Assembly comparison via KAD profiling

22 August 2019

The analysis is to compare k-mer profiles of two sets of sequences, asumming these are two genome assemblies for the same genome. The two assemblies are *a0* and *a2*. From KAD data, k-mers with different copies in the two assemblies were extracted. These k-mers have different KAD values in the two KAD sets. For each KAD set, KAD counts per bin using the bin size as 0.05 were then determined. These counts were plotted to visualize the differentiation of k-mer profiles of the two assemblies.

Run environment and KAD script

R environment:
x86_64-pc-linux-gnu, x86_64, linux-gnu, x86_64, linux-gnu, , 3, 5.1, 2018, 07, 02, 74947, R, R version 3.5.1 (2018-07-02), Feather Spray

KAD script:
KADcompare.pl 0.10

INPUT

assemblies
1. asm1: a0
2. asm2: a2

counts of k-mer in KAD bins
- KAD file: ./result/result_4_kad.txt

KAD profiles

KAD Profiles of unequal KADs

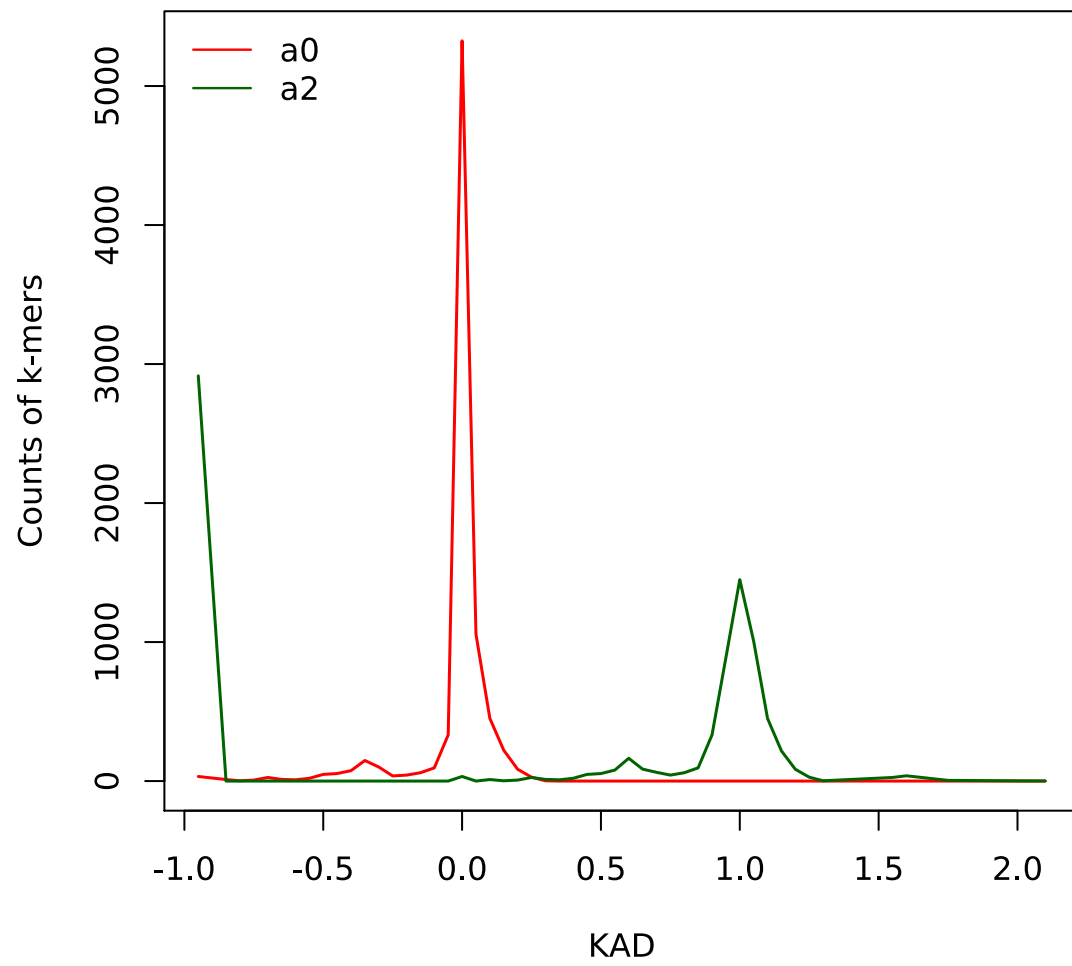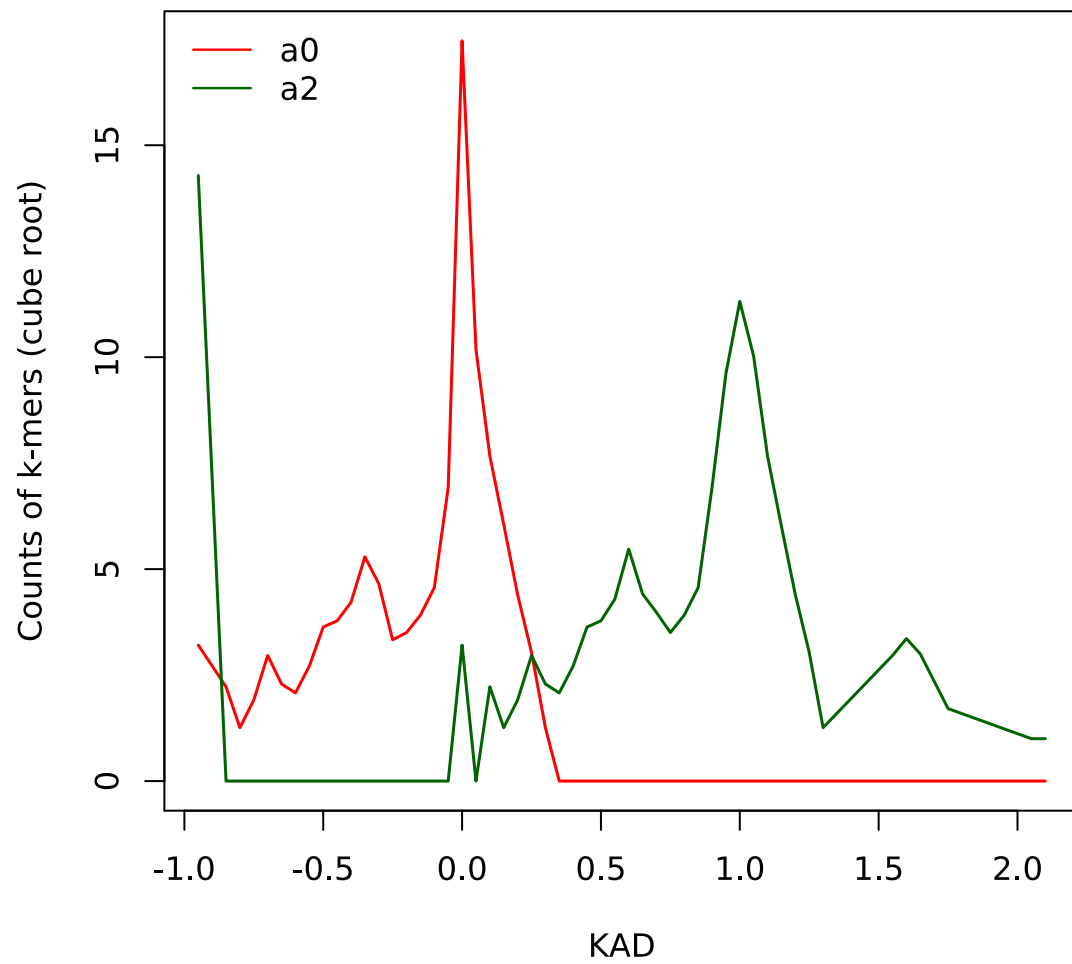Fig 1. KAD profiles

Fig 2. KAD profiles

In the output directory, a0_2, both figures were separately generated in PDF format.