

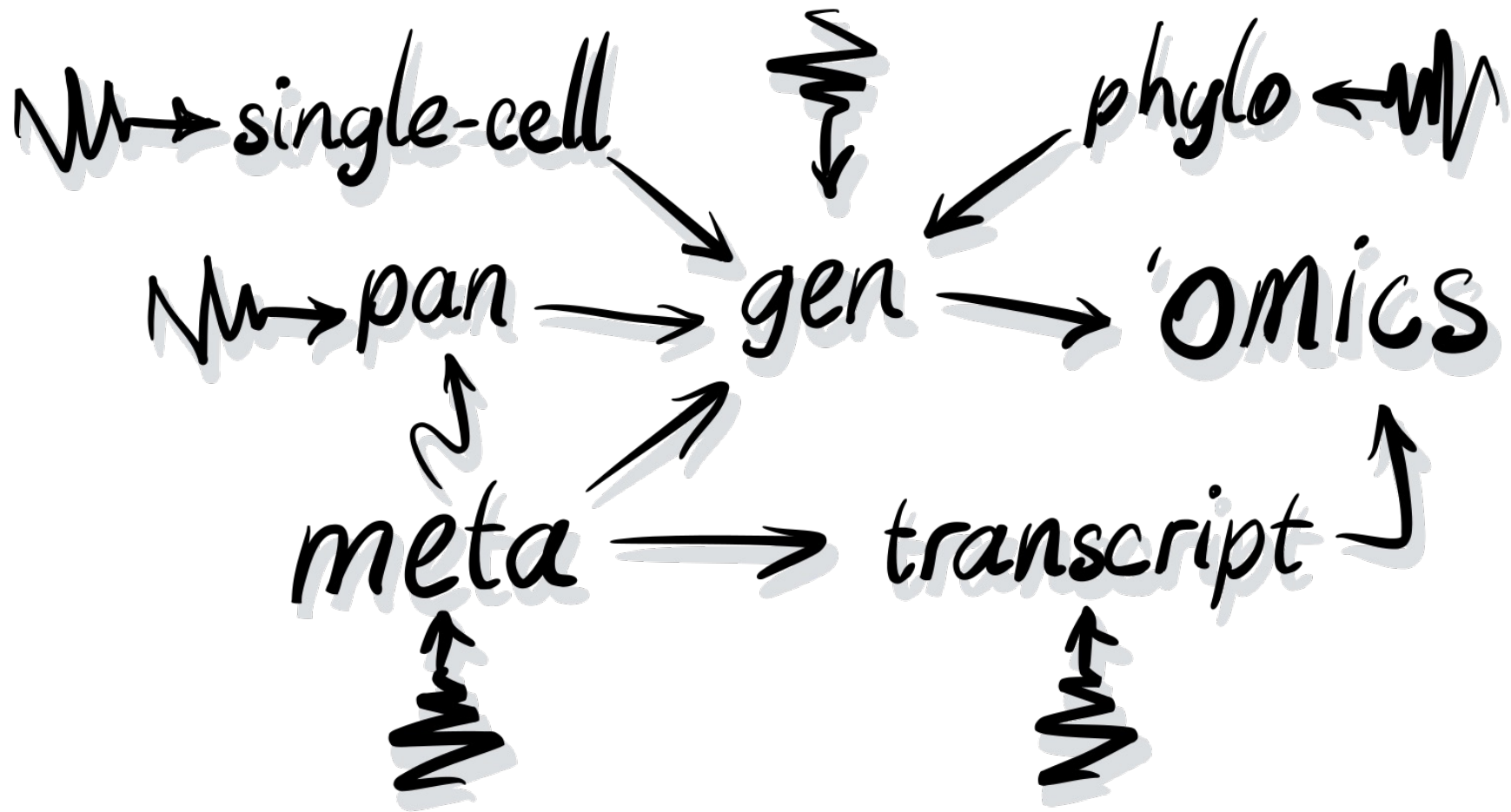


An introduction to
anvi'o

What is anvi'o?



Anvi'o is an open-source, community-driven software platform for integrated 'omics





200,000+ lines-of-code

150,000+ words of tutorials

Robust and flexible design

Integrated visualization

Community-driven

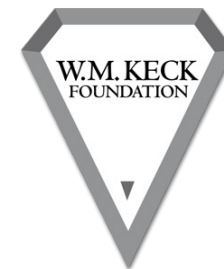
Anvi'o today



Language	Files	Lines	Blanks	Comments	Code	Complexity
Python	374	115635	10712	5434	99489	10150
JavaScript	28	17348	2589	811	13948	2870
Markdown	261	10358	3263	0	7095	0
HTML	8	2783	207	25	2551	0
Shell	25	2666	318	91	2257	79
CSS	11	2587	379	181	2027	0
Go Template	8	1683	235	0	1448	0
Plain Text	54	1009	119	0	890	0
R	1	150	28	8	114	19
YAML	2	116	12	4	100	0
JSON	1	96	0	0	96	0
Dockerfile	1	88	19	29	40	10
Autoconf	1	7	0	0	7	0
Total	775	154526	17881	6583	130062	13128

Estimated Cost to Develop (organic) **\$4,481,808**
Estimated Schedule Effort (organic) 24.320153 months
Estimated People Required (organic) 16.372043

SIMONS
FOUNDATION



Marine Biological Laboratory, Woods Hole (USA)



MBL
Woods Hole



THE UNIVERSITY OF
CHICAGO

State of affairs in 'omics in 2014



Interactive environments for life scientists to perform in-depth analyses of 'omics data.

Easy ways to combine datasets of different nature (i.e., putting metagenomics together with metatranscriptomics or pangenomics into one intuitive display).

Ability to share / communicate findings that emerge from complex data.

Access to reasonable data structures for mathematical or algorithmic mining of complex 'omics data without having to deal with boring steps of common workflows.



ivagljiva

1,240 commits 169,299 ++ 161,042 --



ekiefl

1,581 commits 87,710 ++ 43,682 --



ShaiberAlon

891 commits 25,349 ++ 18,028 --



matthewlawrenceklein

181 commits 2,616 ++ 1,723 --



semiller10

192 commits 31,829 ++ 21,575 --



mschecht

158 commits 7,379 ++ 4,714 --



isaacfink21

88 commits 3,059 ++ 1,316 --



qclayssen

89 commits 744,606 ++ 6,389 --



Jessica-Pan

308 commits 28,365 ++ 24,127 --



mahmoudyousef98

40 commits 3,519 ++ 1,386 --



blankenberg

8 commits 375 ++ 309 --



mooreryan

49 commits 4,633 ++ 1,510 --



gokmen

10 commits 274 ++ 116 --



watsonar

8 commits 16 ++ 17 --



AstrobioMike

4 commits 16 ++ 9 --



efogarty11

4 commits 57 ++ 37 --



farukuzun

8 commits 497 ++ 194 --




tdelmont

1 commit 1 ++ 0 --





Community-led, integrated, reproducible multi-omics with anvi'o

A. Murat Eren , [Evan Kiefl](#), [Alon Shaiber](#), [Iva Veseli](#), Samuel E. Miller, [Matthew S. Schechter](#), [Isaac Fink](#), [Jessica N. Pan](#), [Mahmoud Yousef](#), [Emily C. Fogarty](#), Florian Trigodet, [Andrea R. Watson](#), [Özcan C. Esen](#), [Ryan M. Moore](#), [Quentin Clayssen](#), [Michael D. Lee](#), [Veronika Kivenson](#), [Elaina D. Graham](#), [Bryan D. Merrill](#), Antti Karkman, Daniel Blankenberg, John M. Eppley, Andreas Sjödin, Jarrod J. Scott, Xabier Vázquez-Campos, Luke J. McKay, [Elizabeth A. McDaniel](#), Sarah L. R. Stevens, Rika E. Anderson, Jessika Fuessel, Antonio Fernandez-Guerra, Lois Maignien, Tom O. Delmont & Amy D. Willis

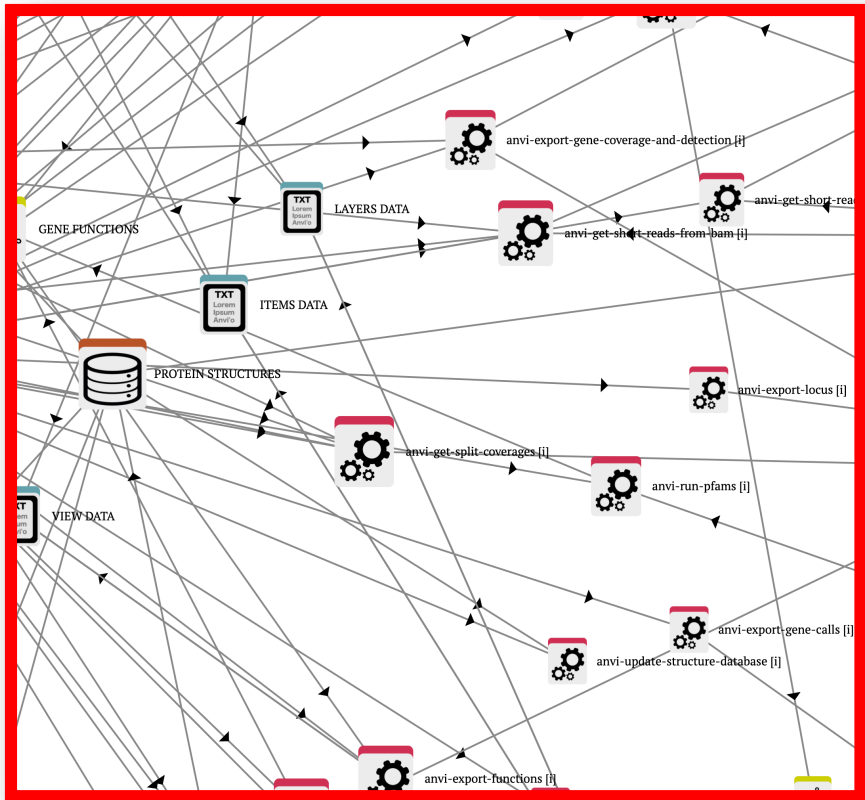
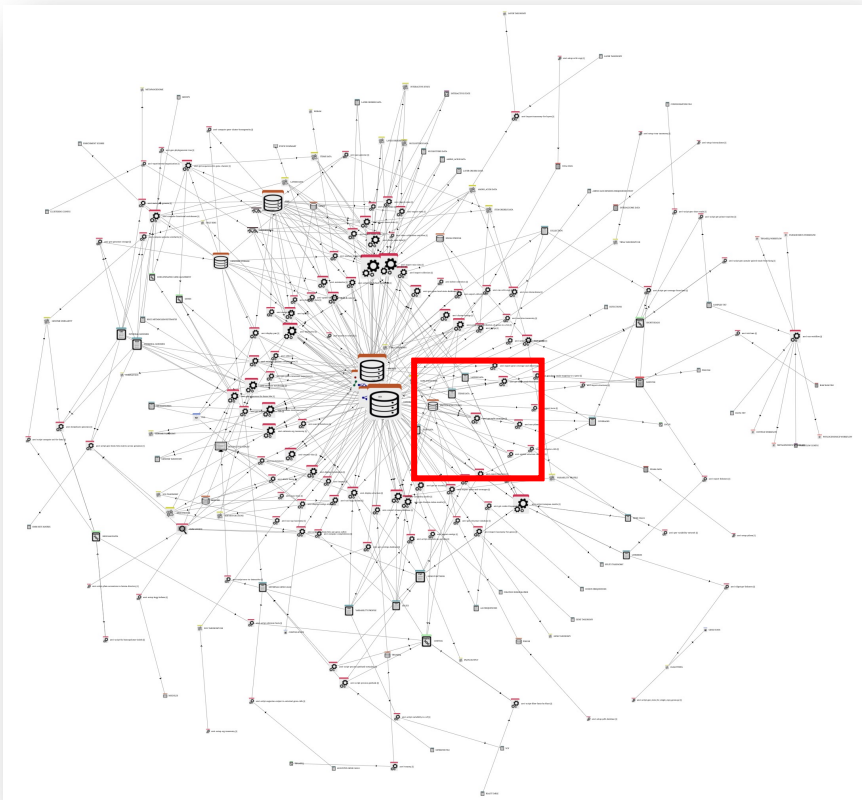
A LEGO-like architectural design



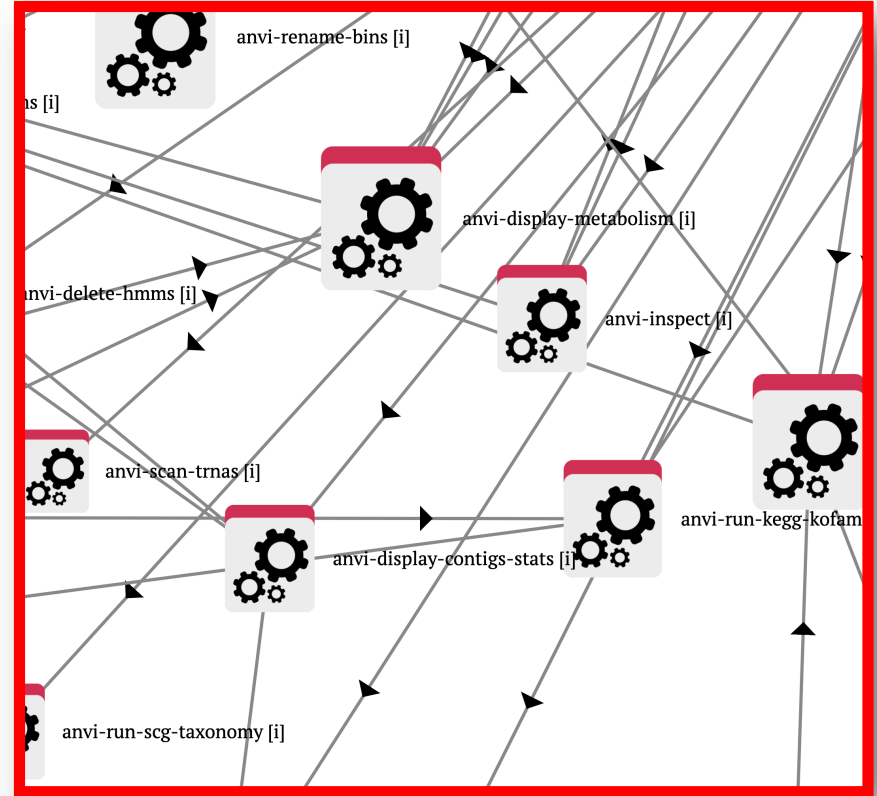
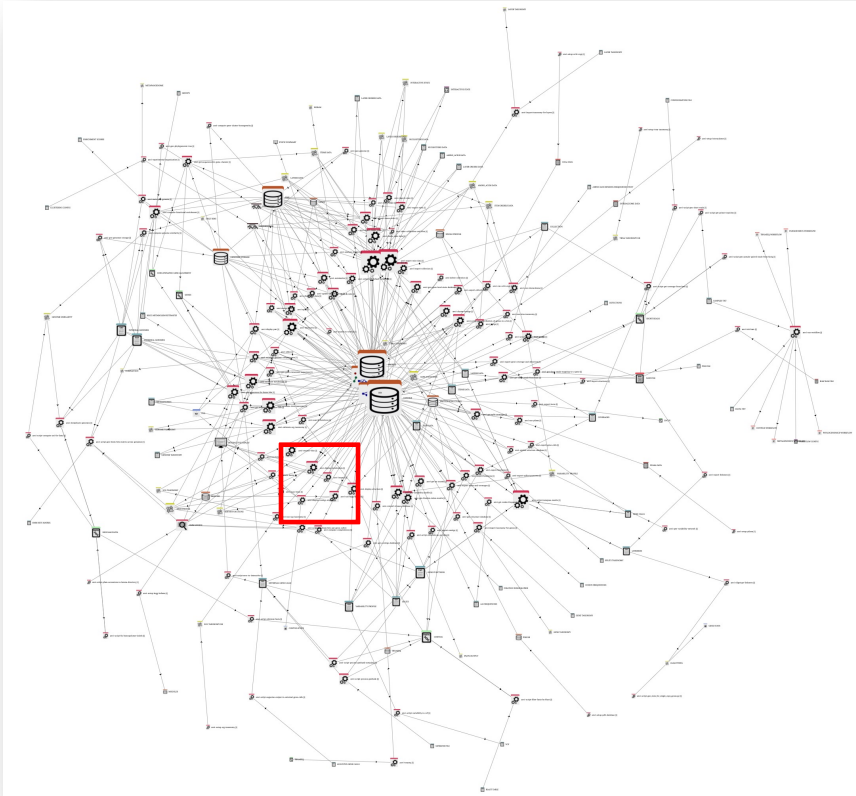
```
$ ls */anvi-*
bin/anvi-3dev
bin/anvi-agglomerate
bin/anvi-analyze-synteny
bin/anvi-cluster-contigs
bin/anvi-compute-ani
bin/anvi-compute-completeness
bin/anvi-compute-gene-cluster-homogeneity
bin/anvi-compute-genome-similarity
bin/anvi-db-info
bin/anvi-delete-collection
bin/anvi-delete-hmms
bin/anvi-delete-misc-data
bin/anvi-delete-state
bin/anvi-dereplicate-genomes
bin/anvi-display-contigs-stats
bin/anvi-display-metabolism
bin/anvi-display-pan
bin/anvi-estimate-genome-completeness
bin/anvi-estimate-genome-taxonomy
bin/anvi-estimate-metabolism
bin/anvi-estimate-scg-taxonomy
bin/anvi-estimate-trna-taxonomy
bin/anvi-experimental-organization
bin/anvi-export-collection
bin/anvi-export-contigs
bin/anvi-export-functions
bin/anvi-export-gene-calls
bin/anvi-export-gene-coverage-and-detection
bin/anvi-export-items-order
bin/anvi-export-locus
bin/anvi-export-misc-data
bin/anvi-export-splits-and-coverages
bin/anvi-export-splits-taxonomy
bin/anvi-export-state
bin/anvi-export-structures
bin/anvi-export-table
bin/anvi-gen-contigs-database
bin/anvi-gen-fixation-index-matrix
bin/anvi-gen-gene-consensus-sequences
bin/anvi-gen-gene-level-stats-databases
bin/anvi-gen-genomes-storage
bin/anvi-gen-network
bin/anvi-gen-phylogenomic-tree
bin/anvi-gen-structure-database
bin/anvi-gen-variability-matrix
bin/anvi-gen-variability-network
bin/anvi-gen-variability-profile
bin/anvi-get-aa-counts
bin/anvi-get-codon-frequencies
bin/anvi-get-enriched-functions-per-pan-group
bin/anvi-get-sequences-for-gene-calls
bin/anvi-get-sequences-for-gene-clusters
bin/anvi-get-sequences-for-hmm-hits
bin/anvi-get-short-reads-from-bam
bin/anvi-get-short-reads-mapping-to-a-gene
bin/anvi-get-split-coverages
bin/anvi-help
bin/anvi-import-collection
bin/anvi-import-functions
bin/anvi-import-items-order
bin/anvi-import-misc-data
bin/anvi-import-state
bin/anvi-import-taxonomy-for-genes
bin/anvi-import-taxonomy-for-layers
bin/anvi-init-bam
bin/anvi-inspect
bin/anvi-interactive
bin/anvi-matrix-to-newick
bin/anvi-update-mcg-classifier
bin/anvi-merge
bin/anvi-merge-bins
bin/anvi-meta-pan-genome
bin/anvi-migrate
bin/anvi-oligotype-linkmers
bin/anvi-pan-genome
bin/anvi-profile
bin/anvi-push
bin/anvi-refine
bin/anvi-rename-bins
bin/anvi-report-linkmers
bin/anvi-run-hmms
bin/anvi-run-interacdome
bin/anvi-run-kegg-kofams
bin/anvi-run-ncbi-cogs
bin/anvi-run-pfams
bin/anvi-run-scg-taxonomy
bin/anvi-run-trna-taxonomy
bin/anvi-run-workflow
bin/anvi-scan-trnas
bin/anvi-search-functions
bin/anvi-self-test
bin/anvi-setup-interacdome
bin/anvi-setup-kegg-kofams
bin/anvi-setup-ncbi-cogs
bin/anvi-setup-pdb-database
bin/anvi-setup-pfams
bin/anvi-setup-scg-taxonomy
bin/anvi-setup-trna-taxonomy
bin/anvi-show-collections-and-bins
bin/anvi-show-misc-data
bin/anvi-split
bin/anvi-summarize
bin/anvi-trnaseq
bin/anvi-update-db-description
bin/anvi-update-structure-database
bin/anvi-upgrade
sandbox/anvi-script-calculate-pn-ps-ratio
sandbox/anvi-script-checkm-tree-to-interactive
sandbox/anvi-script-compute-ani-for-fasta
sandbox/anvi-script-estimate-genome-size
sandbox/anvi-script-filter-fasta-by-blast
sandbox/anvi-script-gen-CPR-classifier
sandbox/anvi-script-gen-distribution-of-genes-in-a-bin
sandbox/anvi-script-gen-help-pages
sandbox/anvi-script-gen-hmm-hits-matrix-across-genomes
sandbox/anvi-script-gen-programs-network
sandbox/anvi-script-gen-programs-vignette
sandbox/anvi-script-gen-pseudo-paired-reads-from-fastq
sandbox/anvi-script-gen-scg-domain-classifier
sandbox/anvi-script-gen-short-reads
sandbox/anvi-script-gen_stats_for_single_copy_genes.R
sandbox/anvi-script-gen_stats_for_single_copy_genes.py
sandbox/anvi-script-gen_stats_for_single_copy_genes.sh
sandbox/anvi-script-get-collection-info
sandbox/anvi-script-get-coverage-from-bam
sandbox/anvi-script-get-hmm-hits-per-gene-call
sandbox/anvi-script-get-short-reads-matching-something
sandbox/anvi-script-merge-collections
sandbox/anvi-script-predict-CPR-genomes
sandbox/anvi-script-process-genbank
sandbox/anvi-script-process-genbank-metadata
sandbox/anvi-script-reformat-fasta
sandbox/anvi-script-run-eggno-mapper
sandbox/anvi-script-run-functional-enrichment-stats
sandbox/anvi-script-snvs-to-interactive
sandbox/anvi-script-tabulate
sandbox/anvi-script-transpose-matrix
sandbox/anvi-script-variability-to-vcf
sandbox/anvi-script-visualize-split-coverages
```

```
$
```

