test1

March 26, 2024

Contents

Introduction **2**

Sample Summary **2**

Sequencing and Integration Summary Table **2**

Abundance Frequency Plots **3**

AAV ITR Breakpoint Plots **6**

Gene distribution dot plots **10**

Methods **11**

Introduction

The attached report describes results of analysis of integration site distributions and relative abundance for samples from gene therapy trials.

Sample Summary

pull info from gtsp database for summary table.

Sequencing and Integration Summary Table

Presented below are summary details of each sample. The Unique Molecular Identifiers (UMIs) estimate the number of unique molecules queried. Unique Sites refers to the number of unique integration sites detected. Chao1 estimates the total population sizes of integration sites in the sample.

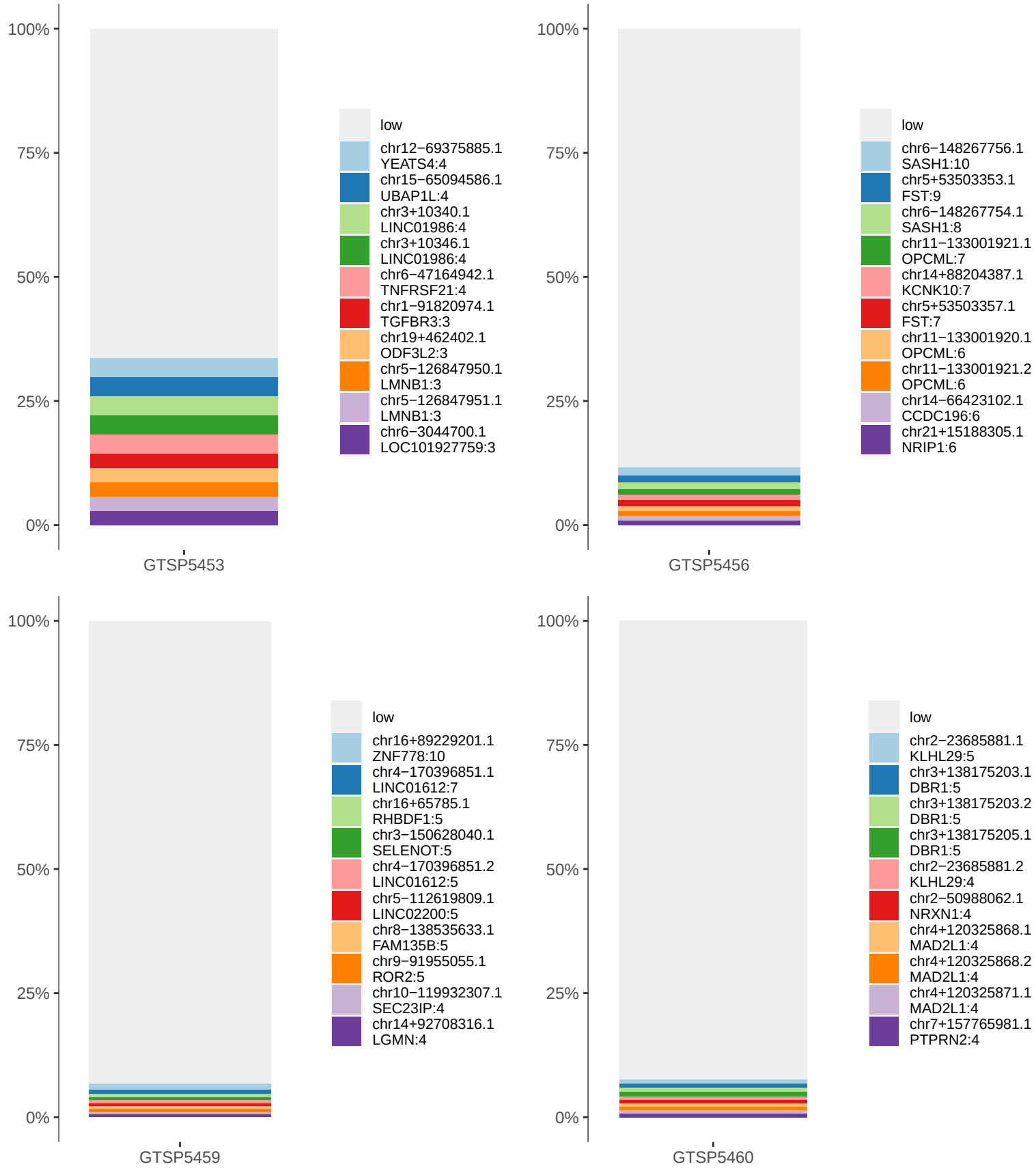
Prelimary work with negative controls support the use of a filtering parameter in such that sites only supported by one read are considered artifactual and removed.

