

eDentity Sequencer Qc

12/11/24

1.1.1 Percentage of bases \geq Q30

Metric: Calculate the fraction of bases whose Phred score are greater or equal to Q30.

Illumina	Elements
0.92	0.93

1.1.2 Read Length Retention after Q30 Trimming (Percentage).

Metric: $(\text{Average length after trimming} / \text{original read length}) \times 100$

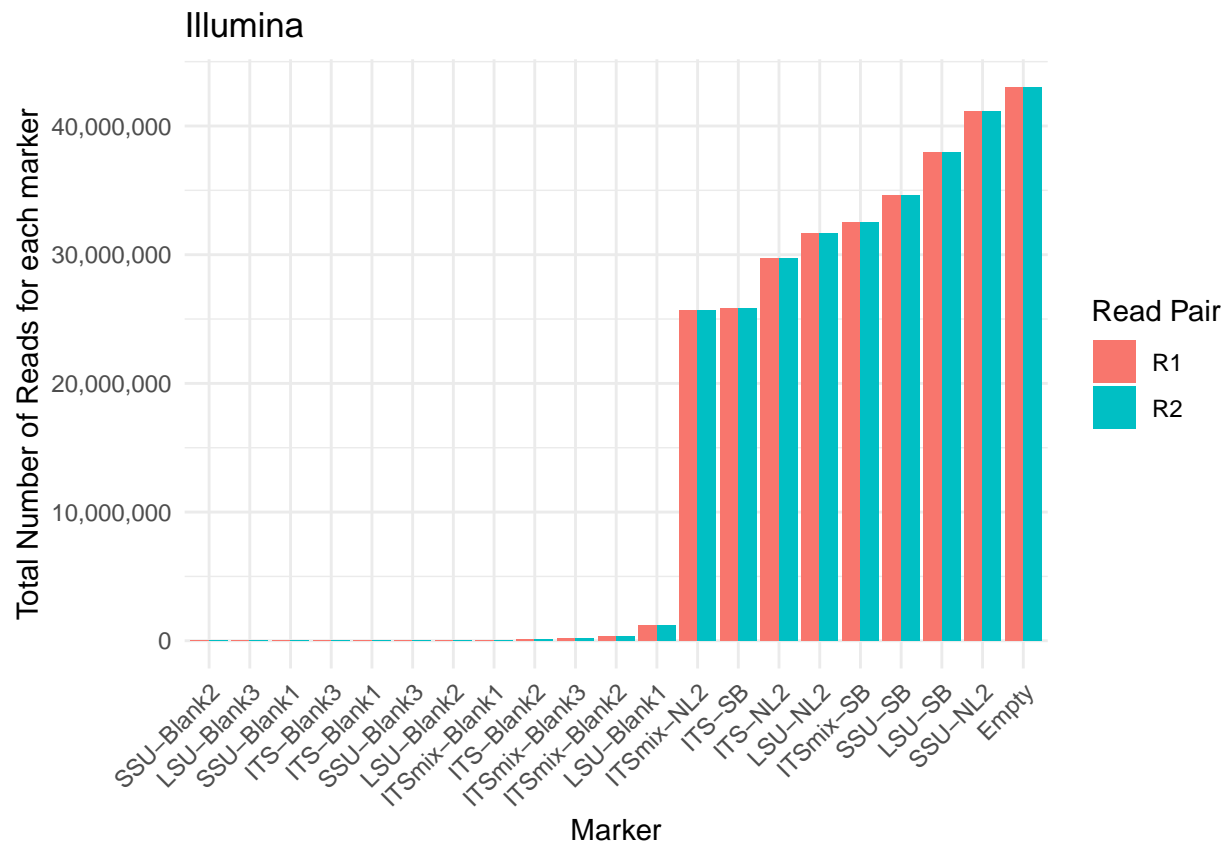