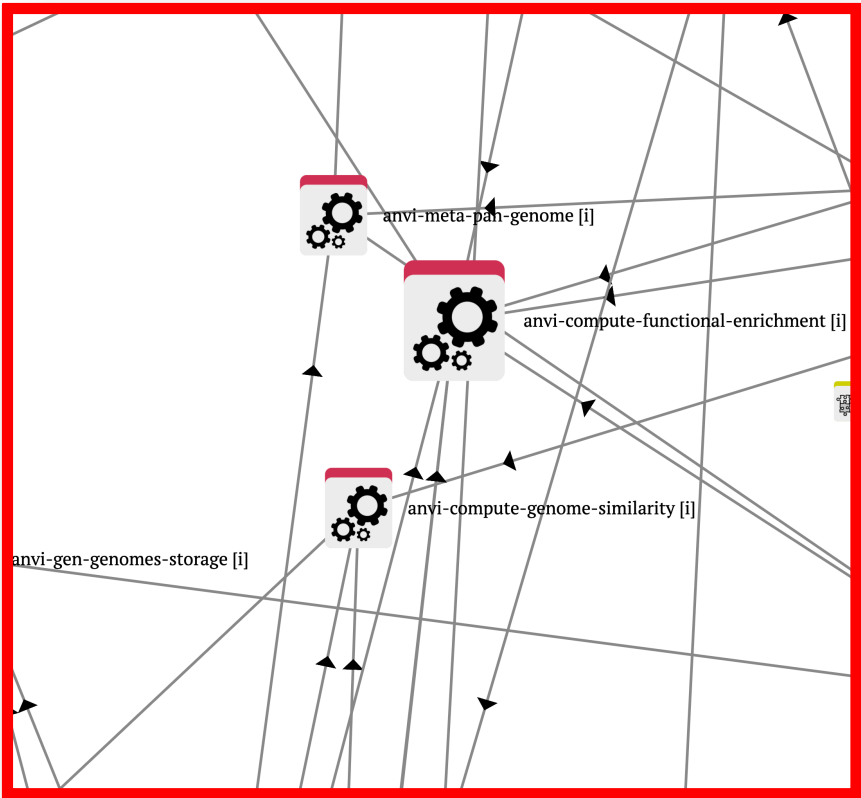
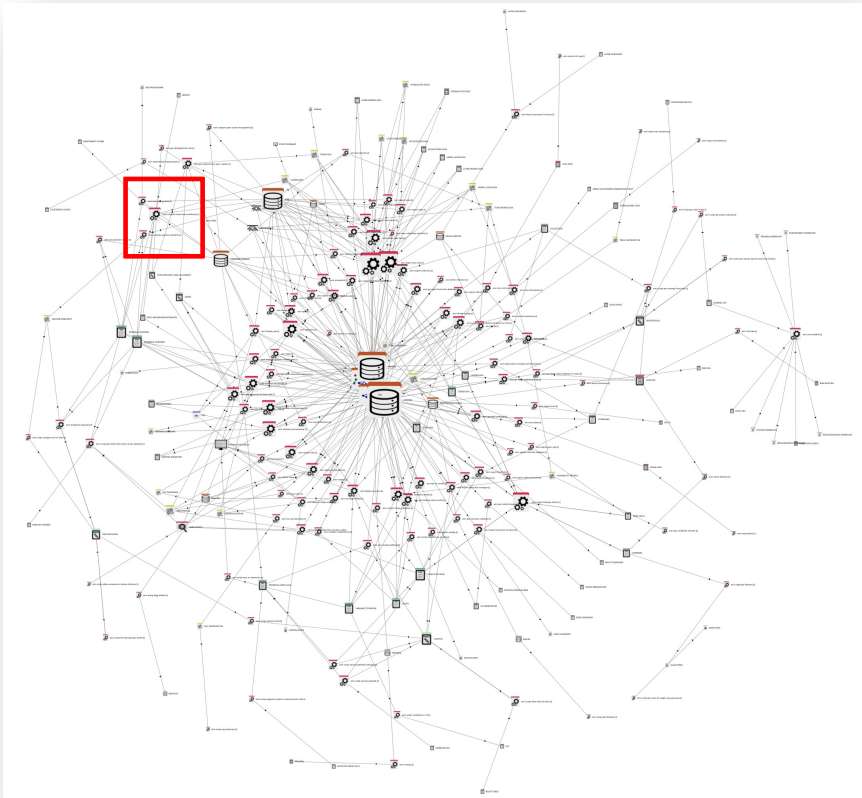
Programs and concepts connected

