Hw 6 Questions
Mateusz Grobelny

1.
Looking at major parameters, going from a kmer size of 31 to 49 increased the
N50 increased however the total length of the assembly decreased getting closer to the
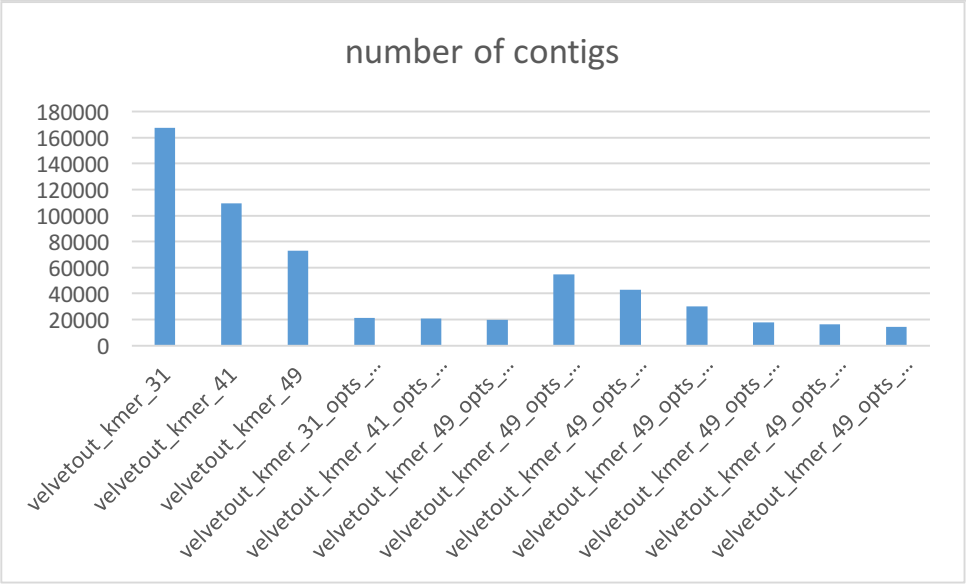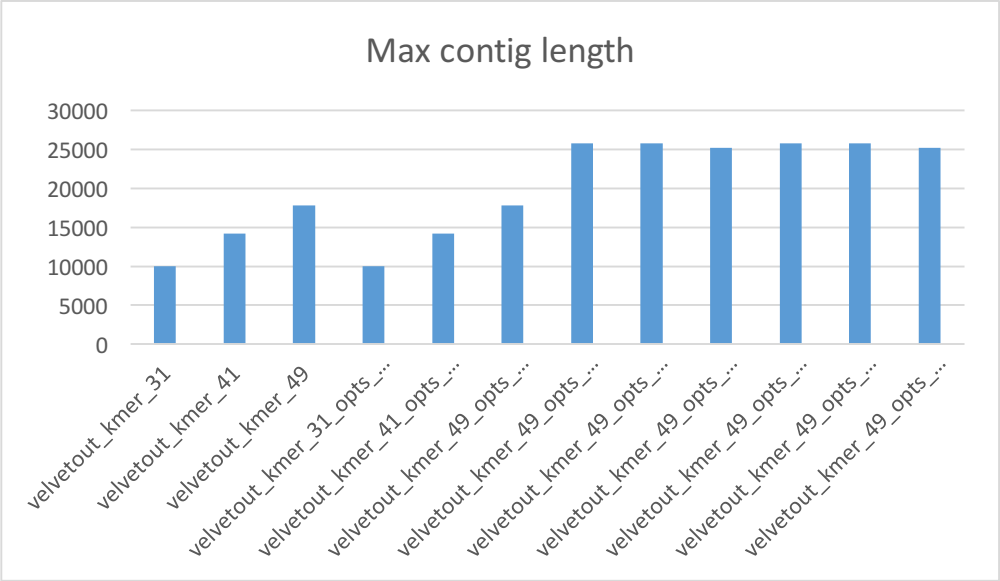32mb size of chrom 4.