Illumina	Elements
94.67	94.47

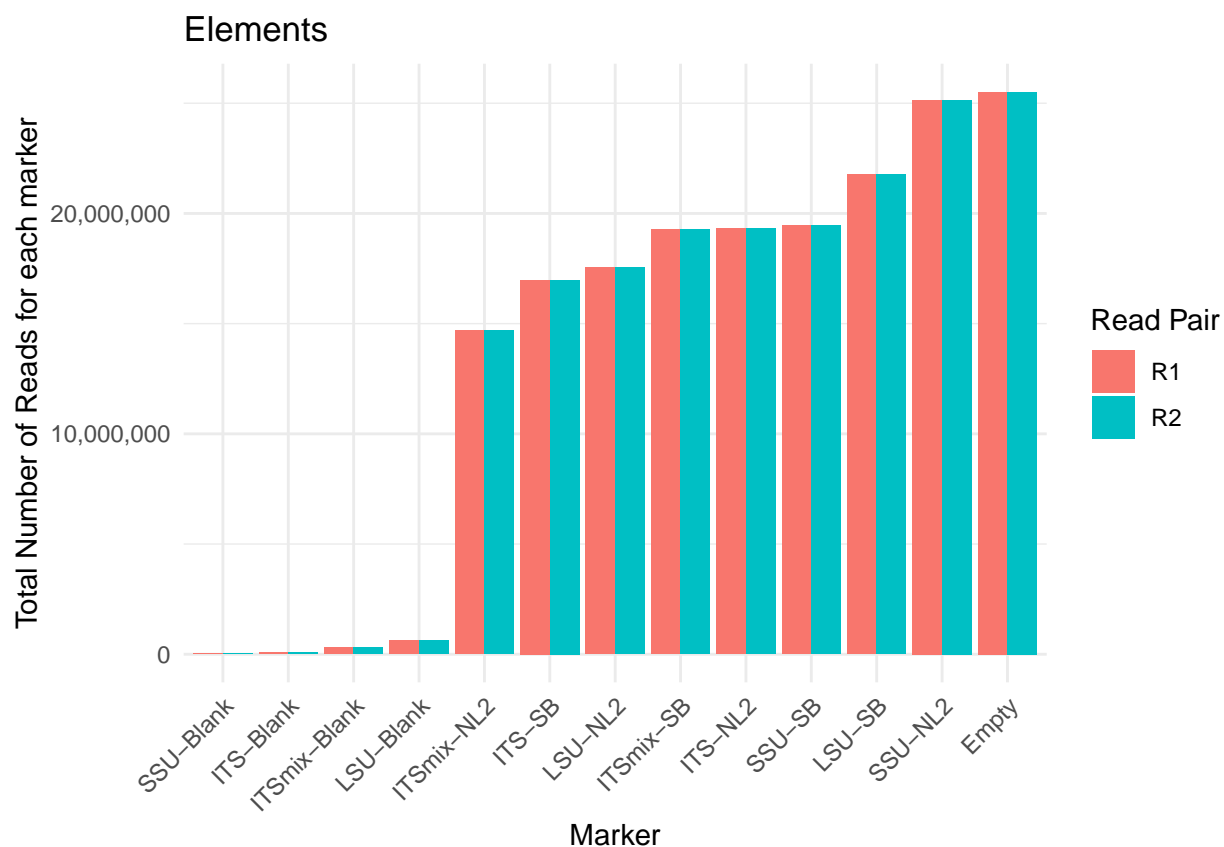
Total Number of reads.

This section looks at the total number of reads as received from the sequencer

total_reads	total_files	average_reads	sequencer	read_pair
442,555,140	1,537	287,934.4	illumina	R1
442,555,140	1,537	287,934.4	illumina	R2
260,002,493	1,541	168,723.2	elements	R1
260,002,493	1,541	168,723.2	elements	R2

