

Category	Subcategory	Scripts*	Brief description	Remote access**	Input***	Output***	
Sequence similarity search	Statistics	BlastTab.seqdepth_ZIP.pl	Average sequencing depth of reference sequences corrected by Zero-Inflated Poisson.	-	BLAST, FastA	TSV	
		BlastTab.advance.bash	Advance of a partial sequence similarity search result.	-	BLAST, FastA	Text	
		BlastTab.sumPerHit.pl	Weighted sums per reference sequence.	-	BLAST, TSV	TSV	
		BlastTab.seqdepth.pl	Average and median sequencing depth of reference sequences.	-	BLAST, FastA	TSV	
		BlastTab.seqdepth_nomedian.pl	Average sequencing depth of reference sequences with small RAM footprint.	-	BLAST, FastA	TSV	
		BlastPairwise.AAsubs.pl	Counts the different AA substitutions in the best hit.	-	BLAST-Pairwise	TSV	
		BlastTab.recplot2.R	Recruitment plots.	-	BLAST, TSV	Graphic, R:enve.recplot2	
	RecPlot2.find_peaks.R	Identifies peaks in the population histogram potentially identifying subpopulation mixtures.	-	R:enve.recplot2	R:enve.recplot2.findPeaks		
	Manipulation	BlastTab.topHits_sorted.rb	Reports top-N hits (requires sorted queries to reduce RAM footprint).	-	BLAST	BLAST	
		BlastTab.cogCat.rb	Replaces identifiers by categories from COG (Cluster of Orthologous Groups of proteins).	-	BLAST	BLAST	
		BlastTab.filter.pl	Extracts a subset of hits filtering by query or subject.	-	BLAST, TSV	BLAST	
		BlastTab.addlen.rb	Appends extra columns to a similarity search result with the lengths of the sequences.	-	BLAST, FastA	BLAST	
		BlastTab.catsbj.pl	Generates a list of hits from a similarity search concatenating the subject sequences.	-	BLAST, FastA	TSV	
		BlastTab.pairedHits.rb	Identifies hits of paired-reads.	-	BLAST	TSV	
		BlastTab.subsample.pl	Extracts hits from a set of query sequences.	-	BLAST, FastA	BLAST	
		BlastTab.best_hit_sorted.pl	Reports best hits (requires sorted queries to reduce RAM footprint).	-	BLAST	BLAST	
		BlastTab.taxid2taxrank.pl	Replaces NCBI taxonomy IDs by names at a given taxonomic rank.	-	BLAST	BLAST	
		BlastTab.kegg_peg2path_rest.pl	Replaces peptide identifiers by pathways from KEGG.	KEGG	BLAST	BLAST	
	BlastTab.kegg_peg2ko_rest.pl	Replaces peptide identifiers by KOs from KEGG.	KEGG	BLAST	BLAST		
	Execution	aa_i.rb	Calculates the Average Amino acid Identity between two genomes.	NCBI	FastA	Text, TSV, SQLite3	
		ani.rb	Calculates the Average Nucleotide Identity between two genomes.	NCBI	FastA	Text, TSV, SQLite3	
		rbm.rb	Finds the reciprocal best matches between two sets of sequences.	-	FastA	Text, TSV	
	Sequence analyses	Statistics	FastA.gc.pl	Calculates the G+C content of a set of sequences.	-	FastA	TSV
			FastA.length.pl	Calculates the length of a set of sequences.	-	FastA	TSV
FastA.qlen.pl			Calculates the quartiles of the length in a set of sequences.	-	FastA	Text	
FastA.N50.pl			Calculates the N50 (or any other N*) value of a set of sequences.	-	FastA	Text	
FastQ.toFastA.awk			Translates a FastQ file into FastA.	-	FastQ	FastA	
FastA.tag.rb		Easy-to-parse tagged sequences from FastA files.	-	FastA	FastA		
FastQ.tag.rb		Easy-to-parse tagged sequences from FastQ files.	-	FastQ	FastQ		
FastA.rename.pl		Renames a set of sequences in FastA format.	-	FastA, TSV	FastA		
FastQ.split.pl		Splits a large FastQ file into several evenly-distributed smaller FastQ files.	-	FastQ	FastQ		
