Ecography

ECOG-05049
Marques, V., Guérin, P.-É., Rocle, M., Valentini, A., Manel, S., Mouillot, D. and Dejean, T. 2020. Blind assessment of vertebrate taxonomic diversity across spatial scales by clustering environmental DNA metabarcoding sequences. – Ecography doi: 10.1111/ecog.05049

Supplementary material

Appendix 1

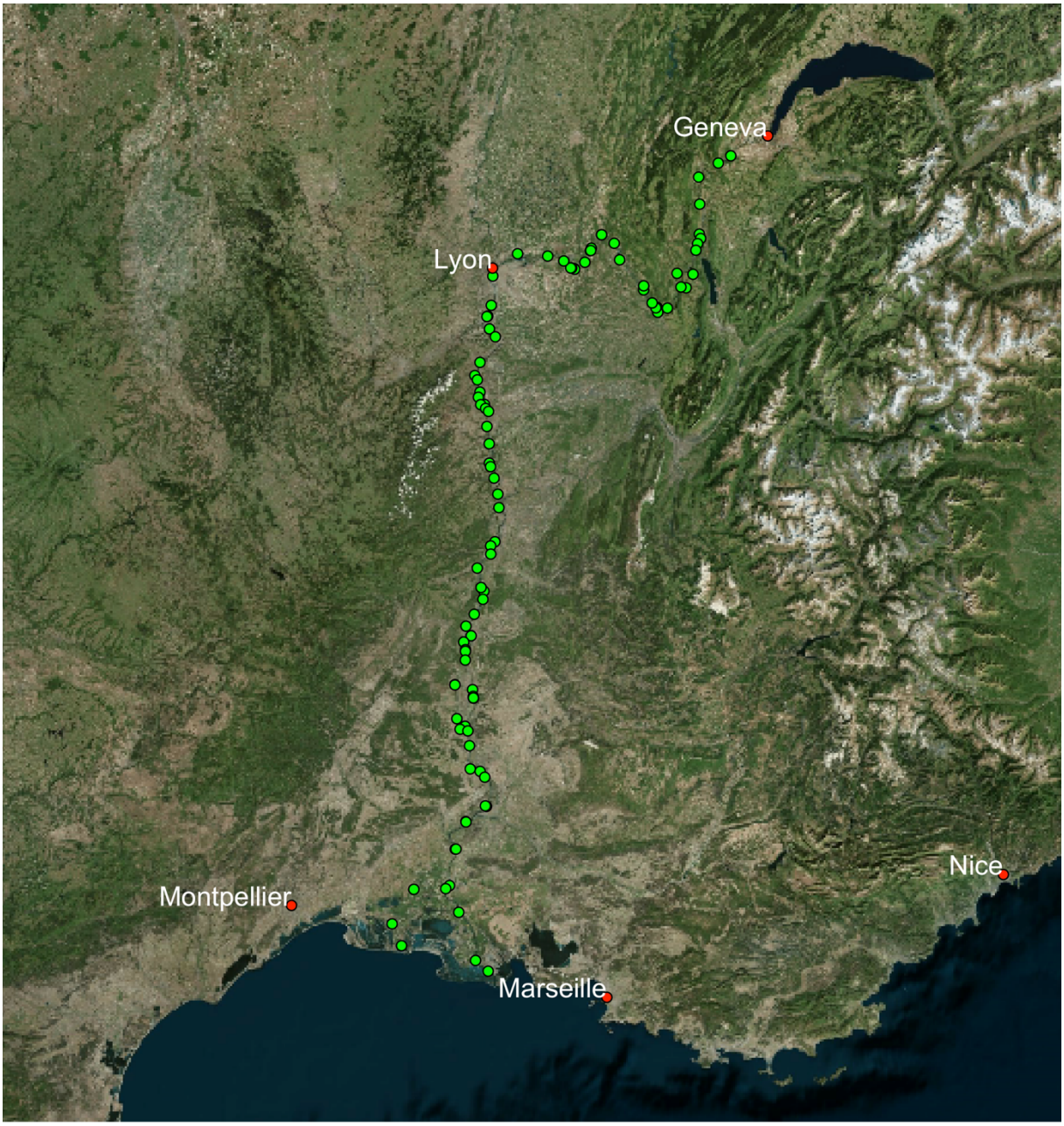

Fig. A1 Sampling map for the 196 samples, in 103 distinct sites, with a mean of 2 samples per site.

