

A Survey of Viral Integration Sites in Hepatocellular Carcinoma

Katie L. Hsia¹, Shaolei Lu²

1) Warren Alpert Medical School 2) Lifespan Physician Group

Abstract

Viral insertions in the human genome have been shown to increase instability and encourage carcinogenesis. Genetic information from viral Hepatitis has been found in hepatocellular carcinoma (HCC). Studies have not previously investigated what other viral integrations are found in HCC. This study found viral information in nearly all of the samples, which could possibly be contributed to contamination. Ignoring contaminants, Hepatitis was found in nearly half of the viral insertions. Older patients with a history of Hepatitis had a greater number of insertions.

Introduction

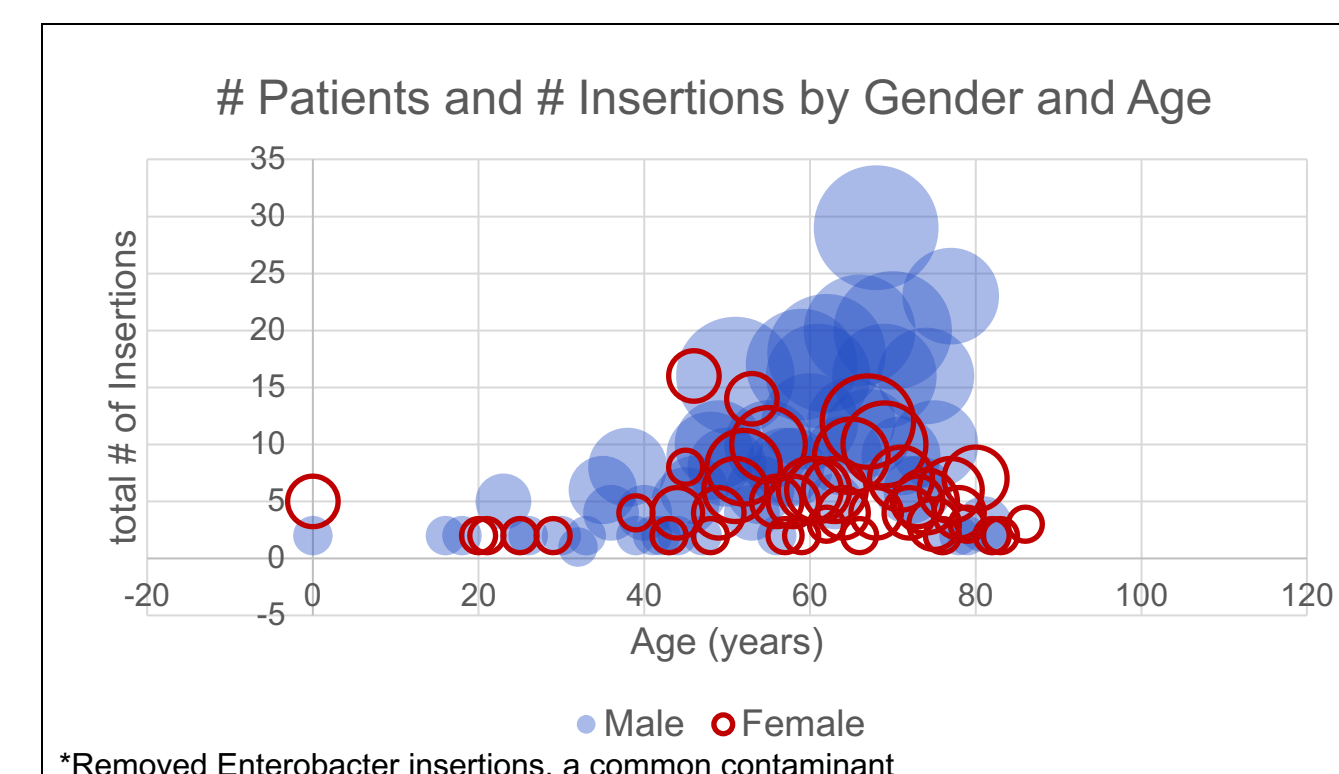
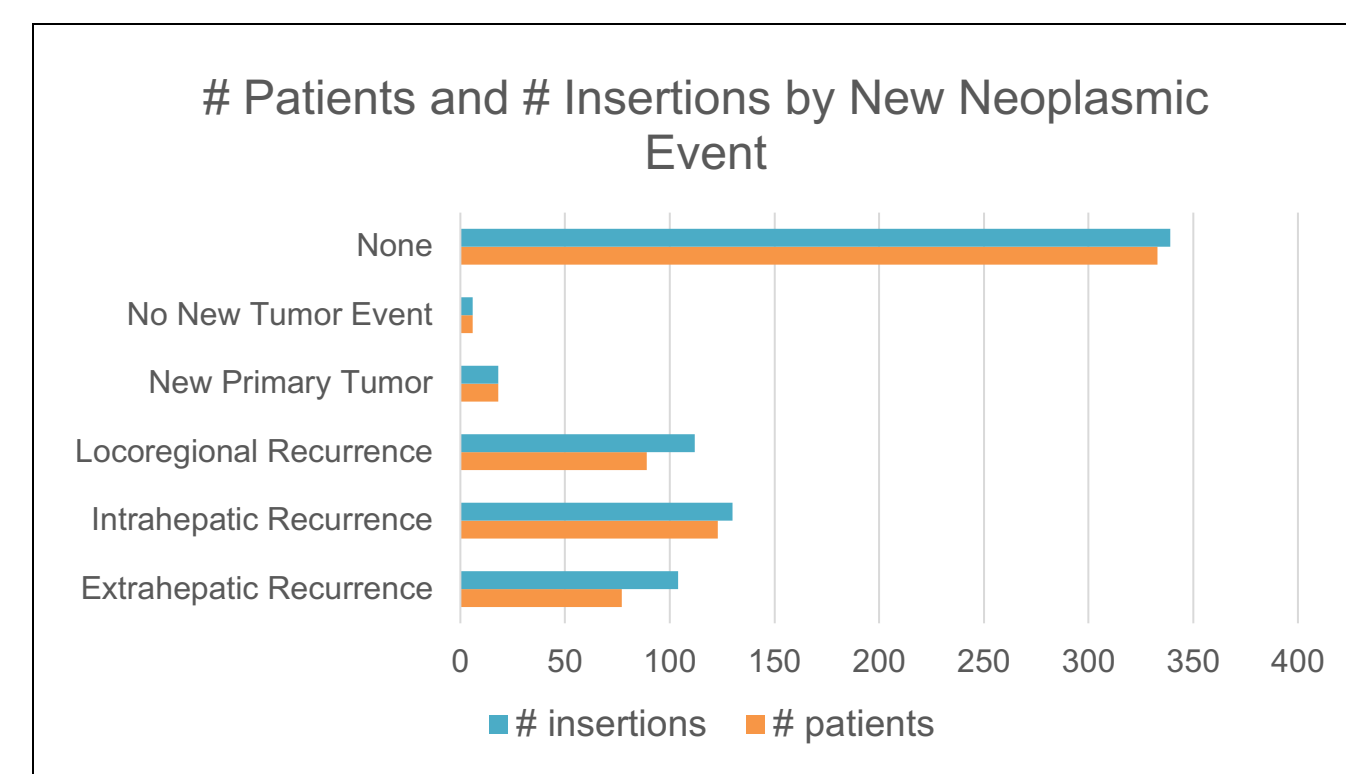
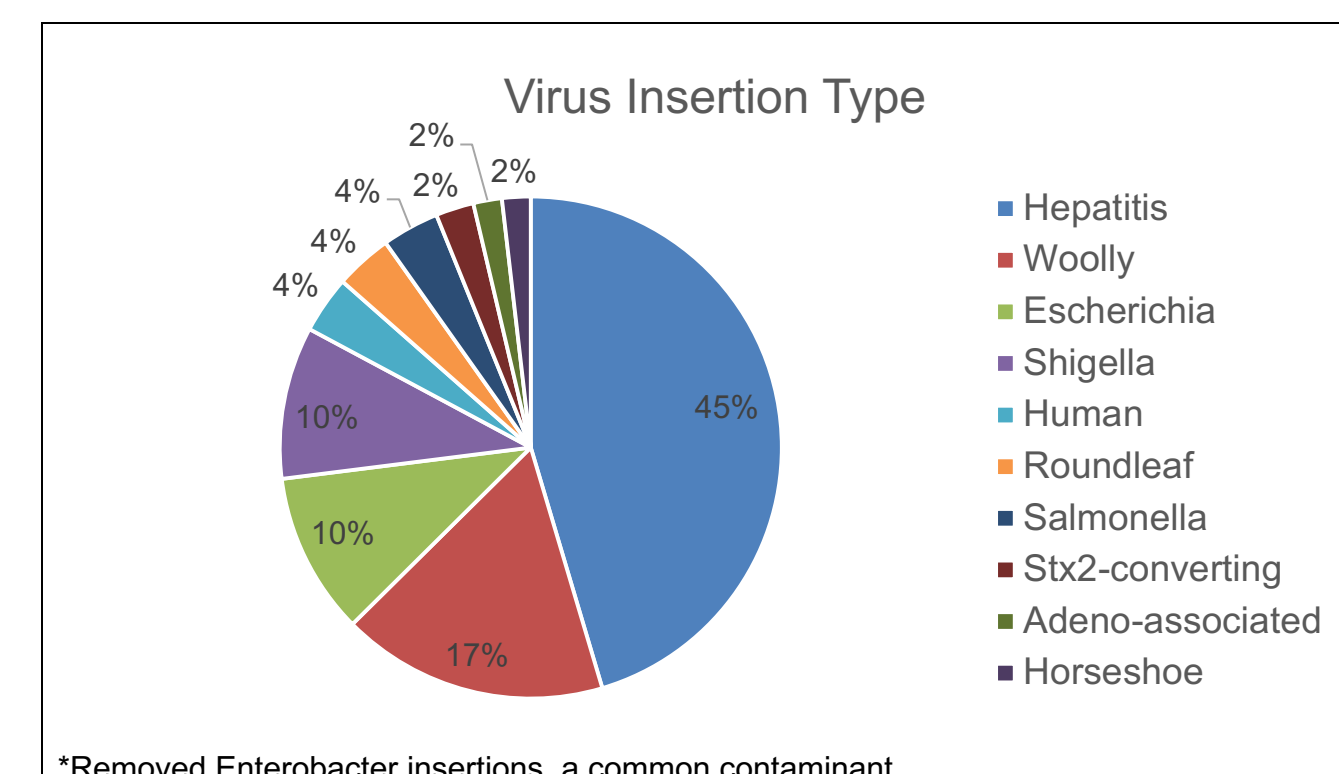
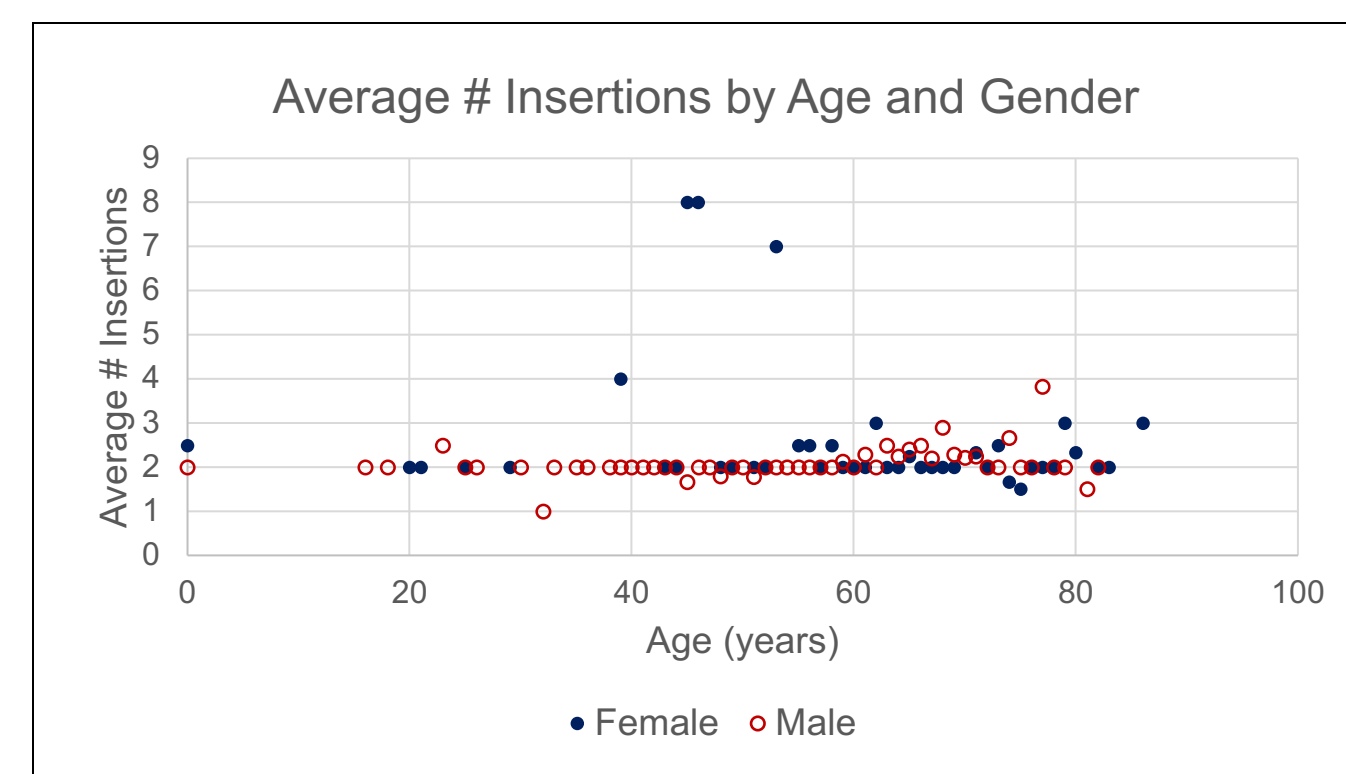
Viral interference has been shown to play a role in the progression of HCC. HCC is the 5th most common cancer and the 3rd deadliest cancer. The incidence of HCC has tripled since the 1980s and is expected to increase with the associated peak of the Hepatitis C virus, and the rise of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis. The integration of Hepatitis B viral material has been found in HCC.