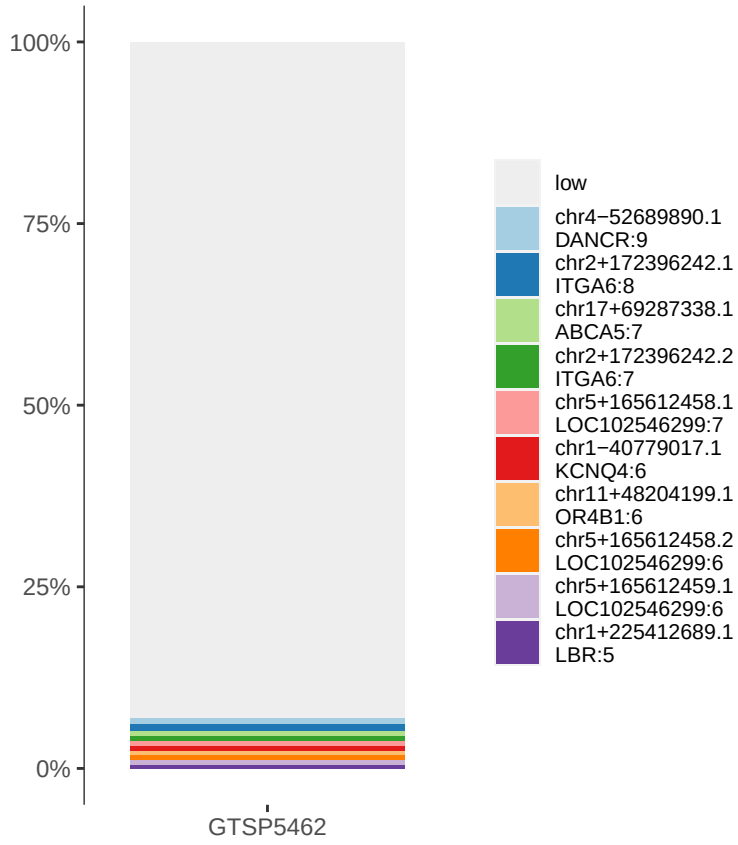
Table 1: integration summary table

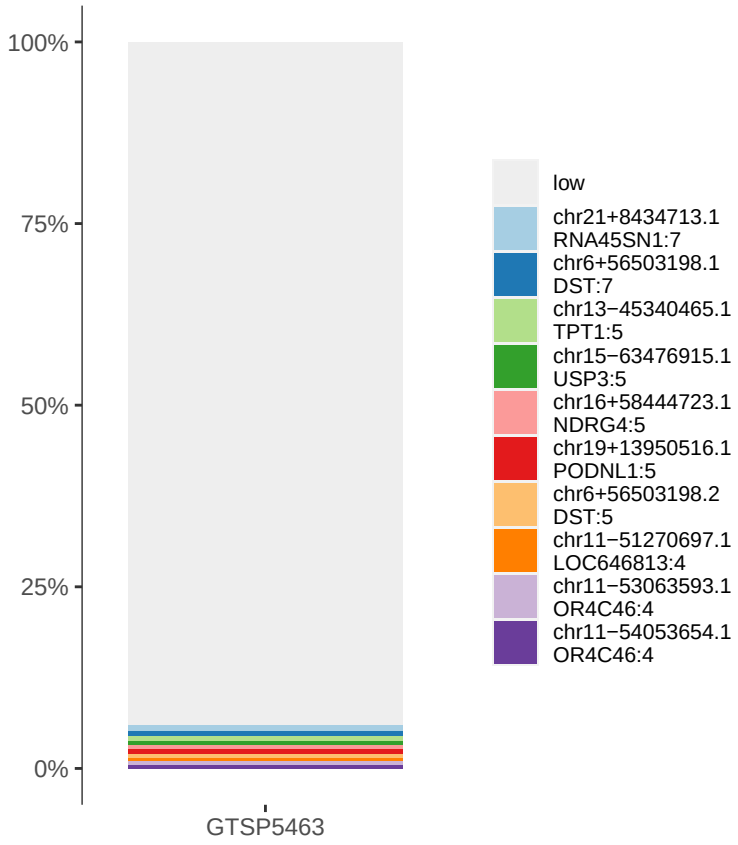
sample	subject	total reads	Unique Sites	inferred cell	Chao1
GTSP5453	pExp2_8902	7627	68	104	180.8000
GTSP5456	pExp2_8905	444847	345	623	917.5000
GTSP5459	pExp2_8910	404014	513	818	1110.9677
GTSP5460	pExp2_8911	450612	429	579	1126.6800
GTSP5462	pExp2_8916	184739	597	970	1295.3036
GTSP5463	pExp2_8918	167473	599	851	1326.6080
GTSP5465	pExp4_9889	139845	257	696	567.1622
GTSP5467	pExp4_9891	236374	344	1118	644.5849
GTSP5468	pExp4_9883	248734	301	623	485.3636
GTSP5470	pExp4_9888	335405	276	576	927.0370
PositiveControl_231103	na_positive_control	80191	27	1165	27.3750

