

# Supplemental Figure S5

A

Insertion site : chr6: 62,429,042



5' - CCTAGGTTGTGTGAGTTCTCTCACCTGATACATTATTTAAGAAAAGCAAAGACTTCCATTTAAATATTTTGCCTATTATACCAATCTA - 3'  
 3' - GGATCCAACACACTCAAGAGAGTGGACTATGTAATAAAATTCCTTTTCGTTTCTGAAGGTAAATTTATAAAAACGGATAATATGGTTAGAT - 5'

Insertion structure:

5' - TCTCTCACCTGATACATTATTTAAGAAAAGCAAAGAgT (A)<sub>51</sub> AAGAAAAGCAAAGACTTCCATTTAAATATTTTG - 3'  
 3' - AGAGAGTGGACTATGTAATAAAATTCCTTTTCGTTTCTca (T)<sub>51</sub> TTCCTTTTCGTTTCTGAAGGTAAATTTATAAAAC - 5'

TSD: 16nt

5961

L1 EN motif : TT/AAGA

B

Insertion site : chr4: 61,822,453



5' - AAATGCAGATCCTAAAATTTTGTGGAAGTAAAAAAAATGCAtA ATA GAAGCATAAAGTGAGGGGATTGTCTTTACAGATATCA - 3'  
 3' - TTTACGTCTAGGATTTTAAAAACACCTTCATTTTTTTTTTACGTTT TAT CTTCGTATTTCACTCCCCTAAACAGAAATGTCTATAGT - 5'

genomic deletion

Insertion structure:

5' - AGATCCTAAAATTTTGTGGAAGTAAAAAAAATGCAtA (A)<sub>26</sub> GAAGCATAAAGTGAGGGGATTGTCTTTACAGAT - 3'  
 3' - TCTAGGATTTTAAAAACACCTTCATTTTTTTTTTACGTaT (T)<sub>26</sub> CTTCGTATTTCACTCCCCTAAACAGAAATGTCTA - 5'

TSD: none

5902

L1 EN motif : TA/GAAG

C

Insertion site : chr14: 85,294,576



5' - CATTGTCTCGTTTGAGTAATTAGTTTTTCCATATTTAAAATAGTTGTGAATATCTGGAGATACTATGCAAATACCATGCAAATTGATT - 3'  
 3' - GTAACAGAGCAAACCTCATTAATCAAAAAGGTATAAATTTTATCAACACTTATAGACCTCTATGATACGTTTATGGTACGTTTAACTAA - 5'

Insertion structure:

5' - TAGTTTTTCCATATTTAAAATAGTTGTGAAGTGTCIGTT (A)<sub>49</sub> TTTAAAATAGTTGTGAATATCTGGAGATAC - 3'  
 3' - ATCAAAAAGGTATAAATTTTATCAACACTT CACACACAA (T)<sub>49</sub> AATTTTATCAACACTTATAGACCTCTATG - 5'

TSD: 16nt

untemplate nt

5690

L1 EN motif : AT/TTAA

D

Insertion site : chr4: 147,188,057



5' - GAATGAAATGCACCTACTTTTGTGAGTTTTTGGATTTTTTAAAAATAAAGAATAAAAATAATATAAAAATGTCTGGTATATAGCAAGTGTT - 3'  
 3' - CTTACTTTACGTGGATGAAACAGTCAAAAACCTAAAAATTTTTATTCTTATTTTTATTATATTTTACAGACCATATATCGTTCACAA - 5'

Insertion structure:

5' - ACTTTGTGAGTTTTTGGATTTTTTAAAAATAAAG (A)<sub>14</sub> AAAAATAAAGAATAAAAATAATATAAAAATGTCTGGT - 3'  
 3' - TGAACAGTCAAAAACCTAAAAATTTTTATTCT (T)<sub>14</sub> TTTTTATTCTTATTTTTATTATATTTTACAGACCA - 5'

TSD: 10nt

5614

L1 EN motif : TT/AAAA

**E**

Insertion site : chr11: 101,194,358

chr 11 

5' - AATACCTGCTTGCTACAAAGAAGCAGAAAAAATCTAAAGT**AGA**AAATAGAAAGTCACAATAATCCCATGACCCAGATATATAACCACT - 3'

3' - TTATGGACGAACGATGTTTCTTCGTCCTTTTTTTAGATTTCAT**CT**TTTATCTTCAGTGTATTAGGGTACTGGGTCTATATATTGGTCA - 5'

Insertion structure:

5' - GCTTGCTACAAAGAAGCAGAAAAAATCTAAAGT**AGA**  (A)<sub>68</sub> **AGAAAAATAGAAAGTCACAATAATCCCATGACCCA** - 3'

3' - CGAACGATGTTTCTTCGTCCTTTTTTTAGATTTCAT**CT**  (T)<sub>68</sub> **TCTTTTATCTTCAGTGTATTAGGGTACTGGGT** - 5'

5648

TSD: 3nt

L1 EN motif : GT/AGAA

**F**

Insertion site : chr4: 67,639,716

chr 4 

5' - TTAAATATTCTATTTTATAAAACATGAATTATGCAT**AAAAAATTTCTGCTCC**TAAACAAAAGTATGGTAATTTTCTAAAGAATAAAC - 3'

3' - AATTTATAAGATAAAATATTTTGTACTTAATACGT**TTTTTTAAAGACGAGG**ATTTGTTTTCATACCATTAAAAGATTTCTATTG - 5'

Insertion structure:

5' - TTATAACATGAATTATGCAT**AAAAAATTTCTGCTCC**  (A)<sub>57</sub> **AAAAAATTTCTGCTCC**TAAACAAAAGTATGGT - 3'

3' - AATATTGTACTTAATACGT**TTTTTTAAAGACGAGG**  (T)<sub>57</sub> **TTTTTTAAAGACGAGG**ATTTGTTTTCATACCA - 5'

5734

TSD: 16nt

L1 EN motif : AT/AAAA

**G**

Insertion site : chr5: 103,499,178

chr 5 

5' - ATTTTATAAAGCAAAACAATAATTTTCTAATAAT**AAAACCTGCAAC**CATATGAAAAATAAACTGGAGGAGAAAATTCAAATCAT - 3'

3' - TAAAATATTTGTTTTGTTGATTAAAAAGATTATTA**TTTGGACGTTG**TATACTTTTATTTGACCTCCTCTTTAAGTTTAGTA - 5'

Insertion structure:

5' - AAAACAATAATTTTCTAATAAT**AAAACCTGCAAC**  (A)<sub>121</sub> **AAAACCTGCAAC**CATATGAAAAATAAACTGGAGAAA - 3'

3' - TTTTGTGATTAAAAAGATTATTA**TTTGGACGTTG**  (T)<sub>121</sub> **TTTGGACGTTG**TATACTTTTATTTGACCTCCTT - 5'

5977

TSD: 12nt

L1 EN motif : AT/AAAA

**H**

Insertion site : chr8: 88,407,755

chr 8 

5' - AAAGTTTTTAAAAGTAGTTCTGAGTATAATAGAAGT **ATAAAAAATATT** **AAAAGTATTGTTATCTTTAAAACAGTGGTACTAAACCT** - 3'

3' - TTTCCAAAATTTTTCATCAAGACTCATATTATCTTCA **TATTTTTATAA** **TTTTTCATAACAATAGAAATTTTGTCCACCATGATTTGGA** - 5'

genomic deletion 

Insertion structure:

5' - AAAGTTTTTAAAAGTAGTTCTGAGTATAATAGAAGT  (A)<sub>83</sub> **AAAAGTATTGTTATCTTTAAAACAGTGGTAA**TAA - 3'

3' - TTTCCAAAATTTTTCATCAAGACTCATATTATCTTCA  (T)<sub>83</sub> **TTTTTCATAACAATAGAAATTTTGTCCACCAT**tATTT - 5'

5988

TSD: none

L1 EN motif : TT/AAAA

**Supplemental Figure S5. Structures of tumor-specific L1 insertions in HCC patients with a history of alcoholism.**

The chromosomal location and nucleotide position of each insertion is shown; if the insertion was found within a gene, the gene name is so indicated. The genomic empty site DNA is depicted above, with the region that ultimately gave rise to the TSDs highlighted in blue. First strand endonuclease cleavage position is indicated with a red arrowhead and second strand cleavage position is indicated with a black arrowhead; the EN cleavage motif is underlined. Genomic sequence deletions are boxed and labeled as such. The insertion structure is depicted below; TSDs are highlighted in blue, the L1 is depicted as rectangles, 3' UTR sequence is depicted in red, and ORF2 sequence is depicted in white. Untemplated nucleotides are so labeled. (A)<sub>n</sub>/(T)<sub>n</sub> indicate poly(A) tail lengths. For each insertion, the TSD length and EN motif are summarized.