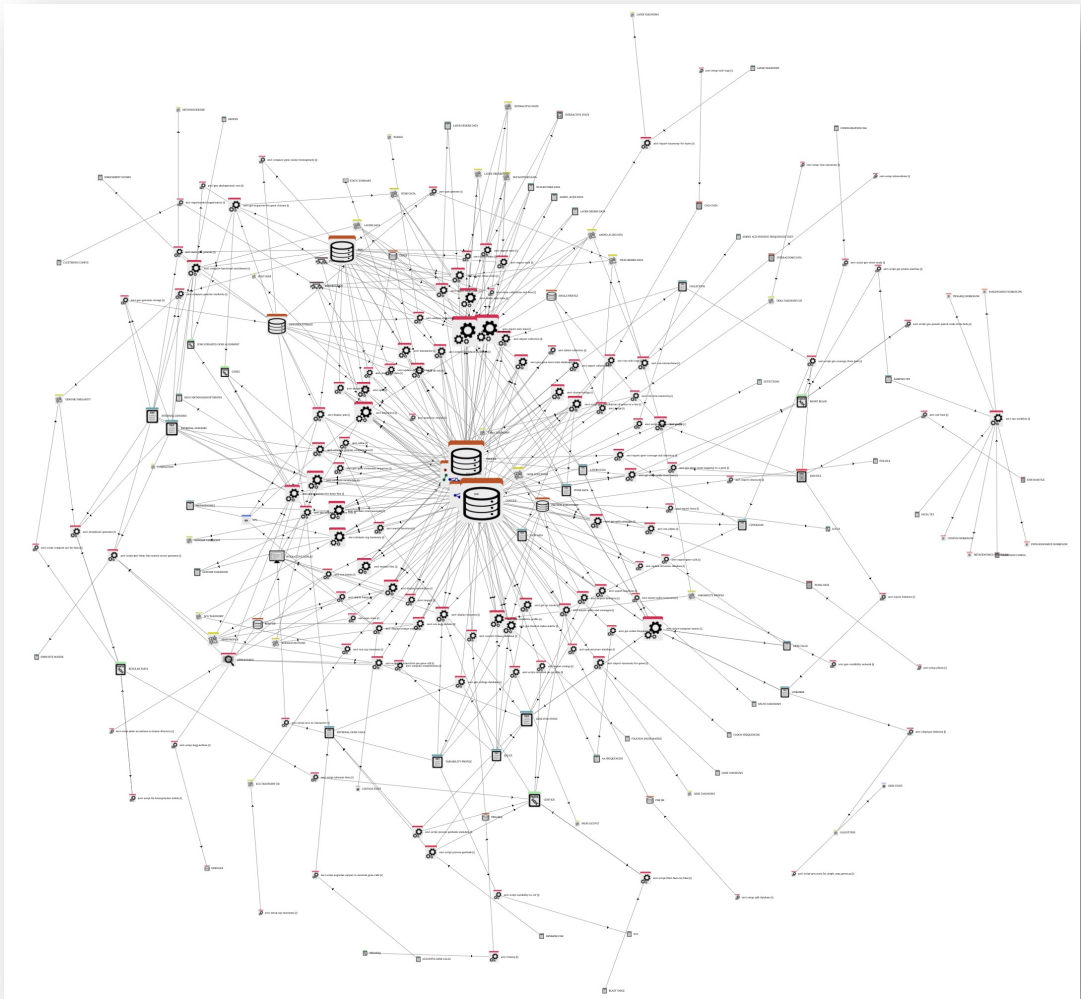


Programs and concepts connected



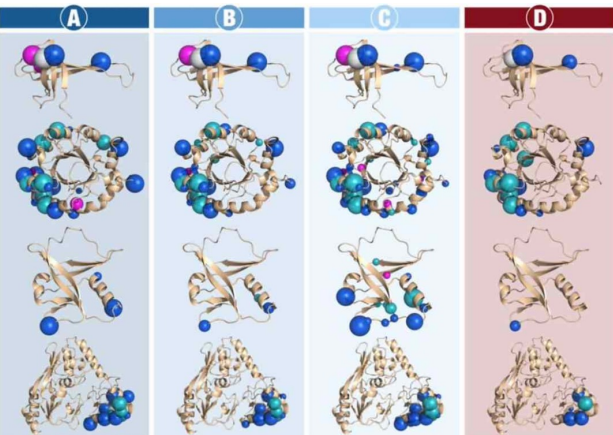
Programs and concepts connected





Proteotypes

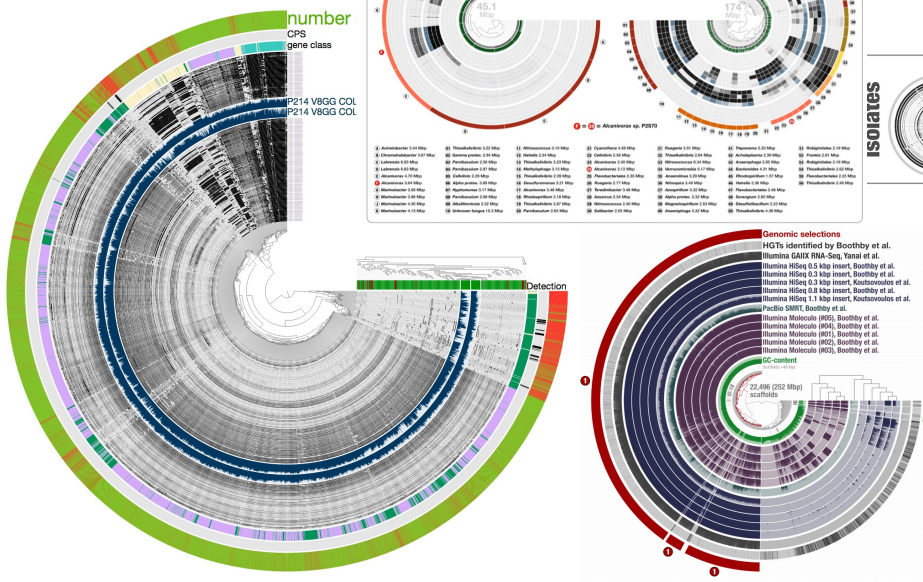
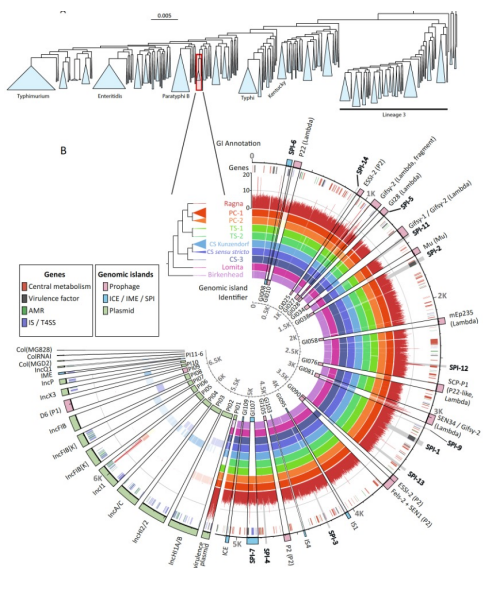
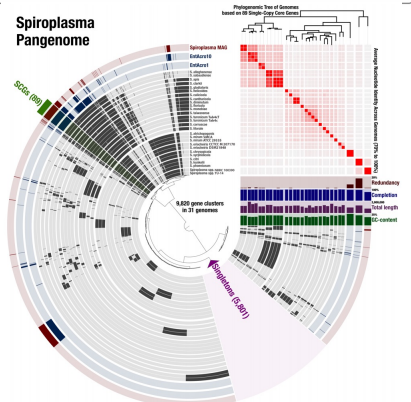
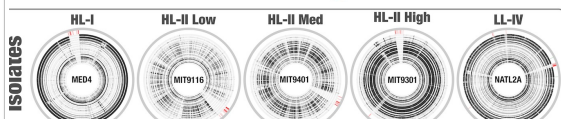
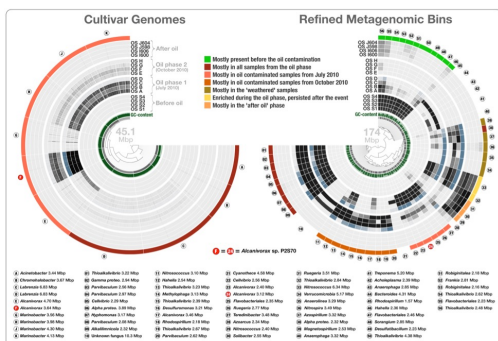
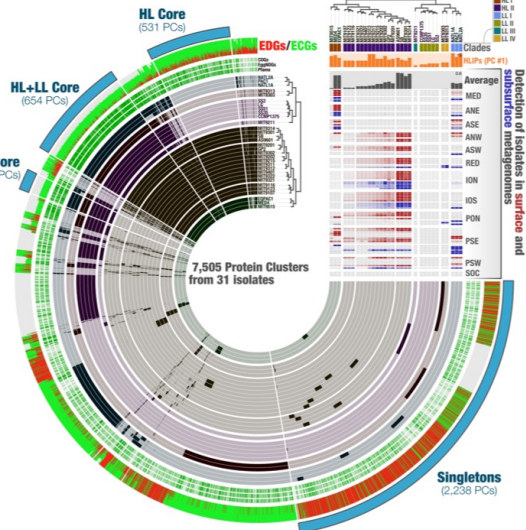
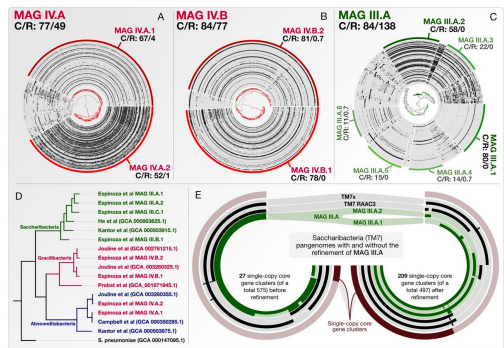
Gene 1270
Ribosomal protein S17