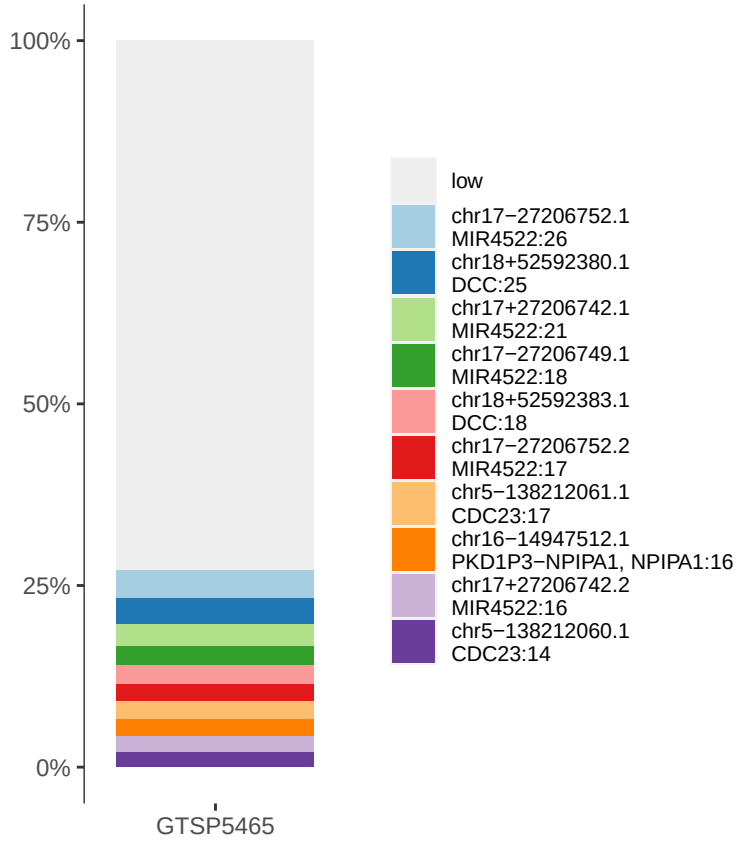
Abundance Frequency Plots

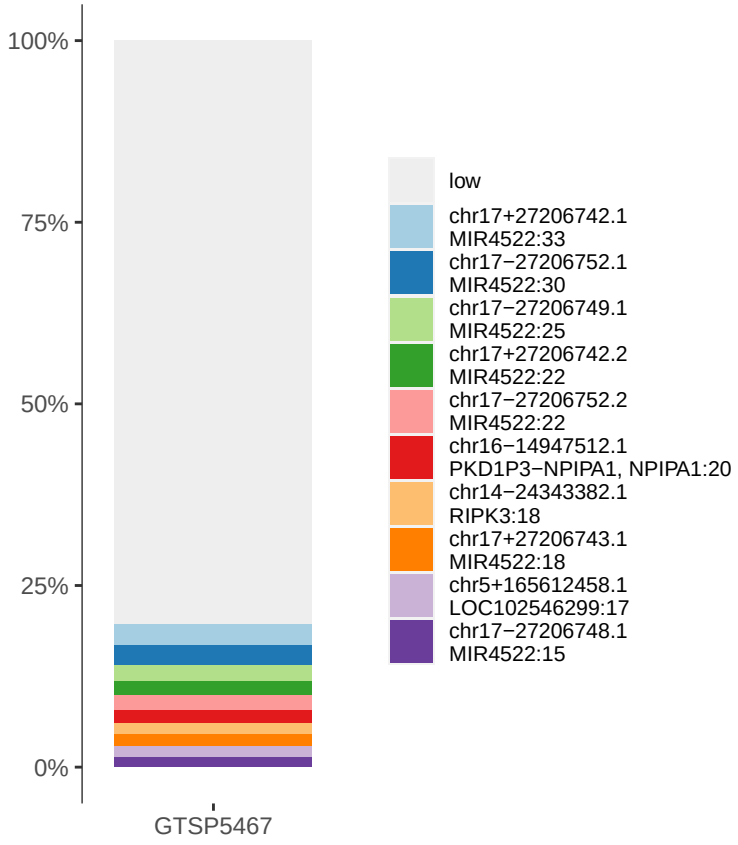
Abundance

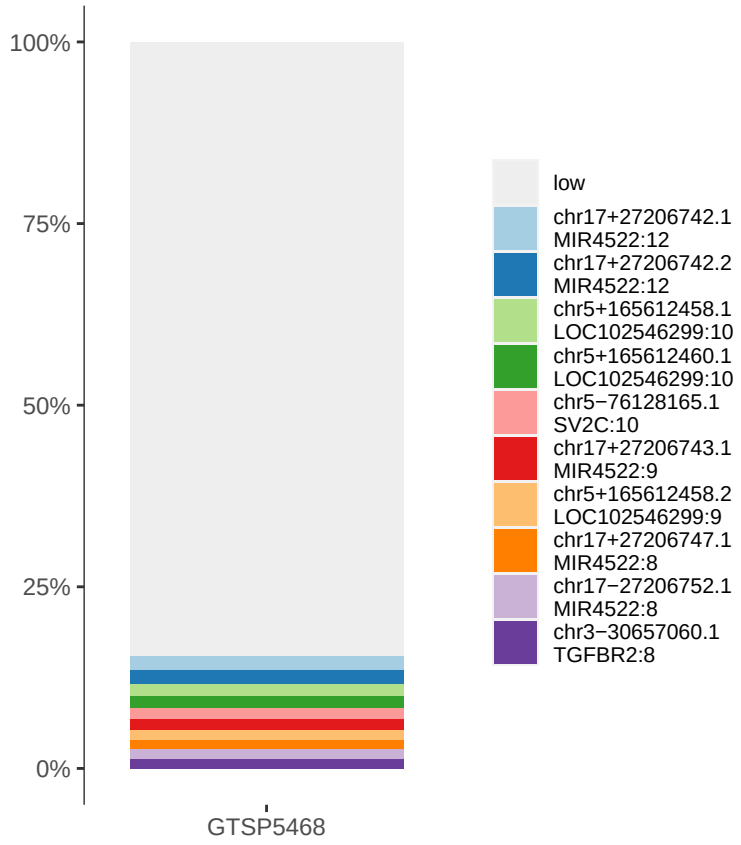


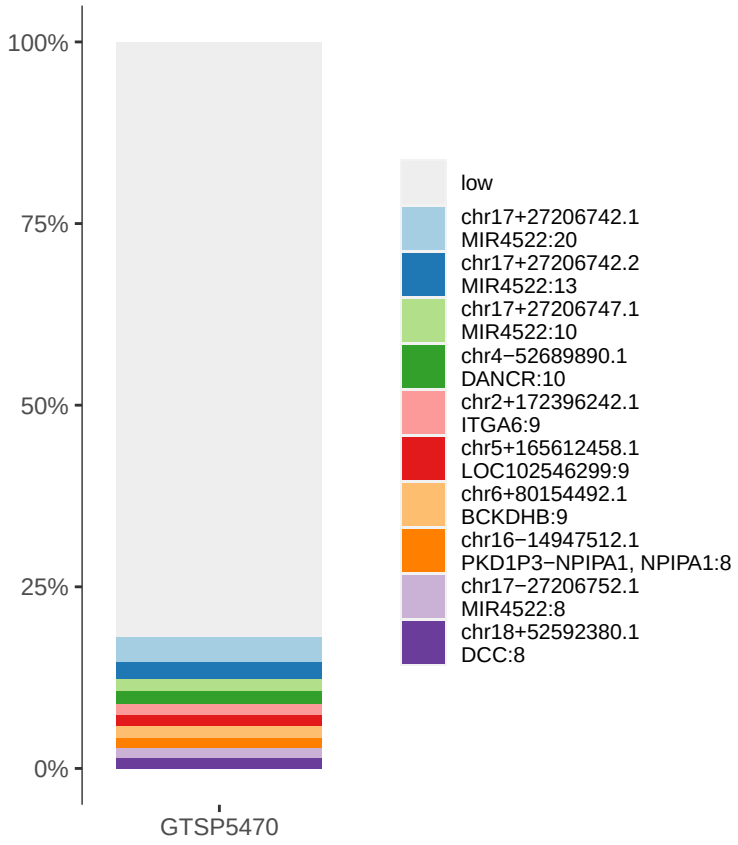

100%
75%
50%
25%
0%
GTSP5462
low
chr4−52689890.1
DANCR:9
chr2+172396242.1
ITGA6:8
chr17+69287338.1
ABCA5:7
chr2+172396242.2
ITGA6:7
chr5+165612458.1
LOC102546299:7
chr1−40779017.1
KCNQ4:6
chr11+48204199.1
OR4B1:6
chr5+165612458.2
LOC102546299:6
chr5+165612459.1
LOC102546299:6
chr1+225412689.1
LBR:5

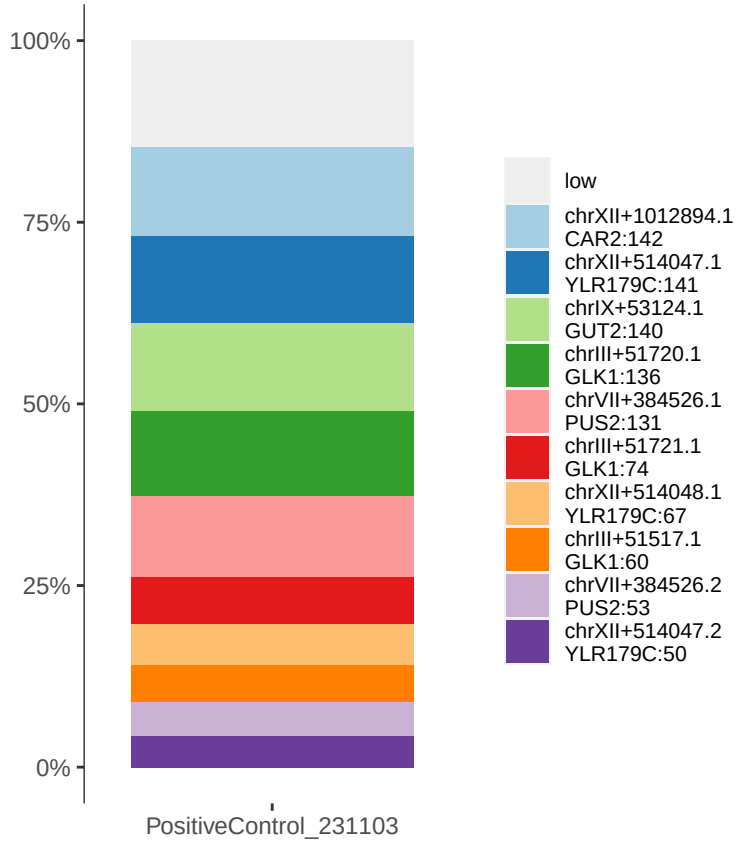

100%
75%
50%
25%
0%
GTSP5463
low
chr21+8434713.1
RNA45SN1:7
chr6+56503198.1
DST:7
chr13−45340465.1
TPT1:5
chr15−63476915.1
USP3:5
chr16+58444723.1
NDRG4:5
chr19+13950516.1
PODNL1:5
chr6+56503198.2
DST:5
chr11−51270697.1
LOC646813:4
chr11−53063593.1
OR4C46:4
chr11−54053654.1
OR4C46:4


