Transcriptome Analysis of Wnt Signaling Pathway in Pediatric Hepatocellular Carcinoma

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

<table>
<thead>
<tr>
<th>Citation</th>
<th>Rabbani, Naveed. 2018. Transcriptome Analysis of Wnt Signaling Pathway in Pediatric Hepatocellular Carcinoma. Doctoral dissertation, Harvard Medical School.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:41973497">http://nrs.harvard.edu/urn-3:HUL.InstRepos:41973497</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
Scholarly Report submitted in partial fulfillment of the MD Degree at Harvard Medical School

Date: 26 February 2018

Student Name: Naveed Rabbani, BS

Scholarly Report Title: Transcriptome analysis of Wnt signaling pathway in pediatric hepatocellular carcinoma

Mentor Name and Affiliation: Khashayar Vakili, MD, Department of Surgery, Boston Children’s Hospital, Boston, MA

Collaborators with Affiliation: Ugur Ayturk, PhD, Department of Genetics, Harvard Medical School, Boston, MA. Matthew Warman, MD, Department of Genetics, Harvard Medical School, Boston, MA. Ghazaleh Sadri-Vakili, PhD, MassGeneral Institute of Neurodegenerative Disease, Massachusetts General Hospital, Boston, MA.
Abstract

TITLE: Transcriptome analysis of Wnt signaling pathway in pediatric hepatocellular carcinoma

Naveed Rabbani¹, Ugur Ayturk, PhD²,³, Michael LaQuaglia, MD¹, Matthew Warman, MD²,³, Ghazaleh Sadri-Vakili, PhD⁴, Khashayar Vakili, MD¹

1. Department of Surgery, Boston Children’s Hospital and Harvard Medical School, Boston, MA
2. Department of Orthopedic Surgery, Boston Children’s Hospital, Boston, MA
3. Department of Genetics, Harvard Medical School, Boston, MA
4. MassGeneral Institute of Neurodegenerative Disease, Massachusetts General Hospital, Boston, MA

Background: Pediatric hepatocellular carcinoma (HCC) is associated with high mortality partly due to the lack of effective chemotherapy. Therefore, identification of therapeutic targets is critical. While Wnt signaling has been widely implicated in adult HCC tumorigenesis, often with underlying cirrhosis, comparable studies in pediatric HCC, which typically occurs without cirrhosis, are lacking. We therefore sought to examine whether the Wnt pathway is also altered in pediatric HCC.

Methods: Next-generation sequencing of mRNA from HCC tumor and non-neoplastic liver tissue from 3 pediatric patients (1.5, 13, and 13 years old) without underlying cirrhosis was performed. Sequenced reads were mapped to the human genome. Normalized expression levels were calculated and used to compare normal tissue to tumor. P-values were calculated by fitting the data to a negative binomial distribution and correcting for multiple hypothesis testing.

Results: Transcriptome analysis demonstrated a statistically significant increase in the expression of key pro-proliferative proteins of the canonical Wnt pathway, including WNT1, WNT10B, and Wnt-receptor FZD10, in the tumor compared to non-neoplastic tissue. Additionally, there was a significant decrease in several pathway inhibitors including SFRP5, TBL1Y, and WNT11 (p-values < 1.0x10⁻⁴). Sequenced reads from one tumor sample also revealed a previously-reported gain-of-function p.S45F CTNNB1 mutation.

Conclusion: Our results demonstrate a significant increase in the expression of pro-proliferative genes alongside a concomitant decrease in the expression of inhibitory genes of the canonical Wnt pathway in pediatric HCC tumors. These findings suggest that targeting the Wnt pathway may provide a potential therapeutic target for the treatment of pediatric HCC.
# Table of Contents

Introduction ................................................................................................................................. 4
Student Role ................................................................................................................................. 4
Methods ...................................................................................................................................... 5
Results ......................................................................................................................................... 6
Discussion ................................................................................................................................. 7
Acknowledgments ....................................................................................................................... 8
References .................................................................................................................................... 8
Tables and Figures ....................................................................................................................... 10
Introduction

Hepatocellular carcinoma (HCC) is a common cause of cancer-related mortality in adults and the second-most common liver malignancy in children.\(^1\) In children, HCC has a prevalence of 0.5 – 1.0 per million patients and accounts for an estimated 30% of pediatric liver cancers and 0.5% of all pediatric cancers.\(^2\) HCC is an aggressive cancer which responds poorly to chemotherapy and is resectable in only one-third of cases. Furthermore, a significant subset of surgically-treated patients go on to develop tumor recurrence. As a result, pediatric HCC survival rates are below 30%.\(^3,4,5,6,7,8\) Therefore, developing an effective chemotherapeutic approach could be essential in treating this disease.

