

Identification of Mutated Genes in Liver Cancer for Survival Prediction

Javad Razvian

Aug 17, 2019

1 Introduction

Liver cancer is common worldwide, more than 600000 deaths from liver cancer are estimated worldwide each year [5]. This study is a try to answer the first part of the question 2 [4]. In this part, we determine the 3 most frequently mutated genes in liver cancer, namely “TP53”, “TTN”, and “CTNNB1”. The data for Liver Hepatocellular Carcinoma (LIHC) is retrieved from “The Cancer Genome Atlas” (TCGA) with the aid of TCGAbiolinks package in R. It contains mutation and clinical data for liver cancer in addition to around 50 other cancers or their variations.

2 Method

We focused on the analysis of LIHC samples. In particular, we used TCGAbiolinks to download some samples with mutation. In this task, we have used two main packages, namely “maftools”[3] and “TCGAbiolinks”[2]. With the help of the latter we got access to liver cancer data and with use of the former we analyzed the data in order to answer the question. There were some subtle issues for combining the `maf` to the `clinical` data which is explained in the source code. There were four types of data in TCGA, called `muse`, `mutect`, `varscan2`, and `somaticsniper`. We used all of them for analyzing phase, but they are almost alike; in the report we just show the results of `mutect` variant as it contains more data compare to three other types.

3 Results

As it is shown in the fig. 1, “**TP53**”, “**TTN**”, and “**CTNNB1**” are the 3 most frequently mutated genes in the liver cancer. TP53 is the most frequently mutated gene in cancer[1], it also true in liver cancer based on our results.

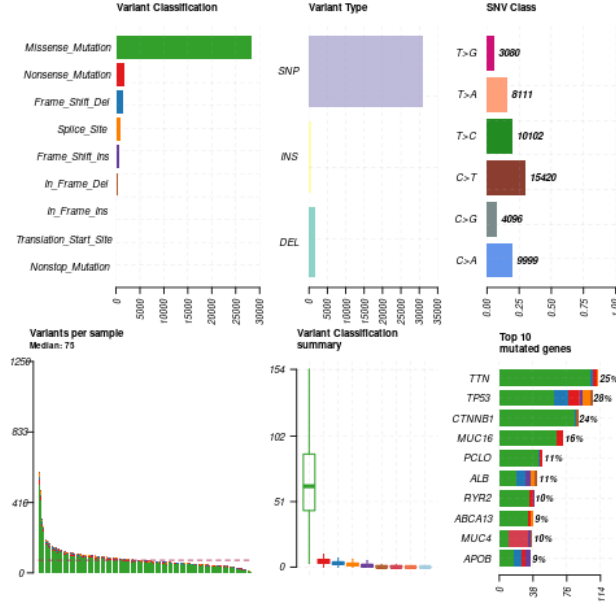


Figure 1: To 10 mutated genes

Which of these genes are more predictive of survival? We can simply answer to this by plotting the Kaplan-Meier survival curve. This plot depicts “time to event” curve. This gives the probability that a patient will survive beyond any specified time. In this case study, event is the patient’s death. The fig. 2 shows the KM plot for all of 3 most frequent mutated genes in Liver Cancer.

The results of fig. 2 are summarized in table 1. It is obvious that “TTN” and “CTNNB1” are more predictive to survive. As plots show their Mutant genes versus Wild Type genes are very similar and the P-values also support this fact that these genes have no significant differences with their WTs. As we know $P\text{-value} < 0.05$ was considered statistically significant. However the KM plot for “TP53” shows that this gene has a great effect on patient’s

genes	P-value	Hazard Ratio
TTN	0.837	1.08
CTNNB1	0.877	1.06
TP53	0.00672	2.31

Table 1: P-values and Hazard Ratio based on KM Plot

death and its P-value (0.00672) confirms this.

