease-free survival in hepatocellular carcinoma by gene expression profiling. *Ann Surg Oncol*. 2013
2. Hoshida Y, Toffanin S, Lachenmayer A, Villanueva A, Minguez B, Llovet JM. Molecular Classification and Novel Targets in Hepatocellular Carcinoma: Recent Advancements. *Semin Liver Dis*. 2010; 30:35–51. [PubMed: 20175032]
3. Hoshida Y, Villanueva A, Llovet JM. Molecular profiling to predict hepatocellular carcinoma outcome. *Expert Rev Gastroenterol Hepatol*. 2009; 3:101–103. [PubMed: 19351279]
4. Villanueva A, Hoshida Y, Toffanin S, Lachenmayer A, Alsinet C, Savic R, Cornella H, et al. New strategies in hepatocellular carcinoma: genomic prognostic markers. *Clin Cancer Res*. 2010; 16:4688–4694. [PubMed: 20713493]
5. Hoshida Y, Moeini A, Alsinet C, Kojima K, Villanueva A. Gene signatures in the management of hepatocellular carcinoma. *Semin Oncol*. 2012; 39:473–485. [PubMed: 22846864]
6. Villanueva A, Hoshida Y, Battiston C, Tovar V, Sia D, Alsinet C, Cornella H, et al. Combining clinical, pathology, and gene expression data to predict recurrence of hepatocellular carcinoma. *Gastroenterology*. 2011; 140:1501–1512. e1502. [PubMed: 21320499]
7. Simon RM, Paik S, Hayes DF. Use of archived specimens in evaluation of prognostic and predictive biomarkers. *J Natl Cancer Inst*. 2009; 101:1446–1452. [PubMed: 19815849]
8. Hoshida Y, Villanueva A, Kobayashi M, Peix J, Chiang DY, Camargo A, Gupta S, et al. Gene expression in fixed tissues and outcome in hepatocellular carcinoma. *N Engl J Med*. 2008; 359:1995–2004. [PubMed: 18923165]
9. Hoshida Y. Nearest Template Prediction: A Single-Sample-Based Flexible Class Prediction with Confidence Assessment. *PLoS One*. 2010; 5:e15543. [PubMed: 21124904]
10. Geiss GK, Bumgarner RE, Birditt B, Dahl T, Dowidar N, Dunaway DL, Fell HP, et al. Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat Biotechnol*. 2008; 26:317–325. [PubMed: 18278033]
11. Ozsolak F, Milos PM. RNA sequencing: advances, challenges and opportunities. *Nat Rev Genet*. 2011; 12:87–98. [PubMed: 21191423]