

AAV Integration Reproducible Report: HelaTopo

June 11, 2023

Contents

Introduction	2
Data Summary	2
Introduction	2
Sample Summary	3
Rearrangement Percent	4
Boolean Rearrangements	4
Summary Table	4
Rearrangements per Length Table	4
Rearrangements per Samples Table	5
Boolean Percent Graph	6
Remnant Plots	7
Methods	9

Introduction

Data Summary

This is where a paragraph about the data we recieve will go. This still needs to be updated such that it can automatically pull in the paragraph given to us, rather than it being manually written here. I will get back to this at some point, for now I am focusing on the main meat of the project.

Bogus control - Positive control -

Introduction

This is where I am going to write a brief summary about what this report is going to do. The main purpose of this report is to have an easily-accessible version of the AAvenger data which can be repeatably-produced and can be distributed to collaborators. The main points here are that the AAvenger version info is recorded in this report along with the date, and the data information/data file.

Sample Summary

This is just an example, but for now I am outputting a table that is from the AAvenger output `anchoReadRearrangements`. Later down the line we can change this if this isn't what we want, it's just an example.

Columns indicate the following: Trial is the trial for which this data was collected. Subject is the treatment or condition. Sample is the GTSP number. UMIs, otherwise known as Unique Molecular Identifiers are a method of quantifying unique DNA samples. More UMIs indicates more unique DNA pieces. Sites is a measure of how many integrations sites are found for a certain condition. Finally, UMIs Rearranged (%) is a measure of the percentage of UMIs which are rearranged.

trial	subject	sample	UMIs	sites	UMIs Rearranged (%)
AAVHelaTopo	Bogus Control	AAVBogusControl_230119	13800	192	6.93%
AAVHelaTopo	Doxorubicin 1	GTSP5335	23	2	34.78%
AAVHelaTopo	Doxorubicin 2	GTSP5336	471	114	2.76%
AAVHelaTopo	Doxorubicin 3	GTSP5337	72506	436	3.82%
AAVHelaTopo	Doxorubicin 4	GTSP5338	12267	50	0.04%
AAVHelaTopo	Etoposide 1	GTSP5339	104	6	40.38%
AAVHelaTopo	Etoposide 2	GTSP5340	6357	140	2.85%
AAVHelaTopo	Etoposide 4	GTSP5342	1	2	0.00%
AAVHelaTopo	No Drug Control 3	GTSP5333	1069	32	7.39%
AAVHelaTopo	Positive Control	AAVPosControl_230119	2346	12	1.58%

Rearrangement Percent

Boolean Rearrangements

length is the total length of DNA sequences. Breaks is the number of breaks found total for each condition. Weighted rearrangements is a measure of rearrangements per length with the number of rearrangements per sequence being taken into account. Boolean rearrangements, on the other hand, only counts whether or not something was rearranged, not the total number of rearrangements

Summary Table

subject	sample	length	breaks
Etoposide 1	GTSP5339	8976	42
Etoposide 2	GTSP5340	221985	200
Etoposide 4	GTSP5342	29	0
Bogus Control	AAVBogusControl_230119	681111	1233
Positive Control	AAVPosControl_230119	119070	113
Doxorubicin 1	GTSP5335	2027	25
Doxorubicin 2	GTSP5336	17762	23
Doxorubicin 3	GTSP5337	2919632	3051
Doxorubicin 4	GTSP5338	410322	10
No Drug Control 3	GTSP5333	37301	89

