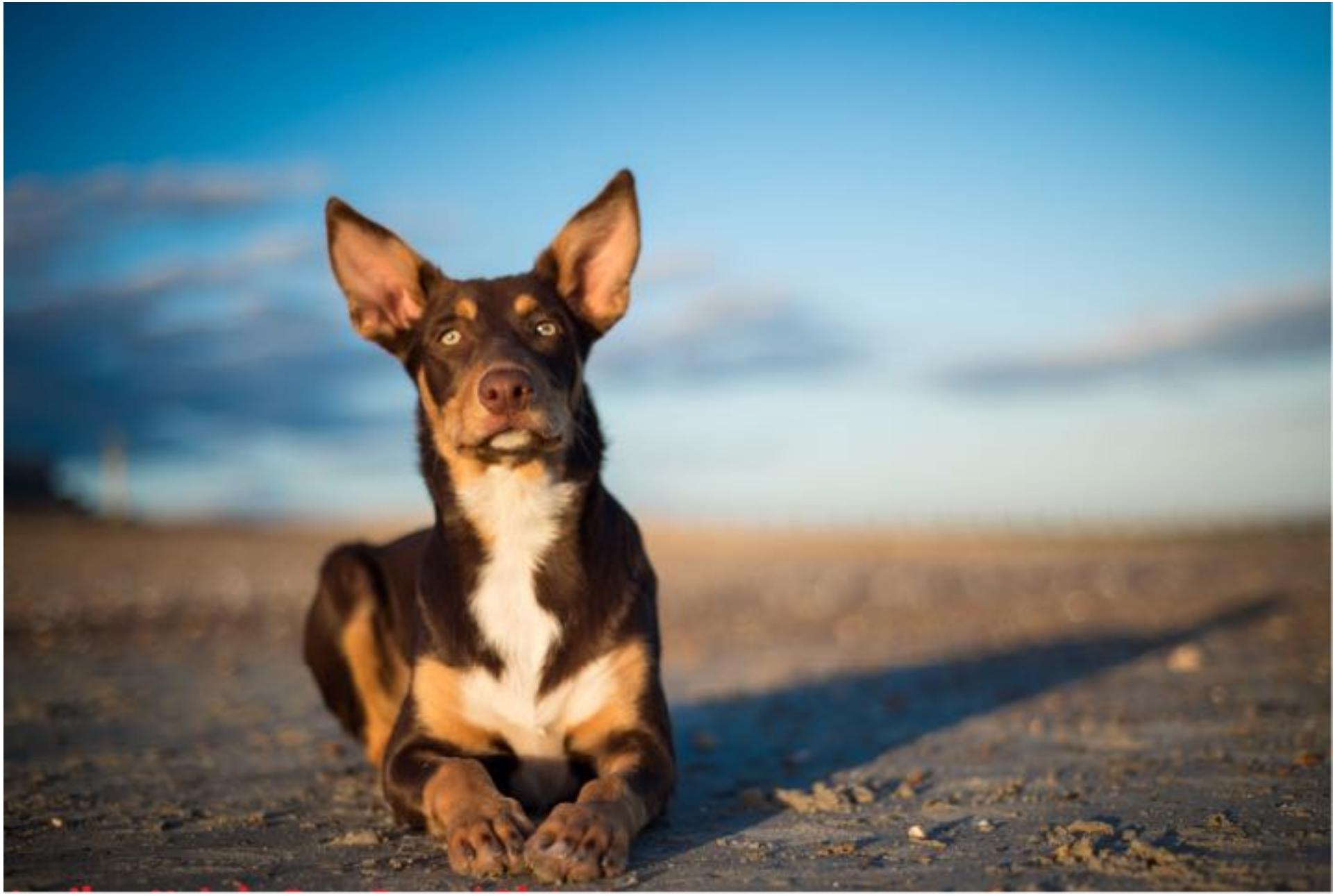


Bulk-sample Metabarcoding without ~~tears~~ PCR

Paul Greenfield, Anthony Chariton, Douglas Yu

Kelpies are a type of ‘working dog’

<https://dogtime.com/dog-breeds/australian-kelpie>



Kelpie: generating full-length ‘amplicons’ from whole-metagenome datasets

Paul Greenfield^{1,2,3}, Nai Tran-Dinh¹ and David Midgley¹

¹ Commonwealth Scientific and Industrial Research Organisation, North Ryde, NSW, Australia

