ace files dealing and split

Reference	SeqQual pipeline / program (prog) name*,¤	usage**	arguments (arg)	description	using programs (other prog)	Shell examples **	Input file(s)	Output folder/file(s)	log file available $
Brousseau et al. 2014	print_source-aceonly.pl	(perl prog arg > out source out) = (A)	(text file with list of folder names, e.g. inputfile) = (B)	(run other prog in batch on a set of folder given as arg , when sourcing the "out" text file) = (C) , creates edit_dir folder and moves *.ace files to it	checkdir_mydata.pl	2.1-ace-only.sh	text file listing the names of one of more folders containing *.ace files	edit_dir/*.ace	X
Brousseau et al. 2014	checkdir_mydata.pl	perl prog		create new folder mydata, if one already exists--> change it to mydata_current_date&time first				empty folder mydata	
"	print_source-ace_changename.pl	(A)	(B)	(C)	change_ace_name.pl	2.3-ace-qual.sh			X
Brousseau et al. 2014	change_ace_name.pl	perl prog		change file with extension *.ace.1 if any to file with extension *.ace in the working folder					
El mujtar et al. 2014	changeMIRAace1.pl	perl prog arg	original ace file name from MIRA with particular features	Produces intermediate ace file for input into the program changeMIRAace2.pl		2.4-ace-qual2ball-combine.sh	*.ace file(s)	*.ace file	
El mujtar et al. 2014	changeMIRAace2.pl	perl prog arg	prelimimary ace file(name) created by changeMIRAace1.pl	Produces final ace file with regular format which marks "poorly aligned" nucleotides as "N". "poorly aligned " means that the nucleotides are located outside the number in the "QA" line.		2.4-ace-qual2ball-combine.sh	*.ace file(s)	*.ace file	
El mujtar et al. 2014	qual2ball_mod.pl	perl prog arg	name.fasta (without the qual extension, but both files needs to be in the working folder,see description)	Produces phdball file from fasta.qual files. (both "name.fasta" and "name.fasta.qual" are needed), modified from original qual2ball.pl script by Tony Travis (http://www.mendeley.com/profiles/tony-travis/)		2.4-ace-qual2ball-combine.sh	*.qual file(s)	*.phd1.ball file	
"	print_source-largeace_split.pl	(A)	(B)	(C) , move old ace file into oldace folder	split-largeace.pl	2.2-split+ace-only, 2.3-ace-qual.sh, 2.4-ace-qual2ball-combine.sh	*.ace file(s)	oldace	X
"	split-largeace.pl	perl prog arg	name of ace file	Split a large (full assembly) ace file in many *.ace files, each ace file only contains one alignment.			*.ace file(s)	many *.ace files	
"	print_source-largephd_split.pl	(A)	(B)	(C)	split-largephd.pl	2.2-split+ace-only, 2.3-ace-qual.sh, 2.4-ace-qual2ball-combine.sh			X
El mujtar et al. 2014	split-largephd.pl	perl prog arg	name of phdball file	Split a large phd1.ball file into many phd files, each phd file only contains the information of one read.			*.phd1.ball file	many *.phd1 files	
Brousseau et al. 2014, El mujtar et al. 2014	print_source-write_aln-aceonly.pl	(A)	(B)	(C) , write alignment(s) from standard ace formatted file, dependent on running print_source-aceonly.pl, moves fasta *.aln files to aln_final, workable for ace file with few thousands alignments, if more, use split-largeace.pl first	write_acealn-nonqual_multinput.pl	2.1-ace-only.sh	text file listing the names of one of more folders containing *.ace files	aln_final/*.aln fasta files	
Brousseau et al. 2014, El mujtar et al. 2014	write_acealn-nonqual_multinput.pl	perl prog	none	look for files with extension ".ace"; Produces multiple fasta files for each ace file without integrating information about quality. Name them "OriginalSuffix_nonqual.aln".			*.ace file(s)	aln/fasta *.aln files	
lang et al.	print_source-acedealing-qual.pl	(A)	(B)	create working directories and files; organize *.ace and *.phd files	checkdir_mydata.pl	2.3-ace-qual.sh , 2.4-ace-qual2ball-combine.sh	text file listing the names of one of more folders containing *.ace and *.phd files	phd_dir, edit_dir	X
"	print_source-write_aln-qual.pl	(A)	(B), phred score	(C) , moves edited fasta alignments into output folder	write_acealn-qual_multinput.pl	2.3-ace-qual, 2.4-ace-qual2ball-combine.sh	see arg, phd/poly files in same folder than ab1 files	aln_final/*.aln fasta files	X
El mujtar et al. 2014	write_acealn-qual_multinput.pl	perl prog arg	phred score	look for files with extension ".ace"; Produces alignment files for each ace file with combining the information from phd file in a new folder, i.e. mask nucleotides which quality is below or equal to the phred quality score given (but keeps it if it is the same than the consensus). Name them "OriginalSuffix_phred*.aln". (* here means phred score nb).		see 2.3-ace-qual.sh for the print_source script calling this one, 2.4-ace-qual2ball-combine.sh	*.ace file(s)	aln/*.aln fasta files	

| El mujtar et al. 2014 | combine_aln_byace.pl | perl prog | Looks for *nonqual.aln (created by write_acealn-nonqual_multinput.pl on the original ace assembly file) and *phred*.aln (created by write_acealn-qual_multinput.pl on both ace & phd files). If names match before extension, a new alignment is made from the *nonqual.aln which already contain the masked nucleotides and where the quality (i.e. "?" below the phd score) is integrated per nucleotide from the corresponding sequences of same length from the phred file. Name the alignment files "OriginalSuffix_combined.aln" and moves them to aln_combined folder | **2.4-ace-qual2ball-combine.sh** | *.aln fasta files, *.ace files | aln_combined/*. aln fasta files |