Impact of insufficient reference data on marine biodiversity assessment in Te Wāhipounamu/Fiordland (Aotearoa/New Zealand)

Paul Czechowski[1,2] and Michael Knapp[1,3]

[1]Department of Anatomy, University of Otago, PO Box 56, Dunedin 9054, Aotearoa / New Zealand;

[2]Helmholtz Institute for Metabolic, Obesity and Vascular Research, Philipp-Rosenthal-Str. 27, 04103 Leipzig, Germany;

[3]Coastal People Southern Skies Centre of Research Excellence, University of Otago, PO Box 56, Dunedin 9054, Aotearoa / New Zealand

Thank you

Michel de Lange[1,2b], Michael Heldsinger[3,4], Anya Kardailsky[1], Will Rayment[4,5], Christopher Hepburn[4,5], Monique Ladds[6], + DoC for funding

[1] Biostatistics Centre, University of Otago, PO Box 56, Dunedin 9054, Aotearoa / New Zealand

[2] Pacific Edge, 87 Saint David Street, Dunedin 9016, Aotearoa / New Zealand

[3] RPS Group, 27-31 Troode Street, West Perth WA 6005, Australia

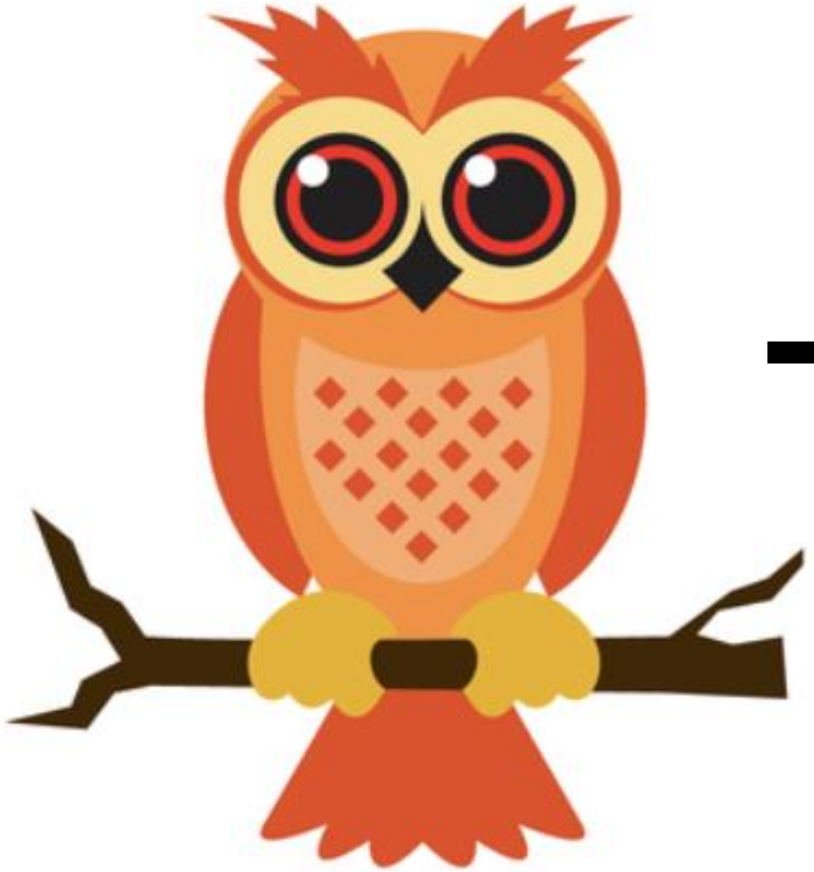
[4] Department of Marine Science, University of Otago, PO Box 56, Dunedin 9054, Aotearoa / New Zealand,

[5] Coastal People Southern Skies Centre of Research Excellence, University of Otago, PO Box 56, Dunedin 9054, Aotearoa / New Zealand

[6] Department of Conservation, PO Box 10 420, Wellington 6143, Aotearoa / New Zealand.

Department of Anatomy
School of Biomedical Sciences

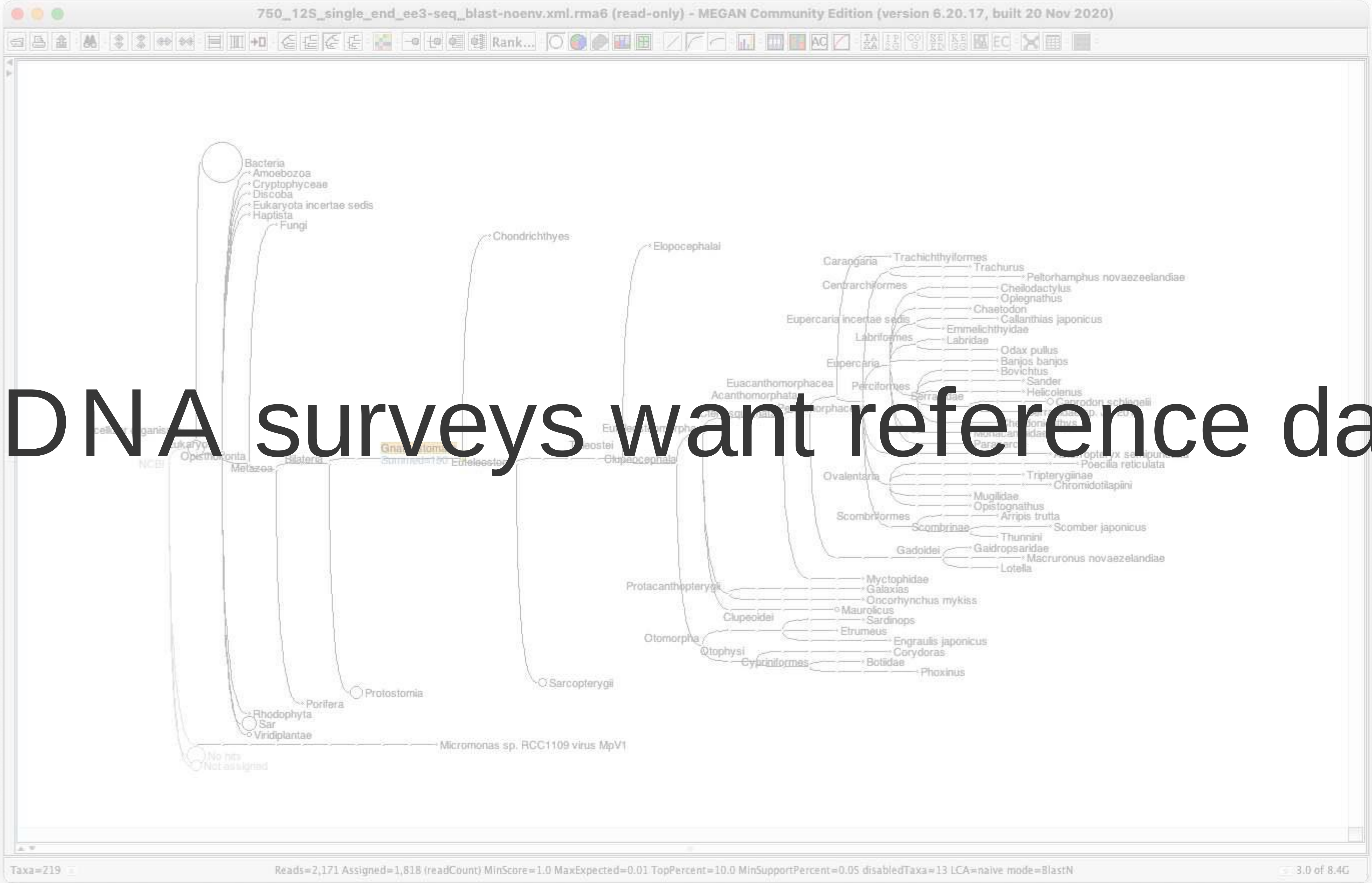
Disclaimer



By Alexander Popkov (engineerontour.com) - Own work, CC BY-SA 4.0,

Department of Anatomy
School of Biomedical Sciences

eDNA surveys want reference data

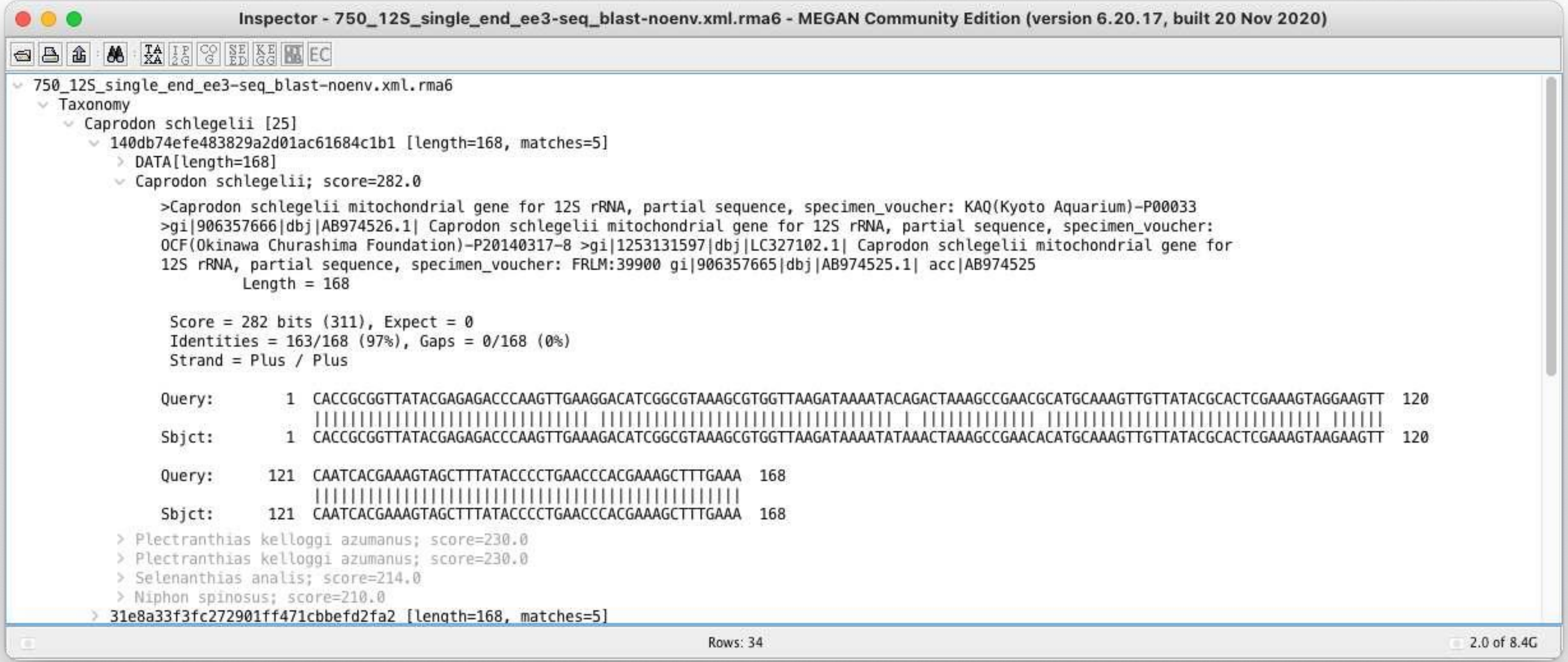


750_12S_single_end_ee3-seq_blast-noenv.xml.rma6 (read-only) - MEGAN Community Edition (version 6.20.17, built 20 Nov 2020)

Taxa=219

Reads=2,171 Assigned=1,818 (readCount) MinScore=1.0 MaxExpected=0.01 TopPercent=10.0 MinSupportPercent=0.05 disabledTaxa=13 LCA=naive mode=BlastN

3.0 of 8.4G



Inspector - 750_12S_single_end_ee3-seq_blast-noenv.xml.rma6 - MEGAN Community Edition (version 6.20.17, built 20 Nov 2020)

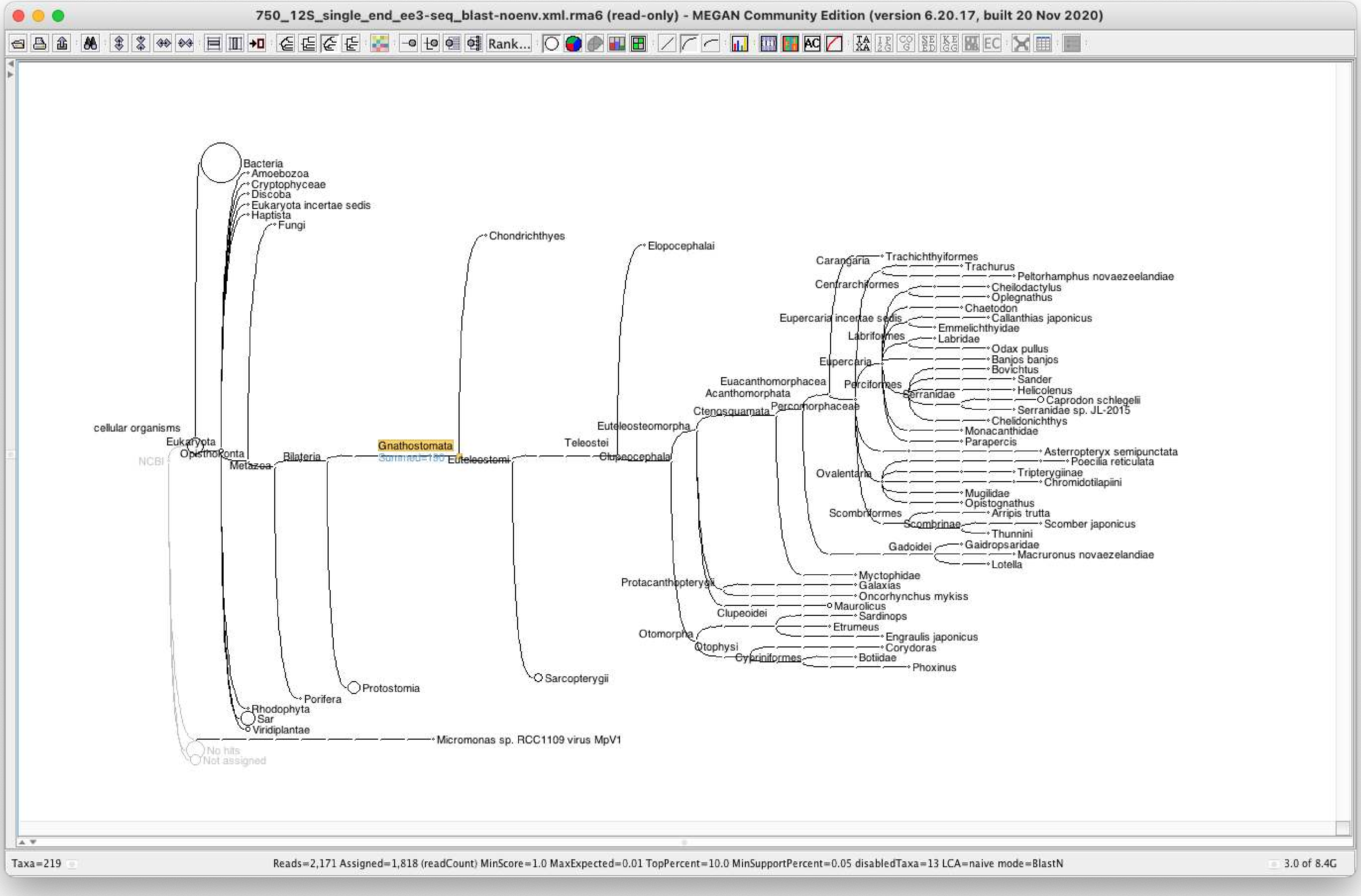
```
750_12S_single_end_ee3-seq_blast-noenv.xml.rma6
  Taxonomy
    Caprodon schlegelii [25]
      140db74efe483829a2d01ac61684c1b1 [length=168, matches=5]
        DATA[length=168]
        Caprodon schlegelii; score=282.0
          >Caprodon schlegelii mitochondrial gene for 12S rRNA, partial sequence, specimen_voucher: KAQ(Kyoto Aquarium)-P00033
          >gi|906357666|dbj|AB974526.1| Caprodon schlegelii mitochondrial gene for 12S rRNA, partial sequence, specimen_voucher:
          OCF(Okinawa Churashima Foundation)-P20140317-8 >gi|1253131597|dbj|LC327102.1| Caprodon schlegelii mitochondrial gene for
          12S rRNA, partial sequence, specimen_voucher: FRLM:39900 gi|906357665|dbj|AB974525.1| acc|AB974525
                   Length = 168

           Score = 282 bits (311), Expect = 0
           Identities = 163/168 (97%), Gaps = 0/168 (0%)
           Strand = Plus / Plus

          Query:       1  CACCGCGGTTATACGAGAGACCCAAGTTGAAGGACATCGGCGTAAAGCGTGGTTAAGATAAAATACAGACTAAAGCCGAACGCATGCAAAGTTGTTATACGCACTCGAAAGTAGGAAGTT  120
                          ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||| | ||||||||||||| ||||||||||||||||||||||||||||||| ||||||
          Sbjct:       1  CACCGCGGTTATACGAGAGACCCAAGTTGAAAGACATCGGCGTAAAGCGTGGTTAAGATAAAATATAAACTAAAGCCGAACACATGCAAAGTTGTTATACGCACTCGAAAGTAAGAAGTT  120

          Query:     121  CAATCACGAAAGTAGCTTTATACCCCTGAACCCACGAAAGCTTTGAAA  168
                          ||||||||||||||||||||||||||||||||||||||||||||||||
          Sbjct:     121  CAATCACGAAAGTAGCTTTATACCCCTGAACCCACGAAAGCTTTGAAA  168
        Plectranthias kelloggi azumanus; score=230.0
        Plectranthias kelloggi azumanus; score=230.0
        Selenanthias analis; score=214.0
        Niphon spinosus; score=210.0
      31e8a33f3fc272901ff471cbbefd2fa2 [length=168, matches=5]

Rows: 34                                                  2.0 of 8.4G
```



