



Robust Biomarker Screening Using Spares Learning Approach for Liver Cancer Prognosis

Aman Chandra Kaushik^{1,2}, Aamir Mehmood^{2,3}, Dong-Qing Wei^{2,3*} and Xiaofeng Dai^{1*}

¹ Wuxi School of Medicine, Jiangnan University, Wuxi, China, ² School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China, ³ Peng Cheng Laboratory, Shenzhen, China

OPEN ACCESS

Edited by:

Yongchun Zuo,
Inner Mongolia University, China

Reviewed by:

Quan Zou,
University of Electronic Science
and Technology of China, China
Nitish Kumar Mishra,
University of Nebraska Medical
Center, United States
Shaoliang Peng,
Hunan University, China

*Correspondence:

Dong-Qing Wei
dqwei@sjtu.edu.cn
Xiaofeng Dai
1281423490@qq.com

Specialty section:

This article was submitted to
Bioinformatics and Computational
Biology,
a section of the journal
Frontiers in Bioengineering and
Biotechnology

Received: 16 December 2019

Accepted: 09 March 2020

Published: 03 April 2020

Citation:

Kaushik AC, Mehmood A,
Wei D-Q and Dai X (2020) Robust
Biomarker Screening Using Spares
Learning Approach for Liver Cancer
Prognosis.
Front. Bioeng. Biotechnol. 8:241.
doi: 10.3389/fbioe.2020.00241

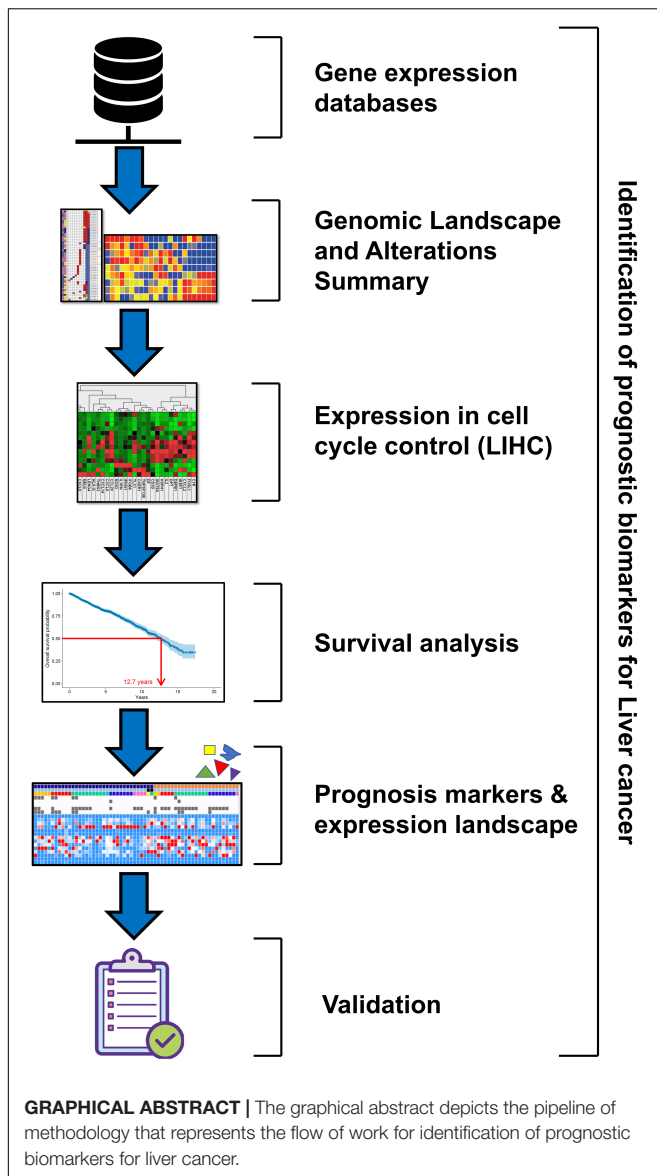
LncRNAs, miRNAs, mRNAs, methylation, and proteins exert profound biological functions and are widely applied as prognostic features in liver cancer. This study aims to identify prognostic biomarkers' signature for liver cancer. Samples with inadequate tumor purity were filtered out and the expression data from different resources were retrieved. The Spares learning approach was applied to select lncRNAs, miRNAs, mRNAs, methylation, and proteins' features based on their differentially expressed groups. The LASSO boosting technique was employed for the predictive model construction. A total of 200 lncRNAs, 200 miRNAs, 371 mRNAs, 371 methylations, and 184 proteins were observed to be differentially expressed. Five lncRNAs, 11 miRNAs, 30 mRNAs, 4 methylations, and 3 proteins were selected for further evaluation using the feature elimination technique. The highest accuracy of 89.32% is achieved as a result of training and learning by Spares learning methodology. Final outcomes revealed that 5 lncRNA, 11 miRNA, 30 mRNA, 4 methylation, and 3 protein signatures could be potential biomarkers for the prognosis of liver cancer patients.

Keywords: biomarkers, spares learning, liver cancer, prognosis, ncRNA

INTRODUCTION

One of the largest organs in the human body is the liver, which is crucial for metabolism and is helpful in detoxification and maintaining homeostasis. Many ailments are concerned with the liver including hepatitis, fibrosis, genetic and metabolic issues, and liver cancer, which is one of the leading causes of cancer-related expiries (Bray et al., 2018; Dooley et al., 2018; Yang et al., 2019). Hepatocellular carcinoma (HCC) can occur in an ailing liver and encompasses numerous molecular cascades (Kanda et al., 2019). It is reported that more than 90% of liver cancers are HCC, which is an extremely assorted form of cancer verified by high-throughput sequencing and gene expression profiling, at both the molecular and histological levels (Calderaro et al., 2019).