100%
75%
50%
25%
0%
GTSP5465
low
chr17−27206752.1
MIR4522:26
chr18+52592380.1
DCC:25
chr17+27206742.1
MIR4522:21
chr17−27206749.1
MIR4522:18
chr18+52592383.1
DCC:18
chr17−27206752.2
MIR4522:17
chr5−138212061.1
CDC23:17
chr16−14947512.1
PKD1P3−NPIPA1, NPIPA1:16
chr17+27206742.2
MIR4522:16
chr5−138212060.1
CDC23:14


100%
75%
50%
25%
0%
GTSP5467
low
chr17+27206742.1
MIR4522:33
chr17−27206752.1
MIR4522:30
chr17−27206749.1
MIR4522:25
chr17+27206742.2
MIR4522:22
chr17−27206752.2
MIR4522:22
chr16−14947512.1
PKD1P3−NPIPA1, NPIPA1:20
chr14−24343382.1
RIPK3:18
chr17+27206743.1
MIR4522:18
chr5+165612458.1
LOC102546299:17
chr17−27206748.1
MIR4522:15

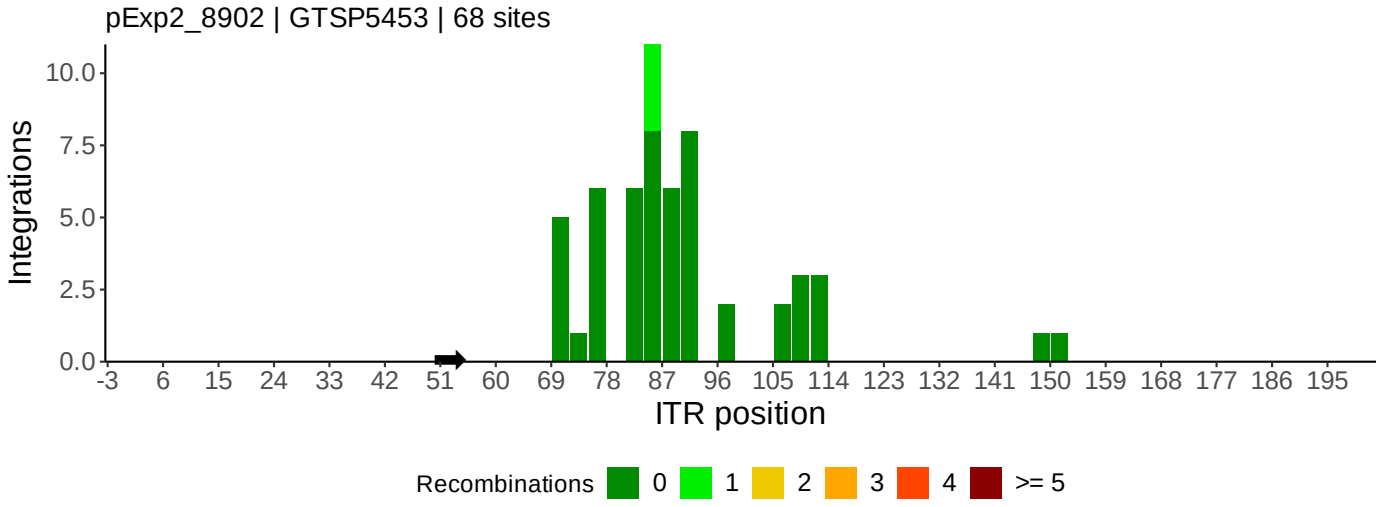

100%
75%
50%
25%
0%
GTSP5468
low
chr17+27206742.1
MIR4522:12
chr17+27206742.2
MIR4522:12
chr5+165612458.1
LOC102546299:10
chr5+165612460.1
LOC102546299:10
chr5−76128165.1
SV2C:10
chr17+27206743.1
MIR4522:9
chr5+165612458.2
LOC102546299:9
chr17+27206747.1
MIR4522:8
chr17−27206752.1
MIR4522:8
chr3−30657060.1
TGFBR2:8

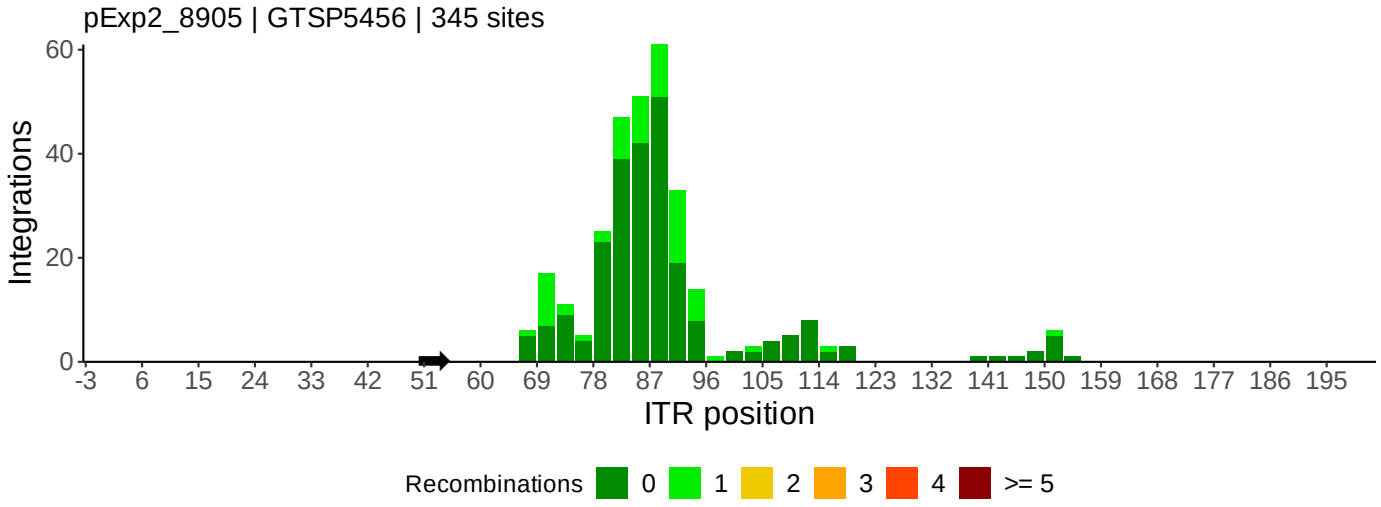

100%
75%
50%
25%
0%
GTSP5470
low
chr17+27206742.1
MIR4522:20
chr17+27206742.2
MIR4522:13
chr17+27206747.1
MIR4522:10
chr4−52689890.1
DANCR:10
chr2+172396242.1
ITGA6:9
chr5+165612458.1
LOC102546299:9
chr6+80154492.1
BCKDHB:9
chr16−14947512.1
PKD1P3−NPIPA1, NPIPA1:8
chr17−27206752.1
MIR4522:8
chr18+52592380.1
DCC:8