However, since HCC is far more common in adults, the majority of our understanding of the pathophysiology of HCC is derived from studies of adult tumors. HCC in adults is generally associated with cirrhosis, liver disease and viral hepatitis.\(^9\) On the other hand, epidemiologic evidence strongly suggests that the risk factors for HCC in children and adolescents differ from those in adults, since pediatric HCC cases are often not associated with liver disease.\(^6,10\) As a result, a more in-depth analysis of the molecular characteristics of pediatric HCC is required to better understand the pathophysiology of the disease and identify possible therapeutic targets.

In this study, we used state-of-the-art biocomputational methods to perform genetic analyses on paired hepatocellular carcinoma and healthy liver tissue samples derived from pediatric patients treated at Boston Children's Hospital. The aim of our study is to identify biological pathways involved in the tumorigenesis of pediatric HCC, with the long-term goal of identifying potential therapeutic targets. To achieve this, we used a combination of RNA-seq transcriptome analysis and whole exome sequencing.

Student Role

Project Conception and Proposal

The research project was conceived by my mentor Khashayar Vakili, MD. Under the supervision of Khashayar Vakili, MD, and collaborator Ugur Ayturk, PhD, I developed a research proposal, which was awarded a summer research stipend from the "Scholars in Medicine Office" at Harvard Medical School and a matching contribution from the Alexandra J. Miliotis Pediatric Oncology Research Fellowship.


**Computational Work**

I performed RNA-seq and whole exome sequencing analysis on 3 pairs of cancer and healthy liver samples from pediatric hepatocellular carcinoma patients at Boston Children’s Hospital. The genetic material had already been extracted from the tissue samples, and next generation sequencing was performed by the hospital’s core facilities. I received the electronic files that result from this process, and began my analysis there, doing the majority of my work on the HMS “Orchestra” computing cluster. The analysis included two parts: whole exome sequencing and RNA-seq. Under the guidance of Ugur Ayturk, PhD, I developed computational methods for performing whole exome analysis to identify driver mutations and RNA-seq analysis to identify signaling abnormalities in cancer tissue samples. I also wrote a guide of these methods to help future lab members apply the same methods to future projects.

**Final Analysis and Preparation for Publication**

I discussed the findings of my computational methods in weekly lab meetings with my mentor and other lab members and collaborators. Based on these discussions, and under the guidance of my mentor and collaborators, I wrote an abstract, which was submitted and accepted to the Massachusetts Chapter of the American College of Surgeons 62nd Annual Meeting. I was the primary author for the abstract, poster, and figures. The publication won a “Poster of Distinction” award and the “Massachusetts Chapter of the American College of Surgeons Commission on Cancer Paper Competition.” The poster’s citation is:


**Methods**

Next-generation sequencing of genomic DNA and messenger RNA from both tumor and healthy tissue was performed in order to identify disease-causing signaling pathway anomalies and somatic mutations in three pediatric patients, aged 1.5, 13, and 13 years old (Figure 1). Fresh-frozen tumor and healthy tissue from resected specimen were used to prepare libraries for whole exome sequencing (WES) and mRNA sequencing (RNA-seq) as previously described.11,12 Over 30 million 50bp and 100bp paired-end reads were generated for RNA-seq and WES respectively using the Illumina HiSeq 2500 platform.
**RNA-seq Bioinformatics Pipeline**

Sequenced RNA-seq paired reads were mapped to GRCh38 using the RNA-Unified Mapper (RUM). Expression levels and fold changes were calculated by comparing RNA-seq data from normal tissue samples against tumor tissue samples. All calculations were made in R using Rsamtools, GenomicFeatures, and edgeR R subroutines. EdgeR was specifically used to normalize differential expression levels using the Trimmed Mean of M values (TMM) method and correct p-values for multiple hypothesis testing. Annotations, gene pathway, and gene ontology analysis were performed also performed in R using the org.Hs.eg.db and goseq subroutines and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways database.

