

Supplemental Figure S3: Localization of vector proviruses to the human genome.

Sequences obtained by shuttle vector rescue are localized to the human genome using BLAT to identify the integration site and nearby genes.

1) ATPAF1

LTR-**chromosome** junction sequence read after shuttle vector rescue

AGTGTGGAAAATCTCTAGCACCTCCTACCTCAGCCTCCTGAGTAGCTGGGACTACAGGCTCACACCAATGC
CCCCAACCAATTTTGTGCATAGTCTACTTGTGTTAGAAGGAAGTCACTGAGTATAGCCATTACCCAAGGAAAGG
AAATTAGGCTCCATCAAAGGGAGGAATATCTTAAAATTTGTGGCAAATTTTAAAACCATCACAGCTGGCT
CTGGGAGTACTTTTCTGACCAAGGTACTATCATAACAGGTTTTATTCAAGTGTATATCTGGTTAGTCAAGA
ACTTAGTTCTCACTCACAAACTGGTTCTTAGAGCTGGGTAAC TATTGCTGTCTTATTCCTCCCTTGAATA

Query sequence localized to chromosome 1 using BLAT

| | | | | | |
|-------------------|------------|------------|------------|------------|----------|
| ggcttgatag | cagaactcat | tctccatcat | ttttcatttt | ttcatagtct | 47102416 |
| tttttttttt | ttaaggggct | cactacatta | cccagacagg | ctcaagcaat | 47102366 |
| <u>CCTCCTACCT</u> | CAGCCTCCTG | AGTAGCTGGG | ACTACAGGCT | CACACCAATG | 47102316 |
| CCCCAACCA | TTTTGTGATA | GTCTACTTGT | TAGAAGGAAG | TCACTGAGTA | 47102266 |
| TAGCCATTAC | CCAAGGAAAG | GGAAATTAGG | CTCCATCAA | GGGAGGAATA | 47102216 |
| TCTTAAAATT | TGTGGGCAA | TTTTAAAACC | ATCACAGCTG | GCTCTGGGAG | 47102166 |

2) TRPM4

LTR-**chromosome** junction sequence read after shuttle vector rescue

AGTGTGGAAAATCTCTAGCAGCCACCACACCTGGCTAATTTTGTAGTTTTAGTAGAGATGGGGTTTTCTCCA
TATTGGTCAGGCTGGTCTCGAACTCCCGACCTCGGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGA
TTACAGGCATGAGCCACTGTGCCAGCCTTTTTTGTATTTTTAGTAGAGATGGGGTTTTCACTATGTTGCC
AGGCTGGTCTCGAACTCCTGGGCTAAAGCTATCTGCCATCTCGGCCTCCCAACGTGCTGGGATTACAGGC
GTGAGCCATTGTGCTTAGGCCAGAAGAATTTTACGGGGGAAATACACACTAAGATAGAAACTCCCCAGATA

Query sequence localized to chromosome 19 using BLAT

| | | | | | |
|-------------------|------------|-------------|------------|-------------|----------|
| gagtgcaata | gtgccatctc | agctcactgc | aacctccgcc | tcccgggctc | 49678431 |
| aagtgattct | cctgectcag | cctceccaagt | agctgggatt | acaggecatgc | 49678481 |
| <u>GCCACCACAC</u> | CTGGCTAATT | TTGTAGTTTT | AGTAGAGATG | GGGTTTCTCC | 49678531 |
| ATATTGGTCA | GGCTGGTCTC | GAACTCCCGA | CCTCGGGTGA | TCCGCCCGCC | 49678581 |
| TCGGCCTCCC | AAAGTGCTGG | GATTACAGGC | ATGAGCCACT | GTGCCAGCC | 49678631 |
| TTTTTTGTAT | TTTTAGTAGA | GATGGGGTTT | CACTATGTTG | CCCAGGCTGG | 49678681 |

3) GCOM1

LTR-**chromosome** junction sequence read after shuttle vector rescue

AGTGTGGAAAATCTCTAGCACATATGAAGAGTCAGGAAAATAGTACCCACTCCTAAGGGAAAAGAGAATCA
ACAGACACCAATCCTGATGTAACCTGATGTTGGAATTTTTCAGACAAAGACTTTAATGCAGCCATCATAAC
AATGATGAGTCAGACAAAGAAAATATGCTTACAATTAATGAAGTAGTAAAAAAAATAGAAAATACAAAA
AGAATAAATGAAAATTTTAAAGAACAGGAAAGTACAATACCTGAATTTTTTTTTTAAAAGTCACTAGATGGA
CTTTAAGACTGAAGAATGGTAAAGGAAAAAAGTCAGTAAAATTTTACAATACAATTTAAATATTAATTAT