Gene therapy has progressed as an effective source of dealing with disease-causing gene imperfections to attain a typical status. The approaches employed to treat illness by gene therapy consist of gene replacement, gene restoration, gene extension, gene muzzling, vaccination, and, currently, gene-editing technology (Alsaggar and Liu, 2016; Chew et al., 2016; Karimian et al., 2019). Thus, the identification of a gene that can be used as a potential biomarker is an important step in the treatment of liver cancer (West et al., 2019; Zheng et al., 2019; Lu et al., 2020; von Felden and Villanueva, 2020).



Statistical approaches like artificial neural networks employing BI-RADS (Baker et al., 1995) and logistic regressions have been used in several reports to improve diagnostic performance. It is most beneficial to use statistical approaches as they enhance the identification of breast cancer, with BI-RADS, as well as along the medical and statistical information concerning infected persons' statistical threat aspects (Chhatwal et al., 2009). Regression processes go through overfitting once the prognostic covariates are involved in a large number. Similar situations lead to the well-fitting of a deterioration mockup into the drilling information. However, this doesn't go parallel with the cases of the real world. Variable selection becomes a necessity in an attempt to get exact predictions associated with covariates of a large number, for instance, BI-RADS qualifiers and statistical data. A very famous fact claims the unfavorability of regular step-by-step assortment methods in case of regression models that have

numerous covariates (Houssami et al., 2004). On the other hand, sparse penalized methods, like the minimum complete reduction and assortment operative (LASSO), together have gained ample consideration. LASSO is a penalized regression technique that approximates the deterioration constants through enhancing the log-similarity purpose (or adding the squared remainders) having restriction that the addition of the total scores of the deterioration constants, $\sum k_j = 1 ||\beta_j||$, is \leq to a positive constant s . LASSO has one of the most fascinating characters that the approximated value belonging to deterioration constants is tenuous in nature, indicating a lot of components that are accurately 0. This proves that unnecessary covariates are automatically deleted by LASSO. It is believed that LASSO has numerous required characteristics that are compulsory for the deterioration mockups with a huge covariate count. Optimization algorithms for the rectilinear deterioration model as well as for general rectilinear mockups are available in large numbers, with good efficiency. As per our information, this work is the first attempt to build a calculated LASSO deterioration mockup that could assist in the diagnosis of breast cancer based on statistical and radiological findings.

This study is aimed to compare the productivity of graphical examination to forecast liver cancer dependent on whether the calculated LASSO deterioration or bit-by-bit calculated (SL) deterioration was employed, along with evaluating the practicality of integrating statistical data into the graphic breakdown for the sake of improving liver cancer diagnosis.

MATERIALS AND METHODS

Gene Expression Databases

The Cancer Genome Atlas (TCGA)¹ catalog can be accessed to gain information regarding alterations in the gene, long non-coding RNAs (lncRNAs), methylations, miRNAs, mRNAs, CAN, mutational expression, and proteins involved in HCC. It is a freely accessible repository at the TCGA (Cerami et al., 2012; Gao et al., 2013). The cancer study "HCC were obtained from TCGA" and information type precedence "Mutation and CNA (DNA copy-number alterations)" were selected before analyzing genomic alterations of cell cycle control in the TCGA data on HCC. This did not require any statement of approval or informed consent for the reason that the information is retrieved from a public repository.

Genomic Alterations Summary

Genomic modifications of cell cycle control via tumor samples were summarized. Genomic modifications inclusive of mutations, CNA (amplifications and homozygous deletions), glyphs, and dye tagging were practiced to summarize the gene expression variations. It was the first step to understand various types of gene signaling in HCC. The shared exclusivity and co-occurrence among cell cycle control were studied as well. Discordant, gene-related happenings linked with a specific cancer are most of the time conflicting in a cluster of tumors,

¹<https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga>

i.e., only single biological happening is expected to occur in every sample of cancer. Another condition is the simultaneous incidence that several genes are changed in each sample (Gao et al., 2013); this was an introductory way to collect information related to various gene signaling in HCC.

Mutations in Cell Cycle Control in HCC

Through the mutations of cell cycle control, the rate and position for all mutations within Pfam protein fields were specified. Colored bars denote an entire extent of cell cycle control proteins and the base of every bar in gray denotes the amino acid count. Protein domains are represented by the boxes colored in red, blue, and green. The lines and points signify the position and rate of genes. The frameshift or nonsense mutations are shown in red, missense mutations are in green, and the black color represents the in-frame deletions (Fang et al., 2015).

Survival Analysis

The survival analysis bears great importance in prognosis to highlight changes in the survival rate. Here, the differences in the overall survival were evaluated via survival analysis among samples having a single or more alteration as that of the inquired genes(s) and also the samples that have no variation.