Gene 2128
Ribulose-phosphate 3-epimerase

Gene 1398
Iron-sulfur cluster biosynthesis protein

Gene 2456
UDP-N-acetylpyruvoyl reductase protein



Genomic selections
HGTs identified by Boothby et al.
Humania GAIIX RNA-Seg, Yanai et al.
Humania H5eq 0.3 kbp insert, Boothby et al.
Humania H5eq 0.3 kbp insert, Boothby et al.
Humania H5eq 0.8 kbp insert, Koshornikova et al.
Humania H5eq 1.1 kbp insert, Koshornikova et al.
Pediococcus SMT, Boothby et al.
Humania Molocato (P2), Boothby et al.
Humania Molocato (P3), Boothby et al.
Humania Molocato (P4), Boothby et al.
Humania Molocato (P5), Boothby et al.
Humania Molocato (P6), Boothby et al.
Humania Molocato (P7), Boothby et al.
Humania Molocato (P8), Boothby et al.
Humania Molocato (P9), Boothby et al.
Humania Molocato (P10), Boothby et al.
Humania Molocato (P11), Boothby et al.
Humania Molocato (P12), Boothby et al.
Humania Molocato (P13), Boothby et al.
Humania Molocato (P14), Boothby et al.
Humania Molocato (P15), Boothby et al.
Humania Molocato (P16), Boothby et al.
Humania Molocato (P17), Boothby et al.
Humania Molocato (P18), Boothby et al.
Humania Molocato (P19), Boothby et al.
Humania Molocato (P20), Boothby et al.
Humania Molocato (P21), Boothby et al.
Humania Molocato (P22), Boothby et al.
Humania Molocato (P23), Boothby et al.
Humania Molocato (P24), Boothby et al.
Humania Molocato (P25), Boothby et al.
Humania Molocato (P26), Boothby et al.
Humania Molocato (P27), Boothby et al.
Humania Molocato (P28), Boothby et al.
Humania Molocato (P29), Boothby et al.
Humania Molocato (P30), Boothby et al.
Humania Molocato (P31), Boothby et al.
Humania Molocato (P32), Boothby et al.
Humania Molocato (P33), Boothby et al.
Humania Molocato (P34), Boothby et al.
Humania Molocato (P35), Boothby et al.
Humania Molocato (P36), Boothby et al.
Humania Molocato (P37), Boothby et al.
Humania Molocato (P38), Boothby et al.
Humania Molocato (P39), Boothby et al.
Humania Molocato (P40), Boothby et al.
Humania Molocato (P41), Boothby et al.
Humania Molocato (P42), Boothby et al.
Humania Molocato (P43), Boothby et al.
Humania Molocato (P44), Boothby et al.
Humania Molocato (P45), Boothby et al.
Humania Molocato (P46), Boothby et al.
Humania Molocato (P47), Boothby et al.
Humania Molocato (P48), Boothby et al.
Humania Molocato (P49), Boothby et al.
Humania Molocato (P50), Boothby et al.
Humania Molocato (P51), Boothby et al.
Humania Molocato (P52), Boothby et al.
Humania Molocato (P53), Boothby et al.
Humania Molocato (P54), Boothby et al.
Humania Molocato (P55), Boothby et al.
Humania Molocato (P56), Boothby et al.
Humania Molocato (P57), Boothby et al.
Humania Molocato (P58), Boothby et al.
Humania Molocato (P59), Boothby et al.
Humania Molocato (P60), Boothby et al.
Humania Molocato (P61), Boothby et al.
Humania Molocato (P62), Boothby et al.
Humania Molocato (P63), Boothby et al.
Humania Molocato (P64), Boothby et al.
Humania Molocato (P65), Boothby et al.
Humania Molocato (P66), Boothby et al.
Humania Molocato (P67), Boothby et al.
Humania Molocato (P68), Boothby et al.
Humania Molocato (P69), Boothby et al.
Humania Molocato (P70), Boothby et al.
Humania Molocato (P71), Boothby et al.
Humania Molocato (P72), Boothby et al.
Humania Molocato (P73), Boothby et al.
Humania Molocato (P74), Boothby et al.
Humania Molocato (P75), Boothby et al.
Humania Molocato (P76), Boothby et al.
Humania Molocato (P77), Boothby et al.
Humania Molocato (P78), Boothby et al.
Humania Molocato (P79), Boothby et al.
Humania Molocato (P80), Boothby et al.
Humania Molocato (P81), Boothby et al.
Humania Molocato (P82), Boothby et al.
Humania Molocato (P83), Boothby et al.
Humania Molocato (P84), Boothby et al.
Humania Molocato (P85), Boothby et al.
Humania Molocato (P86), Boothby et al.
Humania Molocato (P87), Boothby et al.
Humania Molocato (P88), Boothby et al.
Humania Molocato (P89), Boothby et al.
Humania Molocato (P90), Boothby et al.
Humania Molocato (P91), Boothby et al.
Humania Molocato (P92), Boothby et al.
Humania Molocato (P93), Boothby et al.
Humania Molocato (P94), Boothby et al.
Humania Molocato (P95), Boothby et al.
Humania Molocato (P96), Boothby et al.
Humania Molocato (P97), Boothby et al.
Humania Molocato (P98), Boothby et al.
Humania Molocato (P99), Boothby et al.
Humania Molocato (P100), Boothby et al.



Free 'omics workshops for more than 700 scientists over the last three years pre-COVID



Anvi'o as a community in 2023

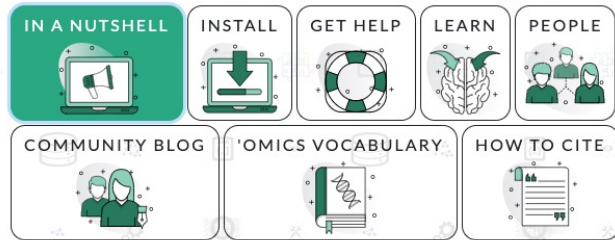


ANVI'O 

[INSTALL](#) [GET HELP](#) [LEARN](#) [PEOPLE](#) [COMMUNITY BLOG](#) ['OMICS VOCABULARY](#) [HOW TO CITE](#)

Integrated multi-omics at scale

An open-source, community-driven **analysis** and **visualization** platform for microbial **'omics**.



Anvi'o as a community in 2023



ANVI'O 

[INSTALL](#) [GET HELP](#) [LEARN](#) [PEOPLE](#) [COMMUNITY BLOG](#) [OMICS VOCABULARY](#) [HOW TO CITE](#)



Anvi'o as a community in 2023



SOURCE

~ 1,500 members, ~ 30,000 messages

The image displays a grid of screenshots from the Anvi'o GitHub repository and community discussions. The central focus is the word "SOURCE" in large, bold, black letters. To the right, text reads "~ 1,500 members, ~ 30,000 messages". The background shows various GitHub issues, pull requests, and forum posts, including discussions about installation, data processing, and community support.

Anvi'o as a community in 2023



Discord Edit View Window Help

Anvi'o Discord questions

Discussion Question

Anvi-estimate-metabolism on large number of samples

drpoo: Hello I have ~300 metagenomes with both contigs and profile dbs for each. I am trying to run anvi-estimate-metabolism in metagenome mode. Unfortunately, the estimated time for the job far exceeds my institution's wall time on our HPC. Looking at the document...

0 17d ago

Question

Visualizing specific gene calls in anvi-interactive

peygadin: ...interactive and so far it seems to be working giving me all the basic information. I have three genes that I have identified in a contig database for one of my metagenomic samples that I would like to highlight in my anvi-interac...

Question

Highlight specific clusters in the p...

ZH: ...gene clusters in 15 genomes using seqtk and maxclust phylogen... using anvi'o with ... commands. anvi'o --contigs .../db --mig... safety .../gen-genomes-storage -e external-genome...

Question

Config Error: The port number X seems to be in use

Nolan: Hi everyone, first message on the Discord so please bear with me! I am running anvio on a remote server, using my laptop as an intermediary to connect to the server. I'm using the command `anvi-run /path/to/server/11/28/visualizing-from-a-server/` to open the anvio i...

3 22d ago

Question

Binning a coassembly with an external program

Kent: Hello all, I'm currently working on a coassembly generated from reads at multiple time points (T0-T7). I've followed the Anvio tutorial for metagenomic workflow until just after running anvi-merge. I'd like to use either concoct or metabat2 for binning but I'm unsure of ...

8 23d ago

Question

anvi-profile config error "The third item..."

Stefan: Hi, using anvio 7.1 in Ubuntu 22.04, I try to make profiles from transcriptomes mapped to bins and the bin contig db (from metagenomes) using the code `anvi-profile -i METAT.bam -c /FOLDER1/FOLDER2/METAG.db -T2 --skip-SNV-profiling --output-dir NAME --min...`

~650 members,
1,000+ threads

Data integration in anvi'o

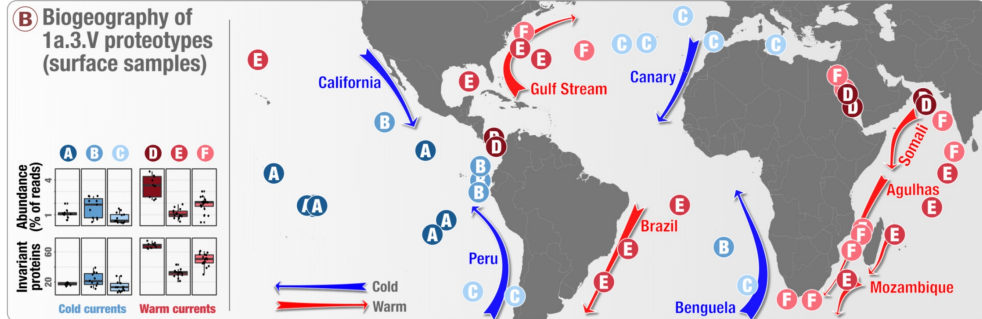
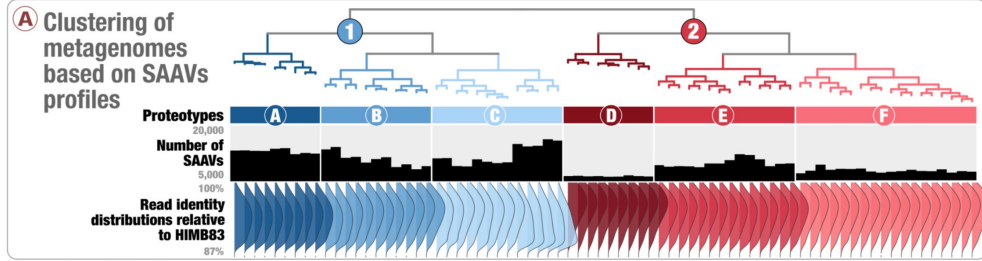


Single-amino acid variants reveal evolutionary processes that shape the biogeography of a global SAR11 subclade