750_12S_single_end_ee3-seq_blast-noenv.xml.rma6 (read-only) - MEGAN Community Edition (version 6.20.17, built 20 Nov 2020)

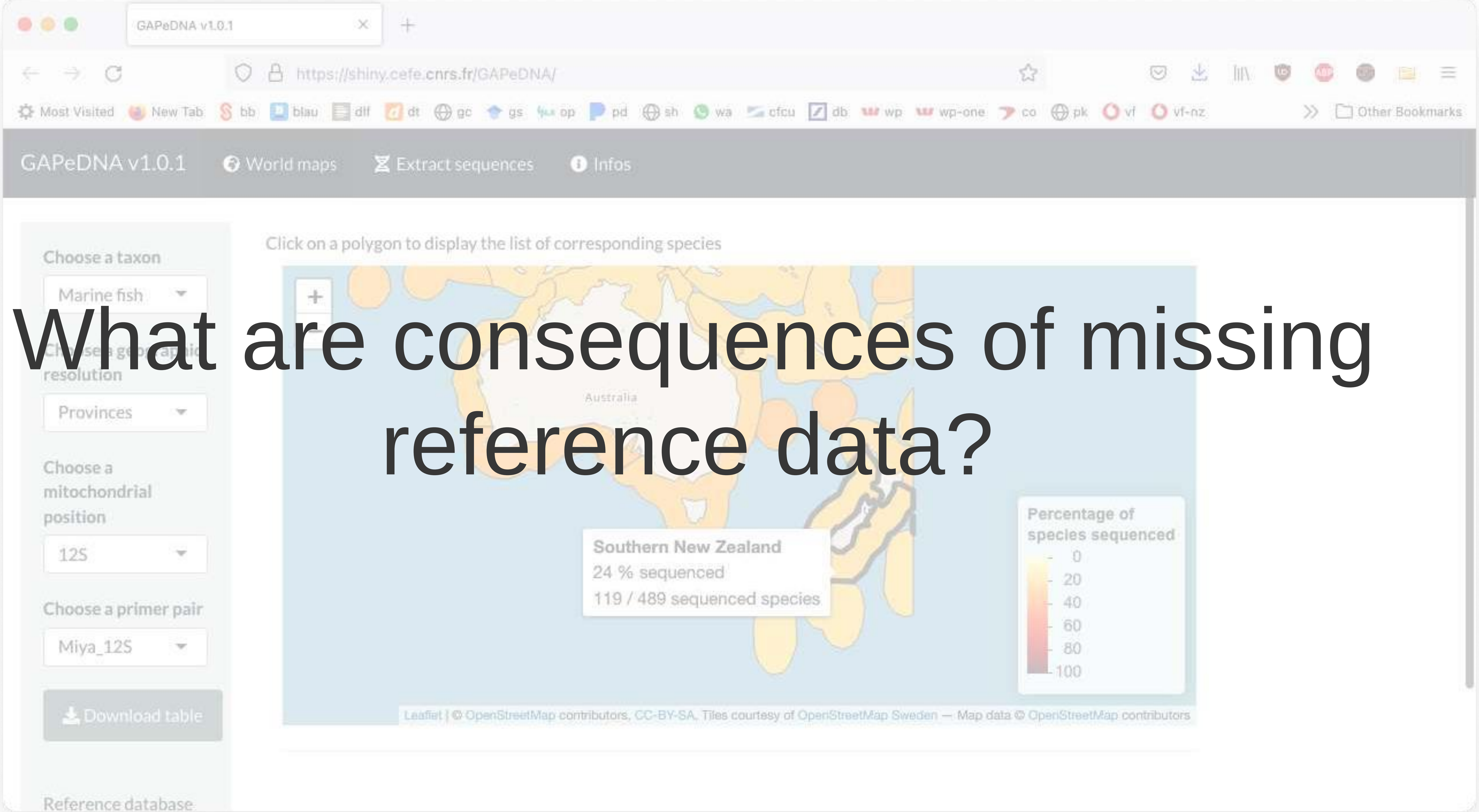
Taxa=219

Reads=2,171 Assigned=1,818 (readCount) MinScore=1.0 MaxExpected=0.01 TopPercent=10.0 MinSupportPercent=0.05 disabledTaxa=13 LCA=naive mode=BlastN

3.0 of 8.4G

Conservation focusses on species

What are consequences of missing reference data?



https://shiny.cefe.cnrs.fr/GAPeDNA/

Click on a polygon to display the list of corresponding species

Choose a taxon: Marine fish

Choose a geographic resolution: Provinces

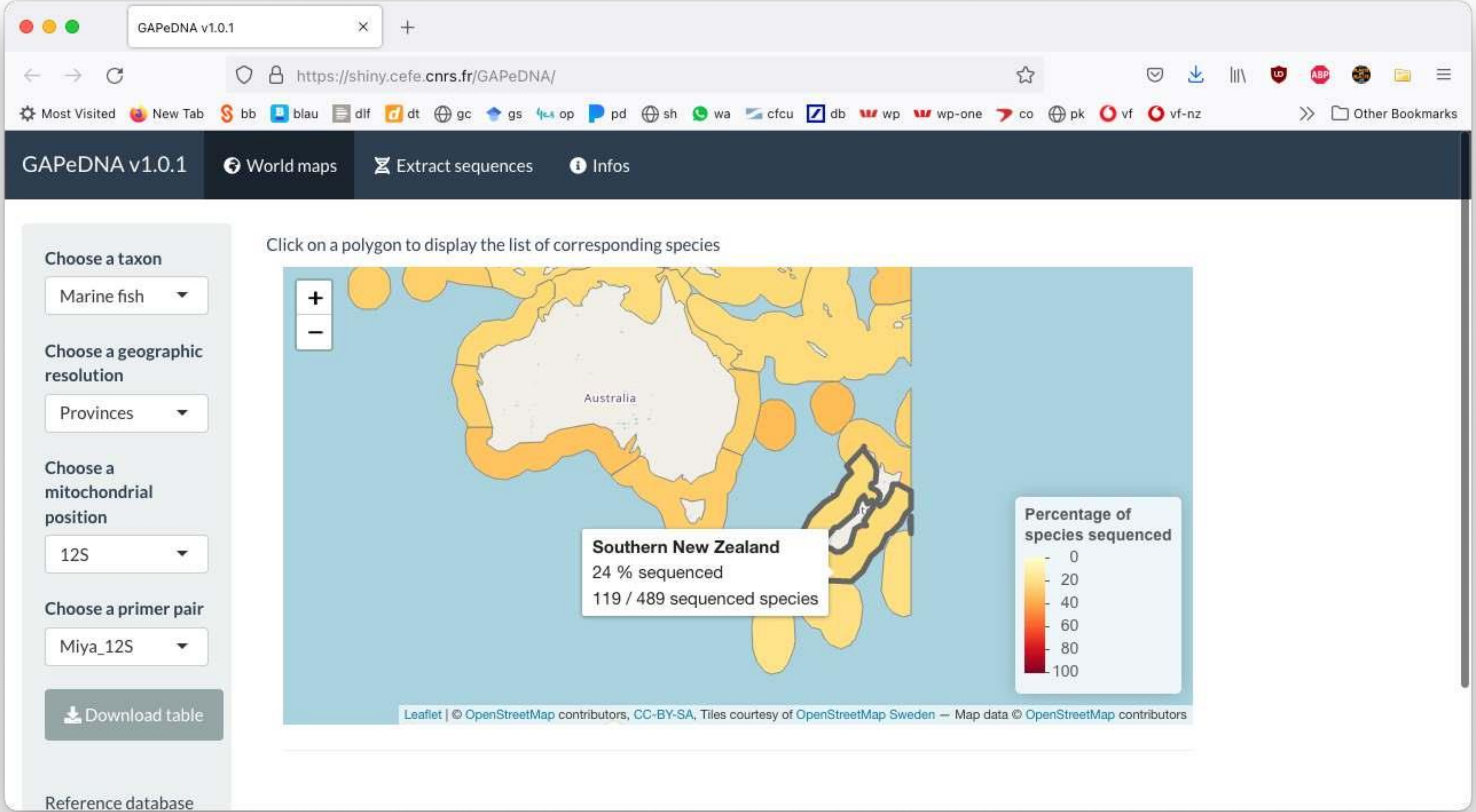
Choose a mitochondrial position: 12S

Choose a primer pair: Miya_12S

Reference database

Southern New Zealand
24 % sequenced
119 / 489 sequenced species

Leaflet | © OpenStreetMap contributors, CC-BY-SA, Tiles courtesy of OpenStreetMap Sweden — Map data © OpenStreetMap contributors



Choose a taxon: Marine fish

Choose a geographic resolution: Provinces

Choose a mitochondrial position: 12S

Choose a primer pair: Miya_12S

Download table

Reference database

Click on a polygon to display the list of corresponding species

Southern New Zealand
24 % sequenced
119 / 489 sequenced species

Leaflet | © OpenStreetMap contributors, CC-BY-SA, Tiles courtesy of OpenStreetMap Sweden — Map data © OpenStreetMap contributors

Data collation

Biodiversity information from Fiordland literature, OBIS, BRUV, eDNA

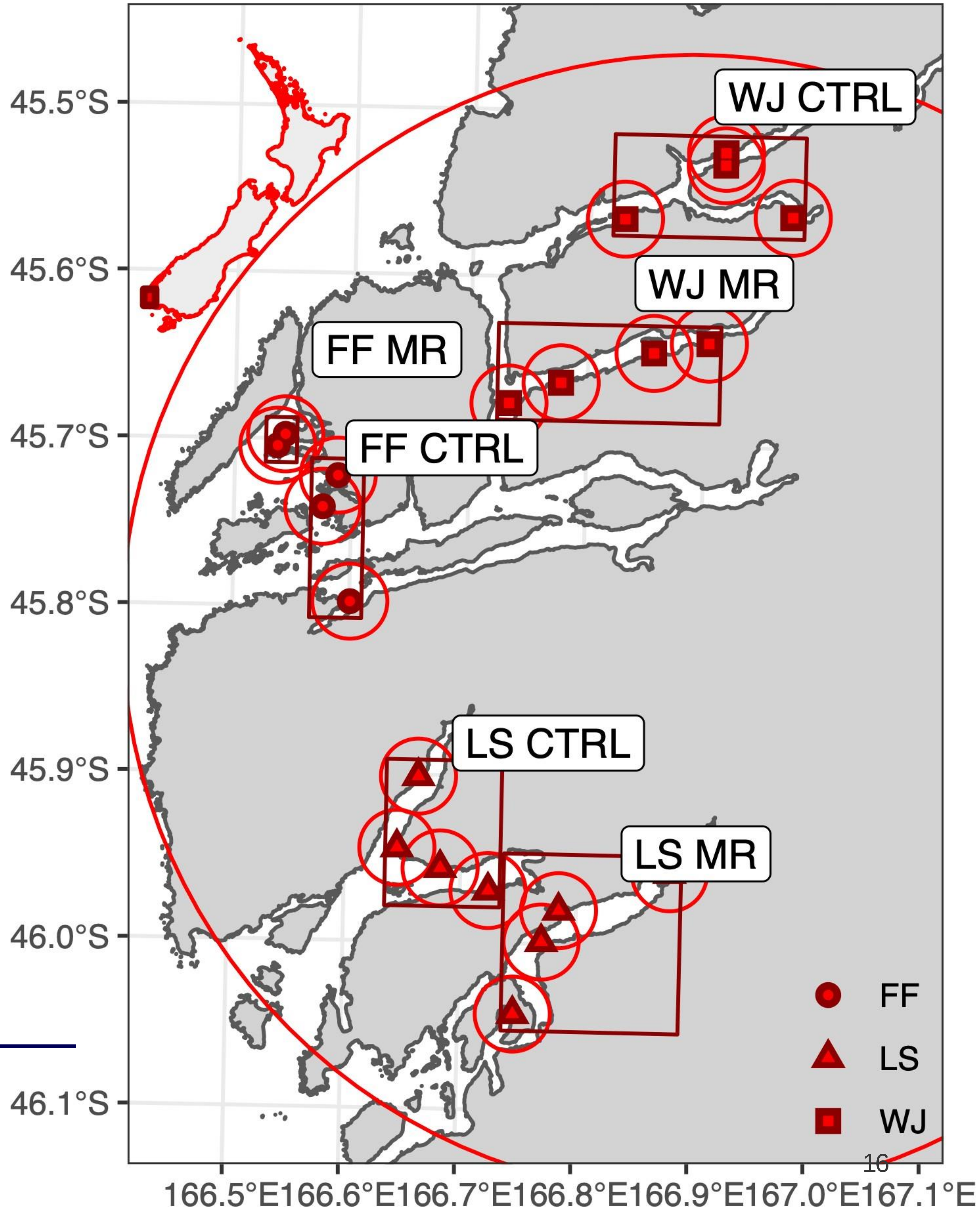
- Concordance of all data sources (order, family, genus, species)
- Sampling effort and species diversity (BRUV, eDNA)
- Good-Turing estimates

Credibility of environmental DNA information

- Compare eDNA & BRUV assignments to literature
- eDNA alignment qualities

Possibility of community analysis of field-based-methods

- ANOSIM of species overlap (i.e. Jaccard distances)
- Regression analysis of alignment parameters




O'NEIL

Biodiversity information from Fiordland literature, OBIS, BRUV, eDNA

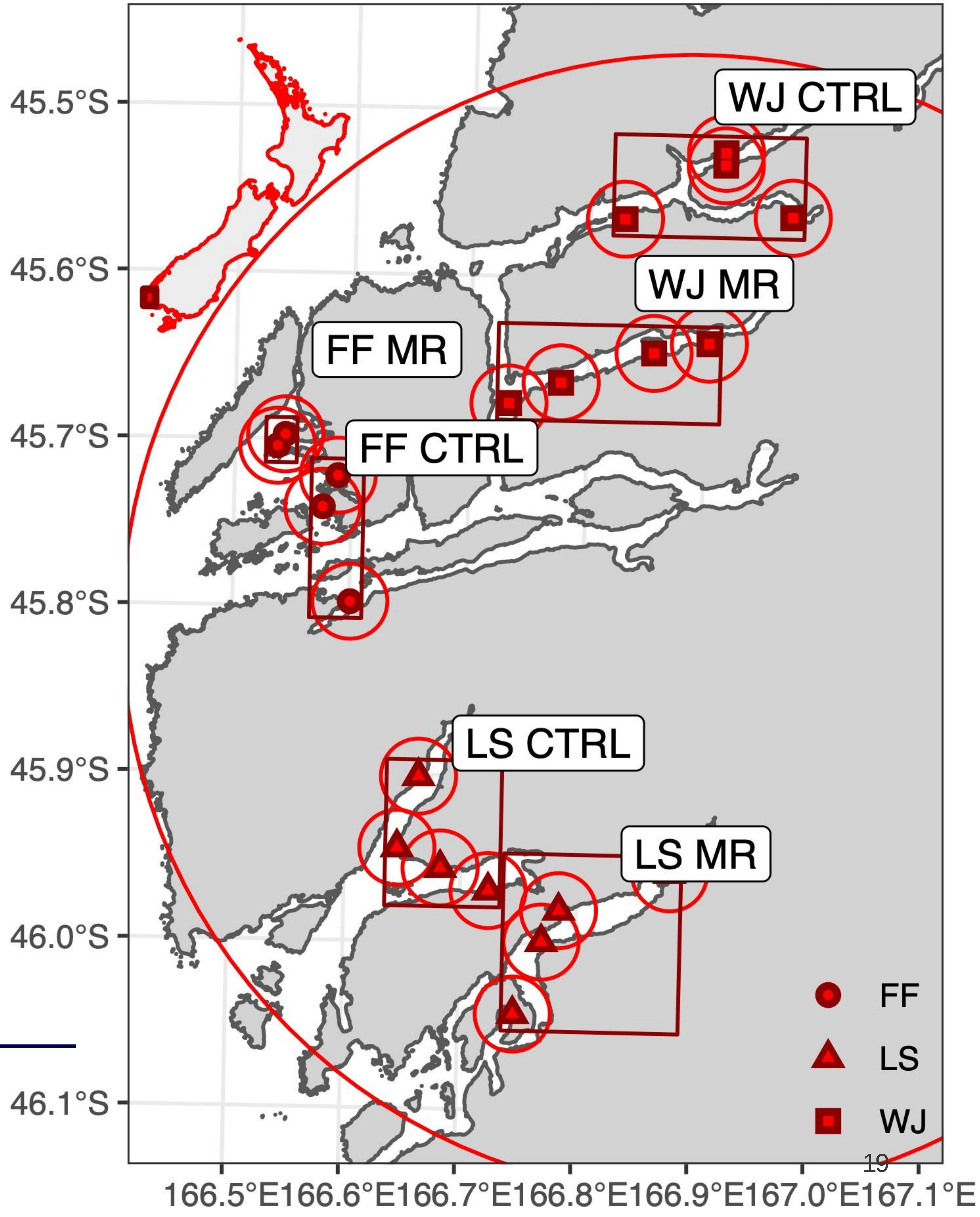
- Concordance of all data sources (order, family, genus, species)
- Sampling effort and species diversity (BRUV, eDNA)
- Good-Turing estimates

Credibility of environmental DNA information

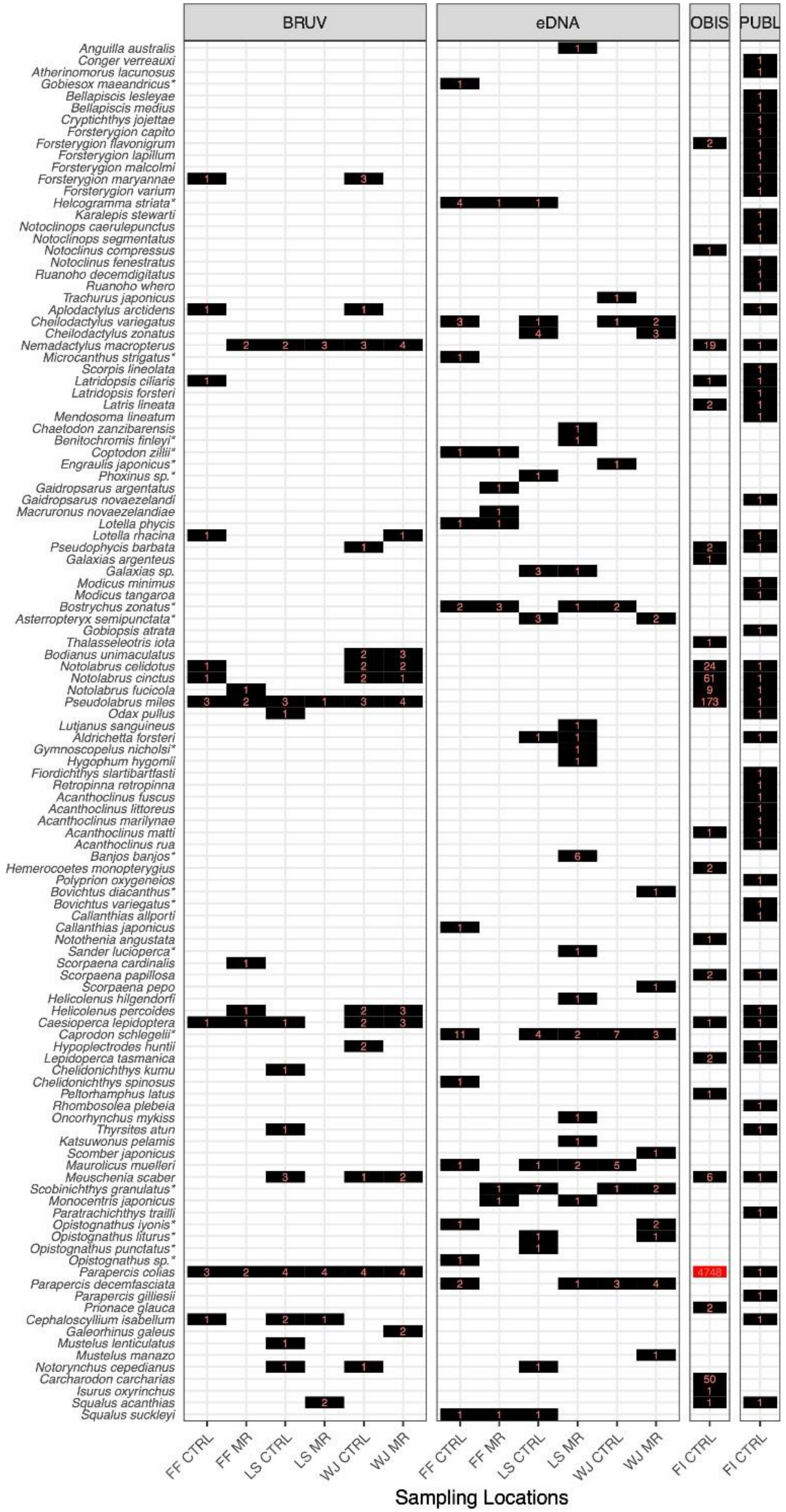
- Compare eDNA & BRUV assignments to literature
- eDNA alignment qualities

Possibility of community analysis of field-based-methods

- ANOSIM of species overlap (i.e. Jaccard distances)
- Regression analysis of alignment parameters



Results and discussion

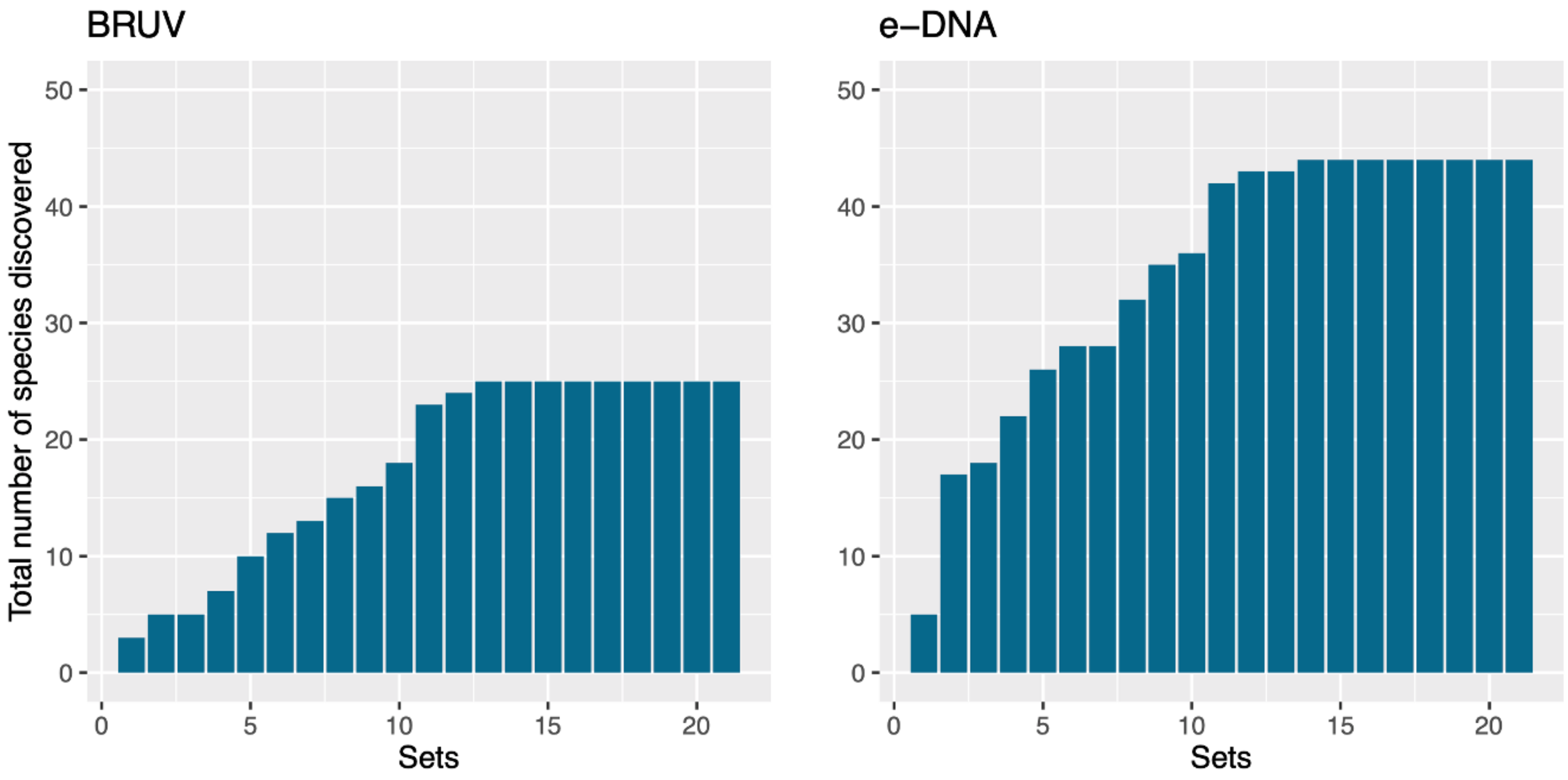


Biodiversity information from Fiordland literature, OBIS, BRUV, eDNA

- 59 species for the Fiordland area
- 44 species from eDNA
- 25 species from BRUV
- 25 species from OBIS

Department of Anatomy
School of Biomedical Sciences

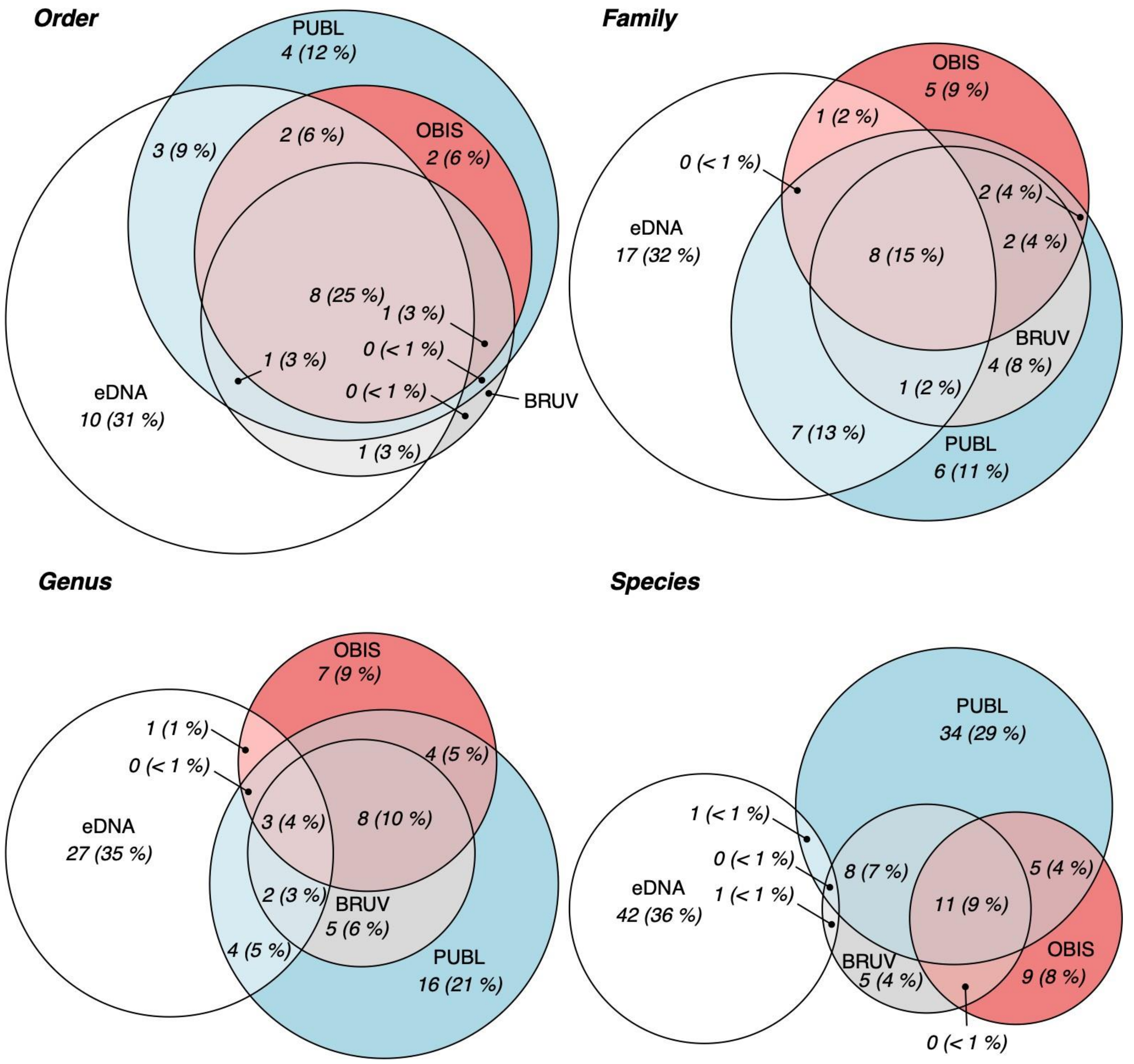
Sampling effort and overall diversity



- 68 species observed across Fiordland literature and OBIS

- accumulation curves suggest exhaustive sampling

- Good - Turing estimation:
  - **eDNA: 60 species**, assuming 27% missed after 44 observations
  - **BRUV: 26 species**, assuming 7% missed after 25 observations



Concordance of literature, OBIS, BRUV, eDNA

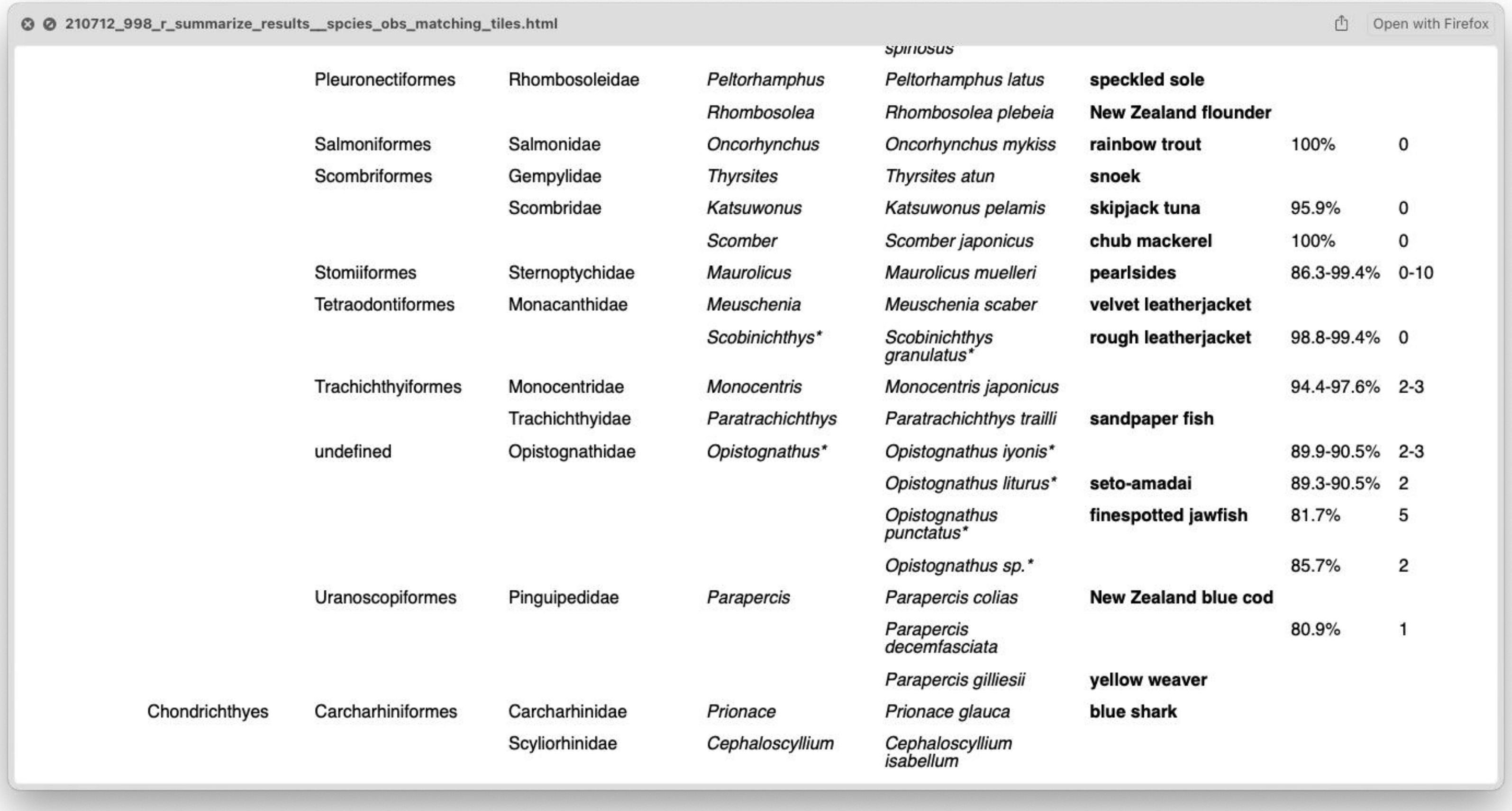
eDNA

- *Notorynchus cepedianus* (broadnose sevengill shark) also in BRUV
- *Aldrichetta forsteri* (yellow-eye mullet) also in literature

BRUV

- 8 species in literature
- 11 species in literature and OBIS

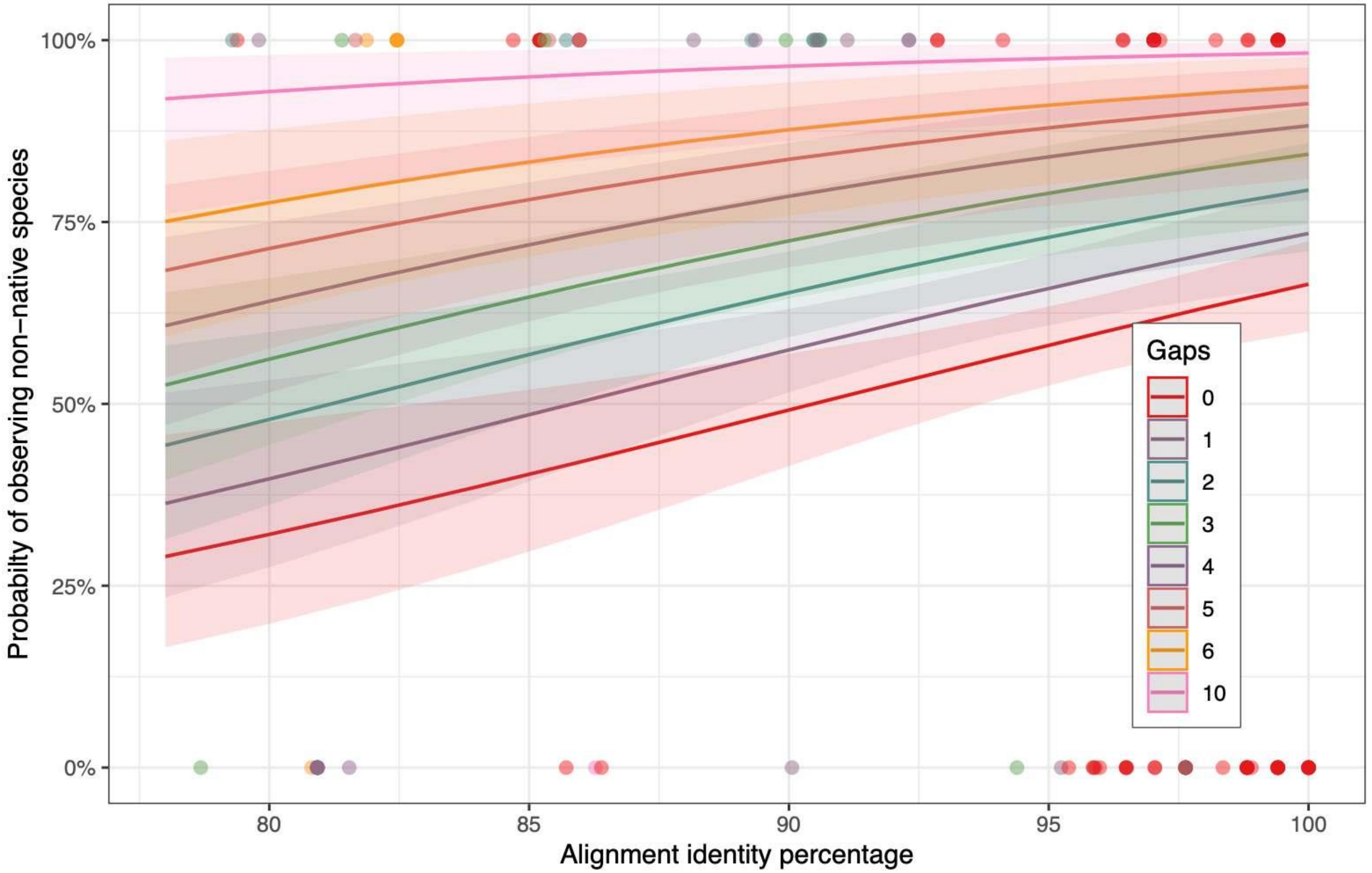
Alignment qualities and indigenous taxa

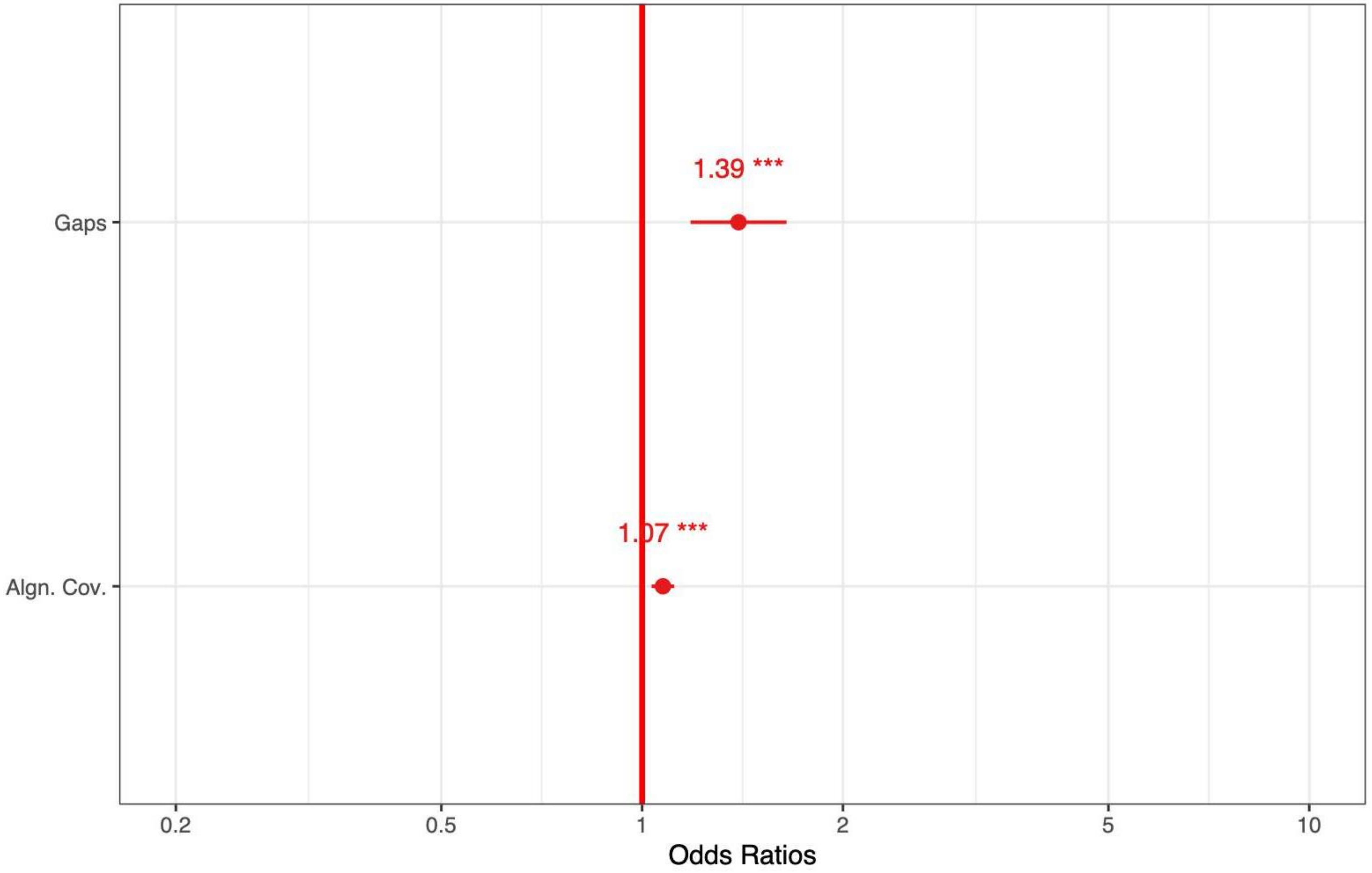


				spinosus			
	Pleuronectiformes	Rhombosoleidae	*Peltorhamphus*	*Peltorhamphus latus*	**speckled sole**		
			Rhombosolea	*Rhombosolea plebeia*	**New Zealand flounder**		
	Salmoniformes	Salmonidae	*Oncorhynchus*	*Oncorhynchus mykiss*	**rainbow trout**	100%	0
	Scombriformes	Gempylidae	*Thyrsites*	*Thyrsites atun*	**snoek**		
		Scombridae	*Katsuwonus*	*Katsuwonus pelamis*	**skipjack tuna**	95.9%	0
			Scomber	*Scomber japonicus*	**chub mackerel**	100%	0
	Stomiiformes	Sternoptychidae	*Maurolicus*	*Maurolicus muelleri*	**pearlsides**	86.3-99.4%	0-10
	Tetraodontiformes	Monacanthidae	*Meuschenia*	*Meuschenia scaber*	**velvet leatherjacket**		
			*Scobinichthys**	*Scobinichthys granulatus**	**rough leatherjacket**	98.8-99.4%	0
	Trachichthyiformes	Monocentridae	*Monocentris*	*Monocentris japonicus*		94.4-97.6%	2-3
		Trachichthyidae	*Paratrachichthys*	*Paratrachichthys trailli*	**sandpaper fish**		
	undefined	Opistognathidae	*Opistognathus**	*Opistognathus iyonis**		89.9-90.5%	2-3
				*Opistognathus liturus**	**seto-amadai**	89.3-90.5%	2
				*Opistognathus punctatus**	**finespotted jawfish**	81.7%	5
				*Opistognathus sp.**		85.7%	2
	Uranoscopiformes	Pinguipedidae	*Parapercis*	*Parapercis colias*	**New Zealand blue cod**		
				Parapercis decemfasciata		80.9%	1
				Parapercis gilliesii	**yellow weaver**		
Chondrichthyes	Carcharhiniformes	Carcharhinidae	*Prionace*	*Prionace glauca*	**blue shark**		
		Scyliorhinidae	*Cephaloscyllium*	*Cephaloscyllium isabellum*			

- 44 species - only 6 with flawless alignments
- 19 species assignments among eDNA (43%) could not be confirmed to be of New Zealand origin
- these 19 eDNA-derived species were also not found using BRUV, in Fiordland literature, or OBIS data