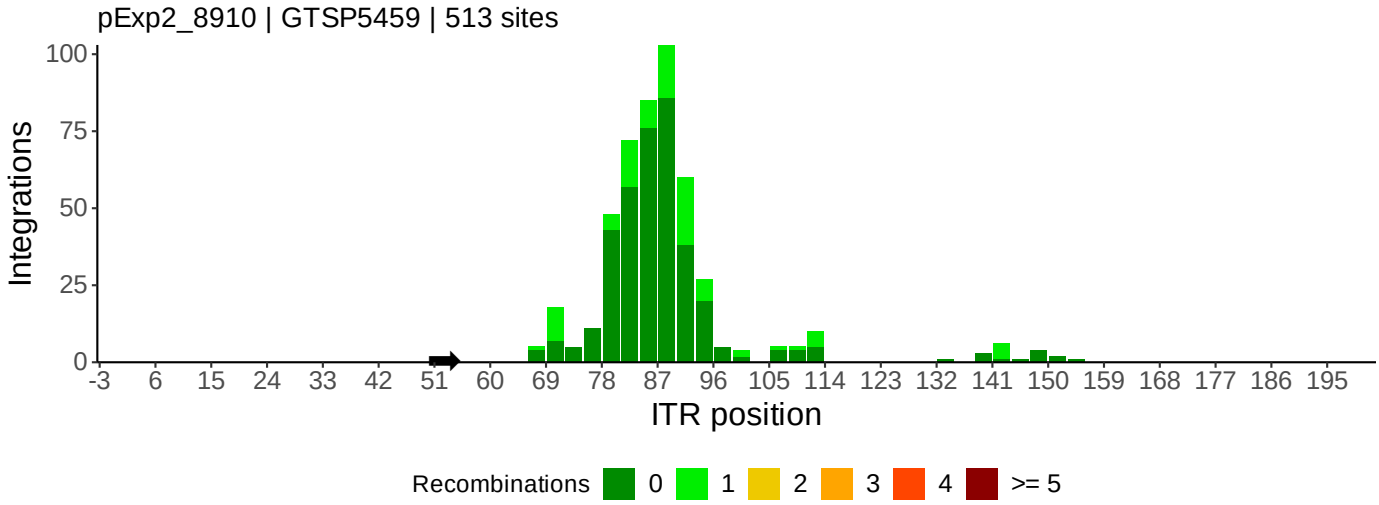

100%
75%
50%
25%
0%
PositiveControl_231103
low
chrXII+1012894.1
CAR2:142
chrXII+514047.1
YLR179C:141
chrIX+53124.1
GUT2:140
chrIII+51720.1
GLK1:136
chrVII+384526.1
PUS2:131
chrIII+51721.1
GLK1:74
chrXII+514048.1
YLR179C:67
chrIII+51517.1
GLK1:60
chrVII+384526.2
PUS2:53
chrXII+514047.2
YLR179C:50

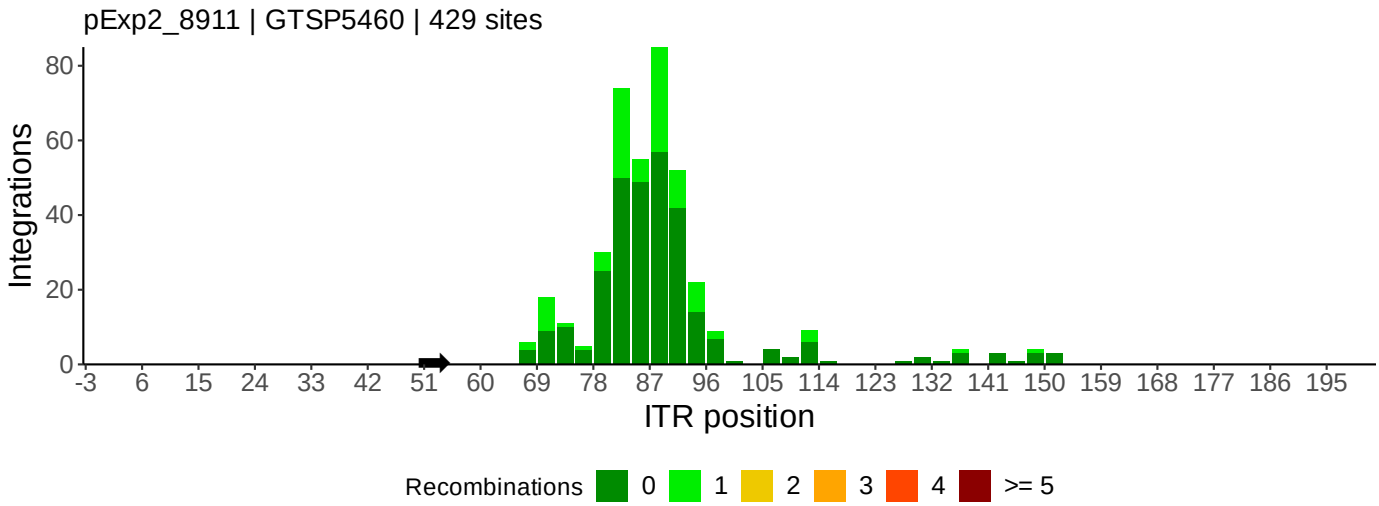
AAV ITR Breakpoint Plots

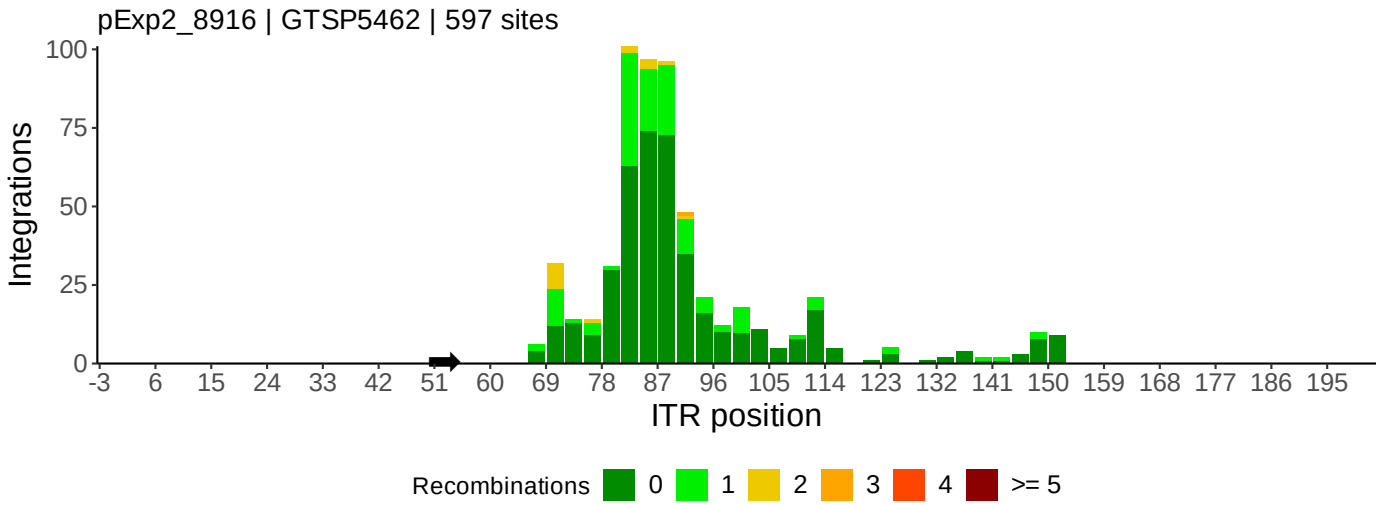
These plots are a visual representation of the ITR integration breakpoint. The x-axis indicates the point of departure from the expected ITR, where zero refers to the annotated start of the D-region. The arrow indicates the location within the ITR of the sequencing primer and the dotted lines indicate the tips of the B and C dumbells. Colors indicate the number of rearrangements in the ITR remnant.

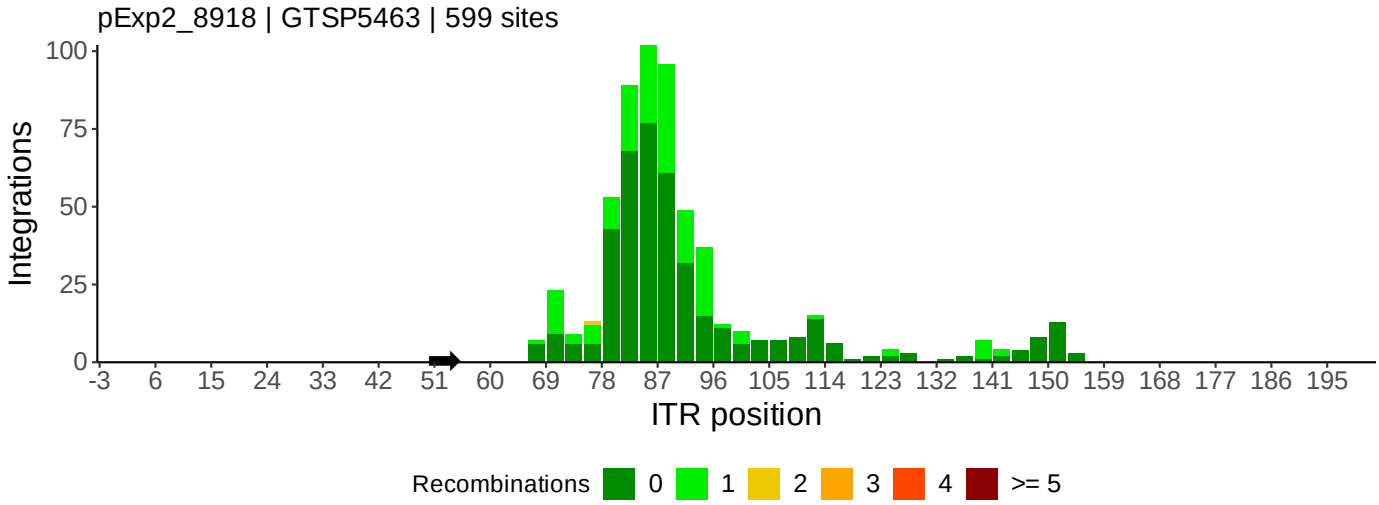


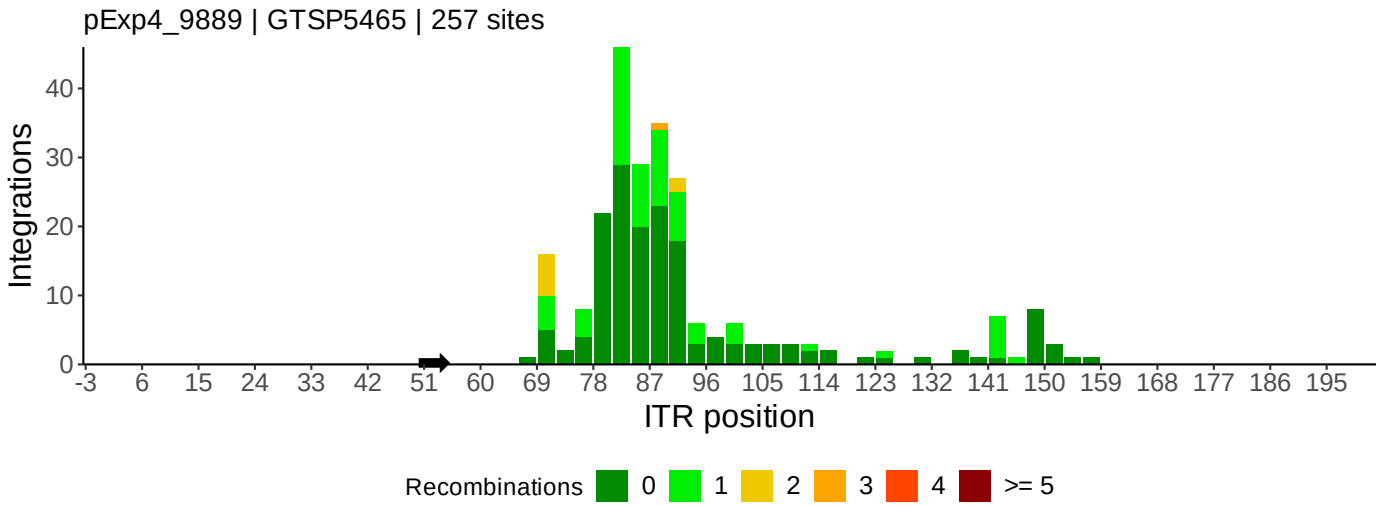


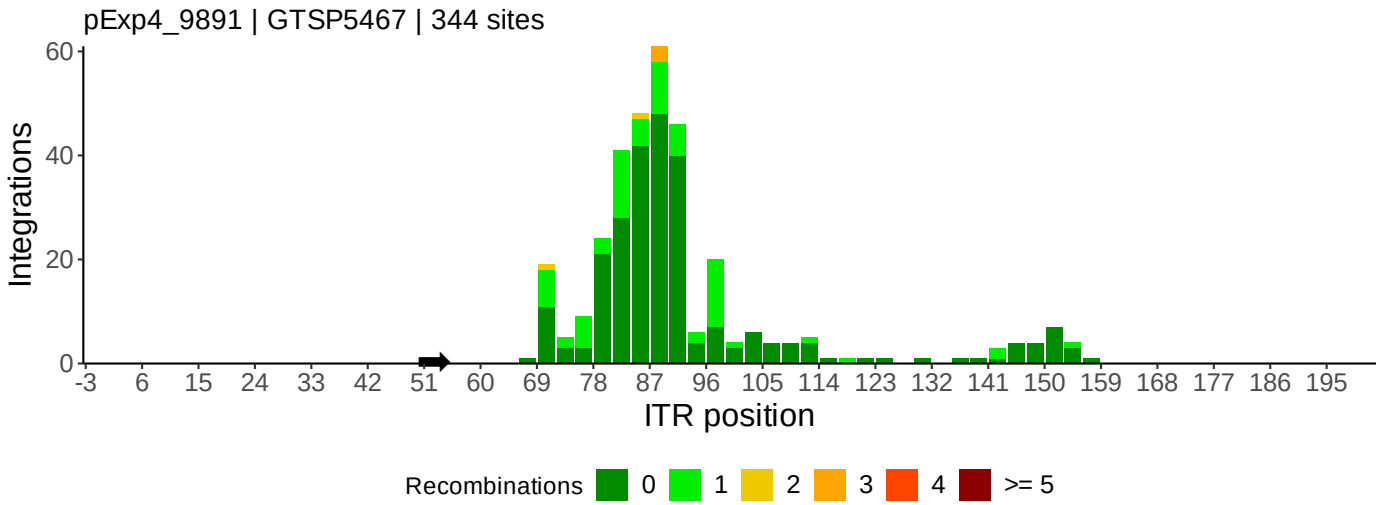

pExp2_8910 | GTSP5459 | 513 sites
Integrations
100
75
50
25
0
-3 6 15 24 33 42 51 60 69 78 87 96 105 114 123 132 141 150 159 168 177 186 195
ITR position
Recombinations 0 1 2 3 4 >= 5

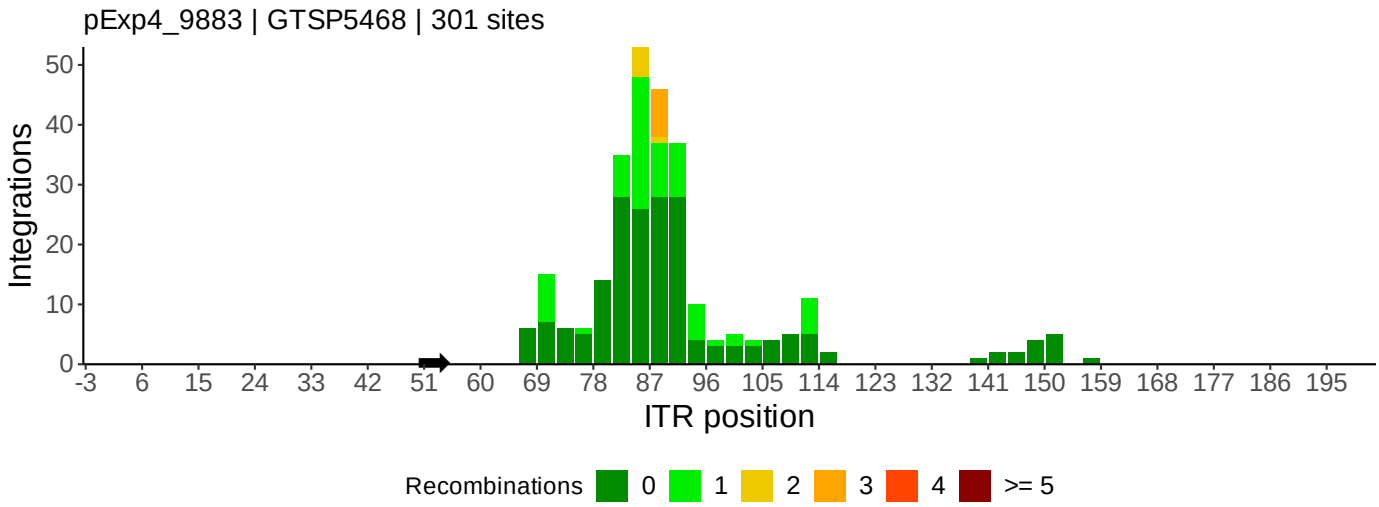

pExp2_8911 | GTSP5460 | 429 sites
Integrations
80
60
40
20
0
-3 6 15 24 33 42 51 60 69 78 87 96 105 114 123 132 141 150 159 168 177 186 195
ITR position
Recombinations 0 1 2 3 4 >= 5

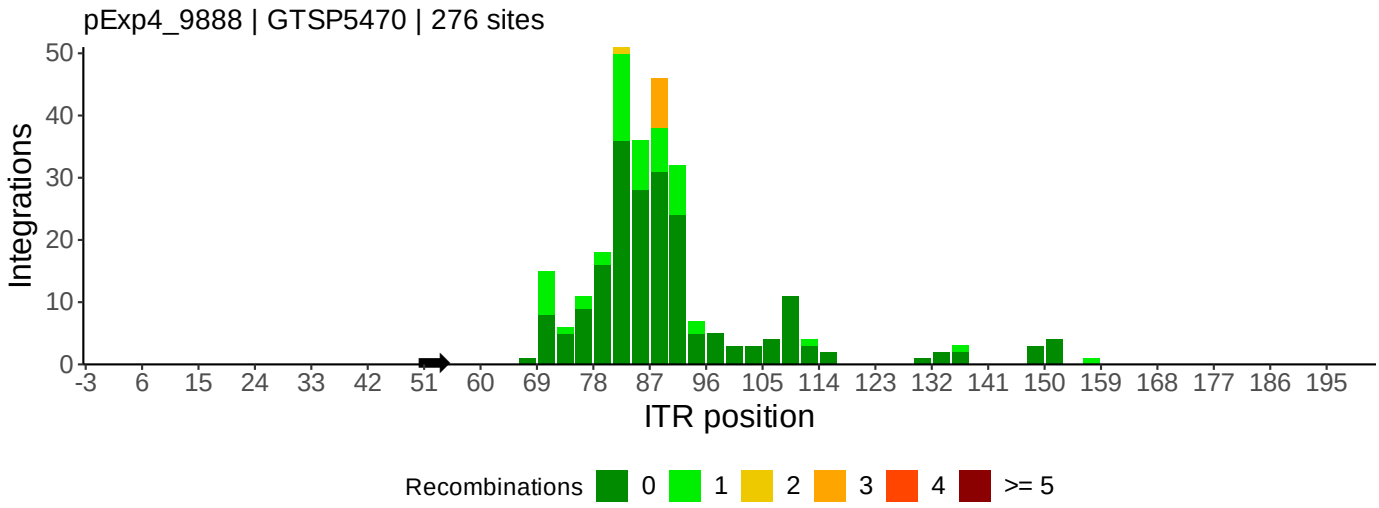

pExp2_8916 | GTSP5462 | 597 sites
Integrations
100
75
50
25
0
-3 6 15 24 33 42 51 60 69 78 87 96 105 114 123 132 141 150 159 168 177 186 195
ITR position
Recombinations 0 1 2 3 4 >= 5