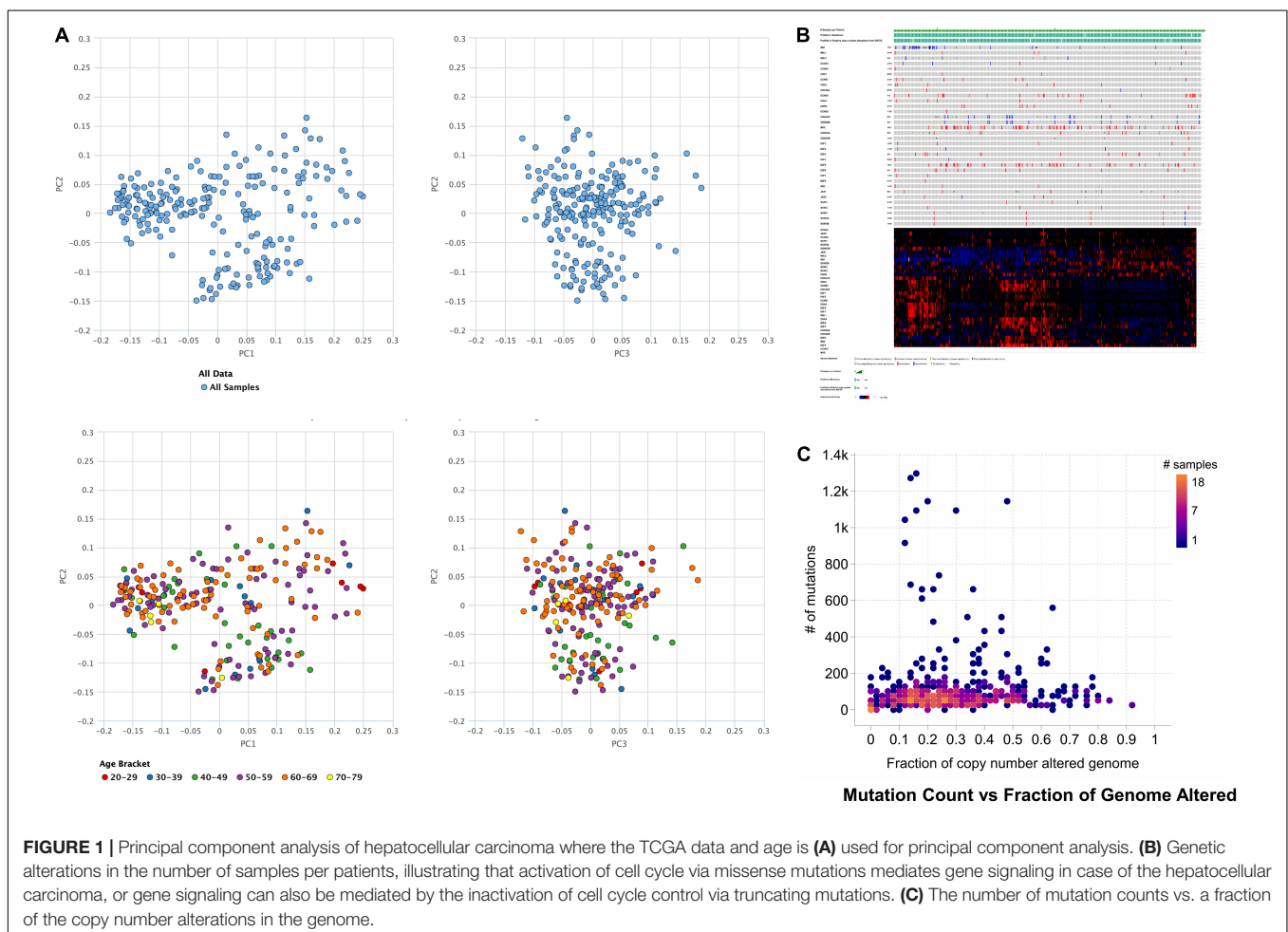
Statistics

To carry out correlation analysis, a scattered graph of lncRNAs, methylations, miRNAs, mRNAs, CAN, mutational expression, or protein level in every sample was presented. The Kaplan–Meier approach having log-rank tests are carried out for comparing global and healthy survival of HCC that have at least a single modification or lack any adjustment within the inquired gene(s). Samples with up-regulation were recognized by the verge of $Z > 2$ (mean expression over 2 SDs). The standard was fixed at 0.05.

Acquisition of Patient's Data

lncRNAs appeared as potential features in the field of oncology. RNA-seq data are obtained from TCGA while the exploration of lncRNAs in cancer is provided by an open-access web app “TANRIC.” The TANRIC (The Atlas of ncRNA In Cancer)² allows rapid and intuitive analyses of lncRNAs in the framework of experimental and other molecular information. Through TANRIC, a high amount of lncRNAs were identified with probable biomedical implication, where the majority of them shows

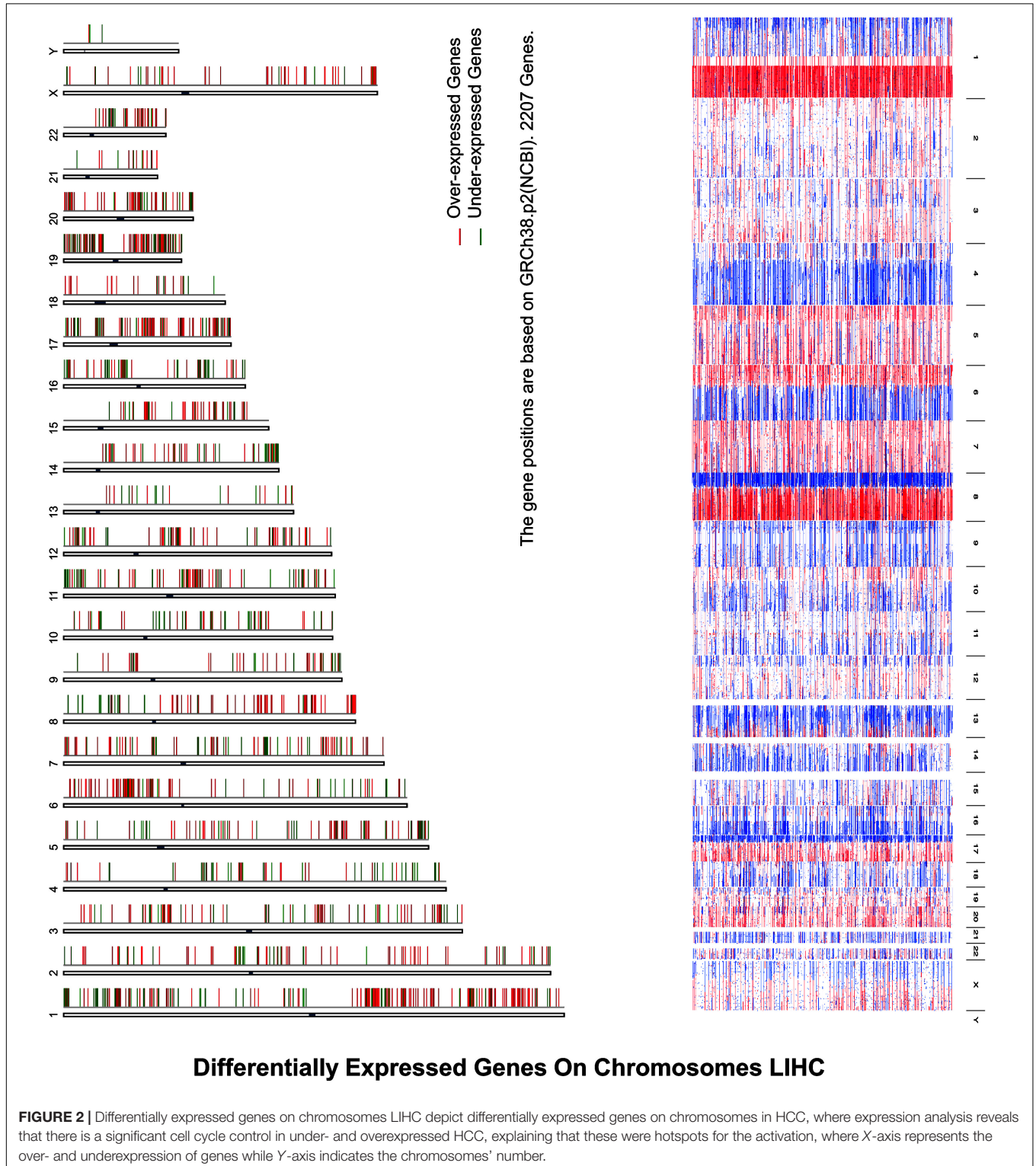
²https://ibl.mdanderson.org/tanric/_design/basic/index.html