Query sequence localized to human genome chromosome 15 using BLAT

```
caggcaggat gtagttgggt ctggctaaat tcattgcctt ttaaaacaag 57964177
aaccacaaca agcactcaca atgtccagct tataatccaa aattatgtga 57964127
CATATAAAGA GTCgGGAAAA TAGTACCCAC TCCTAAGGGA AAAGAGAATC 57964077
AACAGACACC AATCCTGATG TAACTaTGAT GTTGGAAATTT TCAGACAAAG 57964027
ACTTTAATGC AGCCATCATA ACAATGATcA GTCAGACAAA GAAAAATATG 57963977
CTTACAATTA ATGAAGTAGT AAAAAAATA GAAAATACAA AAAAGAACTA 57963927
```

4) PTRF

LTR-**chromosome** junction sequence read after shuttle vector rescue

```
AGTGTGGAAAATCTCTAGCAGGTTGTTTTAGTGTTCATTTCAATTTTGATACACAGAATTAGAATAGCAT
CCAGATGTGGGTCTGTTACAGCTAGACTACTAGATCCTTCAAAATCCAAGTACTAGTATGTCTATTTAAAT
ACCATAAGATCACATTGGCTAGTTACAATGGTTGGTTTGTGGGTTACTTAAAAATCAACTAAAATTTT
TTTTTTTTGAGATGGAGTTTTGCTCTTGTTCCTAGGCTGGAATGCAATGACACAATCTTGGCTCACTGCC
ACCTCTGCCTCCAGGTTCAAGCAATCCCCTGCCTTAGCCTCCTGAGTAGCTGGGATTACAGGCATGTGC
```

Query sequence localized to human genome chromosome 17 using BLAT

```
ctctcctaac ctcaactcctg ttgcctgcat cttcttgctg agcaaaaatat 40512915
tcaaggtcct caactcctca caccctgggt gtccctccct ggatgtggtt 40512865
GGTTGTTTTTA GTGTTCCATT TCAATTTTGA TACACAGAAT TAGAATAGCA 40512815
TCCAGATGTG GGTCTGTTAC AGCTAGACTA CTAGATCCTT CAAAATCCAA 40512765
GTACTAGTAT GTCTATTTAA ATACCATAAG ATCACATTGG CTAGTTACAA 40512715
TGGTTGGTTT GTGGGTTACT TAAAATCAA CTAAAATTct TTTTTTTTTT 40512665
```

5) MEX3D

```
AGTGTGGAAAATCTCTAGCAGATTCCCCTTGTGAGCCTCCATGTCTGAAGCCAGTCTTTTGTCACTTGACC
TCGGCAGTGATGCCCCATTGCTCATGCTGCGTTCTACTCATTAGAAGCCTGTGCCCGGCCGGCGGGTGG
CTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCCGGCGGATCACGAGGTGAGGAGATTGAGACCAT
CCTGGCTAACACGGTGAAACCCCATCTCTACTAAAAATGCAAAAATTAGCCAGGCGCAGTGACGGCGGCC
TGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATTGCTTGAACCCGGGAGGTGGAGATTGCAGTGAG
```

Query sequence localized to human genome chromosome 19 using BLAT

```
cttaatthttg tntagagatg ggggtctcaact ctgthgteta ggetggtctc 1600669
aaactcctag gctcaagtga gcctcctgcc tcaacctccc aaagtgtctg 1600619
GATTCCCCTT GTGAGCCTCC ATGTCTGAAG CCAGTCTTTT GTCCTTGAC 1600569
CTCGGCAGTG ATGCCCCATT GCTCATGCTG CGTTCTACTC ATTAGAAGCC 1600519
TGTGCCCGGC CGGGCGCGGT GGCTCACGCC TGTAATCCCA GCACTTTGGG 1600469
AGGCCGAGGC GGGCGGATCA CGAGGTCAGG AGATTGAGAC CATCCTGGCT 1600419
```