² School of Biological Sciences, Macquarie University, Australia

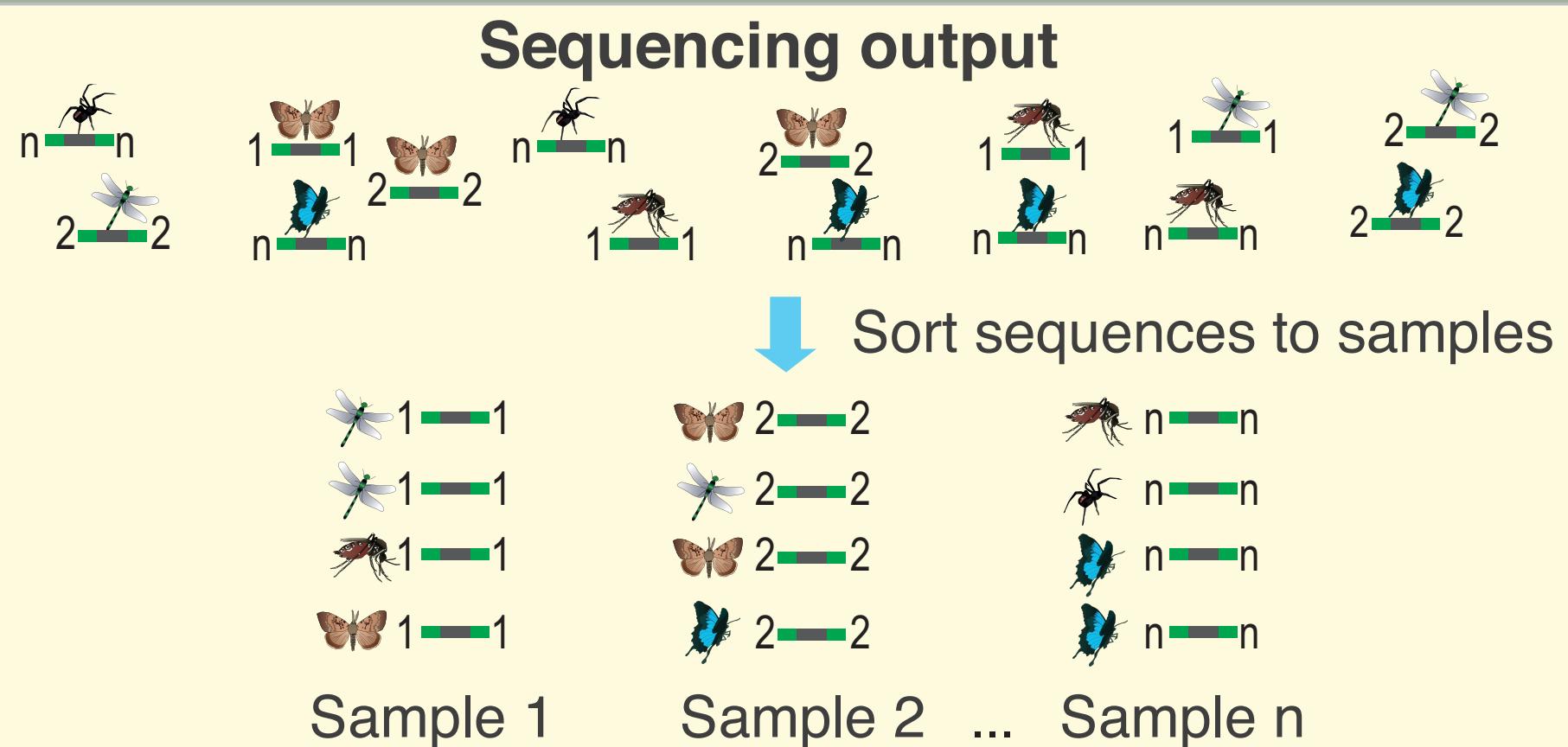
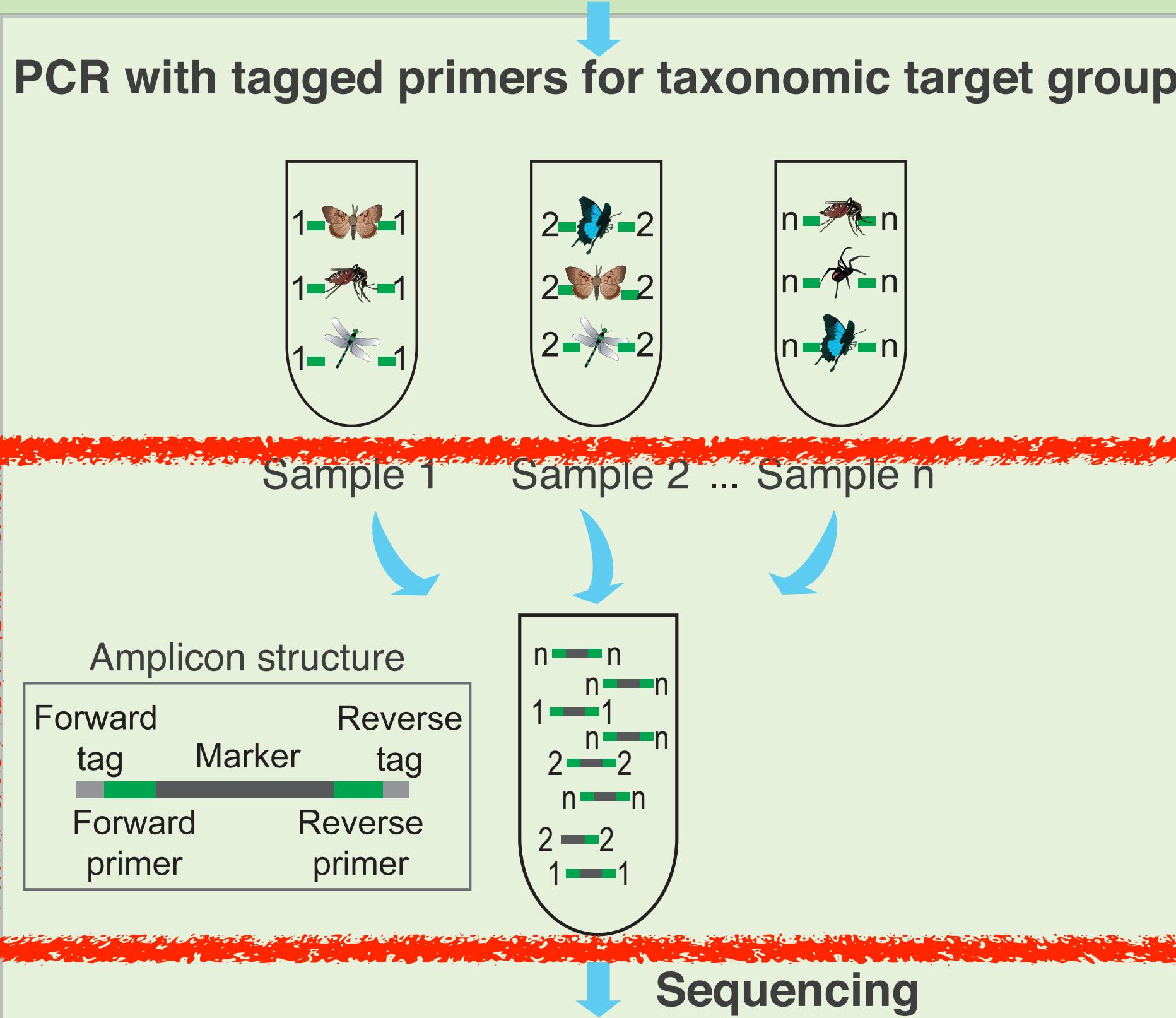
³ School of Information Technologies, University of Sydney, Australia

Published 30 January 2019, PeerJ 6:e6174.

Metabarcoding

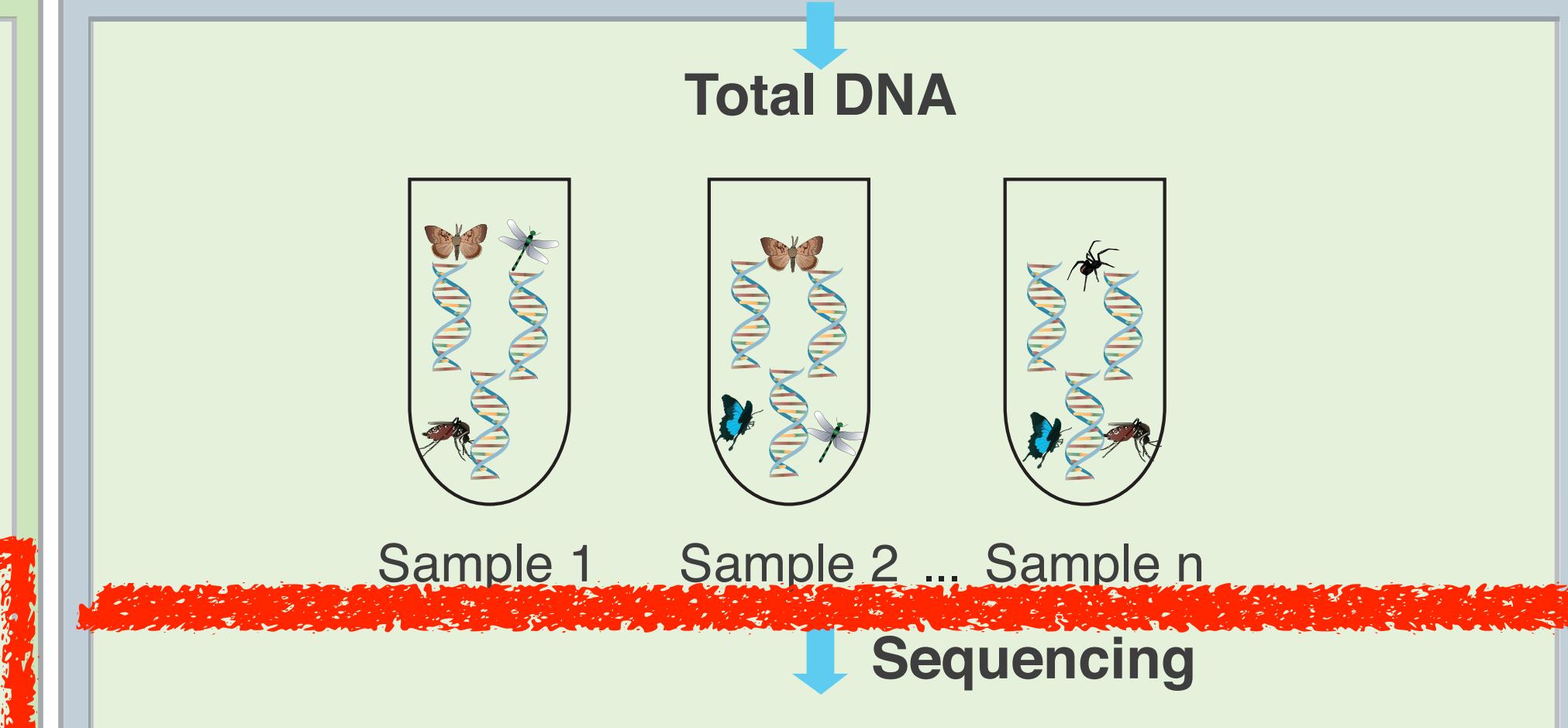
Only sequence specific gene regions (markers) for specific taxonomic groups (e.g the Insecta)

PCR step

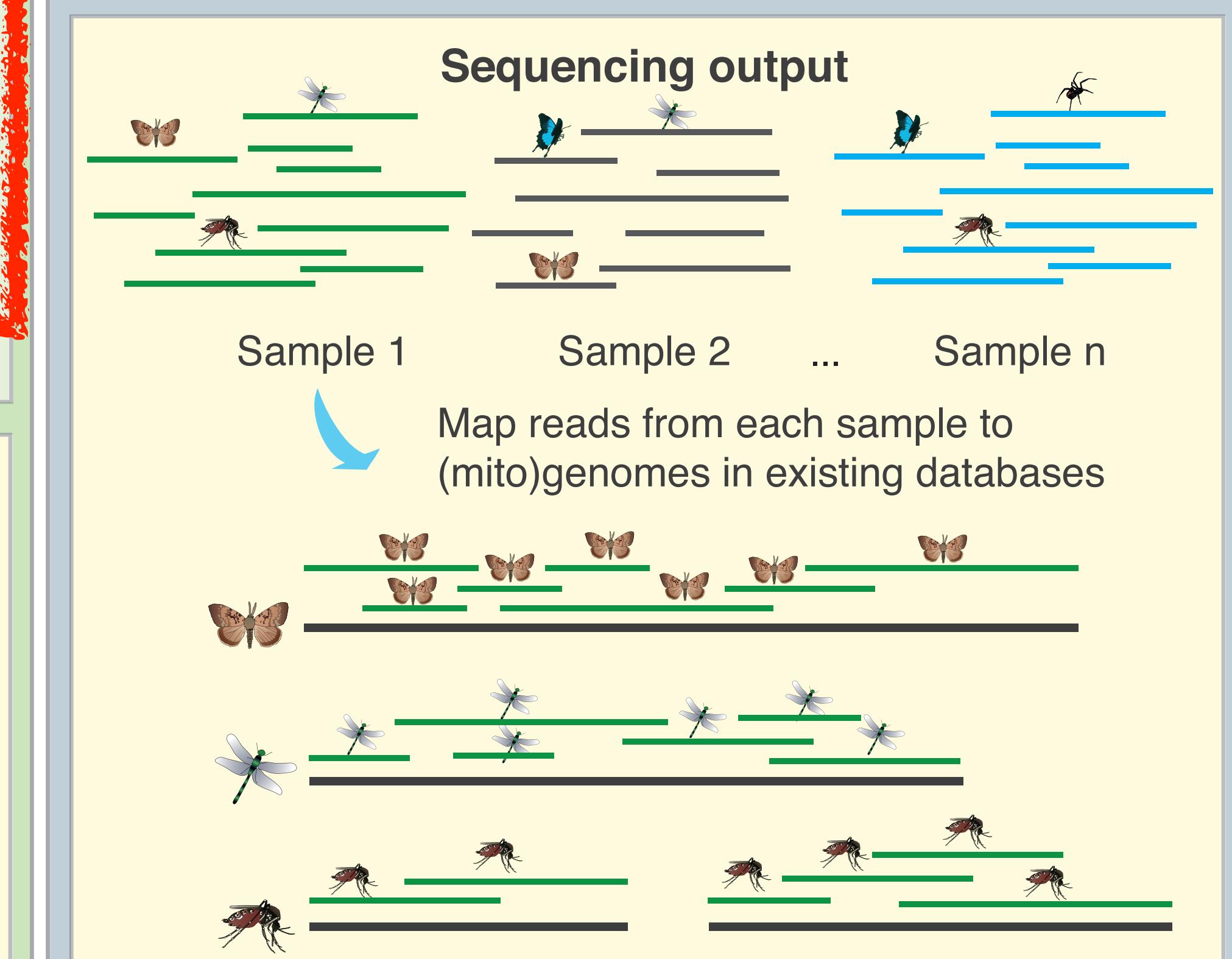


Meta-/mitogenomics

Sequence total DNA



No PCR step :-)



Kelpie: generating full-length ‘amplicons’ from whole-metagenome datasets

Paul Greenfield^{1,2,3}, Nai Tran-Dinh¹ and David Midgley¹

¹Commonwealth Scientific and Industrial Research Organisation, North Ryde, NSW, Australia

²School of Biological Sciences, Macquarie University, Australia

³School of Information Technologies, University of Sydney, Australia

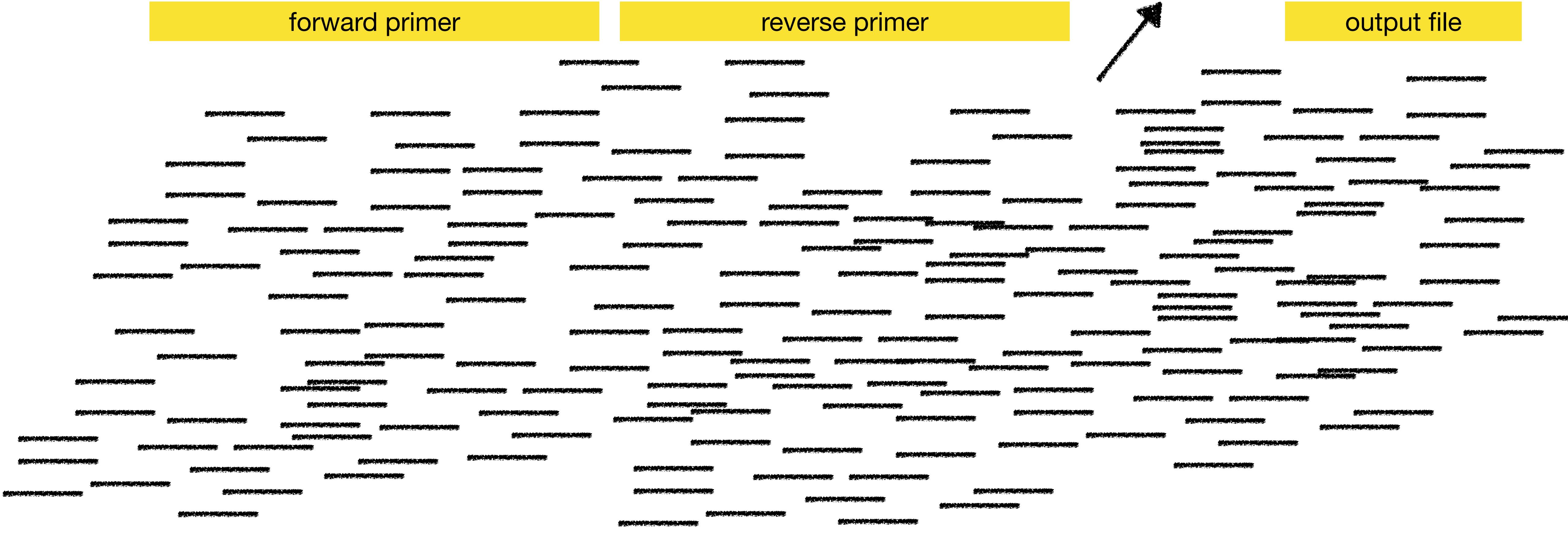
$$\frac{COI \text{ barcode}}{\text{Whole genome}} = \frac{1}{10000} = 0.01 \%$$

```
Kelpie -f GGWACWGGWTGAACWGTWTAYCCYCC -r TANACYTCNGGRTGNCCRAARAAYCA INPUTS.fq output_amplicons.fa
```

forward primer

reverse primer

output file



Kelpie: generating full-length ‘amplicons’ from whole-metagenome datasets

Paul Greenfield^{1,2,3}, Nai Tran-Dinh¹ and David Midgley¹

¹Commonwealth Scientific and Industrial Research Organisation, North Ryde, NSW, Australia

²School of Biological Sciences, Macquarie University, Australia

³School of Information Technologies, University of Sydney, Australia

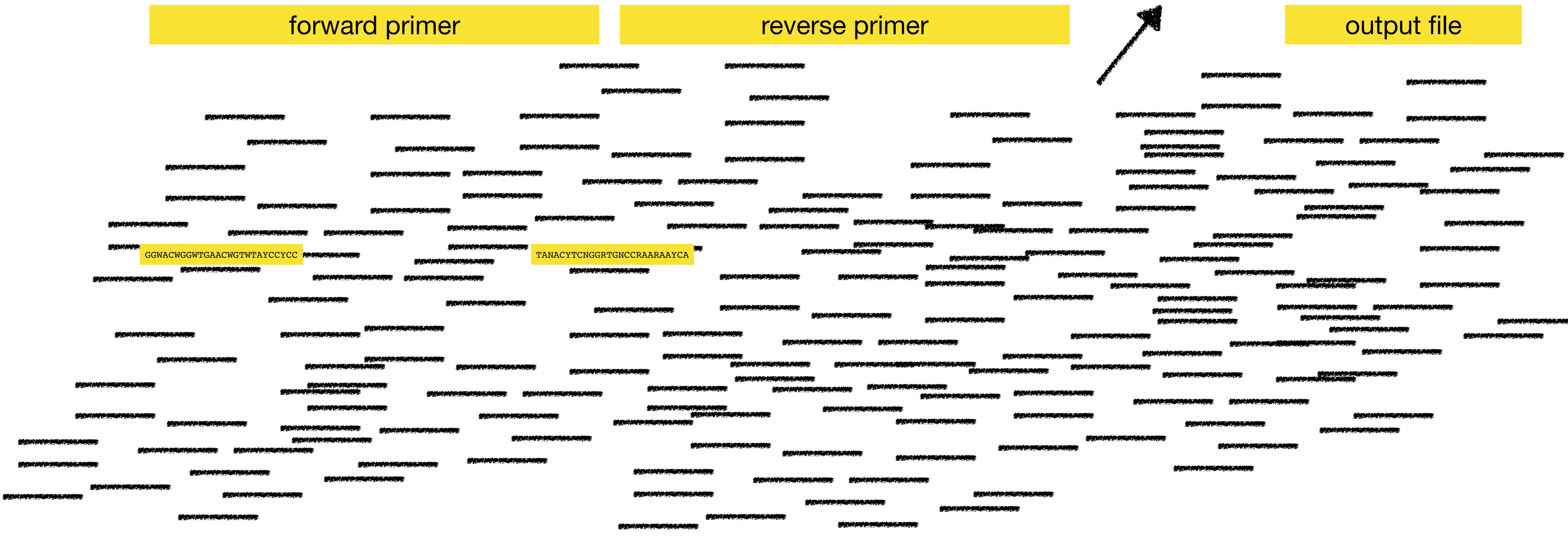
$$\frac{COI \text{ barcode}}{\text{Whole genome}} = \frac{1}{10000} = 0.01 \%$$

```
Kelpie -f GGWACWGGWTGAACWGTWTAYCCYCC -r TANACYTCNGGRTGNCCRAARAAYCA INPUTS.fq output_amplicons.fa
```

forward primer

reverse primer

output file



Kelpie: generating full-length ‘amplicons’ from whole-metagenome datasets

Paul Greenfield^{1,2,3}, Nai Tran-Dinh¹ and David Midgley¹

¹ Commonwealth Scientific and Industrial Research Organisation, North Ryde, NSW, Australia

² School of Biological Sciences, Macquarie University, Australia

³ School of Information Technologies, University of Sydney, Australia

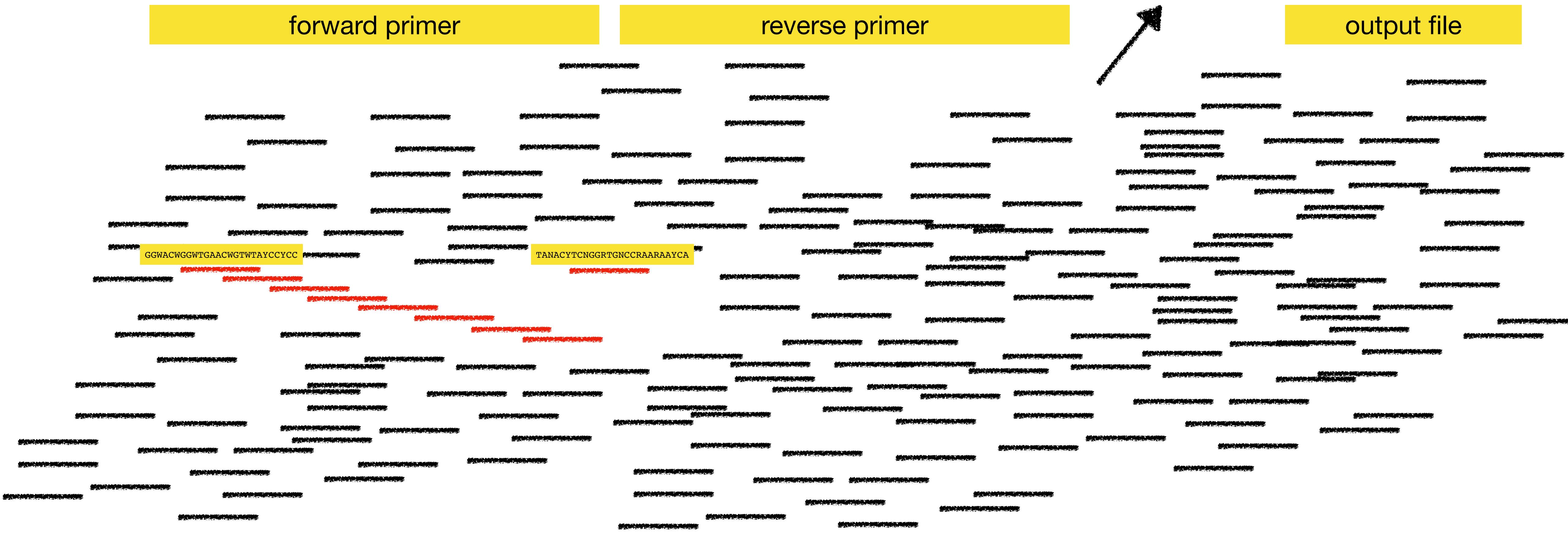
$$\frac{COI \text{ barcode}}{\text{Whole genome}} = \frac{1}{10000} = 0.01 \%$$

```
Kelpie -f GGWACWGGWTGAACWGTWTAYCCYCC -r TANACYTCNGGRTGNCCRAARAAYCA INPUTS.fq output_amplicons.fa
```

forward primer

reverse primer

output file



Kelpie: generating full-length ‘amplicons’ from whole-metagenome datasets

Paul Greenfield^{1,2,3}, Nai Tran-Dinh¹ and David Midgley¹

¹ Commonwealth Scientific and Industrial Research Organisation, North Ryde, NSW, Australia

² School of Biological Sciences, Macquarie University, Australia

³ School of Information Technologies, University of Sydney, Australia

$$\frac{COI \text{ barcode}}{\text{Whole genome}} = \frac{1}{10000} = 0.01\%$$

```
Kelpie -f GGWACWGGWTGAACWGTWTAYCCYCC -r TANACYTCNGGRTGNCCRAARAAYCA INPUTS.fq output_amplicons.fa
```

forward primer

reverse primer

output file



Kelpie: generating full-length ‘amplicons’ from whole-metagenome datasets

Paul Greenfield^{1,2,3}, Nai Tran-Dinh¹ and David Midgley¹

¹ Commonwealth Scientific and Industrial Research Organisation, North Ryde, NSW, Australia

² School of Biological Sciences, Macquarie University, Australia

³ School of Information Technologies, University of Sydney, Australia

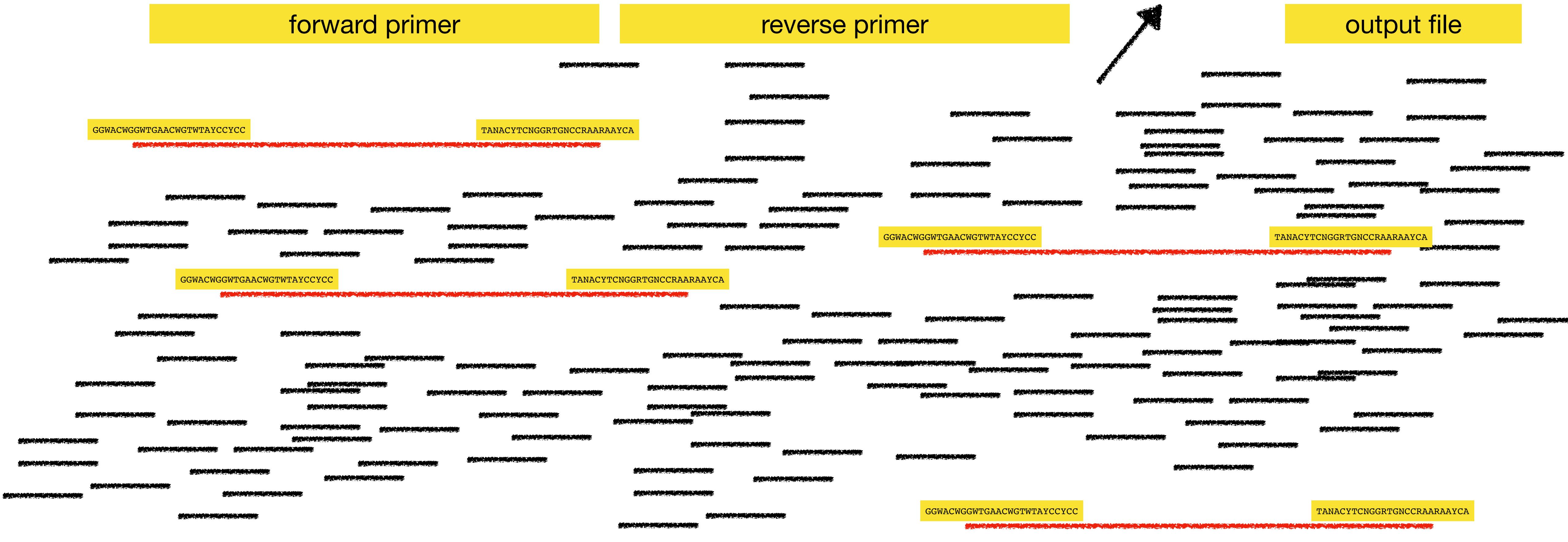
$$\frac{COI \text{ barcode}}{\text{Whole genome}} = \frac{1}{10000} = 0.01\%$$

```
Kelpie -f GGWACWGGWTGAACWGTWTAYCCYCC -r TANACYTCNGGRTGNCCRAARAAYCA INPUTS.fq output_amplicons.fa
```

forward primer

reverse primer

output file



Kelpie: generating full-length ‘amplicons’ from whole-metagenome datasets

Paul Greenfield^{1,2,3}, Nai Tran-Dinh¹ and David Midgley¹

¹ Commonwealth Scientific and Industrial Research Organisation, North Ryde, NSW, Australia

² School of Biological Sciences, Macquarie University, Australia

³ School of Information Technologies, University of Sydney, Australia

$$\frac{COI \text{ barcode}}{\text{Whole genome}} = \frac{1}{10000} = 0.01 \%$$

```
Kelpie -f GGWACGGWTGAACWGTWTAYCCYCC -r TANACYTCNGGRTGNCCRAARAAYCA INPUTS.fq output_amplicons.fa
```

output_amplicons.fa

GGWACGGWTGAACWGTWTAYCCYCC TANACYTCNGGRTGNCCRAARAAYCA
GGWACGGWTGAACWGTWTAYCCYCC TANACYTCNGGRTGNCCRAARAAYCA

This is the equivalent of a metabarcoding output

6 Original Core Steps of Metabarcoding

1. DNA extraction
2. PCR
3. (Illumina) library prep
4. (Illumina) sequencing
5. OTU clustering
6. taxonomic assignment

Kelpie: generating full-length ‘amplicons’ from whole-metagenome datasets

Paul Greenfield^{1,2,3}, Nai Tran-Dinh¹ and David Midgley¹

¹ Commonwealth Scientific and Industrial Research Organisation, North Ryde, NSW, Australia

² School of Biological Sciences, Macquarie University, Australia

³ School of Information Technologies, University of Sydney, Australia

$$\frac{COI \text{ barcode}}{\text{Whole genome}} = \frac{1}{10000} = 0.01\%$$

```
Kelpie -f GGWACGGWTGAACWGTWTAYCCYCC -r TANACYTCNGGRTGNCCRAARAAYCA INPUTS.fq output_amplicons.fa
```

output_amplicons.fa

GGWACGGWTGAACWGTWTAYCCYCC TANACYTCNGGRTGNCCRAARAAYCA
GGWACGGWTGAACWGTWTAYCCYCC TANACYTCNGGRTGNCCRAARAAYCA

This is the equivalent of a metabarcoding output

6 Core Steps of Kelpie Metabarcoding

1. DNA extraction

PCR

2. (Illumina) library prep

3. (Illumina) shotgun sequencing

4. Kelpie targeted assembly

5. OTU clustering

6. taxonomic assignment

Kelpie: generating full-length ‘amplicons’ from whole-metagenome datasets

Paul Greenfield^{1,2,3}, Nai Tran-Dinh¹ and David Midgley¹

¹ Commonwealth Scientific and Industrial Research Organisation, North Ryde, NSW, Australia

² School of Biological Sciences, Macquarie University, Australia

³ School of Information Technologies, University of Sydney, Australia

$$\frac{COI \text{ barcode}}{\text{Whole genome}} = \frac{1}{10000} = 0.01 \%$$

```
Kelpie -f GGWACGGWTGAACWGTWTAYCCYCC -r TANACYTCNGGRTGNCCRAARAAYCA INPUTS.fq output_Leray.fa
```

```
Kelpie -f CCHGAYATRGCHTTYCCCHCG -r TCDGGRTGNCCRAARAAYCA INPUTS.fq output_BF3BR2.fa
```

```
Kelpie -f CGGTTGGGGTGACCTCGGA -r GCTGTTATCCCTAGGGTAACt INPUTS.fq output_16Sbact.fa
```

Advantages of Kelpie Metabarcoding

1. Easy to use multiple PCR primers
arthropods, 18S, 16S bacteria
vertebrates, plants, ITS, functional genes, SNPs, ...
2. Avoid problems caused by physical PCR
no PCR sequence error, runaway, primer dimers
no Illumina length limit (**can use Folmer primers**)
lower contamination risk
no need to optimise PCR conditions (cycles, T_m)
3. Sequence variants more likely to be real variants
4. **Returns primer regions as well as amplicons**

Disadvantages of Kelpie Metabarcoding

1. **Requires large shotgun-sequencing datasets (but this is getting cheaper)**
2. **Cannot find very low-biomass species**
 1. ~~low-abundance, small-bodied species in bulk samples~~
 2. ~~eDNA/iDNA samples like fish eDNA in water, bloodmeal from a leech~~
 3. ~~scats? we don't know yet~~

Kelpie: generating full-length ‘amplicons’ from whole-metagenome datasets

Paul Greenfield^{1,2,3}, Nai Tran-Dinh¹ and David Midgley¹

¹ Commonwealth Scientific and Industrial Research Organisation, North Ryde, NSW, Australia

² School of Biological Sciences, Macquarie University, Australia

³ School of Information Technologies, University of Sydney, Australia

$$\frac{COI \text{ barcode}}{\text{Whole genome}} = \frac{1}{10000} = 0.01 \%$$

```
Kelpie -f GGWACGGWTGAACWGTWTAYCCYCC -r TANACYTCNGGRTGNCCRAARAAYCA INPUTS.fq output_Leray.fa
```

```
Kelpie -f CGGTTGGGGTGACCTCGGA -r GCTGTTATCCCTAGGGTAACt INPUTS.fq output_16Sbact.fa
```

```
Kelpie -f CCHGAYATRGCHTYCCHCG -r TCDGGRTGNCCRAARAAYCA INPUTS.fq output_BF3BR2.fa
```

Advantages of Kelpie Metabarcoding

1. Easy to use multiple PCR primers
arthropods, 18S, 16S bacteria
~~vertebrates, plants, ITS, functional genes, SNPs, ...~~
2. Avoid problems caused by standard PCR
no PCR sequencing bias, no runaway, primer dimers
no Illumina sequencing limit (**can use Folmer primers**)
low sequencing dilution risk
no need to optimise PCR conditions (cycles, T_m)
3. Sequence variants more likely to be real variants
4. Future version to report primer regions

Fewer false positives

Disadvantages of Kelpie Metabarcoding

1. Requires large shotgun-sequencing datasets
 1. 2.5B read pairs for £6500 + £1200 for NovaSeq 6000 + library prep
 2. 60 samples → £128 / sample
2. Cannot find very low-abundance species
 1. low-abundance, e.g. small-bodied species in bulk samples
 2. eDNA from bulk samples like fish eDNA in water, blood from a leech
3. scats? we don't know yet

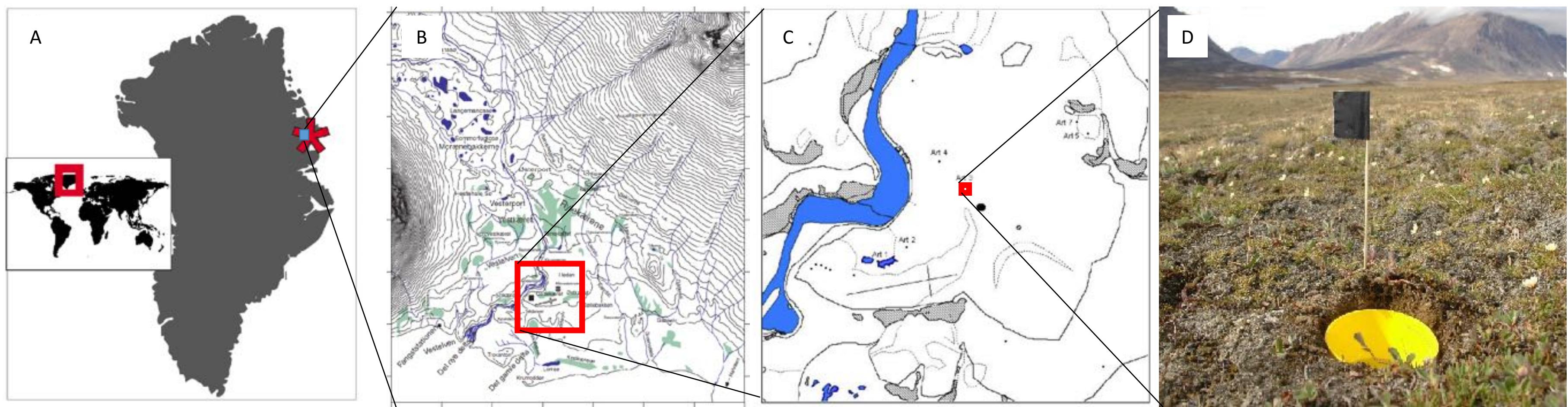
More false negatives

Zackenberg Research Station, Greenland



SPIKEPIPE: A metagenomic pipeline for the accurate quantification of eukaryotic species occurrences and intraspecific abundance change using DNA barcodes or mitogenomes

Yinqui Ji^{1*} | Tea Huotari^{2*} | Tomas Roslin^{2,3} | Niels Martin Schmidt^{4,5} |
Jixin Wang¹ | Douglas W. Yu^{1,6,7} | Otso Ovaskainen^{8,9}





- entire aboveground arthropod community **~375 species**
- **>760,000 arthropods** collected in weekly samples and multiple pan traps from **1996-2013** (and ongoing)
- We have **375 reference barcodes**, which took years to compile.
- **Can we get the barcodes from the samples themselves?**

Strategy:

Use Kelpie to generate a barcode reference dataset from the samples and

Use bwa/minimap2 read-mapping to detect barcodes in each sample



Kelpie: generating full-length ‘amplicons’ from whole-metagenome datasets

Paul Greenfield^{1,2,3}, Nai Tran-Dinh¹ and David Midgley¹

¹ Commonwealth Scientific and Industrial Research Organisation, North Ryde, NSW, Australia
² School of Biological Sciences, Macquarie University, Australia
³ School of Information Technologies, University of Sydney, Australia

DOI: 10.1111/1755-0998.13057

RESOURCE ARTICLE

MOLECULAR ECOLOGY RESOURCES

SPIKEPIPE: A metagenomic pipeline for the accurate quantification of eukaryotic species occurrences and intraspecific abundance change using DNA barcodes or mitogenomes

Yinqui Ji^{1*} | Tea Huotari^{2*} | Tomas Roslin^{2,3} | Niels Martin Schmidt^{4,5} | Jiaxin Wang¹ | Douglas W. Yu^{1,6,7} | Otso Ovaskainen^{8,9}



Reads from one sample mapped to a reference barcode



OTU
Coleoptera_Diatercidae_COI_SPIKE_0_8
Coleoptera_Mordellidae_COI_SPIKE_0_4
Leptodactyla_Bombycidae_Bombyx_mori_COI_SPIKE_0_2
Araeidae_Uropyhiidae_Hilaria_veratris_BOLD:ANG5689
Araeidae_Lycosidae_Pardosa_glaucalis_BOLD:AAA951
Diptera_Anthomyiidae_Zaprimea_frontrata_tundrica_BOLD:AA61723
Diptera_Muscidae_Dryomyza_segnis_BOLD:AAU7654
Diptera_Muscidae_Sollogonina_imiquishi_BOLD:AAU9104
Diptera_Muscidae_Sollogonina_mugistoma_BOLD:AAU9046
Diptera_Muscidae_Sollogonina_sanchpauli_BOLD:AAU9109
Diptera_Sciandidae_Camptochetae_gladiator_BOLD:AAU9262
Diptera_Sciandidae_Scaptosciara_atromaria_BOLD:AAU9520
Diptera_Syphidae_Platycheirus_cafnatus_BOLD:ABY7191
Hymenoptera_Ichneumonidae_Campoplex_horstmanni_estrata_BOLD:AAH1523
Diptera_Agyrtidae_Chromatomyia_pudicella_BOLD:AC48845
Hymenoptera_Ichneumonidae_Stenomacrus_micropennis_BOLD:AAH1490
Leptodactyla_Nymphalidae_Boloria_chryne_BOLD:AAU2067
Hymenoptera_Braconidae_Protopanteles_julivipes_BOLD:ABY9539
Diptera_Syphidae_Helophilus_gracilis_BOLD:AAH1982
Diptera_Ichneumonidae_Petalia_senae_BOLD:AAU25252
Hymenoptera_Ichneumonidae_Atractodes_sp_BOLD:ABZ7717
Diptera_Syphidae_Eupredes_punctifer_nf_punctatus_BOLD:AAU2484
Hymenoptera_Ichneumonidae_Stenomacrus_sp_37ERO_BOLD:AAC8798
Hymenoptera_Ichneumonidae_Pimpla_sordidir_BOLD:AAH1503
Lepidoptera_Notodontidae_Syniphis_vestimentaria_BOLD:AAU7102
Diptera_Muscidae_Sollogonina_arcticola_BOLD:ABW4722
Hymenoptera_Ichneumonidae_Hyposoter_migicus_BOLD:AAU9767
Leptodactyla_Culicidae_Culiseta_nebulosa_BOLD:AAU3447
Diptera_Calliphoridae_Cynomyia_sp_BOLD:AAU868
Diptera_Chironomidae_Smittia_sp_BOLD:ACP4114
Hymenoptera_Ichneumonidae_Campodeorus_jhuratus_BOLD:ACA1844
Diptera_Calliphoridae_Protophormia_attrita_BOLD:AAV6379
Leptodactyla_Geomyzidae_Entephria_kid_wulata_pulata_BOLD:AAU9361
Diptera_Ichneumonidae_Eoistia_thula_BOLD:ACF5729
Hymenoptera_Ichneumonidae_Atractodes_sp_BOLD:AAH2141
Hymenoptera_Ichneumonidae_Glypta_anthon_BOLD:AAH1501
Diptera_Syphidae_Helophilus_japonicus_BOLD:ACE4225
Hymenoptera_Ichneumonidae_Atractodes_sp_BOLD:AAU4528
Diptera_Chironomidae_Limnohyas_andrenaei_BOLD:ANM5308
Diptera_Sciandidae_genus_sp_BOLD:ACP4364
Hymenoptera_Ichneumonidae_Coelchnemonops_occidentalis_BOLD:AAH1869
Diptera_Chironomidae_Diamesa_bennetti_BOLD:AAU0255
Diptera_Syphidae_Panstratus_tarsatus_BOLD:AAC1634
Hymenoptera_Apidae_Bombus_hypoxanthus_BOLD:ACP6863
Diptera_Chironomidae_Smittia_sp_BOLD:AAU3817
Hymenoptera_Braconidae_Cotesia_sp_BOLD:ACE5164
Hymenoptera_Ichneumonidae_Campoplex_horstmanni_BOLD:ABE213
Hymenoptera_Ichneumonidae_Pterostigmas_sp_BOLD:AAU0823
Diptera_Chironomidae_Limnohyas_borealis_BOLD:ACM4348
Hymenoptera_Apidae_Bombus_polaris_BOLD:AAU2051
Hymenoptera_Ichneumonidae_Gelis_masticolar_BOLD:AAH2118
Diptera_Sciandidae_lyceniella_sp_BOLD:AAU26074
Hymenoptera_Ichneumonidae_genus_sp_AAH1506
Diptera_Muscidae_Sollogonina_eulizzi_BOLD:AAU9576
Diptera_Scalophaeidae_Scalophaea_furcata_BOLD:AAU0222
Diptera_Chironomidae_Oreocadus_rivulus_BOLD:AAU1171
Diptera_Anthomyiidae_Zaprimea_dilissa_BOLD:AAU2441
Araeidae_Ectyphidae_Emphytina_borealis_BOLD:ACK5561
Araeidae_Uropyhiidae_Erigone_arctica_BOLD:AAU6851
Diptera_Thomisidae_Xysticus_deichmanni_BOLD:ACE1100
Diptera_Agyrtidae_Phylomyza_aquilonia_BOLD:ABW5539
Diptera_Anthomyiidae_Zaprimea_occidentalis_BOLD:ABZ1211
Diptera_Chironomidae_Limnohyas_asquamatae_BOLD:ANB1720
Diptera_Chironomidae_Mesaphorura_ligni_BOLD:AAU3517
Diptera_Chironomidae_Parapheromadarius_impenitus_BOLD:ACM4201
Diptera_Chironomidae_Smittia_edwardsi_BOLD:AAU6749
Diptera_Chironomidae_Smittia_sp_BOLD:AAU9125
Diptera_Chironomidae_Tanytarsus_niger_BOLD:AAU7095
Diptera_Culicidae_Aedes_ingratus_BOLD:AAU3750
Diptera_Muscidae_Limnophora_groenlandica_BOLD:ACG5873
Diptera_Muscidae_Sollogonina_cerutata_BOLD:ACC7752
Diptera_Muscidae_Sollogonina_corsata_BOLD:AAU5038
Diptera_Muscidae_Sollogonina_micans_BOLD:AAU6168
Diptera_Muscidae_Sollogonina_puberula_BOLD:ACM519
Diptera_Mycetophilidae_Scopula_nitita_BOLD:ANG4892
Diptera_Sciandidae_Lycocella_modesta_BOLD:ABU5287
Diptera_Sciandidae_lyceniella_sp_BOLD:ARY5735
Diptera_Sciandidae_lyceniella_sp_BOLD:ACD1078
Diptera_Syphidae_Platycheirus_greenandensis_BOLD:AA4195
Hymenoptera_Braconidae_Cotesia_sp_BOLD:AAU6099
Hymenoptera_Braconidae_Protopanteles_julivipes_BOLD:ACP7221
Hymenoptera_Eulophidae_Aprostocetus_mitotela_BOLD:ABY8710

- test with samples from 2003, 2004, 2005
- **81 OTUs** known to be in these samples

DOI: 10.1111/1755-0998.13057

RESOURCE ARTICLE

MOLECULAR ECOLOGY
RESOURCES

SPIKEPIPE: A metagenomic pipeline for the accurate quantification of eukaryotic species occurrences and intraspecific abundance change using DNA barcodes or mitogenomes

Yinqui Ji^{1*} | Tea Huotari^{2,*} | Tomas Roslin^{2,3} | Niels Martin Schmidt^{4,5} | Jixin Wang¹ | Douglas W. Yu^{1,6,7} | Otso Ovaskainen^{8,9}