robust associations with the already formed therapeutic goals and biomarkers across the cell lines or tumor types. We retrieved lncRNA, miRNA, mRNA, methylation, and protein expression data from TANRIC (Li et al., 2015) of all the TCGA liver cancer patients (Ciriello et al., 2015). The

corresponding clinical data are retrieved from Genomic Data Commons (GDC)³.

³<https://gdc.cancer.gov/>



Purity estimation was performed for the patients using consensus purity estimate and the Clonal Heterogeneity Analysis Tool (Li et al., 2012; Li and Li, 2014). Patients were filtered out with purity estimators below 60%.

Feature Identification for lncRNAs, miRNAs, mRNAs, Methylations, and Proteins

For the identification of promising discriminative lncRNAs, miRNAs, mRNAs, methylations, and proteins of survival groups, the R limma¹⁴ package was used to identify promising discriminative biomarkers by analyzing the differential expression of lncRNAs, miRNAs, mRNAs, methylations, and proteins.

Feature Selection of lncRNAs, miRNAs, mRNAs, Methylations, and Proteins

The differentially expressed lncRNAs, miRNAs, mRNAs, methylations, and proteins were used as input features for

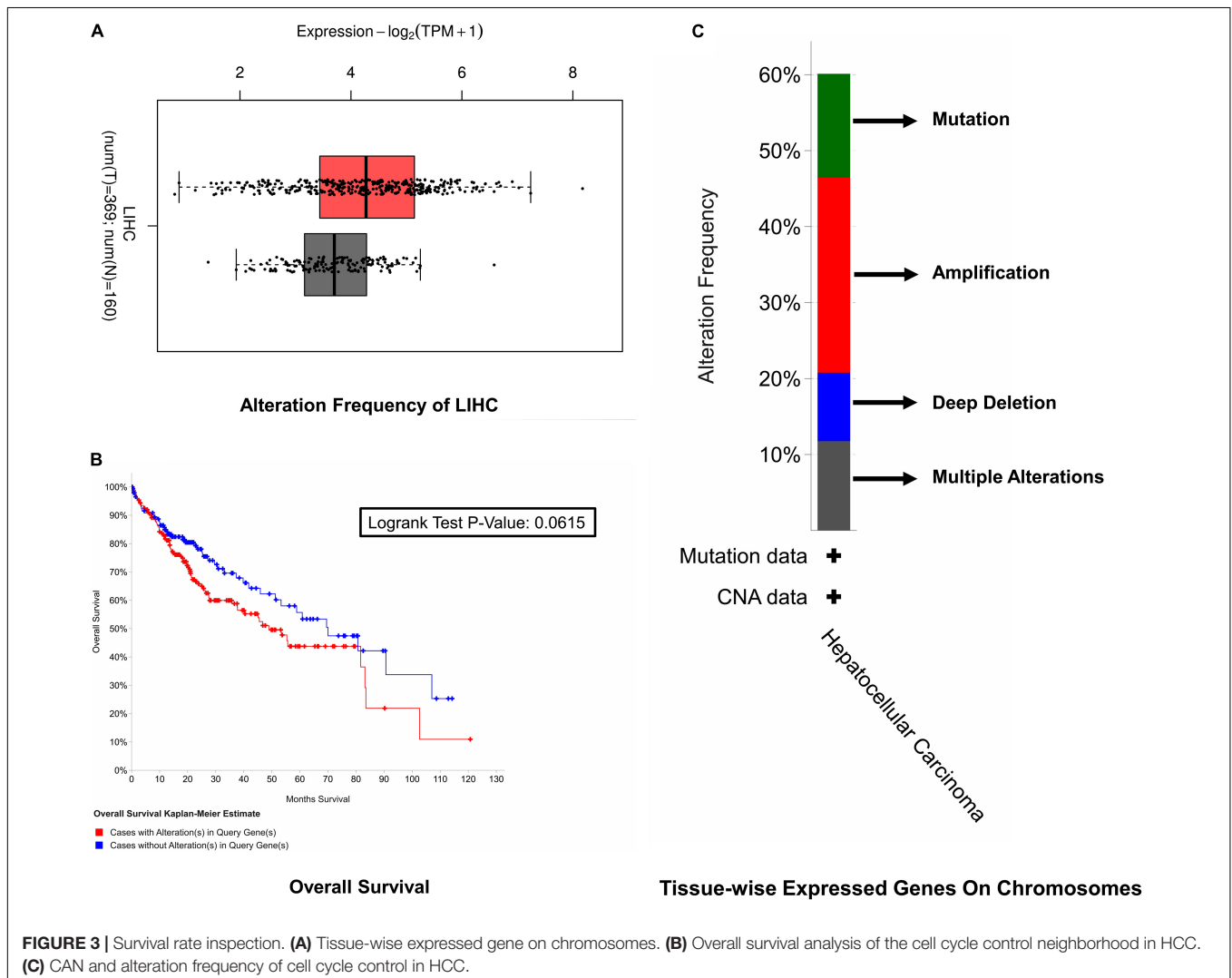
predictive modeling. Spares learning was applied to select features. The Spares learning and LASSO method were ranked by features based on specific importance.

Predictive Modeling and Expression Landscape

We used Spares learning and LASSO to construct the predictive model of survival groups. LASSO is a powerful ensemble learning method that has achieved state-of-the-art performance in many biomedical tasks.

$$g_i = G_{-l}b_i + \varepsilon_i, \text{ subject to } |b_i| \leq C_i, \quad (1)$$

where b_i is the coefficient of expressions other than RNA's i , $|\cdot|$ is an L-1 norm, and the residual is denoted as ε_i . The j_{th} coefficient element in b_i indicates a regulatory relationship from RNAs j to RNAs i (with a direction) in the linear model, where zero shows non-relationship between them. In contrast with correlation-based RNA regulatory networks, linear regression-based RNA regulatory networks can capture the main effects of multiple



RNAs. Correlation-based RNA regulatory networks may fail to infer RNA regulation if the correlation is not significantly high and if multiple RNAs regulate simultaneously. The coefficient vector b_i for RNAs i is used for constructing the adjacency matrix B of the group-specific RNAs regulatory network, i.e., $B = \{b_1, \dots, b_p\}^T$. Then, Eq. (1) can be optimized by:

$$\|g_i - G_{-i}b_i\|^2 + \lambda \|b_i\|, \quad (2)$$

where λ is a hyper-parameter for sparsity regularization, and $\|\cdot\|^2$ is an L-2 norm of a vector.

Algorithm

- 1: $I \in \{1, \dots, p\}$ do
- 2: $b_i^T =$ Random LASSO in equation (2)
- 3: $b_i^C =$ Random LASSO in equation (2)
- 4: end for
- 5: $B^T = \{b_1^T, \dots, b_p^T\}$
- 6: $B^C = \{b_1^C, \dots, b_p^C\}$

RESULTS AND DISCUSSION

In this study, data from the TCGA Cancer Genomics have been used to explore, visualize, and analyze the genetic and medical features of alterations in cell cycle control found in cases of HCC from databases of TCGA. As per our knowledge, this study is the opening data mining approach that tends to discover the existing connection among modifications occurring in control of cell cycle and patients' prognosis. A lot of conclusions in this study are coherent with the previously reported data. Remarkably, we detected in our study that alterations in the cell cycle control mostly exist in HCC. Variations in these genetic factors are on autonomous cascades to HCC and are in an uncommon fashion of increasing gene changes. Although no cell cycle control was linked with any of the survival events (disease-free and global survival) in this work, it provides us with a fresh perspective to concurrently investigate biological modifications and medical features through information exploration.

Genomic Landscape and Alterations

Summary

Based on obtained outcomes, it was observed that the majority of the cases undergo alteration in the cell cycle control, and nearly all of them were missense mutations. Others incorporated deep deletions and few amplifications. However, the rest of the cases remained had modifications in the cell cycle control that comprises most of the truncating and missense mutations. The shared exclusivity analysis implies that events that occurred in cell cycle control were liable to occur again in HCC as shown in **Figure 1** through principal component analysis.

Expression in Cell Cycle Control in HCC

Inspection of the expression analysis reveals that there is a significant cell cycle control in under- and overexpressed

TABLE 1 | Biomarkers list of lncRNAs, miRNAs, mRNAs, methylations, and proteins in liver cancer with their estimated coefficient score.

ncRNAs	Biomarkers	Prediction Scores
lncRNAs	ENSG00000234577 1	-0.009981348
	ENSG00000236753 1	0.002278461
	ENSG00000257883 1	-0.000943499
	ENSG00000260920 1	0.423349833
	ENSG00000273059 1	-1.47647639
miRNAs	has-mir-139	-0.000141116
	has-mir-148a	-2.26E-07
	has-mir-23c	-0.00822489
	has-mir-3156-1	-0.076147329
	has-mir-3680	0.062526769
	has-mir-3682	0.001235262
	has-mir-3910-2	0.008941494
	has-mir-3924	-0.035235217
	has-mir-548v	0.006351925
	has-mir-592	0.000212203
mRNAs	has-mir-632	0.107689301
	ACSM4	0.212644291
	C15orf33	-0.004560181
	CAD	4.21E-05
	CDCA8	0.000189659
	CHAC2	0.000861851
	DYNC1L1G6PD	5.19E-05
	G6PD	6.03E-06
	GOT2	-2.16E-06
	GRPEL2	0.000536562
	KLRB1	-8.62E-05
	LDHA	7.57E-07
	MYCN	0.000171789
	MYOD1	0.00620387
	NBPF22P	0.004514996
	PDE6A	0.014750229
	PSD4	-2.48E-05
	PSMD1	352E-05
	RPL21P44	0.00651285
	SEPHS1	0.000165664
SLC1A7	2.08E-05	
SNAR-C4	0.101384012	
SNORD94	-0.009345186	
TAF3	0.00158546	
TMEM69	3.91E-05	
TTC26	0.000160117	
TXNRD1	1.10E-05	
UCK2	0.000240815	
UTP11L	0.00051883	
YARS	9.13E-06	
ZNF643	1.88E-05	
Methylations	Cg02603140	0.0886803
	Cg04104820	-0.612239
	Cg15744128	-0.119538
	Cg24153171	-1.46113
Proteins	Acetyla-Tubulin-Lys40	0.041212138
	B-Raf	0.102643389
	PAI-1	0.114385869

HCC, explaining that these were hotspots for the activation as illustrated in **Figure 2**.