**WES Bioinformatics Pipeline**

Sequenced WES paired reads were mapped to the Genome Reference Consortium Human Genome Build 38 (GRCh38) using the Burrows-Wheeler Aligner (BWA). PCR duplicates were removed from each WES dataset using Picard MarkDuplicates. Reads around areas of known variation (determined by Single Nucleotide Polymorphism Database GRCh38 Build 144) were locally realigned using GATK as suggested by “GATK: Best Practices.” A list of single nucleotide variants and insertions and deletions were compiled using Samtools and VarScan2. The variants were annotated with ANNOVAR. Those variants that yielded either stop-gain or non-synonymous changes in the protein-coding region of the genome were identified and analyzed. Circular plots were created with the resulting genetic data using the CIRCOS software package.

**Results**

RNA-Seq differential expression analysis of paired tumor and healthy liver samples from three patients (N=3) revealed several statistically significant abnormalities in the canonical Wnt signaling pathway, which appeared to be the most consistently abnormal signaling pathway among the three samples. Within the Wnt signaling pathway, a statistically significant increase in the expression of the following pro-proliferative genes was observed: \textit{WNT1}, \textit{WNT10B}, and Wnt-receptor \textit{FZD10}, when comparing tumor to non-tumor tissue (Figures 2, 3). Additionally, a statistically significant decrease in the expression of the following anti-proliferative genes was observed: \textit{SFRP5}, \textit{TBL1Y}, and \textit{WNT11} (all p-values < 1.0x10^{-4}).
Furthermore, with our whole exome sequencing pipeline, we identified 5 frameshift and 444 non-synonymous somatic substitution mutations in the first tumor sample, 20 frameshift and 679 non-synonymous somatic substitution mutations in the second tumor sample, and 7 frameshift and 330 non-synonymous somatic substitution mutations in the third sample. While the majority of these mutations could not be clearly linked to the differential expression patterns observed with RNA-Seq, a single nucleotide p.S45F substitution in CTNNB1 was identified in one patient, which is likely directly related to overexpression of the Wnt signaling pathway observed in that sample (Figure 4).

**Discussion**

Transcriptome analysis revealed consistent and statistically significant pro-proliferative expression of the Wnt signaling pathway. This is consistent with what has been previously observed in adult HCC studies as well.\(^{30}\) Furthermore, the CTNNB1 point mutation identified via whole exome sequencing is a previously reported mutation, which has been identified in a number of solid tumor cancers in the COSMIC database, including, most commonly, melanoma and liver cancer.\(^{31}\) Mutations like this, which occur at the 45\(^{th}\) residue of β-catenin, an important regulatory phosphorylation site, have been shown to stabilize the transcription factor, allowing it to uninhibitedly localize to the nucleus and promote cell proliferation.\(^ {32}\)

Our findings suggest that the Wnt signaling pathway is involved in the tumorgenesis of pediatric HCC, as it is believed to do in adult HCC, despite the epidemiologic and phenotypic differences between the adult and pediatric forms of the disease. As such, the Wnt pathway may provide a potential therapeutic target for the treatment of pediatric HCC and its role in this disease warrants further investigation.

This study is limited by a small sample size (N=3) and also in the scope of its analysis. For example, while RNA-sequencing is a powerful tool, it cannot identify post-transcriptional gene regulation mechanisms, such as siRNAs; and RNA expression profiles do not always correlate with gene expression. Furthermore, with whole exome sequencing, we are unable to identify disease-causing variations in the intronic portions of DNA, which account for the large majority of the human genome. Finally, we have limited our DNA analysis to somatic mutations, and have excluded germline mutation analysis for the time being.

Future work should include repeating similar analysis on more samples as they become available, and, as resources allow, expanding the analysis to include intronic DNA and germline mutations. As more samples become available, we may also be able to correlate RNA expression profiles with specific HCC phenotypes.
Acknowledgments

This study was supported by Alex’s Team through the Alexandra J. Miliotis Fellowship in Pediatric Oncology, Harvard Medical School Scholars in Medicine Office, the Tisch Families’ Faculty Development Fund, and the Joshua Ryan Rappaport Surgical Fellowship Fund.

References


**Tables and Figures**

*Figure 1. Schematic overview of RNA-seq and whole exome sequencing analysis methods*
Figure 2. (A) Comparison of normalized mRNA expression level of Wnt pathway components in tumor and non-neoplastic tissue, (B) Heatmap representation of individual samples.

Figure 3. Simplified Wnt pathway showing statistically significantly over-expressed pro-proliferative proteins in red and under-expressed anti-proliferative proteins in blue.
Figure 4. pS45F CTNNB1 Driver Mutation in Patient #3