FastA.revcom.pl		Reverse-complements.	-	FastA	FastA		
FastQ.filter.pl		Extracts a subset of sequences.	-	FastQ	FastQ		
FastA.subsample.pl		Subsamples a set of sequences.	-	FastA	FastA		
FastQ.offset.pl		Generates a FastQ with a given offset for quality.	-	FastQ	FastQ		
FastA.filter.pl		Extracts a subset of sequences.	-	FastA	FastA		
FastA.split.pl		Splits a large FastA file into several evenly-distributed smaller FastA files.	-	FastA	FastA		
FastA.slider.pl		Slices one or more input sequence into several shorter sequences.	-	FastA	FastA		
FastA.filterN.pl		Filters out sequences with large portions of N's and/or homopolymers.	-	FastA	FastA		
FastA.filterLen.pl		Filters a multi-FastA file by length.	-	FastA	FastA		
FastQ.interpose.pl		Interposes sequences from two files into one output file.	-	FastQ	FastQ		
FastA.interpose.pl		Interposes sequences from two files into one output file.	-	FastA	FastA		
Diversity		OTUs	AlphaDiversity.pl	Estimates Rao, Rao-Jost, Shannon, Inverse Simpson, and qD indices of diversity.	-	TSV	TSV
			Chao1.pl	Estimates Chao1 index with 95% confidence interval.	-	TSV	Text
			Table.barplot.R	Creates nice barplots from tab-delimited tables.	-	TSV	PDF
Annotation		Database mapping	BlastTab.taxid2taxrank.pl	Replaces NCBI taxonomy IDs by names at a given taxonomic rank.	-	BLAST	BLAST
	EBIseq2tax.rb		Maps EBI-supported IDs to NCBI taxonomy.	EBI, NCBI	CLI or TSV	TSV	
	RefSeq.download.bash		Downloads a collection of assemblies (or annotations) from NCBI's RefSeq.	NCBI-FTP	CLI	FastA, ...	
	SRA.download.bash		Downloads the set of runs from a project, sample, or experiment in SRA.	EBI	CLI	FastQ	
	gi2Tax.rb		Maps NCBI GIs to NCBI taxonomy.	NCBI	CLI or TSV	TSV	
	M5nr.getSequences.rb		Downloads a set of sequences from M5nr with a given functional annotation.	M5nr	CLI	FastA	
	BlastTab.kegg_peg2path_rest.pl		Replaces peptide identifiers by pathways from KEGG.	KEGG	BLAST	BLAST	
	BlastTab.kegg_peg2ko_rest.pl		Replaces peptide identifiers by KOs from KEGG.	KEGG	BLAST	BLAST	
	GenBank.add_fields.rb		Adds annotations to GenBank files.	-	GenBank, TSV	GenBank	
	MyTaxa.fragByTax.pl		Lists the fragments classified as a given taxon in MyTaxa.	-	MyTaxa	TSV	
	Tables	Table.split.pl	Splits a file with multiple columns into multiple two-columns lists.	-	TSV	TSV	
		Table.filter.pl	Extracts (and re-orders) a subset of rows from a raw table.	-	TSV	TSV	
		Table.round.rb	Rounds numbers in a table.	-	TSV	TSV	
		Table.merge.pl	Merges multiple (two-column) lists into one table.	-	TSV	TSV	
		Table.barplot.R	Creates nice barplots from tab-delimited tables.	-	TSV	PDF	
		Table.df2dist.R	Transforms a dataframe (or coercible object) into a dist object.	-	TSV	TSV	
		ogs.core-pan.rb	Rarefies core and pan-genome in a set of OGS.	-	OGs	TSV, JSON	
		ogs.stats.rb	Estimates some descriptive statistics on a set of OGS.	-	OGs	TSV, JSON	
		HMM.essential.rb	Finds and extracts a collection of essential proteins in Bacteria and Archaea.	-	FastA	TSV, FastA, Text	
		ogs.mcl.rb	Identifies OGS in RBMs between all pairs in a collection of genomes.	-	BLAST	OGs	
	Search	ogs.extract.rb	Extracts sequences of OGS from genomes.	-	OGs, FastA	FastA	
		ogs.annotate.rb	Annotates OGS using one or more reference genomes.	-	OGs, TSV	OGs	
		HMMsearch.extractids.rb	Extracts the sequence IDs and query model form a (multiple) HMMsearch report.	-	HMMsearch	TSV	