Survival Analysis

For the sake of survival rate inspection, Kaplan–Meier plots were used in an order to complete survival analysis in cases of HCC with as well as without cell cycle control overexpression. For the overall survival analysis, mutations in the cell cycle control were found to be concurrent and not linked to a decreased overall

survival ($p = 0.0615$). Likewise, none of the cell cycle control was linked with any of the survival events (**Figure 3**).

Liver Cancer Prognosis Markers and Expression Landscape

It was observed that lncRNAs, miRNAs, mRNAs, methylations, and proteins may exert a more profound biological impact than a single gene by virtue of its intrinsic regulatory nature. Therefore, predictive modeling is also performed for the sake of liver cancer

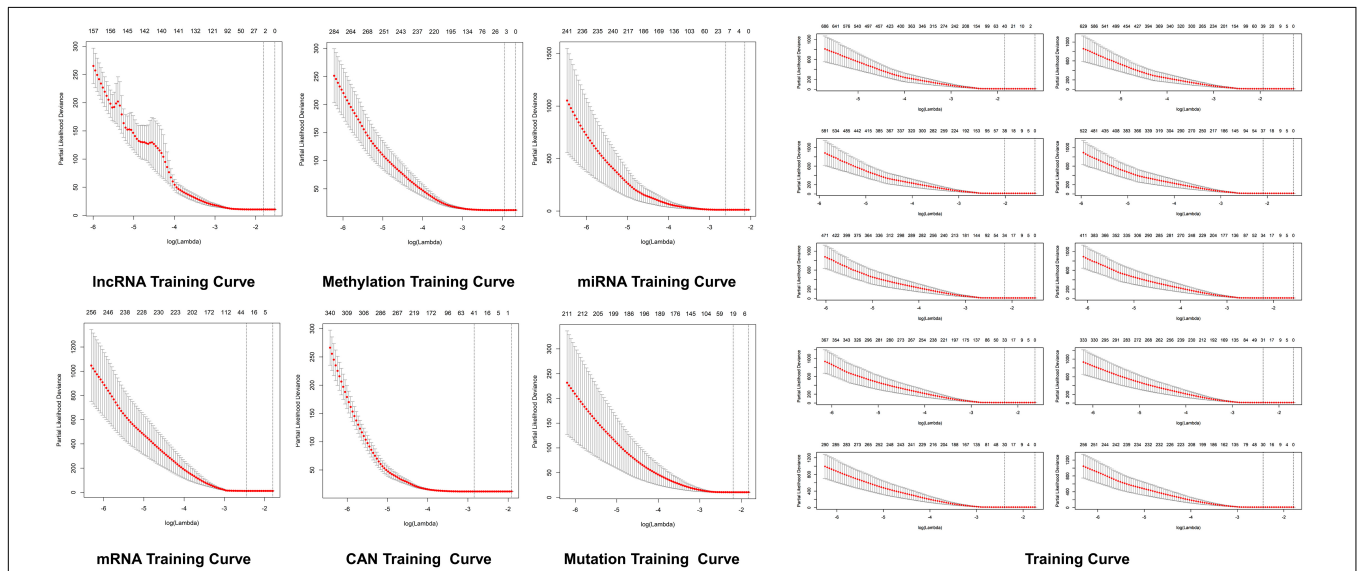


FIGURE 4 | Predictive modeling through LASSO. LASSO with the features weighted training curve in HCC, where optimization equation was resolved through LASSO for RNA implication based on features reputation generation, drilling LASSO with the features biased by reputation, which depend on bootstrapping approach by drilling only a small set of variables instead of drilling the whole variables directly as depicted in figure.