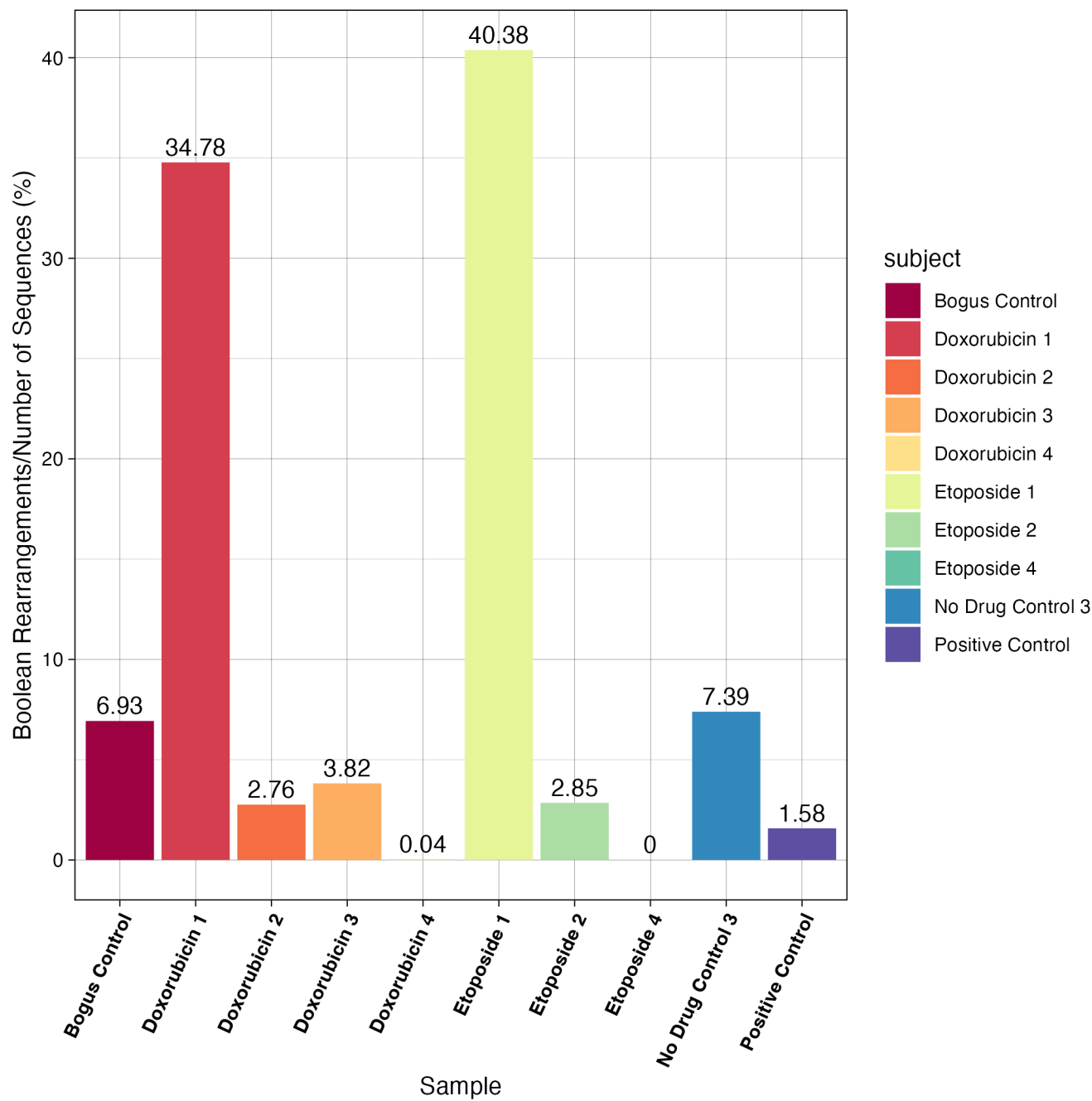
Rearrangements per Length Table

subject	Weighted Rearrangements/Length (%)	Boolean Rearrangements/Length (%)
Etoposide 1	4.68	4.68
Etoposide 2	0.90	0.82
Etoposide 4	0.00	0.00
Bogus Control	1.81	1.40
Positive Control	0.95	0.31
Doxorubicin 1	12.33	3.95
Doxorubicin 2	1.29	0.73
Doxorubicin 3	1.04	0.95
Doxorubicin 4	0.02	0.01
No Drug Control 3	2.39	2.12

Rearrangements per Samples Table

subject	Weighted Rearrangements/Sequences (%)	Boolean Rearrangements/Sequences (%)
Etoposide 1	40.38	40.38
Etoposide 2	3.15	2.85
Etoposide 4	0.00	0.00
Bogus Control	8.93	6.93
Positive Control	4.82	1.58
Doxorubicin 1	108.70	34.78
Doxorubicin 2	4.88	2.76
Doxorubicin 3	4.21	3.82
Doxorubicin 4	0.08	0.04
No Drug Control 3	8.33	7.39

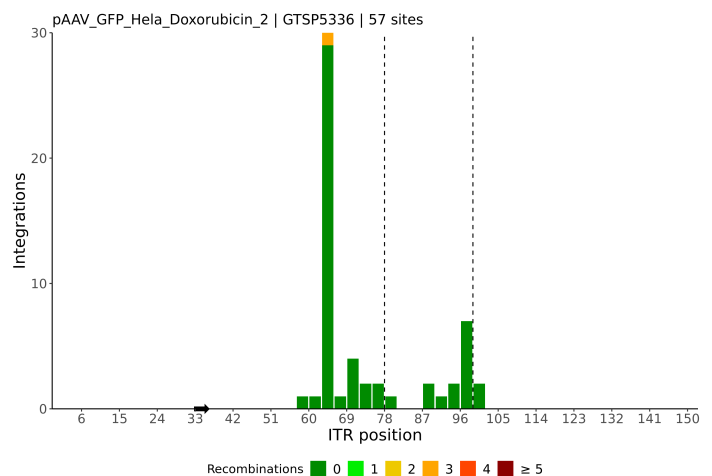
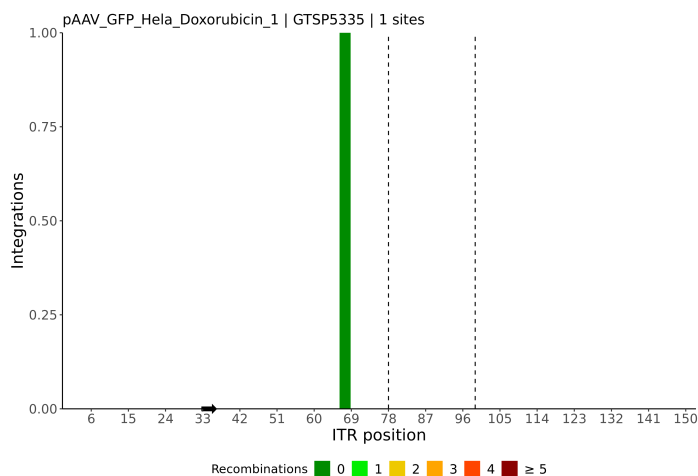
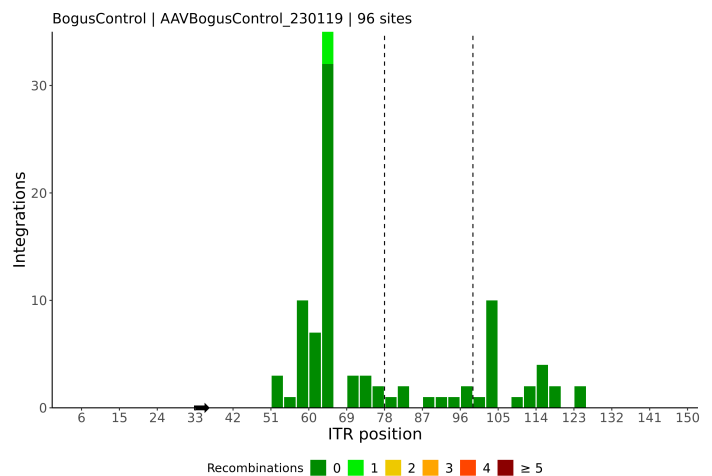
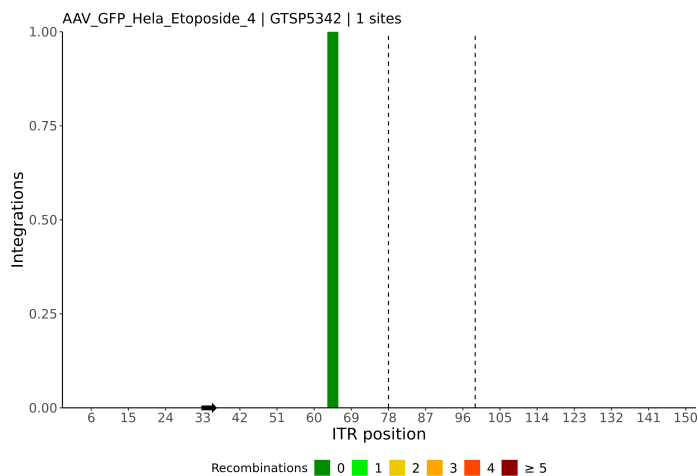
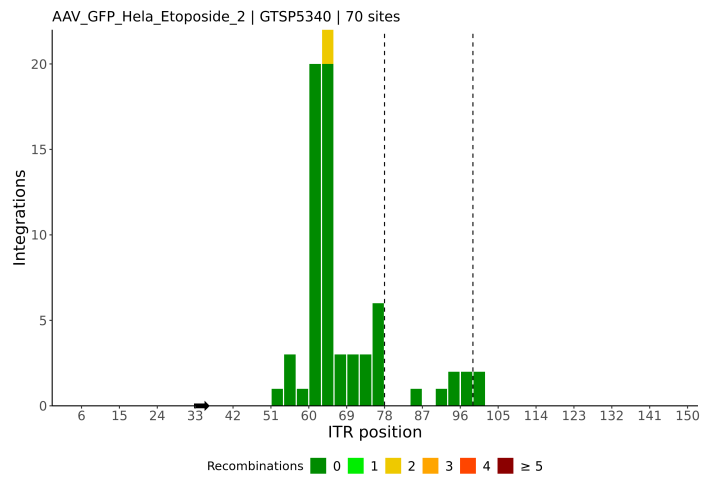
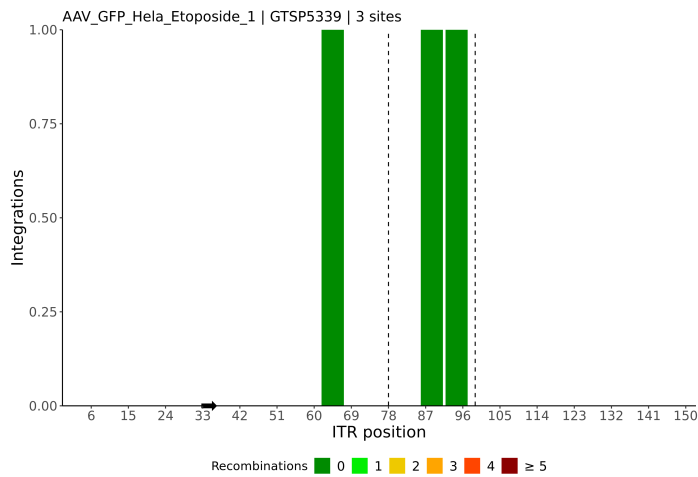
Boolean Percent Graph