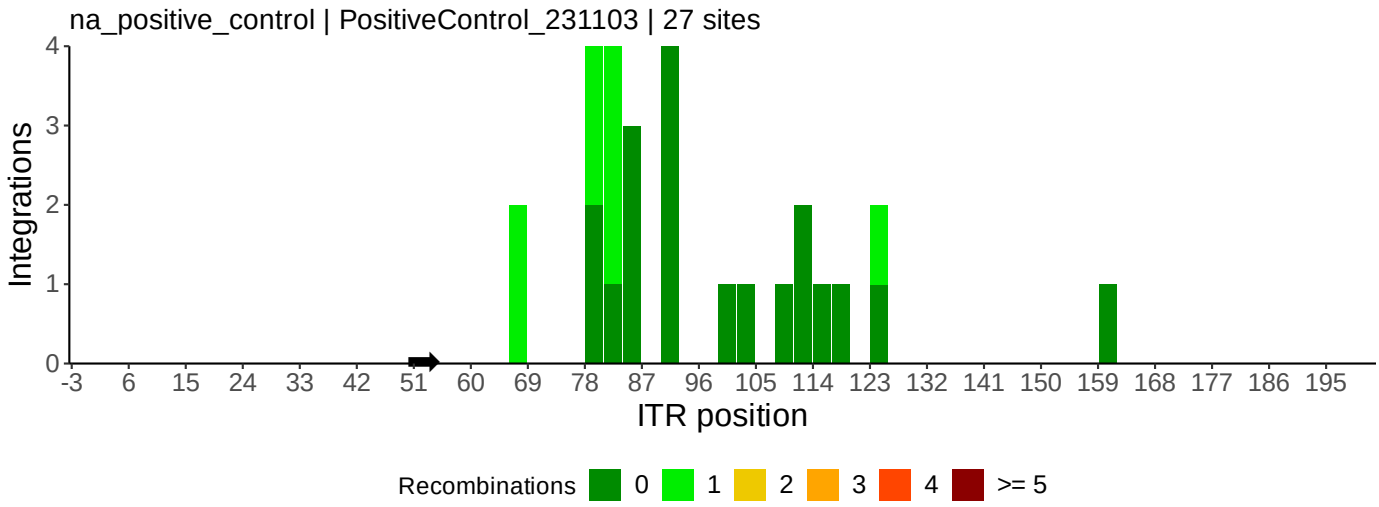

pExp2_8918 | GTSP5463 | 599 sites
Integrations
100
75
50
25
0
-3 6 15 24 33 42 51 60 69 78 87 96 105 114 123 132 141 150 159 168 177 186 195
ITR position
Recombinations 0 1 2 3 4 >= 5


pExp4_9889 | GTSP5465 | 257 sites
Integrations
40
30
20
10
0
-3 6 15 24 33 42 51 60 69 78 87 96 105 114 123 132 141 150 159 168 177 186 195
ITR position
Recombinations 0 1 2 3 4 >= 5


pExp4_9891 | GTSP5467 | 344 sites
Integrations
60
40
20
0
-3 6 15 24 33 42 51 60 69 78 87 96 105 114 123 132 141 150 159 168 177 186 195
ITR position
Recombinations 0 1 2 3 4 >= 5


pExp4_9883 | GTSP5468 | 301 sites
Integrations
0
10
20
30
40
50
-3 6 15 24 33 42 51 60 69 78 87 96 105 114 123 132 141 150 159 168 177 186 195
ITR position
Recombinations 0 1 2 3 4 >= 5


pExp4_9888 | GTSP5470 | 276 sites
Integrations
0
10
20
30
40
50
-3 6 15 24 33 42 51 60 69 78 87 96 105 114 123 132 141 150 159 168 177 186 195
ITR position
Recombinations 0 1 2 3 4 >= 5


na_positive_control | PositiveControl_231103 | 27 sites
Integrations
0
1
2
3
4
-3 6 15 24 33 42 51 60 69 78 87 96 105 114 123 132 141 150 159 168 177 186 195
ITR position
Recombinations 0 1 2 3 4 >= 5

Gene distribution dot plots

Gene distribution for integration sites.

Table 2: gene distribution percentage table

sample	In Exon %	In Gene %
PositiveControl_231103	74.0740741	74.07407
GTSP5467	2.6162791	46.51163
GTSP5465	3.1128405	58.75486
GTSP5470	0.7246377	43.47826
GTSP5468	1.6611296	43.52159
GTSP5456	2.3188406	40.57971
GTSP5459	7.0175439	46.78363
GTSP5462	4.1876047	47.57119
GTSP5463	4.5075125	50.41736
GTSP5460	5.1282051	48.71795
GTSP5453	4.4117647	57.35294

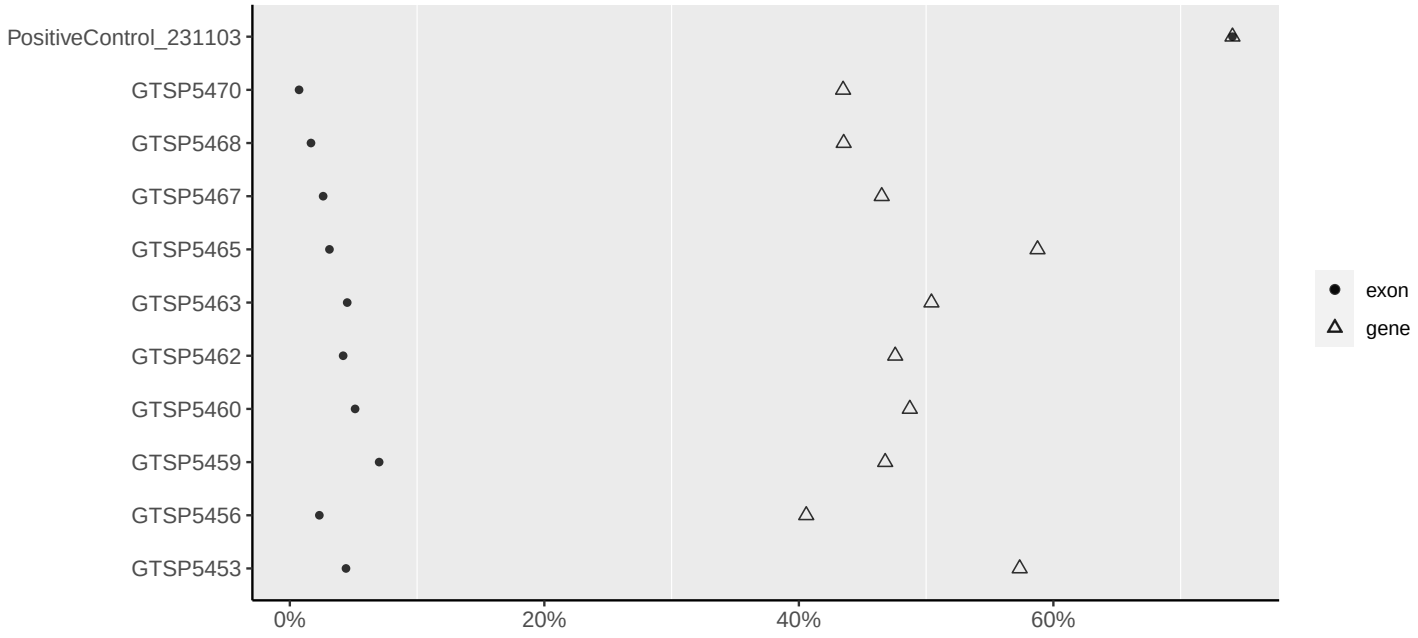


Figure 1: gene distribution scatter plot

Methods

Report Generated on: March 26, 2024
AAVenger Version: 1.1

Modules Called:

- core.R
- mapSiteLeaderSequences.R
- buildAAVremnantPlots.R
- anchorReadRearrangements.R