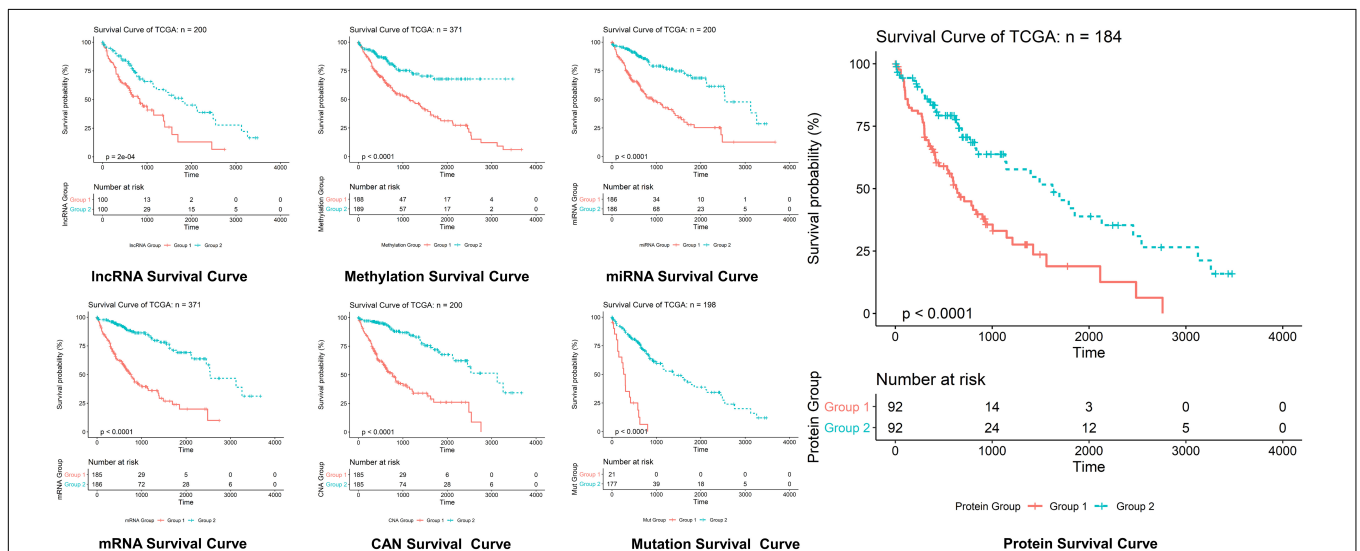


FIGURE 5 | Survival curves prediction through LASSO. Survival curves of lncRNAs, methylations, miRNAs, mRNAs, protein, CAN, and mutation in hepatocellular carcinoma, where optimization equation was resolved through LASSO for RNAs implication based on features reputation generation, drilling LASSO with the features biased by reputation, which depend on bootstrapping approach by drilling only a small set of variables instead of drilling the whole variables directly, and constant approximation is much dependable on every individual training through high-dimensional information as shown here.

prognosis (Table 1) and the expression landscape of survival patients as shown in Figures 4, 5.

Predictive Modeling

The optimization of an equation can be resolved by a LASSO explanation. For reliable RNA implication, we applied the Random LASSO (Lee et al., 2014). This technique is divided into two main steps: (1) features' reputation generation for RNAs and (2) drilling LASSO with the features biased by reputation. This method uses a bootstrapping approach by drilling only a small set of variables instead of drilling the whole variables directly. The constant approximation is much dependable on every individual training through high-dimensional information (Figures 4, 5).

Validation of Biomarkers

Based on literature reports, an initial major component assessment leads to separation. So, the first thing that we asked was whether the whole organization is practical by using a standardized and shared dataset. Hence, to evaluate our data for equality and applicability, a fivefold leave group cross-validation was employed by the use of LASSO and spare learning. All the datasets were examined distinctly, so each dataset was utilized as a drilling set along with group tags that were equitable to the general receptor position, via reference against liver cancer. Consequently, the receptor status of the entire untrained datasets was predicted by the use of attained Spares learning and LASSO model. Nevertheless, the accuracy of the categorization of the patient's data builds by the utilization of Spares learning and LASSO models is high. Reasonable Matthews correlation coefficients and low error rates were 0.1 for the grouping of references against liver cancer. The bootstrapping sampling was modified at the time of model formation for addressing inequity in the class in various datasets for drawing an equal number of samples out of each group.

CONCLUSION

Detection of cancer at an early curable phase and eradicating the tissues can be capable of preventing the expansion of lethal intrusive cancers, which would save countless lives. Presently, it is extensively stated that lncRNAs, mRNAs, and miRNAs could be probable biomarkers for various cancers. Identification of lncRNAs, mRNAs, and miRNAs related to disease adds to the enhancement of understanding of diagnosis and pathogenesis. Therefore, for the investigation of disease association of lncRNAs, mRNAs, and miRNAs, development of numerous potential computational models has been done. Nevertheless, only some studies centered on the identification of lncRNA, mRNA, and miRNA signs for the diagnosis of liver cancer in the early stage. Consequently, in the current

study, we put forward a new classification technique based on lncRNAs, mRNAs, and miRNAs for the categorization of early and advanced phases of liver cancer. The increasing trend in the implementations of machine learning methods and the latest developments of personalized medicine enhanced the forecasting of cancer. For the sake of identifying main aspects that could influence cancer development, recurrence, and survival, different machine learning techniques and algorithms employed for feature selection are globally cast. In general, cancer prediction studies based on machine learning employed expression profiles of mRNA/miRNA, clinical factors, and histological variables as an input for the procedure of cancer prediction. Success in the development of computational models for prediction of cancer rests on comprehending the biological information and shortcomings of the drilling dataset, for instance, minor collection of high-dimensional samples known as "curse of dimensionality." Nevertheless, the over-drilling problem may be overcome through appropriate feature selection and cross-validation approaches. Our findings provide a new vision for exploring biological functions of lncRNAs, miRNAs, mRNAs, methylations, and proteins in liver cancer, and screening novel potential biomarker (lncRNAs, miRNAs, mRNAs, methylations, and proteins) signature could be a biomarker for the prognosis of liver cancer patients. Better performance toward liver cancer was shown by logistic LASSO regression descriptor where significant improvement was seen in predicting liver cancer.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation, to any qualified researcher.

AUTHOR CONTRIBUTIONS

AK, XD, and D-QW designed the experiments. XD, D-QW, and AK performed the entire computational experiments and assisted in writing the manuscript. D-QW and AK analyzed the data and wrote the manuscript. AK, D-QW, XD, and AM read the manuscript and advised on method development. All authors approved the final version of the manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (Grant No. 81972789), the National Science and Technology Major Project (Grant No. 2018ZX10302205-004-002), the Six Talent Peaks Project in Jiangsu Province (Grant No. SWYY-128), the Fundamental Research Funds for the Central Universities (Grant No. JUSRP22011), and Technology Development Funding of Wuxi (Grant No. WX18IVJN017).