The addition of the min contig length of 500, parameter had no effect on the N50
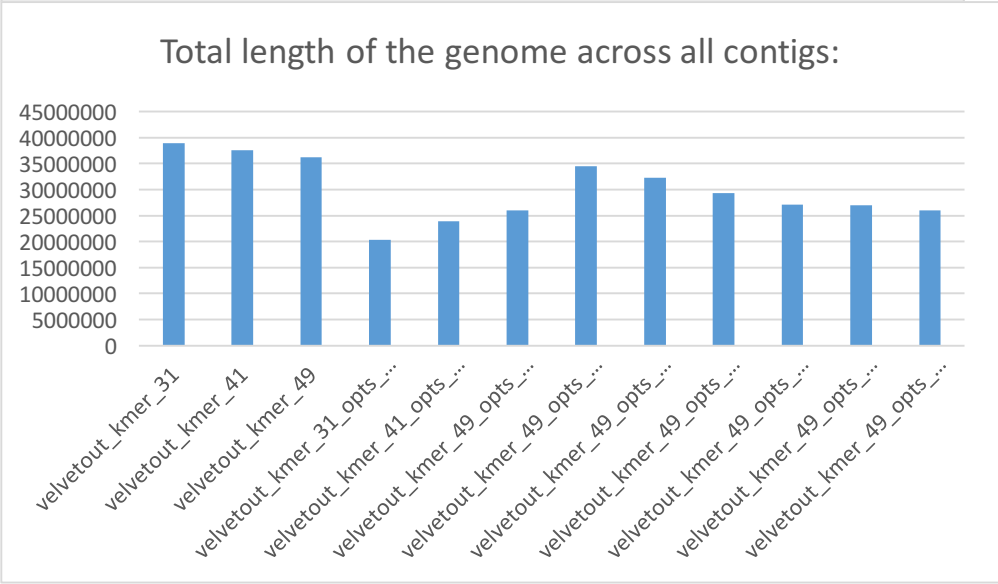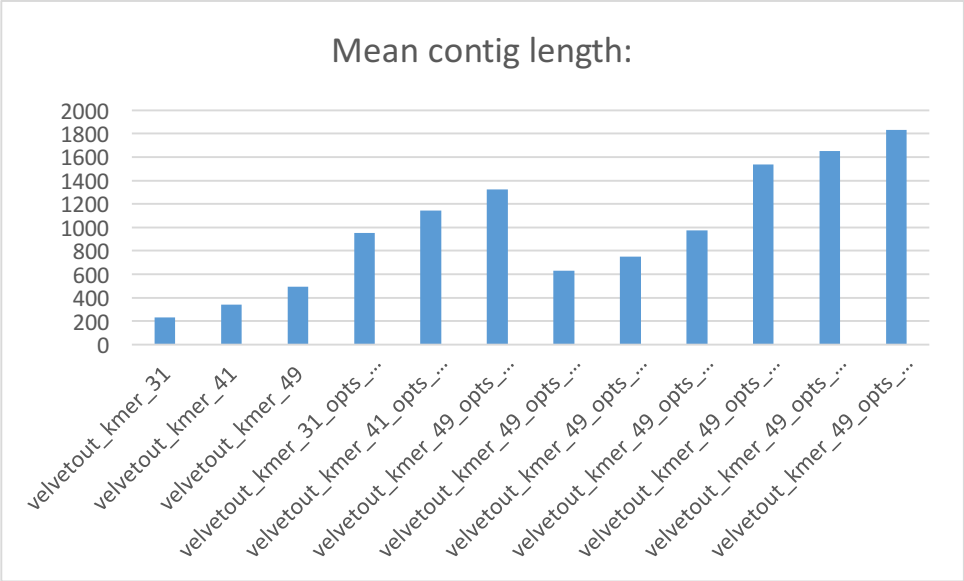or total size of the assembly.

The coverage cutoff parameter had a large impact on the assembly for kmer 49, increasing
the N50 of the assembly and setting the cutoff to auto resulted in the largest N50
of all assembly.

2.With increasing kmer size the contig length distribution gets skewed towards the right,
meaning there are more longer length contigs produced from the assembly. In addition, with
increasing kmer size: the number of contigs goes down, max contig length increases, mean
contig length increases, total length of genome decreases, and the N50 decreases.

3. Increasing coverage cut off reduced contigs with low kmer coverage which typical would
represent sequencing errors. By increasing coverage cutoff your remove paths through the de
Bruijn graph which do not have kmer coverage above that threshold. Accoding to the Velvet
manual Velvetg calculates the auto coverage cutoff by setting it to half the length weighted
median contig coverage depth.

4. Setting the minimum contig length to 500, removes the short contigs from the assembly
reducing contig count. This in turn results in the contig length distribution to be shifted to the
right due to the cut off of shorter contigs.

Max contig length

number of contigs

Mean contig length:

Chart showing "Mean contig length:" (y-axis 0 to 2000) with bars for:
velvetout_kmer_31, velvetout_kmer_41, velvetout_kmer_49, velvetout_kmer_31_opts_..., velvetout_kmer_41_opts_..., velvetout_kmer_49_opts_..., velvetout_kmer_49_opts_..., velvetout_kmer_49_opts_..., velvetout_kmer_49_opts_..., velvetout_kmer_49_opts_..., velvetout_kmer_49_opts_..., velvetout_kmer_49_opts_...

Total length of the genome across all contigs:

Chart showing "Total length of the genome across all contigs:" (y-axis 0 to 45000000) with bars for:
velvetout_kmer_31, velvetout_kmer_41, velvetout_kmer_49, velvetout_kmer_31_opts_..., velvetout_kmer_41_opts_..., velvetout_kmer_49_opts_..., velvetout_kmer_49_opts_..., velvetout_kmer_49_opts_..., velvetout_kmer_49_opts_..., velvetout_kmer_49_opts_..., velvetout_kmer_49_opts_..., velvetout_kmer_49_opts_...

Mean depth of coverage:

N50 of assembly: