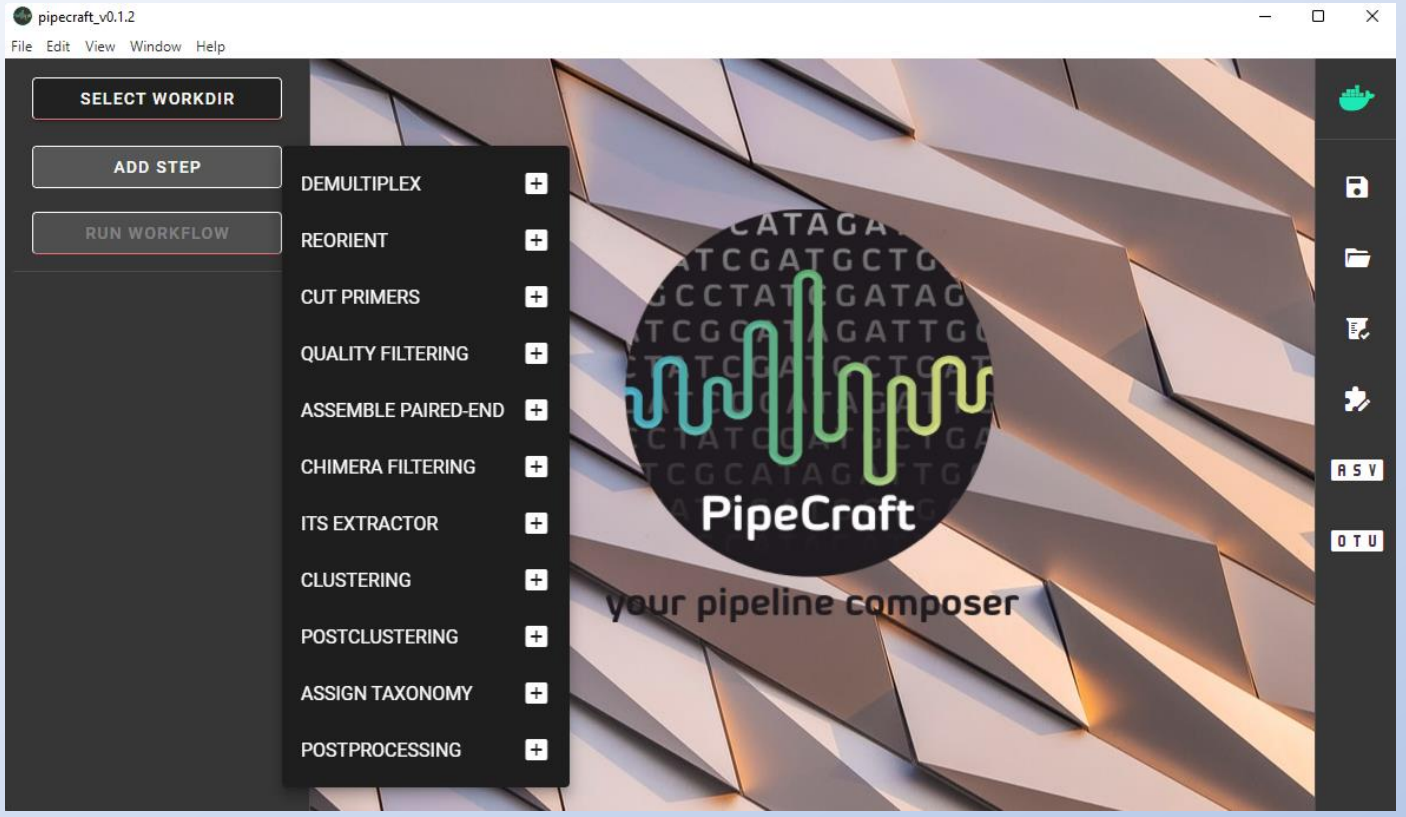
UT SummerSchool ‘Metabarcoding: from lab to bioinformatics’
03.08.2022



Metabarcoding pipeline steps

Sten Anslan <sten.anslan@ut.ee>

Received: 30 December 2016 | Revised: 24 April 2017 | Accepted: 18 May 2017
DOI: 10.1111/1755-0998.12692

RESOURCE ARTICLE

WILEY MOLECULAR ECOLOGY RESOURCES

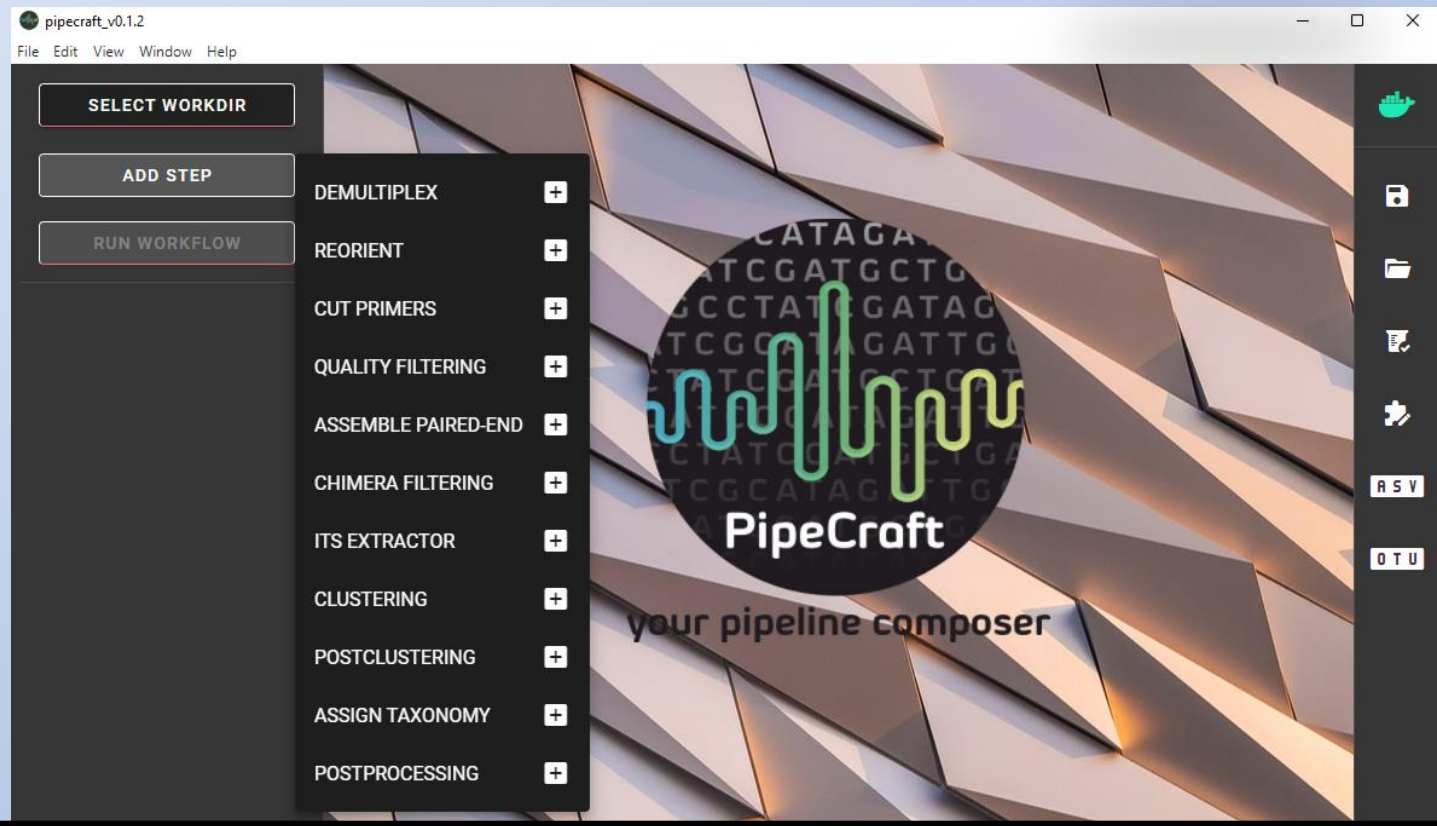
PipeCraft: Flexible open-source toolkit for bioinformatics analysis of custom high-throughput amplicon sequencing data

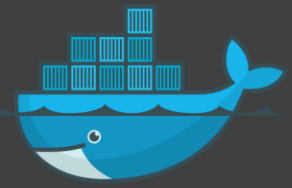
Sten Anslan[1] | Mohammad Bahram[1,2] | Indrek Hiiesalu[1] | Leho Tedersoo[3]

User Guide

https://pipecraft2-manual.readthedocs.io/en/stable/

Martin Metsoja





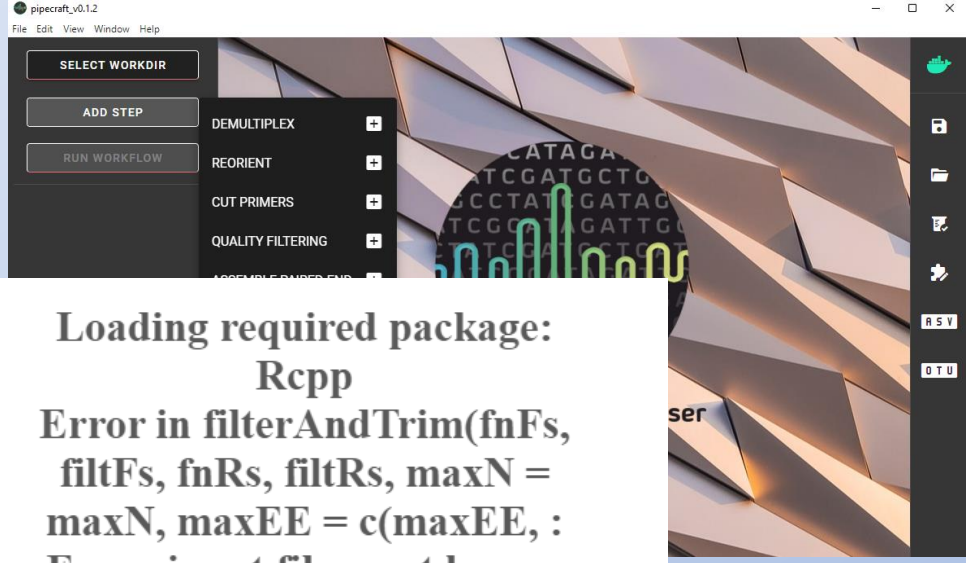
docker

Linux

your pipeline composer

Please report the bugs/errors

(contact details in User Guide)

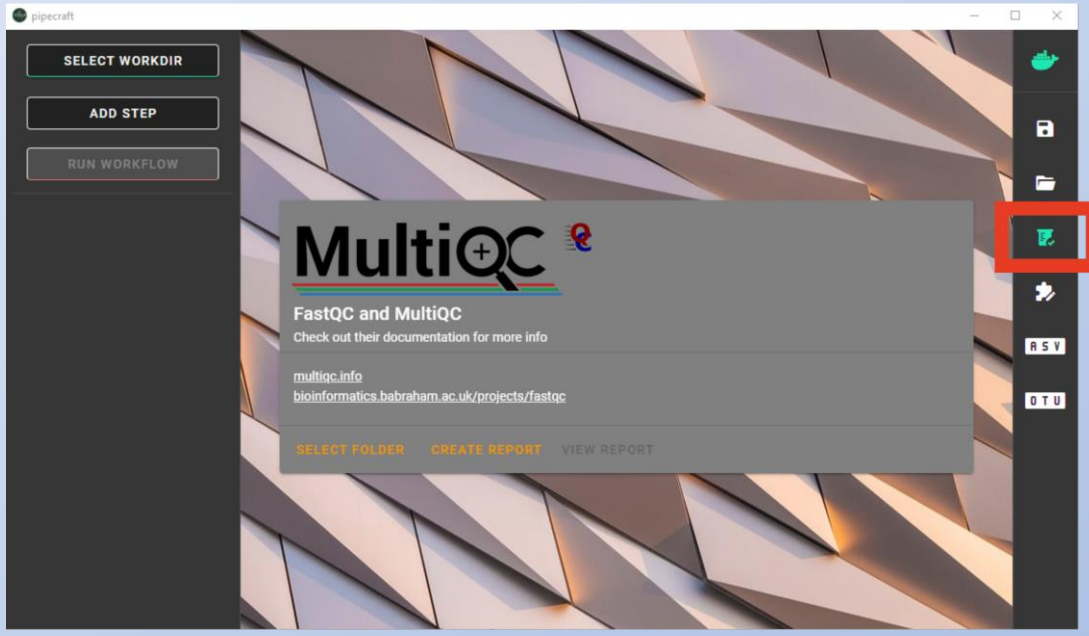


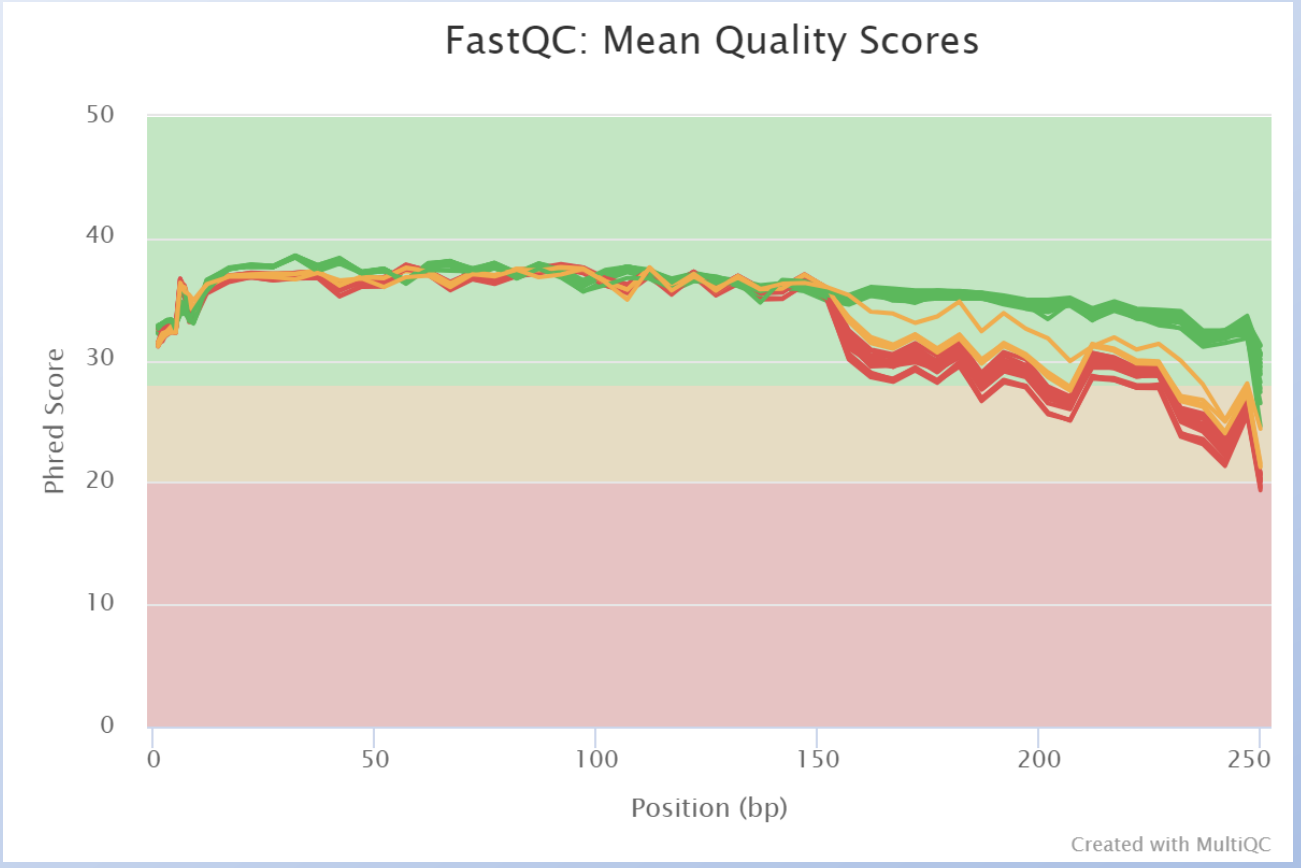
Workflow/software recommendatoins are welcome!


Demultiplexing
Reorient reads; cut primers
Assemble paired-end
Quality filtering
Chimera filtering
ITS extraction
Clustering
Assign taxonomy

Quality check

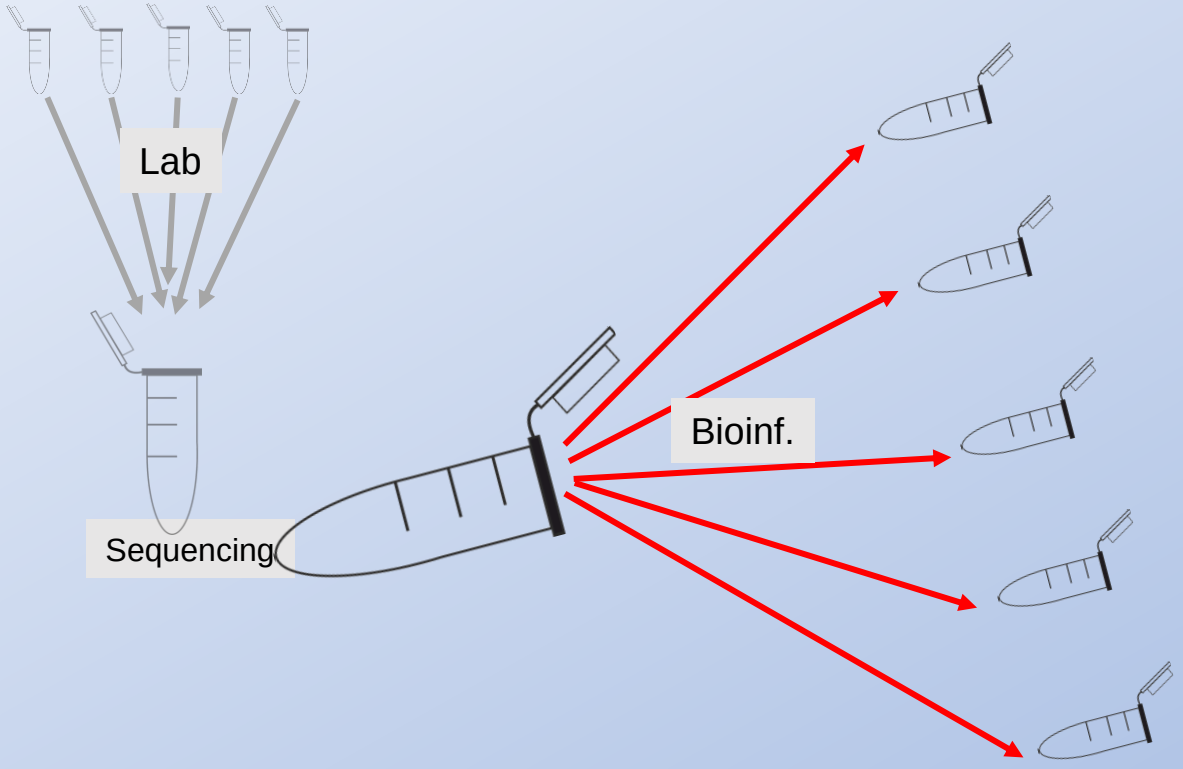
for fastq files

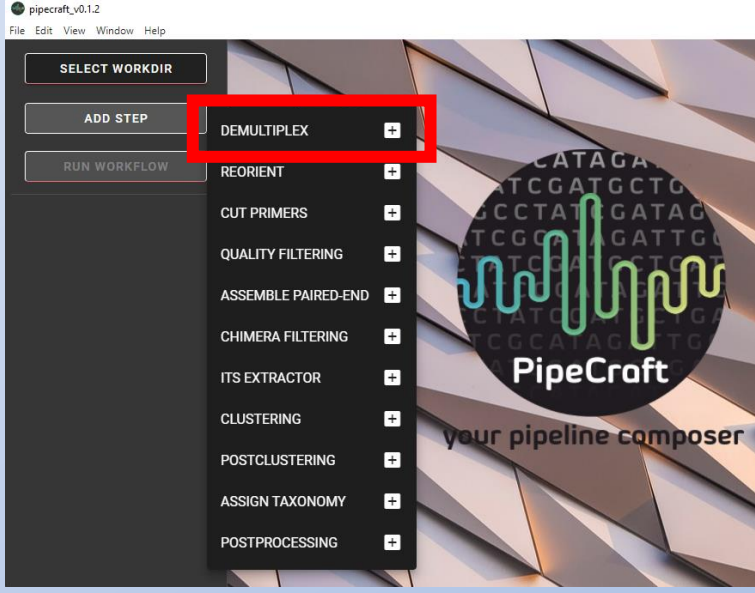




Demultiplexing

Pooled sequences to reads per sample

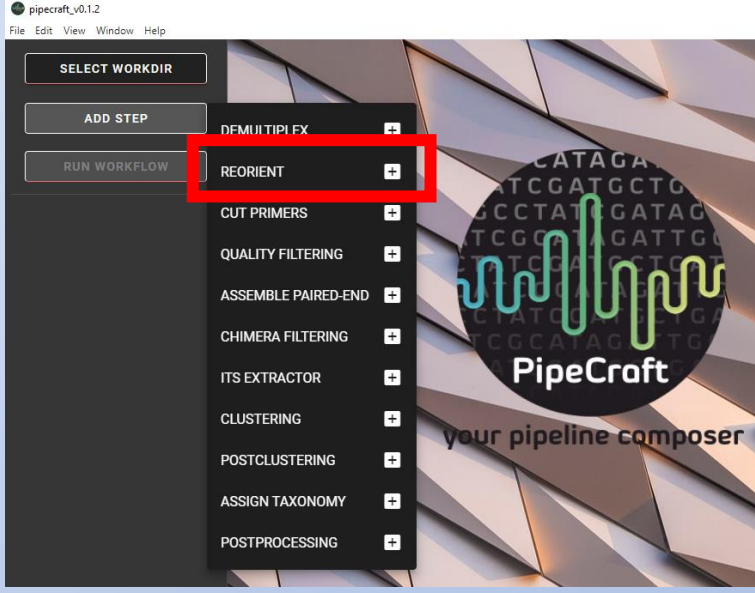




FwdTAG-FwdPRIMER-METABARCODE-RevPRIMER-RevTAG

Reorient reads

based on primer sequences



5’ FwdPRIMER-METABARCODE-RevPRIMER 3’

3’ RevPRIMER-rc.METABARCODE-FwdPRIMER 5’

5’ FwdPRIMER-METABARCODE-RevPRIMER 3’

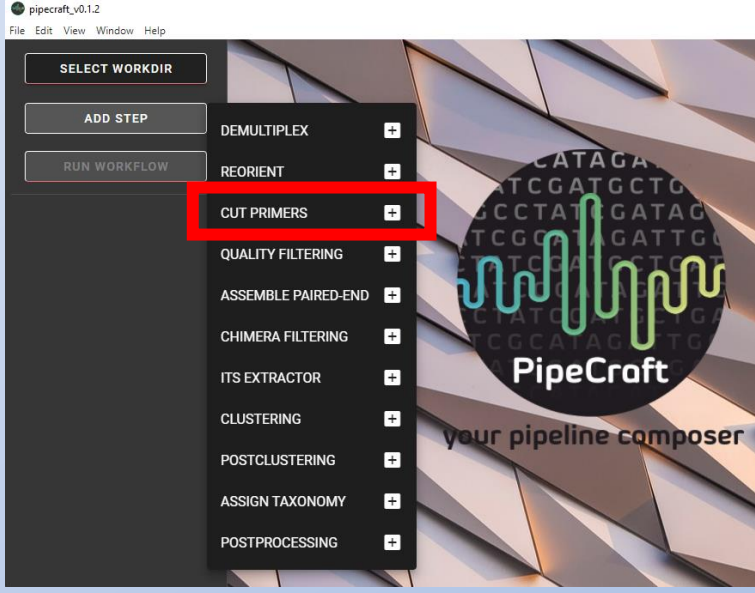
Identical sequences in different orientation may result in **different ASVs/OTUs**

Cut primers

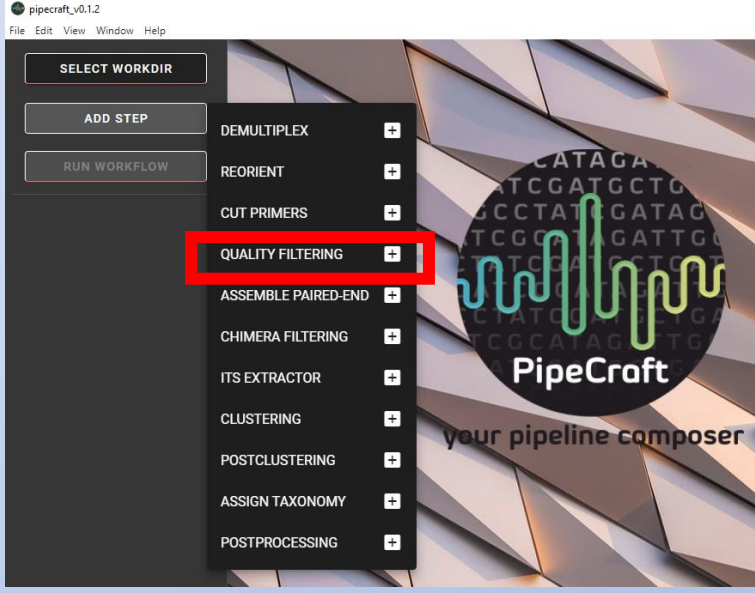
5’ FwdPRIMER-METABARCODE-RevPRIMER 3’

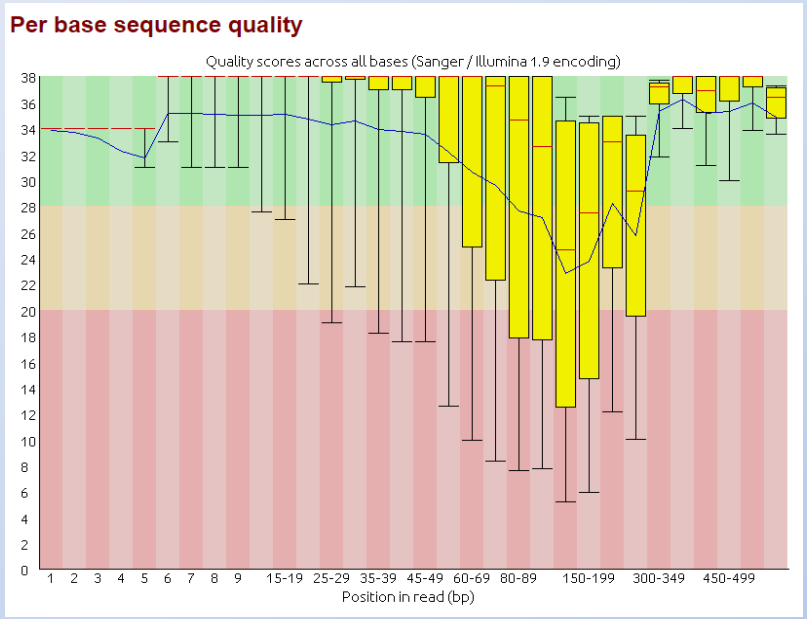
5’ METABARCODE 3’

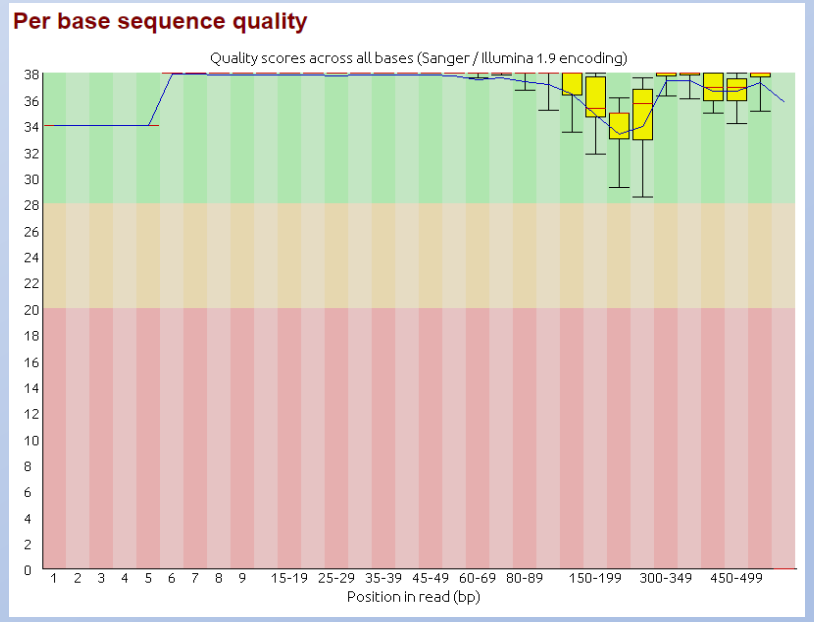
Recommended when generating OTUs or ASVs
(unless ITS Extractor is used)



Quality filtering

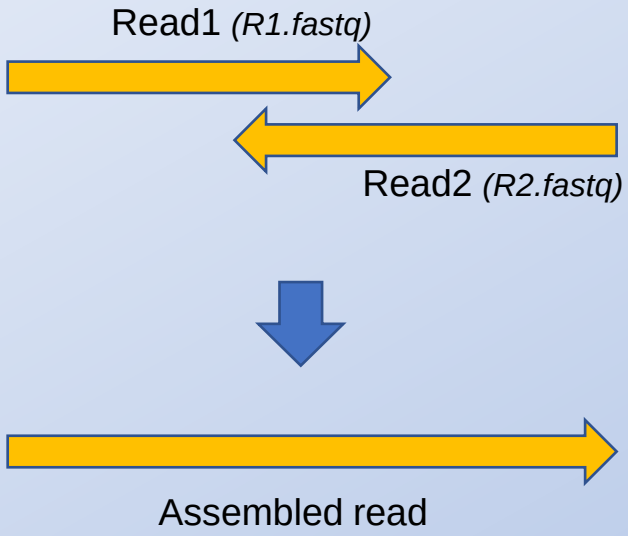


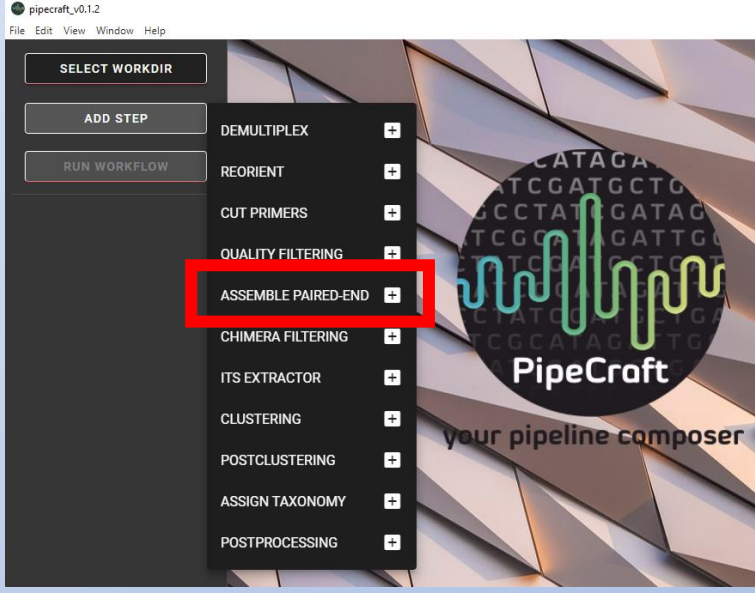




Assemble paired-end reads

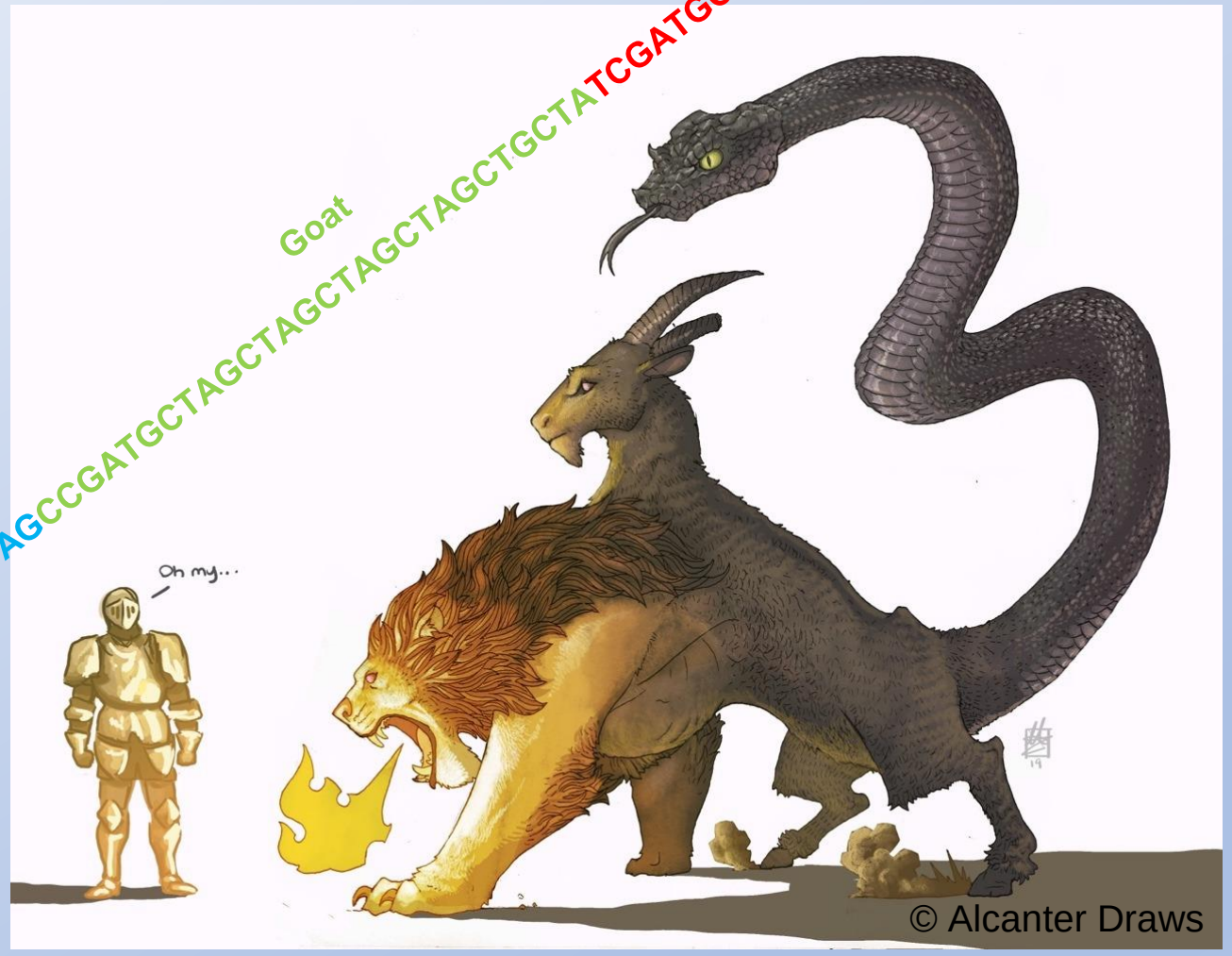
Illumina, MGI-Tech paired-end reads

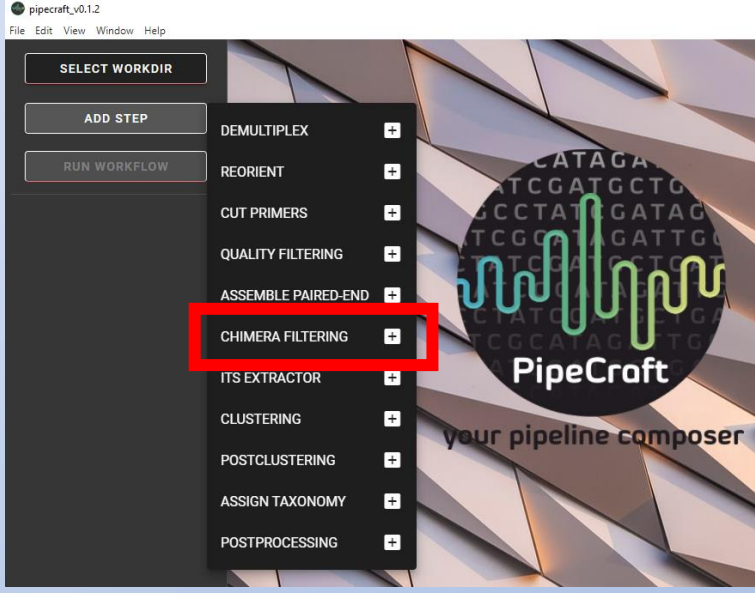


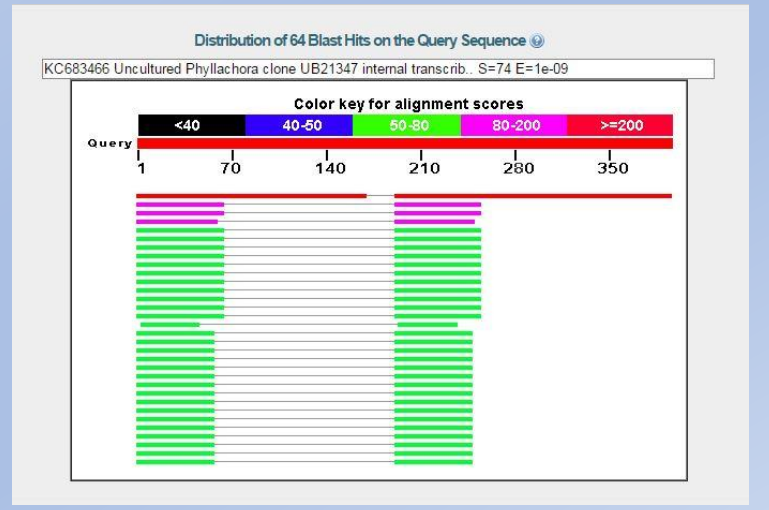


Chimera filtering

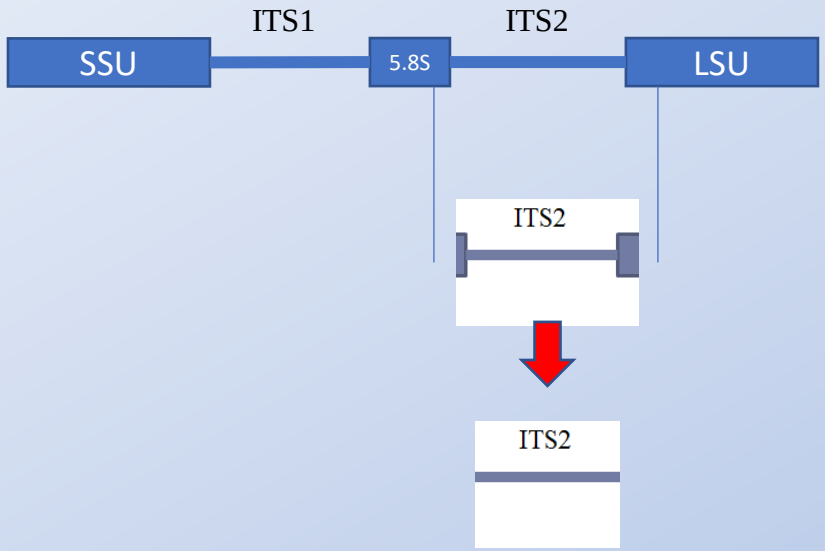
Lion GGAAACGTAGCTAGTAGCGTAGCCGATGCTAGCTAGCTAGCTGCTATCGATGCATTATTAGCGCGACTAGCTAGCA Snake

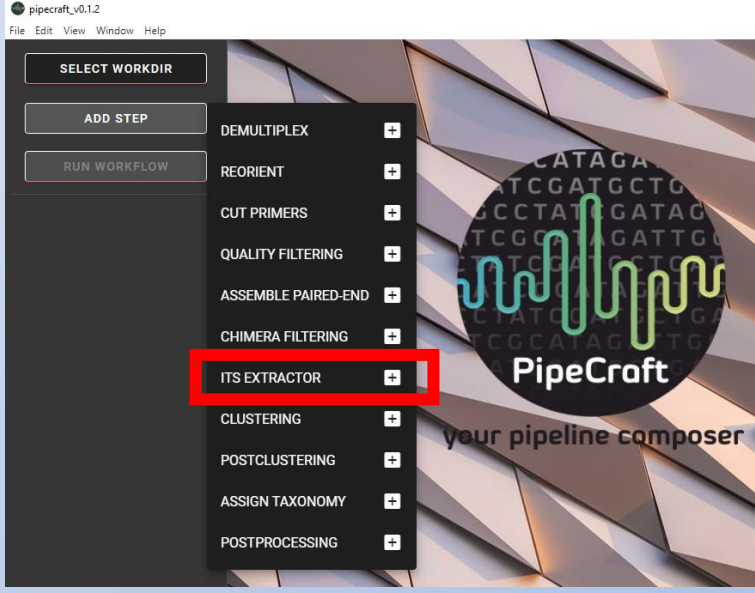




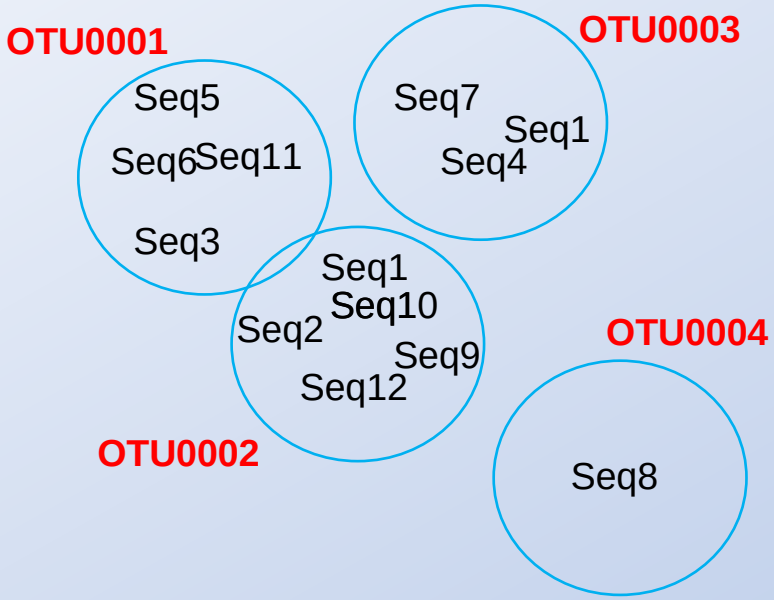


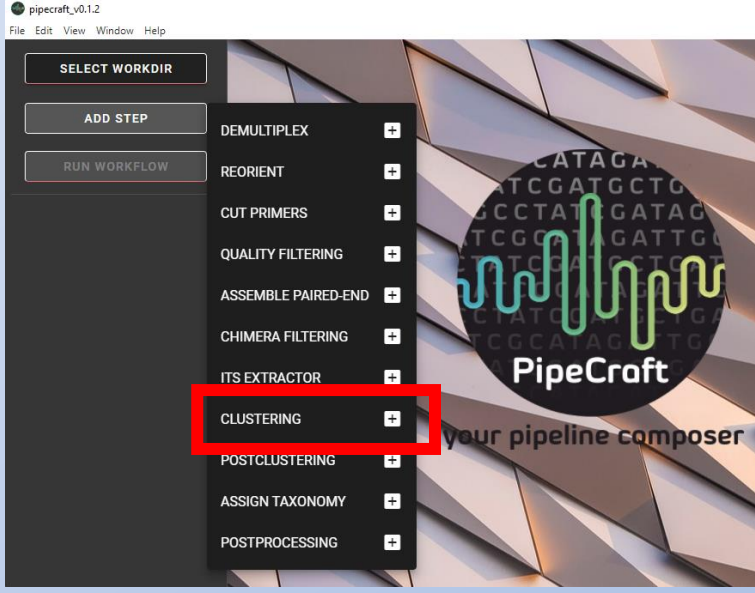
Extract ITS region(s)





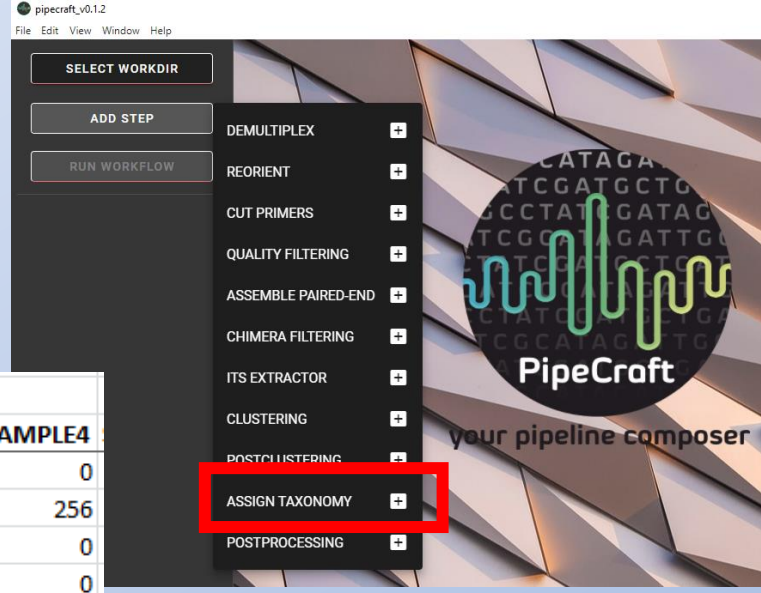
Clustering





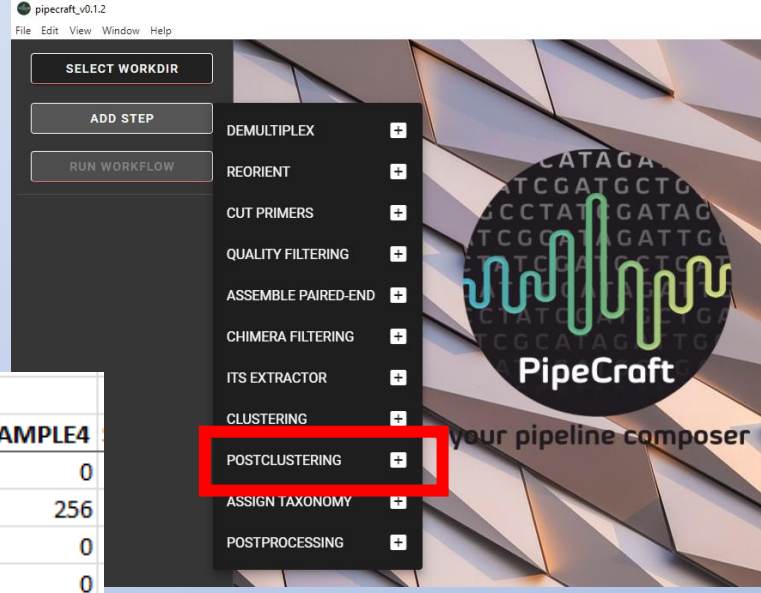
	Group	smpl1		smpl2		smpl3		smpl4		smpl5		smpl6	
2	Otu00001	0	0	0	52	0	0	0	2	0	0	0	
3	Otu00002	11	0	2	0	0	0	5	0	0	83	0	
4	Otu00003	0	0	0	0	15	0	0	0	0	0	88	
5	Otu00004	0	0	0	0	0	0	66	0	0	0	0	
6	Otu00005	0	1	251	10	0	0	2	0	58	0	0	
7	Otu00006	0	0	0	15	0	35	1	0	0	0	0	
8	Otu00007	0	0	0	0	0	0	0	0	0	0	0	
9	Otu00008	11	32	0	0	55	0	1	1	0	0	0	
10	Otu00009	0	1	0	0	0	0	0	0	0	0	0	
11	Otu00010	0	0	0	0	0	0	0	0	0	253	0	
12	Otu00011	0	0	21	0	0	0	0	0	0	0	0	
13	Otu00012	1	25	0	0	15	0	0	3	0	0	0	
14	Otu00013	0	0	0	0	0	0	0	0	0	0	0	
15	Otu00014	0	1	0	0	0	0	1	0	0	0	0	
16	Otu00015	0	0	0	0	0	0	0	0	0	0	0	
17	Otu00016	0	0	0	0	0	0	0	0	0	0	0	
18	Otu00017	1	2	1	0	0	1	3	1	0	257	163	
19	Otu00018	0	1	1	0	0	0	1	1	0	0	0	
20	Otu00019	1	1	0	0	0	0	0	1	0	0	0	

Assign taxonomy



Kingdom	Phylum	Class	Order	Family	Genus	Species	e value	query start	query end	coverage%	id%	SAMPLE1	SAMPLE2	SAMPLE3	SAMPLE4
k__Fungi	p__Basic	c__Ag	o__Th	f__Thel	g__Ton	s__Tomentella_ellisii	1.48E-87	1	220	100.00	100%	10	0	156	0
k__Fungi	p__Asco	c__Pe	o__Pe	f__Tub	g__Tub	s__Tuber_puberulum	5.44E-77	1	196	100.00	100%	0	11	0	256
k__Fungi	p__Basic	c__Ag	o__Ru	f__Russ	g__Rus	s__Russula_pelargonia	5.2E-99	1	246	100.00	100%	55	0	0	0
k__Fungi	p__Basic	c__Ag	o__Ag	f__Cort	g__Cor	s__Cortinarius_sp	2.36E-94	1	190	100.00	100%	0	0	5899	0
k__Fungi	p__Basic	c__Ag	o__Ag	f__Cort	g__Cor	s__Cortinarius_ferrugineov	6.51E-74	1	189	100.00	100%	96	25	0	566
k__Fungi	p__Basic	c__Ag	o__Ru	f__Russ	g__Rus	s__Russula_foetens	1.62E-91	1	229	100.00	100%	0	0	0	0
k__Fungi	p__Basic	c__Ag	o__Ag	f__Inoc	g__Ino	s__Inocybe_melanopus	2.86E-82	1	208	100.00	100%	0	542	0	0
k__Fungi	p__Asco	c__Pe	o__Pe	f__Pyrc	g__Oti	s__Otidea_mirabilis	6.98E-167	1	173	100.00	100%	9	0	0	12
k__Fungi	p__Basic	c__Ag	o__Ag	f__Hyd	g__Lac	s__Laccaria_laccata	4.12E-29	1	94	48.45	100%	0	11	0	0
k__Fungi	p__Basic	c__Ag	o__Ca	f__Can	g__Can	s__Cantharellus_cibarius	2.11E-111	1	340	100.00	100%	9856	2456	0	8856
k__Fungi	p__Basic	c__Ag	o__Ca	f__Can	g__Can	s__Cantharellus_cibarius	4.12E-136	1	340	100.00	100%	9	9	0	8
k__Fungi	p__Basic	c__Ag	o__Ca	f__Can	g__Can	s__Cantharellus_cibarius	4.12E-95	1	340	100.00	98%	6	0	0	3
k__Fungi	p__Basic	c__Ag	o__Ca	f__Can	g__Can	s__Cantharellus_cibarius	3.08E-89	3	338	98.82	99%	2	11	0	1
k__Fungi	p__Basic	c__Ag	o__Ca	f__Can	g__Can	s__Cantharellus_cibarius	5.52E-88	1	340	100.00	100%	1	0	0	5
k__Fungi	p__Basic	c__Ag	o__Th	f__Thel	g__Ton	s__Tomentella_galzinii	5.06E-89	1	216	100.00	99%	0	0	0	25
k__Fungi	p__Basic	c__Ag	o__Bo	f__Bole	g__Lec	s__Leccinum_scabrum	1.3E-120	1	295	100.00	100%	12	0	0	0
k__Fungi	p__Basic	c__Ag	o__Ag	f__Cort	g__Cor	s__Cortinarius_flexipes	3.12E-75	1	192	100.00	100%	0	0	45	0
k__Fungi	p__Basic	c__Ag	o__Th	f__Thel	g__uni	s__Thelephoraceae_sp	5.38E-88	1	221	100.00	100%	0	9	0	965
k__Fungi	p__Basic	c__Ag	o__Ag	f__Inoc	g__Ino	s__Inocybe_pelargonium	3.1E-52	1	181	100.00	95%	56	0	65	0
k__Fungi	p__Basic	c__Ag	o__Ag	f__Inoc	g__Ino	s__Inocybe_rimosa	6.78E-69	1	224	99.00	86%	0	89	0	12

Post-clustering



Kingdom	Phylum	Class	Order	Family	Genus	Species	e value	query start	query end	coverage%	id%	SAMPLE1	SAMPLE2	SAMPLE3	SAMPLE4
k__Fungi	p__Basic	c__Ag	o__Th	f__Thel	g__Ton	s__Tomentella_ellisii	1.48E-87	1	220	100.00	100%	10	0	156	0
k__Fungi	p__Asco	c__Pe	o__Pe	f__Tub	g__Tub	s__Tuber_puberulum	5.44E-77	1	196	100.00	100%	0	11	0	256
k__Fungi	p__Basic	c__Ag	o__Ru	f__Russ	g__Rus	s__Russula_pelargonia	5.2E-99	1	246	100.00	100%	55	0	0	0
k__Fungi	p__Basic	c__Ag	o__Ag	f__Cort	g__Cor	s__Cortinarius_sp	2.36E-94	1	190	100.00	100%	0	0	5899	0
k__Fungi	p__Basic	c__Ag	o__Ag	f__Cort	g__Cor	s__Cortinarius_ferrugineov	6.51E-74	1	189	100.00	100%	96	25	0	566
k__Fungi	p__Basic	c__Ag	o__Ru	f__Russ	g__Rus	s__Russula_foetens	1.62E-91	1	229	100.00	100%	0	0	0	0
k__Fungi	p__Basic	c__Ag	o__Ag	f__Inoc	g__Ino	s__Inocybe_melanopus	2.86E-82	1	208	100.00	100%	0	542	0	0
k__Fungi	p__Asco	c__Pe	o__Pe	f__Pyrc	g__Oti	s__Otidea_mirabilis	6.98E-167	1	173	100.00	100%	9	0	0	12
k__Fungi	p__Basic	c__Ag	o__Ag	f__Hyd	g__Lac	s__Laccaria_laccata	4.12E-29	1	94	48.45	100%	0	11	0	0
k__Fungi	p__Basic	c__Ag	o__Cai	f__Cant	g__Can	s__Cantharellus_cibarius	2.11E-111	1	340	100.00	100%	9856	2456	0	8856
k__Fungi	p__Basic	c__Ag	o__Cai	f__Cant	g__Can	s__Cantharellus_cibarius	4.12E-136	1	340	100.00	100%	9	9	0	8
k__Fungi	p__Basic	c__Ag	o__Cai	f__Cant	g__Can	s__Cantharellus_cibarius	4.12E-95	1	340	100.00	98%	6	0	0	3
k__Fungi	p__Basic	c__Ag	o__Cai	f__Cant	g__Can	s__Cantharellus_cibarius	3.08E-89	3	338	98.82	99%	2	11	0	1
k__Fungi	p__Basic	c__Ag	o__Cai	f__Cant	g__Can	s__Cantharellus_cibarius	5.52E-88	1	340	100.00	100%	1	0	0	5
k__Fungi	p__Basic	c__Ag	o__Th	f__Thel	g__Ton	s__Tomentella_galzinii	5.06E-89	1	216	100.00	99	0	0	0	25
k__Fungi	p__Basic	c__Ag	o__Bo	f__Bole	g__Lec	s__Leccinum_scabrum	1.3E-120	1	295	100.00	100%	12	0	0	0
k__Fungi	p__Basic	c__Ag	o__Ag	f__Cort	g__Cor	s__Cortinarius_flexipes	3.12E-75	1	192	100.00	100%	0	0	45	0
k__Fungi	p__Basic	c__Ag	o__Th	f__Thel	g__uni	s__Thelephoraceae_sp	5.38E-88	1	221	100.00	100%	0	9	0	965
k__Fungi	p__Basic	c__Ag	o__Ag	f__Inoc	g__Ino	s__Inocybe_pelargonium	3.1E-52	1	181	100.00	95%	5	0	65	0
k__Fungi	p__Basic	c__Ag	o__Ag	f__Inoc	g__Ino	s__Inocybe_rimosa	6.78E-69	1	224	99.00	86%	0	89	0	12

Merging co-occurring ‘daughter OTUs’

Assign traits

Published: 19 January 2021

FungalTraits: a user-friendly traits database of fungi and fungus-like stramenopiles

Sergei Põlme, Kessy Abarenkov, ... Leho Tedersoo + Show authors

Fungal Diversity **105**, 1–16 (2020) | Cite this article

GENUS	C	primary_lifestyle	Secondary_life	C	E	P	D	De	A	Ani	Specific_h	Growth_form_ter	Fruitbody_type_	Hymenium	Ectomy	Ectomycorrhiza_linea	primary_phc	secondary_p
Pycnocarpon		epiphyte	litter_saprotroph		foliar_enc		leaf/fruit/s		non-aquatic			filamentous_myceliu	thyrothecium_(tiny_	closed				
Neobueliella		lichen_parasite							non-aquatic			filamentous_myceliu	apothecium_(hyme	smooth				
Neocochlearomyces		plant_pathogen				leaf/fruit/seed_pathogen					Chromolaena	filamentous_mycelium						
Anthodidymella		wood_saprotroph					wood,leaf/		non-aquatic			filamentous_myceliu	perithecium_(hyme	closed				
Aquimassariosphaeria		wood_saprotroph					wood		freshwater			filamentous_myceliu	perithecium_(hyme	closed				
Neolindgomyces		litter_saprotroph					leaf/fruit/seed					filamentous_mycelium						
Crassiperidium		wood_saprotroph					wood		non-aquatic			filamentous_myceliu	perithecium_(hyme	closed				
Aquihelicascus		wood_saprotroph					wood		freshwater			filamentous_myceliu	perithecium_(hyme	closed				
Paratrimmatostroma		litter_saprotroph					leaf/fruit/seed					filamentous_mycelium						
Arezzomyces		wood_saprotroph					wood		non-aquatic			filamentous_myceliu	perithecium_(hyme	closed				
Dlhawksworthia		wood_saprotroph					leaf/fruit/s		non-aquatic			filamentous_myceliu	perithecium_(hyme	closed				
Hydeomyces		wood_saprotroph					wood					filamentous_mycelium						
Piniphoma		wood_saprotroph					wood					filamentous_mycelium						
Pleohelicoon		litter_saprotroph					leaf/fruit/s		non-aquatic		Fagus	filamentous_mycelium						
Typhicola		litter_saprotroph					leaf/fruit/seed				Typha	filamentous_mycelium						
Rubroshiraia		plant_pathogen				leaf/fruit/seed_p			non-aquatic		bamboo	filamentous_mycelium						
Anthosulcatispora		wood_saprotroph					wood		non-aquatic			filamentous_myceliu	perithecium_(hyme	closed				
Liua		wood_saprotroph					wood					filamentous_mycelium						
Discotubeufia		litter_saprotroph					leaf/fruit/seed				Brownea	filamentous_mycelium						
Acantharia		litter_saprotroph					leaf/fruit/se		non-aquatic			filamentous_mycelium						
Clavidisculum		litter_saprotroph					leaf/fruit/seed					filamentous_myceliu	perithecium_(hyme	closed				