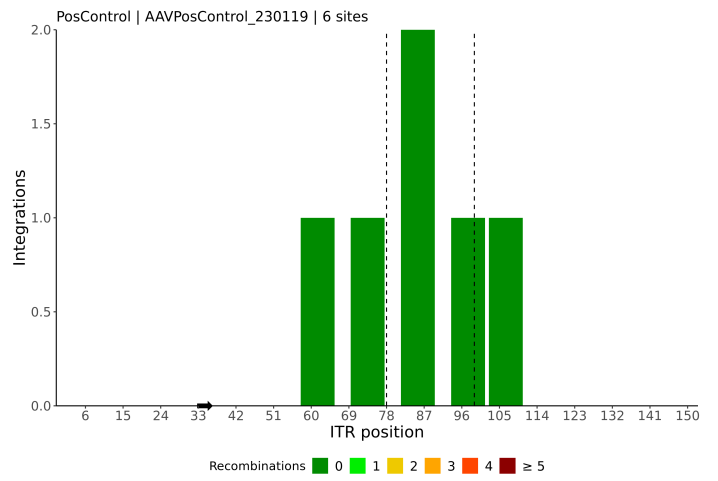
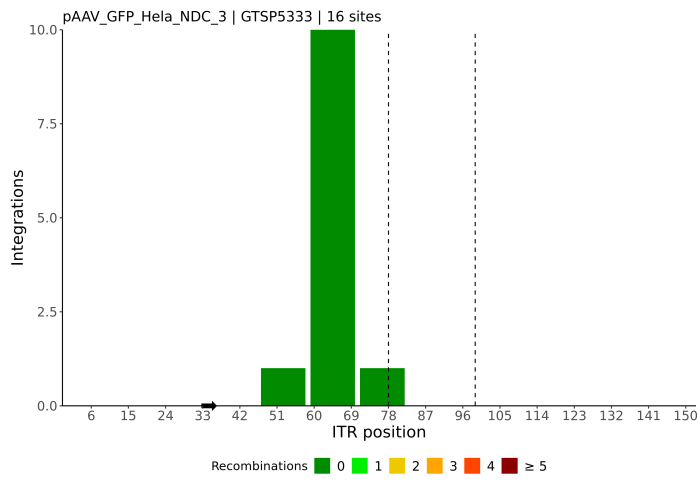
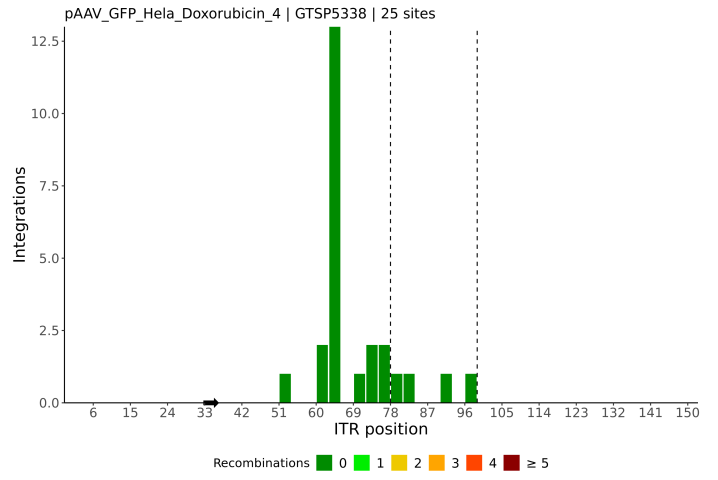
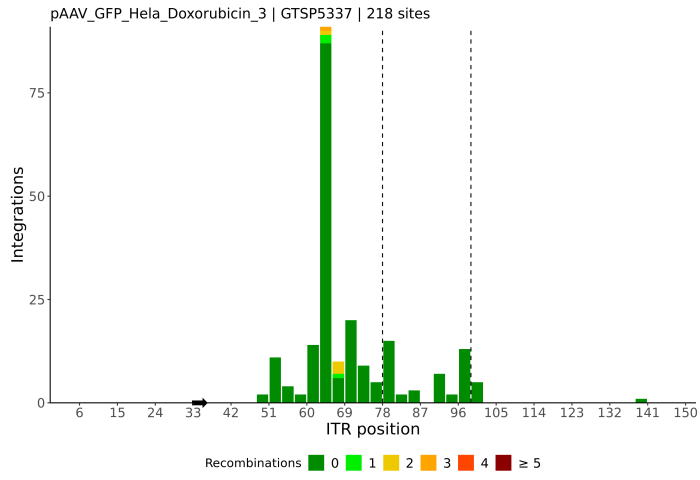


Remnant Plots

Remnant plots are a visual representation of integration by location. The arrow indicates where are read begins, with everything in the direction of the arrow having been sequenced. Integrations are marked by green bars, with colors at the bottom indicating the number of rearrangements. Sites for each condition are listed in the title. Dotted lines indicate the AAV ITRs (inverted terminal repeats).

Dotted lines indicate the AAV ITRs (inverted terminal repeats).





Methods

Report Generated on: June 11, 2023

AAvenger Version: XXXXXXXXXXXXX

Data Version: XXXXXXXXXXX