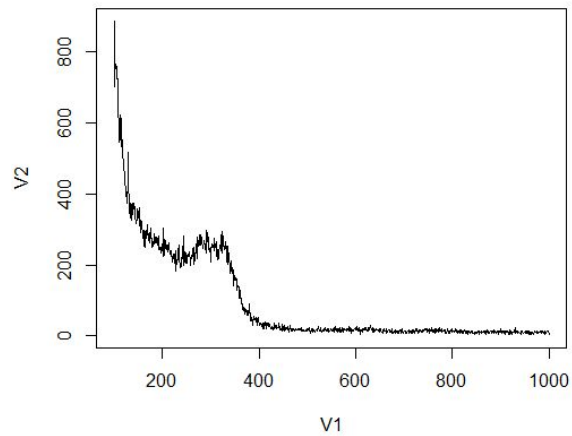


## JellyFish Result

**Setting 1:** -m 31 -t 10 -s 300M

Using: SRR6058604\_scaffold\_06.1P.fastq.gz

Plot [1-1000]:

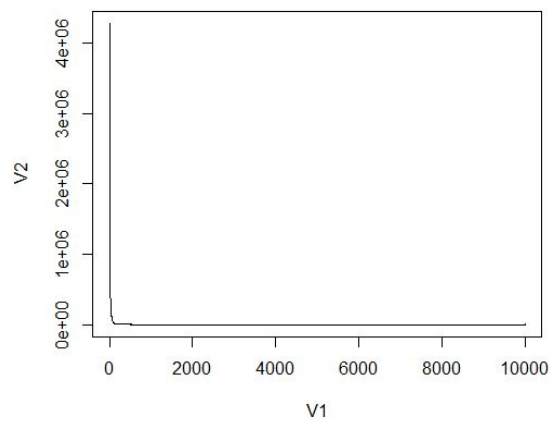


The genome size is 0.3M when k-mer=31.

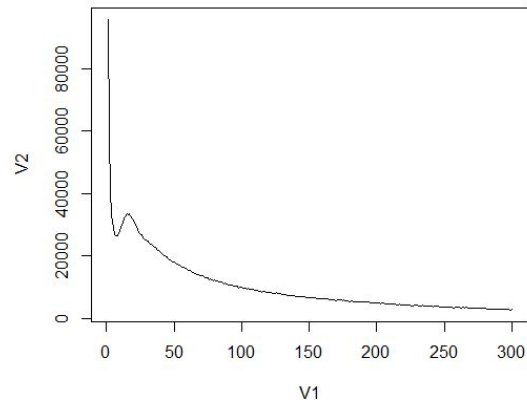
**Setting 2:** -m 11 -t 10 -s 300M

Using: SRR6058604\_scaffold\_06.[1P or 2P].fastq.gz

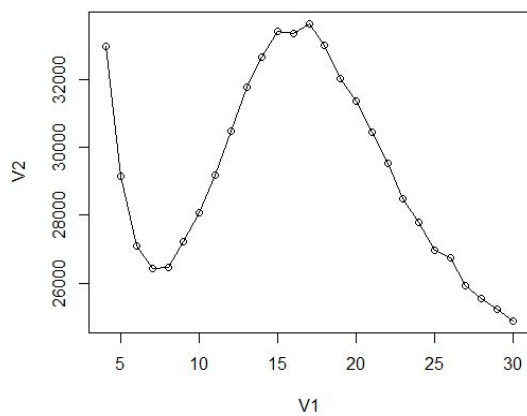
Plot [1-10000]:



Plot [1-300]:



Plot [4-30]:



Total k-mer included: around 1.2G. The genome size is 70.9M when k-mer=11, 91.6M when k-mer=13.

### Summarize:

The length of k-mer seems to affect the result seriously. Larger k-mer seems to filter most of the short reads, it's hard to tell if this could give a more accurate result. My plot is not good, partly due to the read quality (poor quality checked by the FastQC). So the predicted genome sizes are not trustable.

Instead, according to the paper (<https://www.ncbi.nlm.nih.gov/Traces/wgs/NSDW01?display=contigs>), I would use '27M' as predicted genome size for the following CANU assembly.