Curation of the OTU matrix

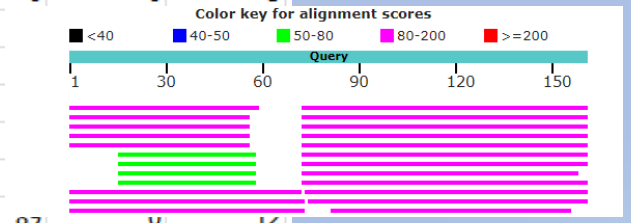
Kingdom	Phylum	Class	Order	Family	Genus	Species								
k__Fungi	p__Basic	c__Ag	o__Th	f__Thel	g__Ton	s__Tomentella_elli							0	156	0
k__Fungi	p__Asco	c__Pe	o__Pe	f__Tub	g__Tub	s__Tuber_puberulu							11	0	256
k__Fungi	p__Basic	c__Ag	o__Ru	f__Russ	g__Rus	s__Russula_pelargo							0	0	0
k__Fungi	p__Basic	c__Ag	o__Ag	f__Cort	g__Cor	s__Cortinarius_sp							0	5899	0
k__Fungi	p__Basic	c__Ag	o__Ag	f__Cort	g__Cor	s__Cortinarius_ferr							25	0	566
k__Fungi	p__Basic	c__Ag	o__Ru	f__Russ	g__Rus	s__Russula_foeten							0	0	0
k__Fungi	p__Basic	c__Ag	o__Ag	f__Inoc	g__Ino	s__Inocybe_melan							542	0	0
k__Fungi	p__Asco	c__Pe	o__Pe	f__Pyrc	g__Oti	s__Otidea_mirabili							0	0	12
k__Fungi	p__Basic	c__Ag	o__Ag	f__Hyd	g__Lac	s__Lacaria_laccata	4.12E-25	1	34	48.45	100%	0	11	0	0
k__Fungi	p__Basic	c__Ag	o__Cai	f__Cant	g__Can	s__Cantharellus_cibarius	2.11E-111	1	340	100.00	100%	9856	2456	0	8856
k__Fungi	p__Basic	c__Ag	o__Cai	f__Cant	g__Can	s__Cantharellus_cibarius	4.12E-136	1	340	100.00	100%	9	9	0	8
k__Fungi	p__Basic	c__Ag	o__Cai	f__Cant	g__Can	s__Cantharellus_cibarius	4.12E-95	1	340	100.00	98%	6	0	0	3
k__Fungi	p__Basic	c__Ag	o__Cai	f__Cant	g__Can	s__Cantharellus_cibarius	3.08E-89	3	338	98.82	99%	2	11	0	1
k__Fungi	p__Basic	c__Ag	o__Cai	f__Cant	g__Can	s__Cantharellus_cibarius	5.52E-88	1	340	100.00	100%	1	0	0	5
k__Fungi	p__Basic	c__Ag	o__Th	f__Thel	g__Ton	s__Tomentella_galzinii	5.06E-89	1	216	100.00	99%	0			
k__Fungi	p__Basic	c__Ag	o__Bo	f__Bole	g__Lec	s__Leccinum_scabrum	1.3E-120	1	295	100.00	100%	12			
k__Fungi	p__Basic	c__Ag	o__Ag	f__Cort	g__Cor	s__Cortinarius_flexipes	3.12E-75	1	192	100.00	100%	0			
k__Fungi	p__Basic	c__Ag	o__Th	f__Thel	g__uni	s__Thelephoraceae_sp	5.38E-88	1	221	100.00	100%	0			
k__Fungi	p__Basic	c__Ag	o__Ag	f__Inoc	g__Ino	s__Inocybe_pelargonium	3.1E-52	1	181	100.00	95%	56			
k__Fungi	p__Basic	c__Ag	o__Ag	f__Inoc	g__Ino	s__Inocybe_rimosa	6.78E-69	1	224	99.00	86%	0	89	0	12

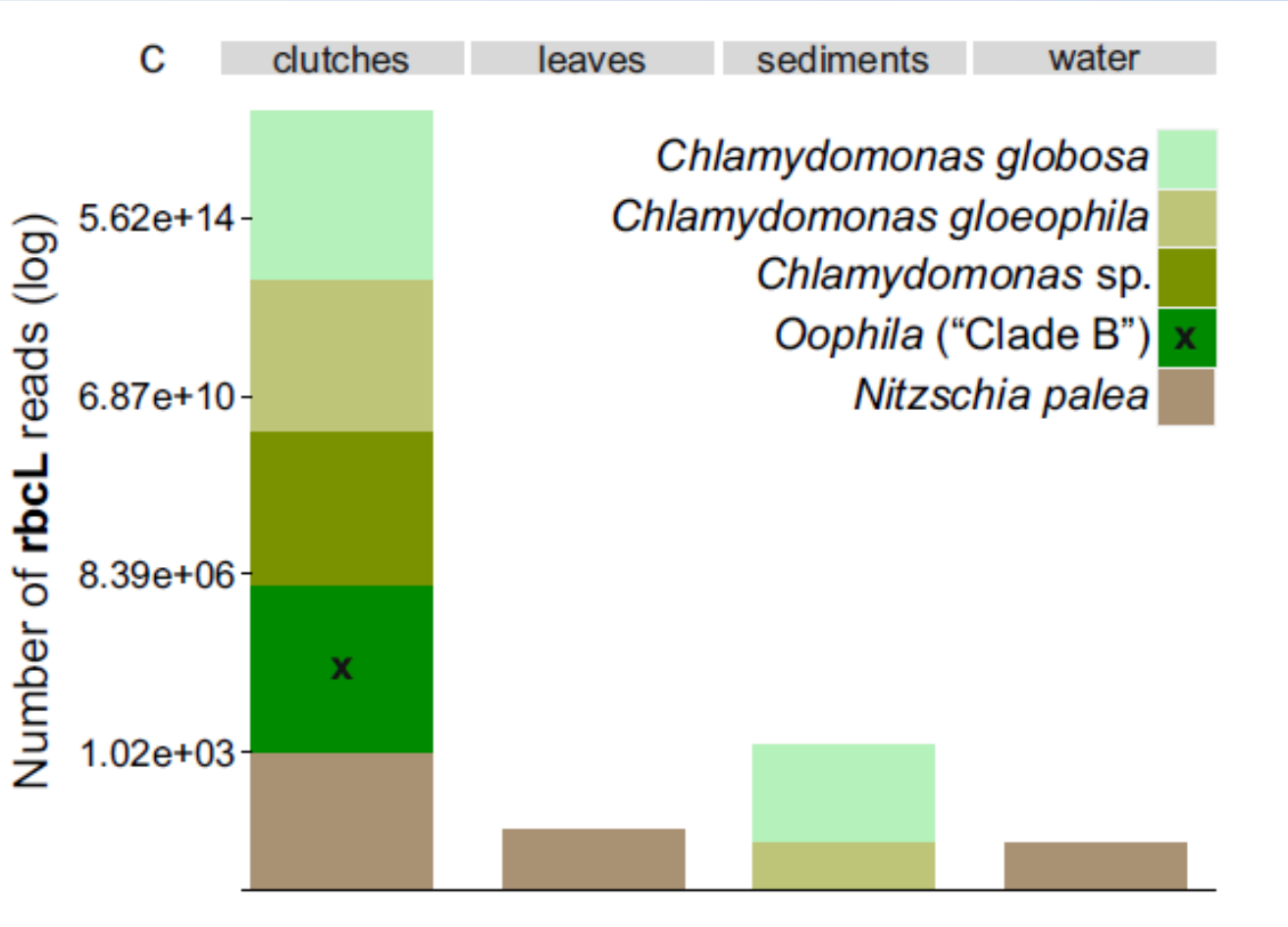
nature COMMUNICATIONS

Article | OPEN

Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates

Tobias Guldberg Frøslev, Rasmus Kjøller, Hans Henrik Bruun, Rasmus Ejrnæs, Ane Kirstine Brunbjerg, Carlotta Pietroni & Anders Johannes Hansen




C
clutches
leaves
sediments
water
Number of rbcL reads (log)
5.62e+14
6.87e+10
8.39e+06
1.02e+03
x
Chlamydomonas globosa
Chlamydomonas gloeophila
Chlamydomonas sp.
Oophila ("Clade B")
x
Nitzschia palea