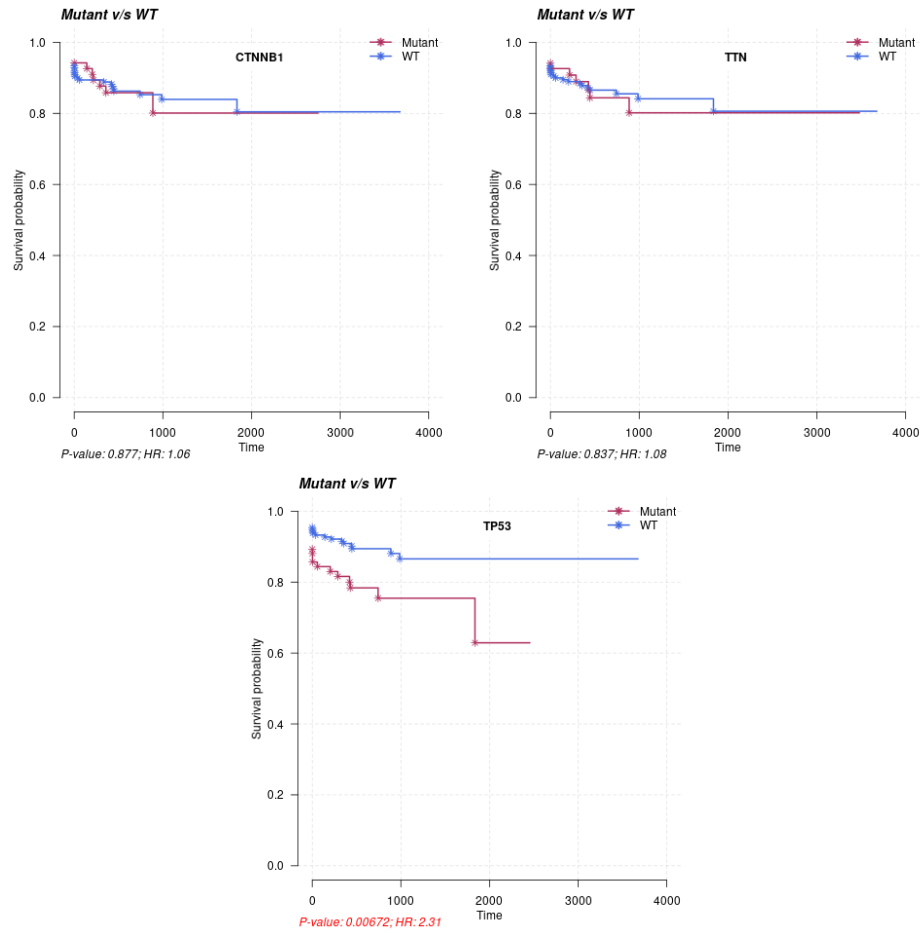


Figure 2: Kaplan-Meier Curve for Survival Analysis

References

- [1] Ungerleider, N. A., Rao, S. G., Shahbandi, A., Yee, D., Niu, T., Frey, W. D., & Jackson, J. G. (2018). Breast cancer survival predicted by TP53 mutation status differs markedly depending on treatment. *Breast Cancer Research*, 20(1), 115.
- [2] TCGAbiolinks: An R/Bioconductor package for integrative analysis with GDC data. Available at <https://bioconductor.org/packages/release/bioc/html/TCGAbiolinks.html>
- [3] Summarize, Analyze and Visualize MAF Files. Available at <https://bioconductor.org/packages/release/bioc/html/maftools.html>
- [4] An R Test. Available at http://oncinfo.org/r_test.
- [5] Rao, C. V., Asch, A. S., Yamada, H. Y. (2016). Frequently mutated genes/pathways and genomic instability as prevention targets in liver cancer. *Carcinogenesis*, 38(1), 2-11.
- [6] Kaplan Meier curve and hazard ratio tutorial, Available at <https://www.youtube.com/watch?v=pxRCpcRupgU>.
- [7] Pancanology Explains Kaplan-Meier Graphs, Available at <https://www.youtube.com/watch?v=KE1tkZmWhqU>.