

Resource Summary Report

Generated by [dkNET](#) on Sep 13, 2026

K-mer-based host screening

DOI:10.17504/protocols.io.ehjbb4n

Type: Protocol

Proper Citation

Bonnie Hurwitz, Ken Youens-Clark 2016. K-mer-based host screening. protocols.io
dx.doi.org/10.17504/protocols.io.ehjbb4n

Protocol Information

URL: <https://dx.doi.org/10.17504/protocols.io.ehjbb4n>

Authors: Bonnie Hurwitz, Ken Youens-Clark

Group: VERVE Net, Hurwitz Lab

Summary: K-mer-based approaches to determining sequence similarity can be faster than traditional use of BLAST. This application uses Jellyfish to index host FASTA files, then uses the mode value (default 2) of the number of matching k-mers (default k=20) from the input sequence to determine whether the sequence is similar enough to the host sequence to be rejected. The output of the app is: "screened" directory containing the sequences from each file that were found to be dissimilar to the host "rejected" directory containing the sequence from each file that were too similar to the host "jf" directory containing Jellyfish indexes of each "host" file (useful for later runs with other files to skip recreating) "kmer" directory containing k-mers of query sequences and ".loc" file showing the number of ".kmer" lines associated to each sequence Code is freely available at Github.

Affiliations: University of Arizona, University of Arizona

Version: 1

Publication Date: 2016

Proper Citation: Bonnie Hurwitz, Ken Youens-Clark 2016. K-mer-based host screening. protocols.io dx.doi.org/10.17504/protocols.io.ehjbb4n

Record Creation Time: 20210329T220522+0000

Record Last Update: 20210329T220852+0000

Data and Source Information

Source: