protocols.io
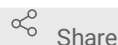
🌐 MIDAS 2 Protocol

miriam.goldman [1,2], chunyu.zhao [3,4]

[1]Data Science and Biotechnology, Gladstone Institutes, San Francisco, CA, USA,;

[2]Biomedical Informatics, University of California San Francisco, San Francisco, CA;

[3]Data Science, Chan Zuckerberg Biohub, San Francisco, CA, USA,;

[4]Data Science and Biotechnology, Gladstone Institutes, San Francisco, CA, USA

Jul 29, 2022

| 1 | Works for me |

| ⌁ Share |

dx.doi.org/10.17504/protocols.io.6qpvr6jzpvmk/v1

miriam.goldman

ABSTRACT

The Metagenomic Intra-Species Diversity Analysis System 2 (MIDAS2) is a scalable pipeline that identifies single nucleotide variants (SNVs) and gene copy number variants (CNVs) in metagenomes using comprehensive reference databases built from public microbial genome collections ("metagenotyping"). MIDAS2 is the first metagenotyping tool with functionality to control metagenomic read mapping filters and to customize the reference database to the microbial community, features that improve the precision and recall of detected variants. Here we present four basic protocols for the most common use cases of MIDAS2, along with two supporting protocols for installation and advanced use. All the steps of metagenotyping, from raw sequencing reads to population genetic analysis, are demonstrated with example data in two downloadable sequencing libraries of single-end metagenomic reads representing a mixture of multiple bacterial species. This set of protocols empowers users to accurately genotype hundreds of species in thousands of samples, providing rich genetic data for studying the evolution and strain-level ecology of microbial communities.

DOI

dx.doi.org/10.17504/protocols.io.6qpvr6jzpvmk/v1

EXTERNAL LINK

https://github.com/czbiohub/MIDAS2, https://midas2.readthedocs.io/en/latest/

COLLECTION CITATION

KEYWORDS

genotyping, strain, microbiome, single nucleotide variant, copy number variant

LICENSE

CREATED

Jul 28, 2022

LAST MODIFIED

Jul 29, 2022

COLLECTION INTEGER ID

67834

ABSTRACT

The Metagenomic Intra-Species Diversity Analysis System 2 (MIDAS2) is a scalable pipeline that identifies single nucleotide variants (SNVs) and gene copy number variants (CNVs) in metagenomes using comprehensive reference databases built from public microbial genome collections ("metagenotyping"). MIDAS2 is the first metagenotyping tool with functionality to control metagenomic read mapping filters and to customize the reference database to the microbial community, features that improve the precision and recall of detected variants. Here we present four basic protocols for the most common use cases of MIDAS2, along with two supporting protocols for installation and advanced use. All the steps of metagenotyping, from raw sequencing reads to population genetic analysis, are demonstrated with example data in two downloadable sequencing libraries of single-end metagenomic reads representing a mixture of multiple bacterial species. This set of protocols empowers users to accurately genotype hundreds of species in thousands of samples, providing rich genetic data for studying the evolution and strain-level ecology of microbial communities.

FILES

BASIC PROTOCOL 1: Species Prescreening
Version 1
by miriam.goldman

BASIC PROTOCOL 2: Download MIDAS Reference Database
Version 1
by miriam.goldman

BASIC PROTOCOL 3: Population Single Nucleotide Variant Calling
Version 1
by miriam.goldman

BASIC PROTOCOL 4: Pan-genome Copy Number Variant Calling
Version 1
by miriam.goldman

Support Protocol 1: Installing MIDAS2
Version 1
by miriam.goldman

Support Protocol 2: Metagenotyping with a Custom Collection of Genomes
Version 1
by miriam.goldman

🐾 protocols.io

3