Probabilities of observing non-native taxonomy assignments


Predicted probabilities of non−native status

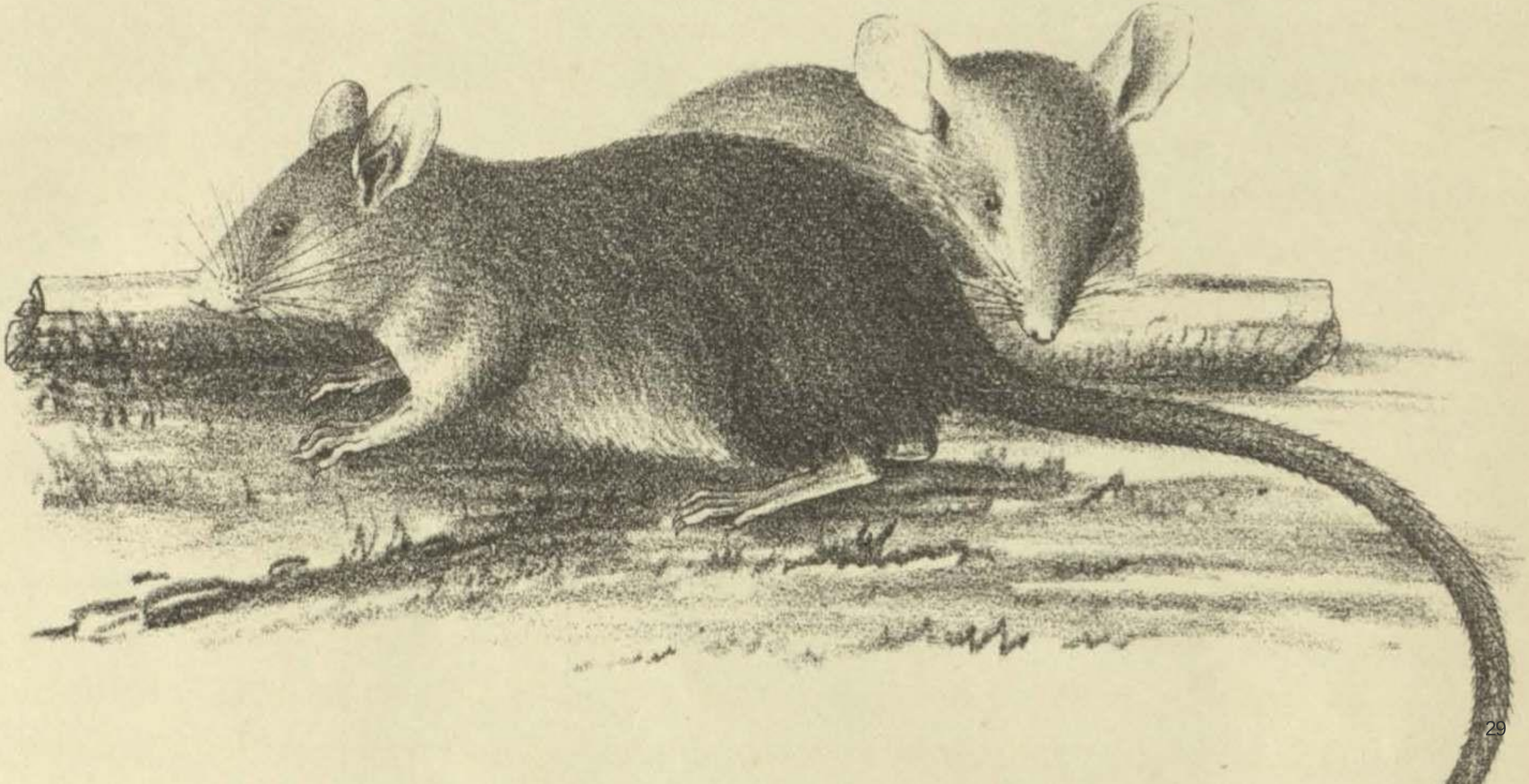

Influence of alignment quality on non−native status

Summary

1. eDNA has big potential for biodiversity surveys

2. Investment into local reference databases is essential to unlock the potential

3. Until exhaustive reference databases are available for target taxa, eDNA biodiversity surveys have to be interpreted with caution.



Conclusions

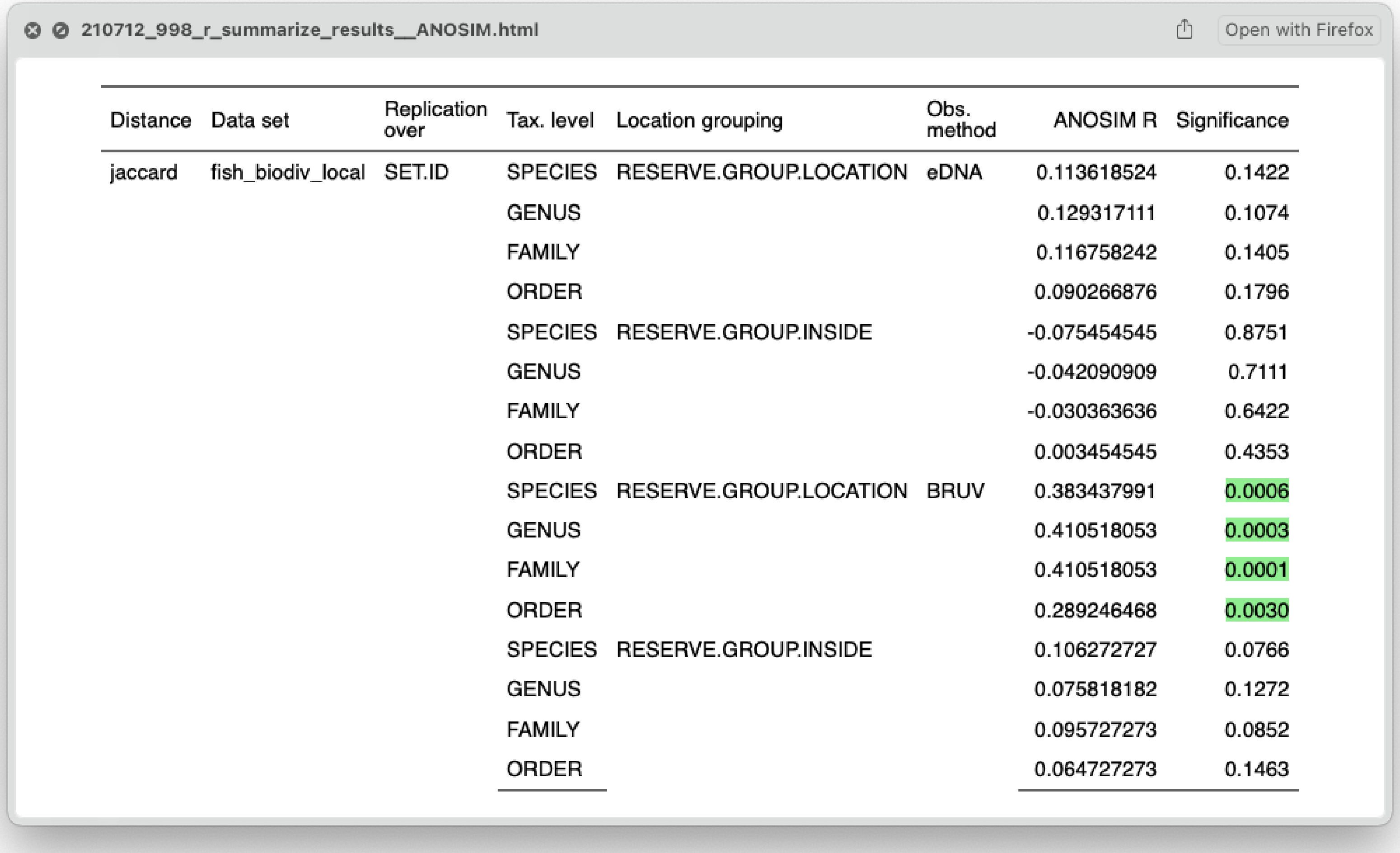
1. Species-based community comparison based on eDNA is somewhat unreliable without already having reference data for all species.

2. If alignment matches aren't perfect, species inferences are challenging.

3. Use of eDNA for biodiversity exploration is limited to the lowest fully described taxonomy-level.

4. Accurate species annotations based on eDNA multi-species surveys require reference data for all target species.

Possibility of Community analysis of field-based-methods



Distance	Data set	Replication over	Tax. level	Location grouping	Obs. method	ANOSIM R	Significance
jaccard	fish_biodiv_local	SET.ID	SPECIES	RESERVE.GROUP.LOCATION	eDNA	0.113618524	0.1422
			GENUS			0.129317111	0.1074
			FAMILY			0.116758242	0.1405
			ORDER			0.090266876	0.1796
			SPECIES	RESERVE.GROUP.INSIDE		-0.075454545	0.8751
			GENUS			-0.042090909	0.7111
			FAMILY			-0.030363636	0.6422
			ORDER			0.003454545	0.4353
			SPECIES	RESERVE.GROUP.LOCATION	BRUV	0.383437991	0.0006
			GENUS			0.410518053	0.0003
			FAMILY			0.410518053	0.0001
			ORDER			0.289246468	0.0030
			SPECIES	RESERVE.GROUP.INSIDE		0.106272727	0.0766
			GENUS			0.075818182	0.1272
			FAMILY			0.095727273	0.0852
			ORDER			0.064727273	0.1463

- Only BRUV data has structure in species overlap
- no difference based on zoning
- eDNA data may fare better once fully resolved