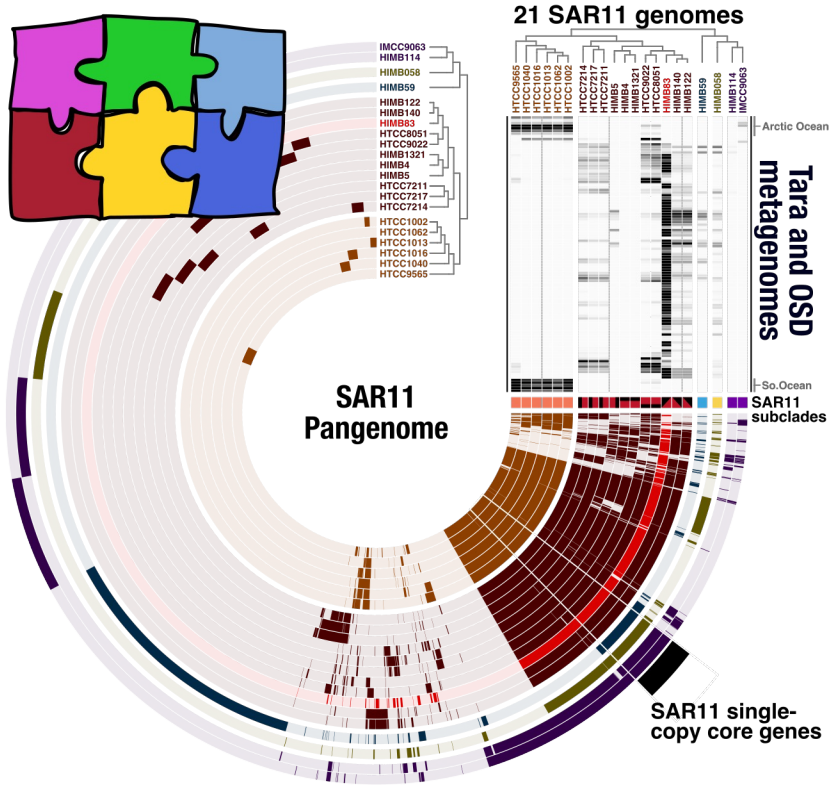
Tom O Delmont^{1†}, Evan Kiefl^{1,2†}, Ozsel Kilinc³, Ozcan C Esen¹, Ismail Uysal³, Michael S Rappé⁴, Steven Giovannoni⁵, A Murat Eren^{1,6*}







21 SAR11 genomes

HTCC9686
HTCC9687
HTCC1016
HTCC1018
HTCC1022
HTCC214
HTCC217
HIM82
HIM82_11
HIM83
HIM83_01
HTCC9022
HIM8140
HIM8122
HIM89
HIM89_01
HIM89_02
HIM89_03
HIM89_04
HIM89_05
HIM89_06
HIM89_07
HIM89_08
HIM89_09
HIM89_10
HIM89_11
HIM89_12
HIM89_13
HIM89_14
HIM89_15
HIM89_16
HIM89_17
HIM89_18
HIM89_19
HIM89_20
HIM89_21
HIM89_22
HIM89_23
HIM89_24
HIM89_25
HIM89_26
HIM89_27
HIM89_28
HIM89_29
HIM89_30
HIM89_31
HIM89_32
HIM89_33
HIM89_34
HIM89_35
HIM89_36
HIM89_37
HIM89_38
HIM89_39
HIM89_40
HIM89_41
HIM89_42
HIM89_43
HIM89_44
HIM89_45
HIM89_46
HIM89_47
HIM89_48
HIM89_49
HIM89_50
HIM89_51
HIM89_52
HIM89_53
HIM89_54
HIM89_55
HIM89_56
HIM89_57
HIM89_58
HIM89_59
HIM89_60
HIM89_61
HIM89_62
HIM89_63
HIM89_64
HIM89_65
HIM89_66
HIM89_67
HIM89_68
HIM89_69
HIM89_70
HIM89_71
HIM89_72
HIM89_73
HIM89_74
HIM89_75
HIM89_76
HIM89_77
HIM89_78
HIM89_79
HIM89_80
HIM89_81
HIM89_82
HIM89_83
HIM89_84
HIM89_85
HIM89_86
HIM89_87
HIM89_88
HIM89_89
HIM89_90
HIM89_91
HIM89_92
HIM89_93
HIM89_94
HIM89_95
HIM89_96
HIM89_97
HIM89_98
HIM89_99
HIM89_100



Data integration in anvi'o



-  Pangenomics
-  Phylogenetics
-  Phylogenomics
-  Metagenomics
-  Population genetics
-  In silico protein biochemistry



Solvent accessibility: ● Exposed ● Intermediate ● Buried ● NA

Data transparency and reproducibility



21 SAR11 genomes

HMCC2063
HIMB114
HIMB059
HIMB59

C20686
C20687
C20688
C20689
C20690
C20691
C20692
C20693
C20694
C20695
C20696
C20697
C20698
C20699
C20700
C20701
C20702
C20703
C20704
C20705
C20706
C20707
C20708
C20709
C20710
C20711
C20712
C20713
C20714
C20715
C20716
C20717
C20718
C20719
C20720
C20721
C20722
C20723
C20724
C20725
C20726
C20727
C20728
C20729
C20730
C20731
C20732
C20733
C20734
C20735
C20736
C20737
C20738
C20739
C20740
C20741
C20742
C20743
C20744
C20745
C20746
C20747
C20748
C20749
C20750
C20751
C20752
C20753
C20754
C20755
C20756
C20757
C20758
C20759
C20760
C20761
C20762
C20763
C20764
C20765
C20766
C20767
C20768
C20769
C20770
C20771
C20772
C20773
C20774
C20775
C20776
C20777
C20778
C20779
C20780
C20781
C20782
C20783
C20784
C20785
C20786
C20787
C20788
C20789
C20790
C20791
C20792
C20793
C20794
C20795
C20796
C20797
C20798
C20799
C20800

Pangenomics
Phylogenetics

 eLife 10.7554/eLife.46497.001

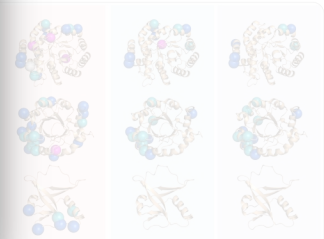
Genetics and Genomics

Code and data availability

The vast majority of analyses relied on the open-source software platform anvi'o v2.4.0 (available from <http://merenlab.org/software/anvio>). The URL <http://merenlab.org/data/sar11-saavs> serves the remaining custom code used in our analyses. We made available (1) SAR11 isolate genomes ([doi:10.6084/m9.figshare.5248945](https://doi.org/10.6084/m9.figshare.5248945)), (2) the anvi'o contigs database and merged profile for SAR11 genomes across metagenomes ([doi:10.5281/zenodo.835218](https://doi.org/10.5281/zenodo.835218)) and the static HTML summary for the mapping results ([doi:10.6084/m9.figshare.5248453](https://doi.org/10.6084/m9.figshare.5248453)), (3) the SAR11 metapangenome ([doi:10.6084/m9.figshare.5248459](https://doi.org/10.6084/m9.figshare.5248459)), single-nucleotide and single-amino acid variant reports for 1a.3.V across 74 TARA Oceans metagenomes ([doi:10.6084/m9.figshare.5248447](https://doi.org/10.6084/m9.figshare.5248447)), and (4) SAAVs overlaid on predicted tertiary structures of 58 core 1a.3.V genes ([doi:10.6084/m9.figshare.5248432](https://doi.org/10.6084/m9.figshare.5248432)). The URL <http://anvi-server.org/p/4Q2TNo> serves an interactive version of the SAR11 metapangenome, and the URL <http://data.merenlab.org/sar11-saavs> serves an interactive web page to investigate the link between SAAVs and predicted protein structures.



Biochemistry



Exposed Intermediate Buried NA

Data transparency and reproducibility



Pangenomics
Phylogenetics

 eLife 10.7554/eLife.46497.001

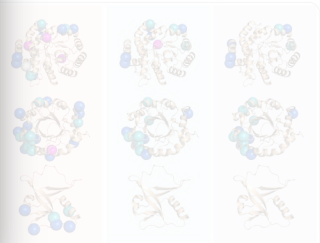
Genetics and Genomics

Code and data availability

The vast majority of analyses relied on the open-source software platform anvi'o v2.4.0 (available from <http://merenlab.org/software/anvio>). The URL <http://merenlab.org/data/sar11-saavs> serves the remaining custom code used in our analyses. We made available (1) SAR11 isolate genomes ([doi:10.6084/m9.figshare.5248945](https://doi.org/10.6084/m9.figshare.5248945)), (2) the anvi'o contigs database and merged profile for SAR11 genomes across metagenomes ([doi:10.5281/zenodo.835218](https://doi.org/10.5281/zenodo.835218)) and the static HTML summary for the mapping results ([doi:10.6084/m9.figshare.5248453](https://doi.org/10.6084/m9.figshare.5248453)), (3) the SAR11 metapangenome ([doi:10.6084/m9.figshare.5248459](https://doi.org/10.6084/m9.figshare.5248459)), single-nucleotide and single-amino acid variant reports for 1a.3.V across 74 TARA Oceans metagenomes ([doi:10.6084/m9.figshare.5248447](https://doi.org/10.6084/m9.figshare.5248447)), and (4) SAAVs overlaid on predicted tertiary structures of 58 core 1a.3.V genes ([doi:10.6084/m9.figshare.5248432](https://doi.org/10.6084/m9.figshare.5248432)). The URL <http://anvi-server.org/p/4Q2TNo> serves an interactive version of the SAR11 metapangenome, and the URL <http://data.merenlab.org/sar11-saavs> serves an interactive web page to investigate the link between SAAVs and predicted protein structures.