		CharTable.classify.rb	Uses a dichotomous key to classify objects parsing a character table.	-	TSV	TSV, Newick	
JPlace.to_jToL.R		Generates jToL-compatible files from a JPlace file.	-	JPlace	TSV, Newick		
JPlace.distances.rb		Extracts the distance of each placed read to a given node.	-	JPlace	TSV		
Newick.prune.R		Automatically prunes a tree, to keep representatives of each clade.	-	Newick	Newick		
Table.df2dist.R		Transforms a dataframe (or coercible object) into a dist object.	-	TSV	TSV		
Other data	Phylogenetic and other distances	TRIBS.test.R	Estimates the empirical difference between the distances in a set of objects and a subset.	-	TSV	R:enve.TRIBStest, Text	
		TRIBS.plot-test.R	Plots an enve.TRIBStest object as distance distributions or the subtraction convolution.	-	R:enve.TRIBStest	Graphic	
		EBIseq2tax.rb	Maps EBI-supported IDs to NCBI taxonomy.	EBI, NCBI	CLI or TSV	TSV	
		Table.barplot.R	Creates nice barplots from tab-delimited tables.	-	TSV	PDF	
		CharTable.classify.rb	Uses a dichotomous key to classify objects parsing a character table.	-	TSV	TSV, Newick	
		gi2Tax.rb	Maps NCBI GIs to NCBI taxonomy.	NCBI	CLI or TSV	TSV	
	Taxonomic	MyTaxa.fragByTax.pl	Lists the fragments classified as a given taxon in MyTaxa.	-	MyTaxa	TSV	
		MyTaxa.seq-taxrank.rb	Generates a simple tabular file from a MyTaxa classification.	-	MyTaxa	TSV	
		Taxonomy.silva2ncbi.rb	Re-formats Silva taxonomy into NCBI-like taxonomy dump files.	-	TSV	TSV, ...	
		Aln.cat.rb	Concatenates several multiple alignments into a single multiple alignment.	-	FastA	FastA, TSV	
		AAsubs.log2ratio.rb	Estimates the log2-ratio of different amino acids in homologous sites.	-	TSV	TSV	
		BlastPairwise.AAsubs.pl	Counts the different AA substitutions in the best hit.	-	BLAST-Pairwise	TSV	
	Alignments	Aln.convert.pl	Pipe-able wrapper for BioPerl's alignment format conversion.	-	...	...	

\* Entries prefixed by "R:" are R functions available at enveomics.R, all other entries are scripts in languages corresponding to the extension. Some entries can be duplicated in different classifications.

\*\* Access to remote services. - indicates no remote access. NCBI-FTP refers to the public FTP server of NCBI. All other values refer to providers of HTTP RESTful APIs (NCBI, EBI, KEGG, M5nr).

\*\*\* Text and Graphic refer to unstructured human-readable values and CLI (Command-Line Interface) refers to short text values. TSV (Tab-Separated Values) refers to tables in raw text with columns separated by tabulations. OGs refers to a particular TSV format for gene IDs organized in Orthologous Groups: genomes in columns and groups in rows, a header with genome names, multiple values (in-paralogs) separated by comma (,), missing values represented by dash (-), and optional comments from sharp (#) to the end of the line. JSON and SQLite3 refer to ad-hoc structures in those respective specifications. BLAST corresponds to the tabular BLAST output (-m 8 in NCBI-BLAST, --outfmt 6 in NCBI-BLAST+) and BLAST-Pairwise to the pairwise output (-m 0 / --outfmt 0). Jplace corresponds to the standard described by Matset et al (2012, PLoS ONE 7(2): e31009). All other values are standard structured formats (FastA, FastQ, Newick, GenBank) or the standard output of external Software (MyTaxa, HMMsearch).