Total number of reads per marker





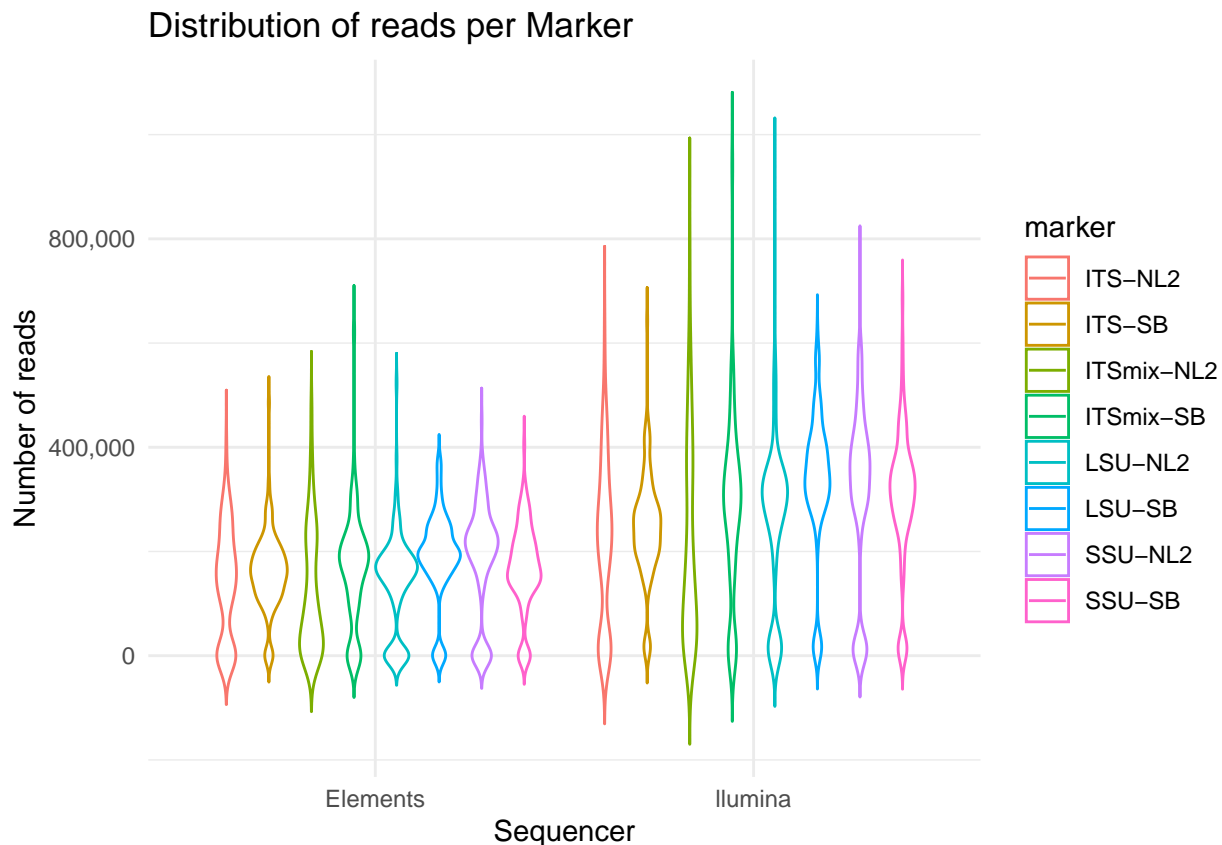
Filter out the Empty, Phix and undetermined reads.

Next we exclude the empty, unassigned as well as sample with “Blank” markers.

Then we look at the number of reads after filtering.

total_reads	total_files	average_reads	sequencer	read_pair
259,220,865	958	270,585.5	illumina	R1
259,220,865	958	270,585.5	illumina	R2
154,178,824	958	160,938.2	elements	R1
154,178,824	958	160,938.2	elements	R2

What is the distribution of reads per marker after filtering?



1.2.1 Homopolymer Accuracy

Homopolymer sequence: A series of consecutive identical bases— in this case four or more consecutive repeats of the same base.

Metric: Average Q score decay per additional base in homopolymer runs (e.g AAAA, TTTT) compared to the Phix reference genome.

Illumina	Elements
0.05	-0.01

1.2.2 Demultiplexing Efficiency (Percentage)

Metric: (Number of reads successfully demultiplexed / Total number of reads) \times 100

	Illumina	Element
Total Reads (without Phix reads)	662,034,609	376,486,325
Demultiplexed Reads	608,699,786	361,464,384
Demultiplexing Efficiency	91.94	96.01

1.2.3 PhiX Control performance

Metric: Assess sequencing quality using spiked-in Phix control

	Illumina	Element
PhiX Error Rate	0.005588	0.007656
PhiX Alignment Rate	0.2520	0.2760
PhiX Coverage Uniformity:		
- Coefficient of Variation:	0.1554	0.1631
- Percentage within $\pm 20\%$ of mean coverage:	0.9332	0.9081
Additional Information:		
- Total Reads:	885,110,280	520,004,986
- Mapped Reads:	223,075,671	143,518,661
- Mean Coverage:	12,410,202.48	7,927,615.76
- Coverage Range:	1,254,247 - 14,897,715	942,836 - 9,820,668

1.2.4 Duplicate Read Rate (Percentage)

Metrics: Percentage of duplicate reads in a standard non-amplified library.

Only reads mapping to Phix were used here under the assumption that Phix is a “standard non-amplified library”

but it appears Phix reads were amplified therefore they may not be a good measure of duplication rate.

	Illumina	Element
Mapped Reads:	223,075,671	143,518,661
Duplicate reads:	219,577,678	138,553,627
Duplicate rate	0.9843	0.9654

1.2.5 GC Bias

Metric: Deviation of coverage in GC-rich regions compared to AT-rich regions

Only reads mapping to Phix were used to compute GC Bias

Illumina	Elements
0.32	0.3