OTU	PlateGH_2003	PlateGH_2004	PlateGH_2005	PlateGH_2003/2004/5
Coleoptera_Batoceridae_COI_SPIKE_0_8	1	1	1	1
Coleoptera_Morulidae_COI_SPIKE_0_4	1	1	1	1
Lepidoptera_Bombycidae_Bombyx_mori_COI_SPIKE_0_2	1	1	1	1
Araeidae_Ulyphantidae_Hilaria_westrimi_BOLD:AAG5689	1	1	1	1
Araeidae_Lycosidae_Pardosa_glacialis_BOLD:AAA9551	1	1	1	1
Diptera_Anthomyidae_Zaphne_franata_tundifica_BOLD:AAG1723	1	1	1	1
Diptera_Musridae_Dyomyza_organica_BOLD:AAJ7864	1	1	1	1
Diptera_Muscidae_Sollogana_bimaculata_BOLD:AAN9104	1	1	1	1
Diptera_Muscidae_Sollogana_megastoma_BOLD:AAP9046	1	1	1	1
Diptera_Muscidae_Sollogana_sanciliensis_BOLD:AMM3109	1	1	1	1
Diptera_Sciadidae_Campiochaeta_cladator_BOLD:AMM262	1	1	1	1
Diptera_Sciadidae_Scaptocheta_stomatia_BOLD:AAH3820	1	1	1	1
Diptera_Syphidae_Platushelius_carninus_BOLD:ABV1791	1	1	1	1
Hymenoptera Ichneumonidae_Camptochela_holstmanni_mitrota_BOLD:AAH1523	1	1	1	1
Diptera_Agromyzidae_Chromatophila_puccinelliae_BOLD:ACAB345	1	1	1	1
Hymenoptera Ichneumonidae_Stenomacrus_micropennatus_BOLD:AAH1490	1	1	1	1
Lepidoptera Nymphalidae_Boloria_chionides_BOLD:AA2057	1	1	1	1
Hymenoptera Braconidae_Protopanteles_fulvipes_BOLD:ABY19539	1	1	1	1
Diptera_Syphidae_Pelephilus_grenlandicus_BOLD:AAH1952	1	1	1	1
Diptera_Tephritidae_Peleteria_aenea_BOLD:AAZ5252	1	1	1	1
Hymenoptera Ichneumonidae_Abraconodes_sp_BOLD:ABZ7717	1	1	1	1
Diptera_Syphidae_Eusarcus_punctifer_rufipunctatus_BOLD:AB2384	1	1	1	1
Hymenoptera Ichneumonidae_Stenomacrus_sp_3ZERO_BOLD:ACB798	1	1	1	1
Hymenoptera Ichneumonidae_Pimpla_sordalis_BOLD:AAH1503	1	1	1	1
Lepidoptera Noctuidae_Sympistis_zetterstedti_BOLD:ABV7102	1	1	1	1
Diptera_Musridae_Sollogana_arctica_BOLD:ABW4777	1	1	1	1
Hymenoptera Ichneumonidae_Hypostole_frigida-Bold:AAU9757	1	1	1	1
Lepidoptera Pieridae_Colias_hedera_BOLD:AAA3417	1	1	1	1
Diptera Calliphoridae_Cynomya_sp_BOLD:ABD868	1	1	1	1
Diptera Chironomidae_Smittia_sp_BOLD:ACM114	1	1	1	1
Hymenoptera Ichneumonidae_Camptochela_lituratus_BOLD:ACA1844	1	1	1	1
Diptera Calliphoridae_Protophermia_ariceps_BOLD:AAH6375	1	1	1	1
Lepidoptera Geometridae_Interchroa_kiduutatae_polyta_BOLD:AAE9861	1	1	1	1
Diptera Tephritidae_Borista_thula_BOLD:ACF5729	1	1	1	1
Hymenoptera Ichneumonidae_Abraconodes_sp_BOLD:AAH2111	1	1	1	1
Hymenoptera Ichneumonidae_Glypta_arctica_BOLD:AAH1501	1	1	1	1
Diptera Syphidae_Helophilus_japonicus_BOLD:ACD4226	1	1	1	1
Hymenoptera Ichneumonidae_Abraconodes_sp_BOLD:AAU14578	1	1	1	1
Diptera Chironomidae_Ulmophyes_anderseni_BOLD:AMM6308	1	1	1	1
Diptera Sciadidae_genus_sp_BOLD:ACP1361	1	1	1	1
Hymenoptera Ichneumonidae_Coelichneumonops_nigricornis_BOLD:AAH1869	1	1	1	1
Diptera Chironomidae_Dimesa_aeraria_BOLD:AMM0255	1	1	1	1
Diptera Syphidae_Panaphus_tenuis_BOLD:AACT1834	1	1	1	1
Hymenoptera Aleyrodidae_Bambus_myopaeus_BOLD:ACP5853	1	1	1	1
Diptera Chironomidae_Antocha_sp_BOLD:AAH1XAH17	1	1	1	1
Hymenoptera Braconidae_Cotesia_sp_BOLD:ACE6464	1	1	1	1
Hymenoptera Ichneumonidae_Camptochela_holstmanni_BOLD:ACD9213	1	1	1	1
Hymenoptera Ichneumonidae_Pleistogaster_sp_BOLD:AAE18724	1	1	1	1
Diptera Chironomidae_Ulmophyes_pachytomus_BOLD:ACM4349	1	1	1	1
Hymenoptera Aleyrodidae_Bambus_nigrae-Bold:AAU2051	1	1	1	1
Hymenoptera Ichneumonidae_Gols_maevicola_BOLD:AAH2118	1	1	1	1
Diptera Sciadidae_Lycoriella_sp_BOLD:AAZ6074	1	1	1	1
Hymenoptera Ichneumonidae_granulatus_sp_BOLD:AAH1NDK	1	1	1	1
Diptera Muscidae_Sollogana_zeteki_BOLD:AA9576	1	1	1	1
Diptera Scathophagidae_Scatophaga_fasciata_BOLD:AAH2HN22	1	1	1	1
Diptera Chironomidae_Othreutes_nivicola_BOLD:AAE1171	1	1	1	1
Diptera Anthomyidae_Zaphne_divisa_BOLD:AAU2441	1	1	1	1
Araeidae Ulyphantidae_Emblyna_borealis_BOLD:ACK5581	1	1	1	1
Araeidae Ulyphantidae_Erigone_arctica_BOLD:AB66851	1	1	1	1
Araeidae Thomomidae_Zygiticus_wiehmanni_BOLD:AAE1H100	1	1	1	1
Diptera Agromyzidae_Phytotrypa_aculeola_BOLD:ABW5539	1	1	1	1
Diptera Anthomyidae_Zaphne_occidentalis_BOLD:ABR21744	1	1	1	1
Diptera Chironomidae_Ulmophyes_acuminatus_BOLD:AAU1720	1	1	1	1
Diptera Chironomidae_Micropsectra_egani_BOLD:AAU3857	1	1	1	1
Diptera Chironomidae_Panamphodoloides_melanurus_BOLD:AAU4701	1	1	1	1
Diptera Chironomidae_Smittia_edwardsi_BOLD:AAU5749	1	1	1	1
Diptera Chironomidae_Antocha_sp_BOLD:ACI9175	1	1	1	1
Diptera Chironomidae_Teniarus_niger_BOLD:AAV7095	1	1	1	1
Diptera Culicidae_Aedes_migrans_migrans_BOLD:AAA3750	1	1	1	1
Diptera Muscidae_Limnohelinus_grenlandicus_BOLD:ACM873	1	1	1	1
Diptera Muscidae_Sollogana_denudata_BOLD:AC7752	1	1	1	1
Diptera Muscidae_Sollogana_dominica_BOLD:AAU1398	1	1	1	1
Diptera Muscidae_Sollogana_micans_BOLD:AAQ1686	1	1	1	1
Diptera Muscidae_Sollogana_puberula_BOLD:ACM549	1	1	1	1
Diptera Mycetophilidae_Scopilia_hirta_BOLD:AAG4892	1	1	1	1
Diptera Sciadidae_Lycoriella_medesta_BOLD:AAZ5287	1	1	1	1
Diptera Sciadidae_Lycoriella_sp_BOLD:AAH1ARY735	1	1	1	1
Diptera Sciadidae_Lycoriella_sp_BOLD:ACT0076	1	1	1	1
Diptera Syphidae_Platushelius_grenlandicus_BOLD:AAU2195	1	1	1	1
Hymenoptera Braconidae_Cotesia_sp_BOLD:AAA6099	1	1	1	1
Hymenoptera Braconidae_Protopanteles_fulvipes_BOLD:ACE7221	1	1	1	1
Hymenoptera Halictidae_Apanteles_mitofasciatus_BOLD:ABX710	1	1	1	1

assembled
in all 3 indiv
years and in
superset

assembled
in 2 indiv
years and in
superset

assembled
in 1 indiv
year and in
superset

assembled
in superset

indiv yr only

Kelpie: generating full-length ‘amplicons’ from whole-metagenome datasets

Paul Greenfield^{1,2,3}, Nai Tran-Dinh¹ and David Midgley¹

¹Commonwealth Scientific and Industrial Research Organisation, North Ryde, NSW, Australia

²School of Biological Sciences, Macquarie University, Australia

³School of Information Technologies, University of Sydney, Australia

DOI: 10.1111/1755-0998.13057

RESOURCE ARTICLE

SPIKEPIPE: A metagenomic pipeline for the accurate quantification of eukaryotic species occurrences and intraspecific abundance change using DNA barcodes or mitogenomes

Yinqiu Ji^{1*} | Tea Huotari^{2*} | Tomas Roslin^{2,3} | Niels Martin Schmidt^{4,5} | Jiaxin Wang¹ | Douglas W. Yu^{1,6} | Otso Ovaskainen^{8,9}

- **81 OTUs known to be in these samples**
- **OTUs assembled by Kelpie_V2**
 - **50/81 = 62% recovery rate, fwhF2, 313 bp**
 - **53/81 = 65% recovery rate, BF3BR2, 418 bp**
 - (only used 3 samples for this test) **With more samples from the same location** (the normal strategy), you should get a larger and larger proportion of the total trapped diversity
 - also, multiple species *partially* assembled by Kelpie
 - I think this will eventually replace amplicon sequencing for bulk samples