Methods

All available genomes from patients with HCC were obtained from The Cancer Genome Atlas. Genomes of 7,500 viruses were obtained from viruSITE. Using VirusFinder2.0, the VERSE algorithm was applied to determine the presence of viral genomic material in the HCC samples.

Results

Viral insertions were found in 1,095 of the 1,212 samples (90%). There were a total of 1,289 viral insertions. Enterobacter comprised 1,126 (87.35%) of the insertions. Of 80 unique samples containing human viruses, there are 3 samples with Human herpesvirus 6B and Human betaherpesvirus 6A, 3 samples with Human adenovirus C, and 74 samples with Hepatitis B. The majority of patients with multiple viral insertions were found with Hepatitis B, Hepatitis C, or no history of primary risk factors. Generally, older patients had a greater number of insertions.



# Patients and # Insertions by TNM Classification and Risk History					
	# patients	# insertions	# patients	# insertions	# patients
T1	322	360	NX	202	258
T2	173	183	N0	461	485
T3	152	172	N1	6	6
T4	20	32			

History	# patients	# insertions
Alcohol consumption	121	128
Hemochromatosis	13	13
Hepatitis B	168	173
Hepatitis C	90	116
No history of primary risk factors	167	205
Non-alcoholic fatty liver disease	30	30
None	36	36
Other	46	50

Conclusions

Viral genomic material was found in over 90% of HCC genomes examined. Many of the viral insertions can be attributed to contamination of the Next Generation Sequencing. It is difficult to determine which viral insertions are genuine and which are merely contaminations. Human-related viruses were found in 6% of the virus-containing samples, with Hepatitis B found in 92% of these. These numbers support claims regarding the importance of viral infection, primarily Hepatitis B, as a risk factor for HCC. It also raises questions about the role of other human viruses as well as the role of viruses from other organisms in the progression of HCC.

The patient population affected skewed older and male. More older patients had insertions and of those patients, they each had a greater number of insertions per patient. Those with extrahepatic reoccurrences had a greater number of insertions on average. Furthermore, those with a history of Hepatitis had a greater number of insertions.

Future Directions

- Learn more about the sites of integration
- Determine if Enterobacter or other non-human viral insertions are contamination
- Explore the model of viral insertion and relation to HCC

Acknowledgments

- Summer Assistantship
- Brown Center for Biomedical Informatics