REFERENCES

- Alsaggar, M., and Liu, D. (2016). "Liver-targeted gene and cell therapies: an overview," in *Gene Therapy and Cell Therapy Through the Liver*, eds S. Terai and T. Suda (Tokyo: Springer), 1–11. doi: 10.1007/978-4-431-55666-4_1
- Baker, J. A., Kornguth, P. J., Lo, J. Y., Williford, M. E., and Floyd, C. E. Jr. (1995). Breast cancer: prediction with artificial neural network based on BI-RADS standardized lexicon. *Radiology* 196, 817–822. doi: 10.1148/radiology.196.3.7644649
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 68, 394–424. doi: 10.3322/caac.21492
- Calderaro, J., Ziol, M., Paradis, V., and Zucman-Rossi, J. (2019). Molecular and histological correlations in liver cancer. *J. Hepatol.* 71, 616–630. doi: 10.1016/j.jhep.2019.06.001
- Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., et al. (2012). *The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data*. Philadelphia: AACR.
- Chew, W. L., Tabebordbar, M., Cheng, J. K., Mali, P., Wu, E. Y., Ng, A. H., et al. (2016). A multifunctional AAV-CRISPR-Cas9 and its host response. *Nat. Methods* 13:868. doi: 10.1038/nmeth.3993
- Chhatwal, J., Alagoz, O., Lindstrom, M. J., Kahn, C. E. Jr., Shaffer, K. A., and Burnside, E. S. (2009). A logistic regression model based on the national mammography database format to aid breast cancer diagnosis. *Am. J. Roentgenol.* 192, 1117–1127. doi: 10.2214/AJR.07.3345
- Ciriello, G., Gatzka, M. L., Beck, A. H., Wilkerson, M. D., Rhie, S. K., Pastore, A., et al. (2015). Comprehensive molecular portraits of invasive lobular breast cancer. *Cell* 163, 506–519. doi: 10.1016/j.cell.2015.09.033
- Dooley, J. S., Lok, A. S., Garcia-Tsao, G., and Pinzani, M. (2018). *Sherlock's Diseases of the Liver and Biliary System*. Hoboken, NJ: John Wiley & Sons.
- Fang, B., Mehran, R. J., Heymach, J. V., and Swisher, S. G. (2015). Predictive biomarkers in precision medicine and drug development against lung cancer. *Chin. J. Cancer* 34, 295–309. doi: 10.1186/s40880-015-0028-4
- Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., et al. (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* 6:p11. doi: 10.1126/scisignal.2004088
- Houssami, N., Irwig, L., Simpson, J. M., McKessar, M., Blome, S., and Noakes, J. (2004). The influence of clinical information on the accuracy of diagnostic mammography. *Breast Cancer Res. Treat.* 85, 223–228. doi: 10.1023/b:brea.0000025416.66632.84
- Kanda, T., Goto, T., Hirotsu, Y., Moriyama, M., and Omata, M. (2019). Molecular mechanisms driving progression of liver cirrhosis towards hepatocellular carcinoma in chronic Hepatitis B and C infections: a review. *Int. J. Mol. Sci.* 20:1358. doi: 10.3390/ijms20061358
- Karimian, A., Azizian, K., Parsian, H., Rafeian, S., Shafiei-Irannejad, V., Kheyrollah, M., et al. (2019). CRISPR/Cas9 technology as a potent molecular tool for gene therapy. *J. Cell. Physiol.* 234, 12267–12277. doi: 10.1002/jcp.27972
- Lee, T. F., Chao, P. J., Ting, H. M., Chang, L., Huang, Y. J., Wu, J. M., et al. (2014). Using multivariate regression model with least absolute shrinkage and selection operator (LASSO) to predict the incidence of xerostomia after intensity-modulated radiotherapy for head and neck cancer. *PLoS One* 9:e89700. doi: 10.1371/journal.pone.0089700
- Li, B., and Li, J. Z. (2014). A general framework for analyzing tumor subclonality using SNP array and DNA sequencing data. *Genome Biol.* 15:473. doi: 10.1186/s13059-014-0473-4
- Li, B., Senbabaoglu, Y., Peng, W., Yang, M.-L., Xu, J., and Li, J. Z. (2012). Genomic estimates of aneuploid content in glioblastoma multiforme and improved classification. *Clin. Cancer Res.* 18, 5595–5605. doi: 10.1158/1078-0432.CCR-12-1427
- Li, J., Han, L., Roebuck, P., Diao, L., Liu, L., Yuan, Y., et al. (2015). TANRIC: an interactive open platform to explore the function of lncRNAs in cancer. *Cancer Res.* 75, 3728–3737. doi: 10.1158/0008-5472.CAN-15-0273
- Lu, Y., Min, Z., Qin, A., Wu, J., Jiang, X., and Qiao, Z. (2020). Role of miR-18a and miR-25 disruption and its mechanistic pattern in progression of liver cancer. *3 Biotech* 10:74.
- von Felden, J., and Villanueva, A. (2020). Role of molecular biomarkers in liver transplantation for hepatocellular carcinoma. *Liver Transpl.* doi: 10.1002/lt.25731
- West, C. A., Black, A. P., and Mehta, A. S. (2019). "Analysis of hepatocellular carcinoma tissue for biomarker discovery," in *Hepatocellular Carcinoma*, ed. Y. Hoshida (Cham: Humana Press), 93–107. doi: 10.1007/978-3-030-21540-8_5
- Yang, J. D., Hainaut, P., Gores, G. J., Amadou, A., Plymoth, A., and Roberts, L. R. (2019). A global view of hepatocellular carcinoma: trends, risk, prevention and management. *Nat. Rev. Gastroenterol. Hepatol.* 16, 589–604. doi: 10.1038/s41575-019-0186-y
- Zheng, Y., Shi, Y., Yu, S., Han, Y., Kang, K., Xu, H., et al. (2019). GTSE1, CDC20, PCNA, and MCM6 synergistically affect regulations in cell cycle and indicate poor prognosis in liver cancer. *Anal. Cell. Pathol.* 2019: 1038069.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Kaushik, Mehmood, Wei and Dai. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.