Biochemistry



● Exposed ● Intermediate ● Buried ● NA

Data transparency and reproducibility



Upload My data



SAR11-TARA-METAPANGENOME

README.txt

S-LLPA-CORE-GENES.txt

SAR11-GENOMES.h5

SAR11-PAN-PAN.db

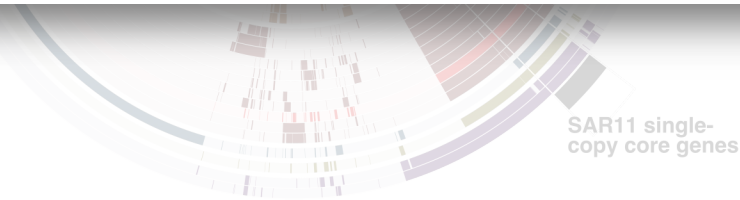
SAR11-PAN-SAMPLES.db

SAR11-TARA-METAPANGENOME.tar.gz (12.32 MB) MD5: 868af86d1b888d12650940a0012253fb

Cite

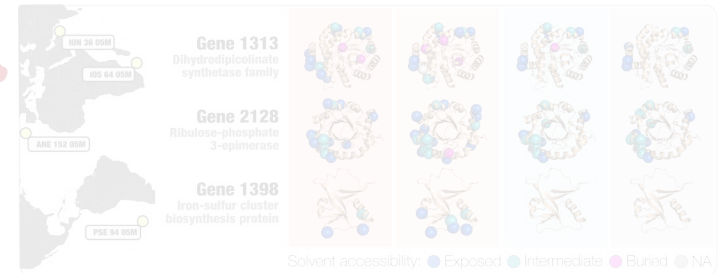
Download (12.32 MB)

Share



Pangenomics
Phylogenetics
Phylogenomics
Metagenomics
Population genetics

In silico protein biochemistry



Data transparency and reproducibility



Upload My data



SAR11-TARA-METAPAN

README.txt

S-LLPA-CORE-GENES

SAR11-GENOMES.b5

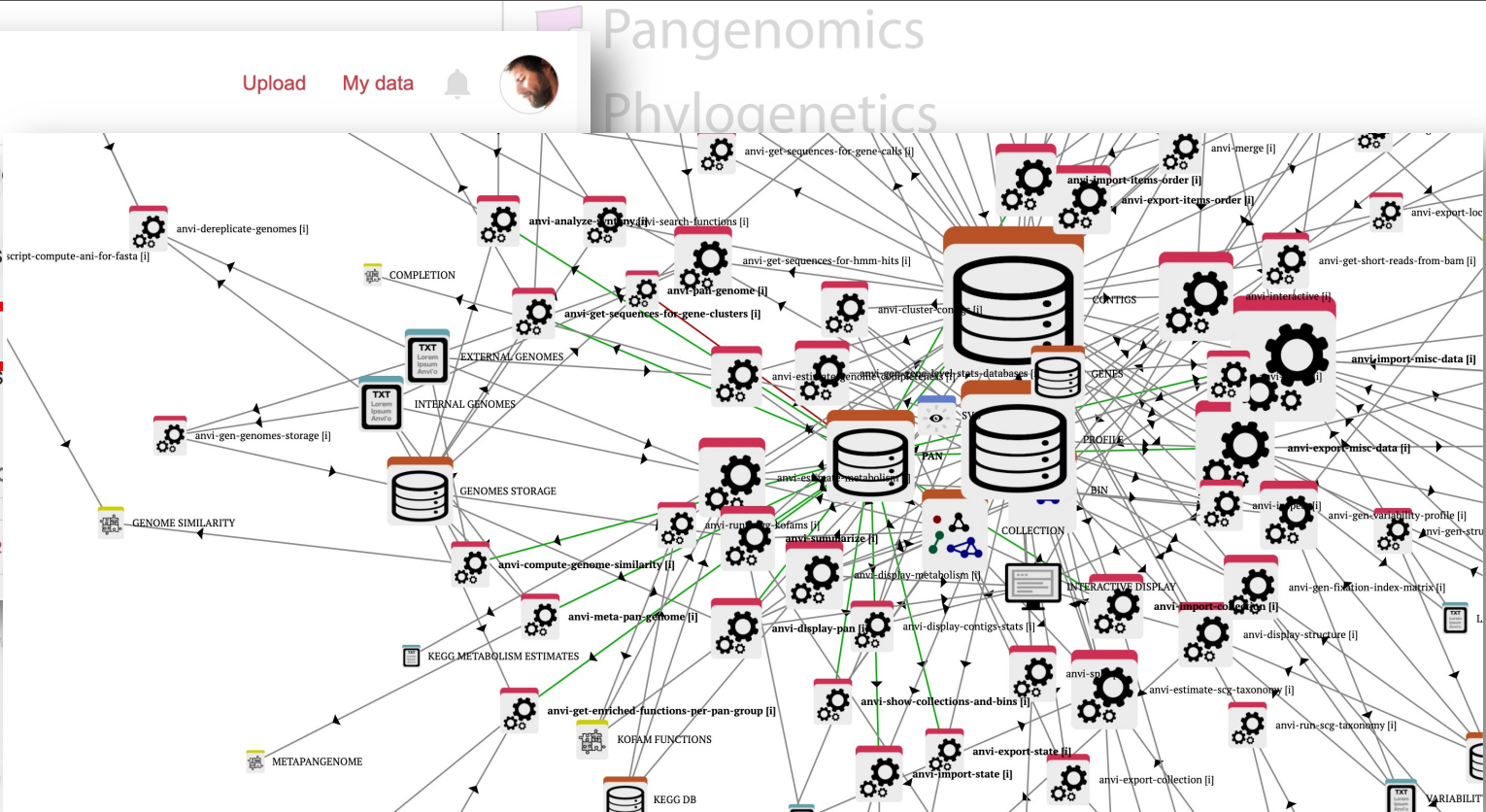
SAR11-PAN-PAN.db

SAR11-PAN-SAMPLES

SAR11-TARA-METAPANENG

Cite

Download (12)



Pangenomics
Phylogenetics

Data transparency and reproducibility



Upload My data



SAR11-TARA-METAPAN

README.txt

S-LLPA-CORE-GENES

SAR11-GENOMES.h5

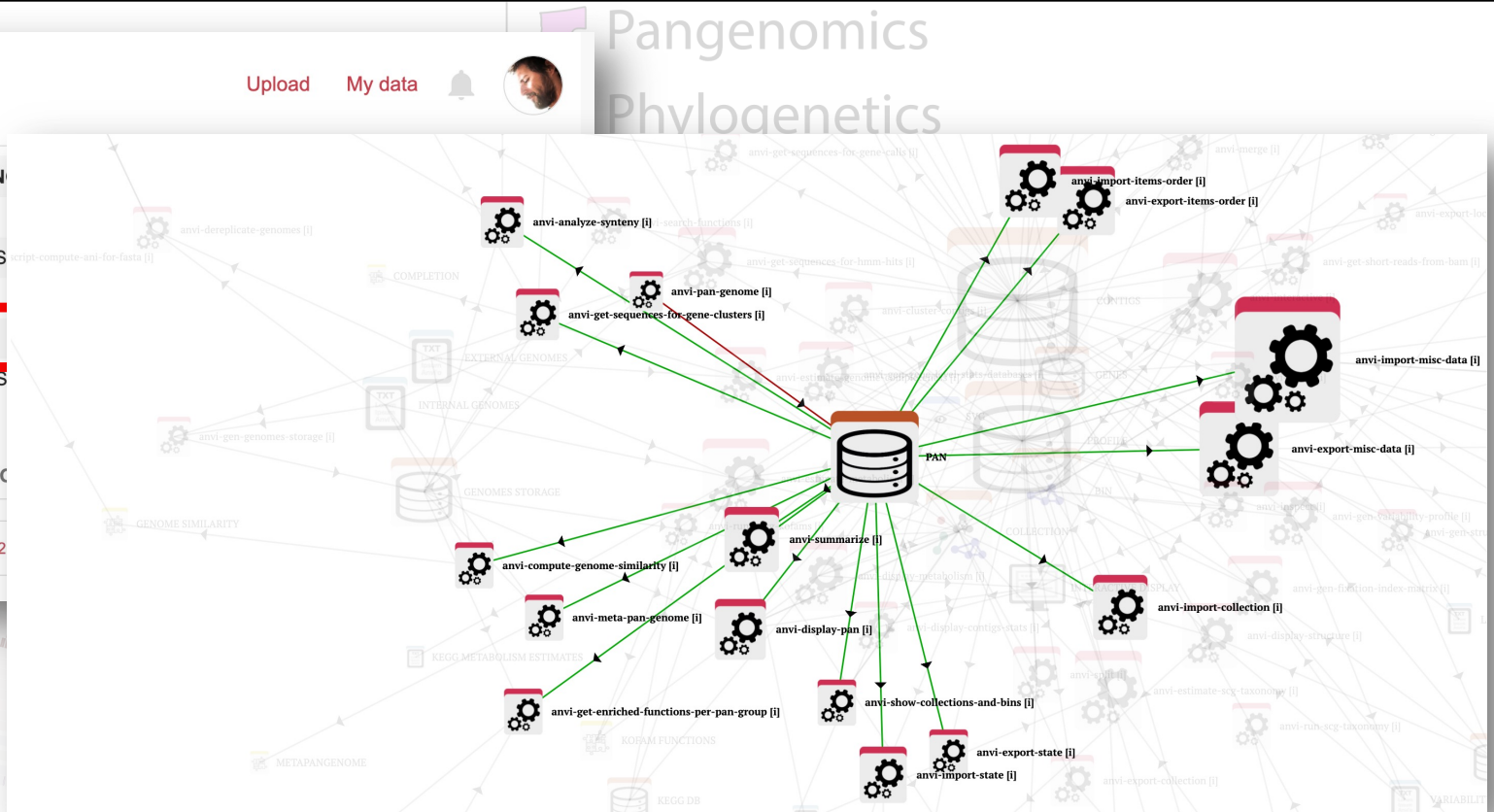
SAR11-PAN-PAN.db

SAR11-PAN-SAMPLES

SAR11-TARA-METAPANGENC

Cite

Download (12)



The new anvi'o help system



```
::: anvi'o v7 :: ~ >>> anvi-interactive --help
```

🔥 Program description:

Start an anvi'o server for the interactive interface

🍷 Can consume:

`profile-db / single-profile-db / contigs-db / genes-db / bin / view-data / dendrogram / phylogeny`

📁 Can provide:

`collection / bin / interactive / svg`

📖 More on `anvi-interactive`:

<https://merenlab.org/software/anvio/help/7/programs/anvi-interactive>

📦 All anvi'o programs and artifacts:

<https://merenlab.org/software/anvio/help/7>

```
::: anvi'o v7 :: ~ >>>
```


The new anvi'o help system



Anvi'o artifacts

Anvi'o artifacts represent **concepts, file types, or data types** anvi'o programs can work with. A given anvi'o artifact can be provided by the user (such as a FASTA file), produced by anvi'o (such as a profile database), or both (such as phylogenomic trees). Anvi'o artifacts link anvi'o programs to each other to build novel workflows.

