

Publication	SeqQual pipeline / program (prog) name	usage*	arguments (arg)	description	using programs	Shell examples (with print_source perl scripts)*	Input file(s)	Output file(s)	new folder created	name of new folder created	Comments	folder where program files should be located**
ace files dealing and split												
Brousseau et al. 2014	print_source-aceonly.pl	perl prog inputfile > source-acedealing.txt	inputfile is the text file listing folders (one or more) by their names where all assembly files *.ace are located	Produce a text file to source in which the checkdir.pl program is being called and the folder structure organized.			text file names inputfile with folder list	text file to source	yes	edit_dir		~/SeqQual
Brousseau et al. 2014	checkdir.pl	source source- acedealing.txt		Needs first print_source-aceonly.pl to run as it is launched from the source-acedealing.txt file above, assumes program is located under the ~/SeqQual/ folder (this can be changed in the code). checks first if a folder mydata exists. If yes, renames it & creates a new one. Looks for assembly files *.ace in the folder or list of folders located below the current folder, and for each ace file, copies them into an edit_dir folder			see description	see description	no	NA		~/SeqQual
El mujtar et al. 2014	change454ace1.pl	perl prog arg	original ace file name	Produces preliminary ace file for input into the program change454ace2.pl to produce a more recent ace file format.			*.ace file(s)	*.ace file	no	NA		~/SeqQual
El mujtar et al. 2014	change454ace2.pl	perl prog arg	preliminary ace file(name) created by change454ace1.pl	Produces final ace file which marks "poorly aligned" nucleotides as "N". "poorly aligned " means that the nucleotides are located outside the number in "QA" line.			*.ace file(s)	*.ace file	no	NA		~/SeqQual
El mujtar et al. 2014	qual2ball_mod.pl	perl prog arg	name.fasta (see comments)	Produces phdball file from fasta.qual files.			*.qual file(s)	*.phd1.ball file	no	NA	Need both "name.fasta" and "name.fasta.qual"/ modified from original qual2ball program by Tony Travis (http://www.mendeley.com/profiles/tony-travis/)	~/SeqQual
El mujtar et al. 2014	separate454ace.pl	perl prog arg	name of ace file	Produces many small ace files from a big ace file. Each ace file only contains one alignment.			*.ace file(s)	many *.ace files	no	NA		~/SeqQual
El mujtar et al. 2014	separate454phd.pl	perl prog arg	name of phdball file	Produces many small phd files from a big phdball file. Each phd file only contains the information of one read.			*.phd1.ball file	many *.phd1 files	no	NA		~/SeqQual
Brousseau et al. 2014, El mujtar et al. 2014	print_source-write_aln- aceonly.pl	perl prog inputfile > source-write_aln.txt	inputfile is the text file listing folders (one or more) by their names where all assembly files *.ace are located	Produces a text file to source in which the write_acealn- nonqual_multinput.pl program is being called.			text file named inputfile with folder list	text file with command lines to source	yes	aln_final		~/SeqQual
Brousseau et al. 2014, El mujtar et al. 2014	write_acealn- nonqual_multinput.pl	perl prog	none	Needs print_source-write_aln-aceonly.pl to run first. Produces multiple fasta files for each ace file without combining the information from phd file in a new folder. Name them "OriginalSuffix_nonqual.aln".			*.ace file(s)	fasta *.aln files (with ? Instead of *.-" or blanks at start & ends of reads not to confuse with true deletions)	yes (no in Brousseau, revolir) aln (check that)			~/SeqQual
El mujtar et al. 2014	write_acealn- qual_multinput.pl	perl prog arg	phred score	Produces alignment files for each ace file with combining the information from phd file in a new folder. Name them "OriginalSuffix_phred*.aln". (* here means phred score nb)			*.ace file(s)	*.aln fasta files	yes	aln		~/SeqQual
El mujtar et al. 2014	combine_aln_byace.pl	perl prog	none	Produces new alignments files in a new folder combining information from files created by write_acealn- nonqual_multinput.pl and write_acealn-qual_multinput.pl, mainly based on the sequence in the ace file. Name the alignment files "OriginalSuffix_combined.aln".			*.aln fasta files, *.ace files	*.aln fasta files	yes	aln_combined	Looks for *nonqual.aln (created by write_acealn- nonqual_multinput.pl) and *phred*.aln (created by write_acealn-qual_multinput.pl). If the name before nonqual.aln is the same as the name before phred*.aln, those two files will be combined to create a new file.	~/SeqQual
El mujtar et al. 2014	combine_aln_byphd.pl	perl prog	none	Produces new alignments files in a new folder combining information from files created by write_acealn- nonqual_multinput.pl and write_acealn-qual_multinput.pl, mainly based on the sequence in the phd file. Name the alignment files "OriginalSuffix_combined.aln".			*.aln fasta files, *.phd1 files	*.aln fasta files	yes	aln_combined	Looks for *nonqual.aln (created by write_acealn- nonqual_multinput.pl) and *phred*.aln (created by write_acealn-qual_multinput.pl). If the name before nonqual.aln is the same as the name before phred*.aln, those two files will be combined to create a new file.	~/SeqQual

*
**

for all scripts, see example shell column
most program locations are assumed to be under home/SeqQual but this can easily be changed in the code
all programs can be used separately with your own batch command lines on fasta files (defined as above)