Listed below a **total of 105 artifacts**.



[pan-db](#) [contigs-db](#) [trnaseq-db](#) [modules-db](#) [structure-db](#) [pdb-db](#) [kegg-data](#) [single-profile-db](#)
[profile-db](#) [genes-db](#) [genomes-storage-db](#)



[fasta](#) [contigs-fasta](#) [trnaseq-fasta](#) [concatenated-gene-alignment-fasta](#) [short-reads-fasta](#) [genes-fasta](#)
[locus-fasta](#)



[configuration-ini](#) [external-gene-calls](#) [protein-structure-txt](#) [samples-txt](#) [fasta-txt](#) [collection-txt](#)
[misc-data-items-txt](#) [misc-data-layers-txt](#) [misc-data-nucleotides-txt](#) [misc-data-amino-acids-txt](#)
[misc-data-layer-orders-txt](#) [misc-data-items-order-txt](#) [linkmers-txt](#) [gene-calls-txt](#) [binding-](#)
[frequencies-txt](#) [functions-txt](#) [functional-enrichment-txt](#) [view-data](#) [layer-taxonomy-txt](#) [gene-](#)
[taxonomy-txt](#) [genome-taxonomy-txt](#) [external-genomes](#) [internal-genomes](#) [metagenomes](#) [coverages-](#)
[txt](#) [detection-txt](#) [variability-profile-txt](#) [codon-frequencies-txt](#) [aa-frequencies-txt](#) [fixation-index-](#)
[matrix](#) [kegg-metabolism](#) [augustus-gene-calls](#) [vcf](#) [blast-table](#) [splits-txt](#) [genbank-file](#) [groups-txt](#)
[splits-taxonomy-txt](#) [hmm-hits-matrix-txt](#) [clustering-configuration](#)



[bam-file](#) [raw-bam-file](#)



[contigs-stats](#) [genes-stats](#)



[svg](#)



[bin](#)



[collection](#)



[hmm-source](#)

[hmm-hits](#) [completion](#) [misc-data-items](#) [misc-data-layers](#) [misc-data-nucleotides](#) [misc-data-amino-](#)
[acids](#) [genome-similarity](#) [misc-data-layer-orders](#) [misc-data-items-order](#) [metapangenome](#) [oligotypes](#)
[functions](#) [kegg-functions](#) [layer-taxonomy](#) [gene-taxonomy](#) [genome-taxonomy](#) [kegg-taxonomy](#) [db](#)

Anvi'o programs

Anvi'o programs perform atomic tasks that can be weaved together to implement complete 'omics workflows. Please note that there may be programs that are not listed on this page. You can type 'anvi-' in your terminal, and press the TAB key twice to see the full list of programs available to you on your system, and type `anvi-program-name --help` to read the full list of command line options.

Listed below a **total of 117 programs**.

anvi-analyze-synteny. Extract ngrams, as in 'co-occurring genes in synteny', from genomes.

[genomes-storage-db](#) [functions](#) [pan-db](#)

[ngrams](#)

anvi-cluster-contigs. A program to cluster items in a merged anvi'o profile using automatic binning algorithms.

[profile-db](#) [contigs-db](#) [collection](#)

[collection](#) [bin](#)

anvi-compute-completeness. A script to generate completeness info for a given list of *splits*.

[contigs-db](#) [splits-txt](#) [hmm-source](#)

anvi-compute-functional-enrichment. This is a driver program for `anvi-script-enrichment-stats`, a script that computes enrichment scores and group associations for annotated entities (ie, functions, KEGG Modules) across groups of genomes or samples..

[kegg-metabolism](#) [groups-txt](#) [misc-data-layers](#) [pan-db](#) [genomes-storage-db](#) [external-](#)
[genomes](#) [internal-genomes](#)

[functional-enrichment-txt](#)

anvi-compute-gene-cluster-homogeneity. Compute homogeneity for gene clusters.

[pan-db](#) [genomes-storage-db](#)


The new anvi'o help system




Anvi'o artifacts

Anvi'o artifacts represent **concepts, file types, or data types** anvi'o programs can work with. A given anvi'o artifact can be provided by the user (such as a FASTA file), produced by anvi'o (such as a profile database), or both (such as phylogenomic trees). Anvi'o artifacts link anvi'o programs to each other to build novel workflows.

Listed below a **total of 105 artifacts**.

 [pan-db](#) [contigs-db](#) [tnaseq-db](#) [modules-db](#) [structure-db](#) [pdb-db](#)
[kegg-data](#) [single-profile-db](#) [profile-db](#) [genes-db](#) [genomes-storage-db](#)

 [fasta](#) [contigs-fasta](#) [tnaseq-fasta](#) [concatenated-gene-alignment-fasta](#) [short-reads-fasta](#) [genes-fasta](#) [locus-fasta](#)

[configuration-ini](#) [external-gene-calls](#) [protein-structure-txt](#) [samples-txt](#) [fasta-txt](#) [collection-txt](#) [misc-data-items-txt](#) [misc-data-layers-txt](#) [misc-data-nucleotides-txt](#) [misc-data-amino-acids-txt](#) [misc-data-layer-orders-txt](#) [misc-data-items-order-txt](#) [linkmers-txt](#) [gene-calls-txt](#) [binding-frequencies-txt](#) [functions-txt](#) [functional-enrichment-txt](#) [view-data](#) [layer-taxonomy-txt](#) [gene-taxonomy-txt](#) [genome-taxonomy-txt](#) [external-genomes](#) [internal-genomes](#) [metagenomes](#) [coverages-txt](#) [detection-txt](#) [variability-profile-txt](#) [codon-frequencies-txt](#) [aa-frequencies-txt](#) [fixation-index-matrix](#) [kegg-metabolism](#) [augustus-gene-calls](#) [vcf](#) [blast-table](#) [splits-txt](#) [genbank-file](#) [groups-txt](#) [splits-taxonomy-txt](#) [hmm-hits-matrix-txt](#) [clustering-configuration](#)

 [bam-file](#) [raw-bam-file](#)

 [contigs-stats](#) [genes-stats](#)

Required or used by

[anvi-analyze-synteny](#) [anvi-compute-functional-enrichment](#) [anvi-compute-gene-cluster-homogeneity](#) [anvi-compute-genome-similarity](#) [anvi-db-info](#) [anvi-delete-misc-data](#) [anvi-delete-state](#) [anvi-display-pan](#) [anvi-export-items-order](#) [anvi-export-misc-data](#) [anvi-export-state](#) [anvi-get-sequences-for-gene-clusters](#) [anvi-import-collection](#) [anvi-import-items-order](#) [anvi-import-misc-data](#) [anvi-import-state](#) [anvi-merge-bins](#) [anvi-meta-pan-genome](#) [anvi-show-collections-and-bins](#) [anvi-show-misc-data](#) [anvi-split](#) [anvi-summarize](#) [anvi-update-db-description](#) [anvi-script-add-default-collection](#)

Description

A pan-db is an anvi'o database that contains **key information associated with your gene clusters**. This is vital for its pangenomic analysis, hence the name. If you want to learn more about the pangenomic workflow in Anvi'o, it has its own tutorial here.

This is the output of the program [anvi-pan-genome](#), which can be run after you've created a [genomes-storage-db](#) with the genomes you want to analyze. That script does the brunt of the pangenomic analysis; it calculates the similarity between all of the genes in your genomes-storage-db, clusters them and organizes the final clusters. All of the results of that analysis are stored in a pan-db.

You can use a pan database to run a variety of pangenomic analyses, including [anvi-compute-genome-similarity](#), [anvi-analyze-synteny](#), and [anvi-compute-functional-enrichment](#). You can also view and interact with the data in a pan-db using [anvi-display-pan](#).

To add additional information to the pangenome display, you'll probably want to use [anvi-import-misc-data](#)

The new anvi'o help system



anvi-db-info [program]

Access self tables, display values

See [program help menu](#) or go

Usage

Displays information about an anvi'o database

This program is particularly useful for debugging a database - to answer questions like "did I run the database correctly?" It can also be very dangerous when used to modify a database, so please proceed with caution.

What information will I see?

All anvi'o databases contain a table of self-critical facts such as the type of the database, how the database was generated, what sort of table exists so that anvi'o can make sure `anvi-db-info` will show you the contents of

The information in the self table will be different. The `self` table will indicate the number of contigs (what gene callers), and which functional annotations it contains mapping information. A `modules-db` (see [modules-db](#)) database and what is the hash value of the c

Modifying database information

We just need to start by saying - you probably shouldn't do this. Manually changing the values in the self table has the potential to break things downstream because it lets you avoid some of anvi'o's internal sanity checks which prevent you from doing things you shouldn't. If you change things and start running into ugly errors, do not be surprised.

That being said, sometimes you just need to live on the edge and do some hacking, and `anvi-db-info` will let you do that. If a programmer sent you here to update a value in the self table or if you are just foraging ahead on your own, this is how you would do it. Let's change the `project_name` value as an example because it is mostly descriptive and seems fairly safe:

```
anvi-db-info --self-key project_name --self-value "test" CONTIGS.db
```

If you run this, you will see a warning telling you what the current value of `project_name` is and what it will be changed to, but the value will not actually be changed just yet. If you are sure you want to do this, you then need to run:

```
anvi-db-info --self-key project_name --self-value "test" CONTIGS.db --just-do-it
```

Omics

Thanks!