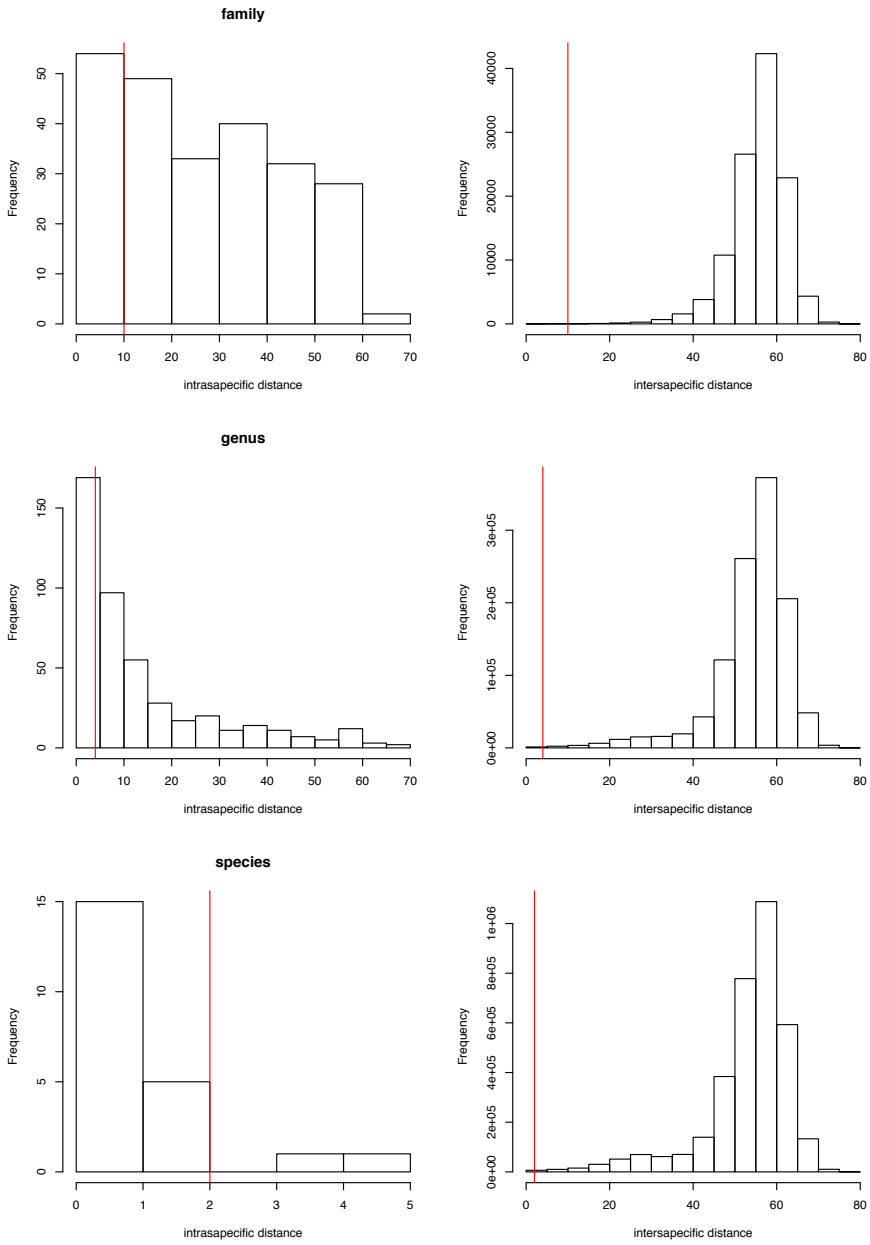

Fig. A2 Fish genetic distance depending on taxa level using the 12S teleo primer, from an *in silico* PCR on all available sequences in the European Nucleotide Archive.

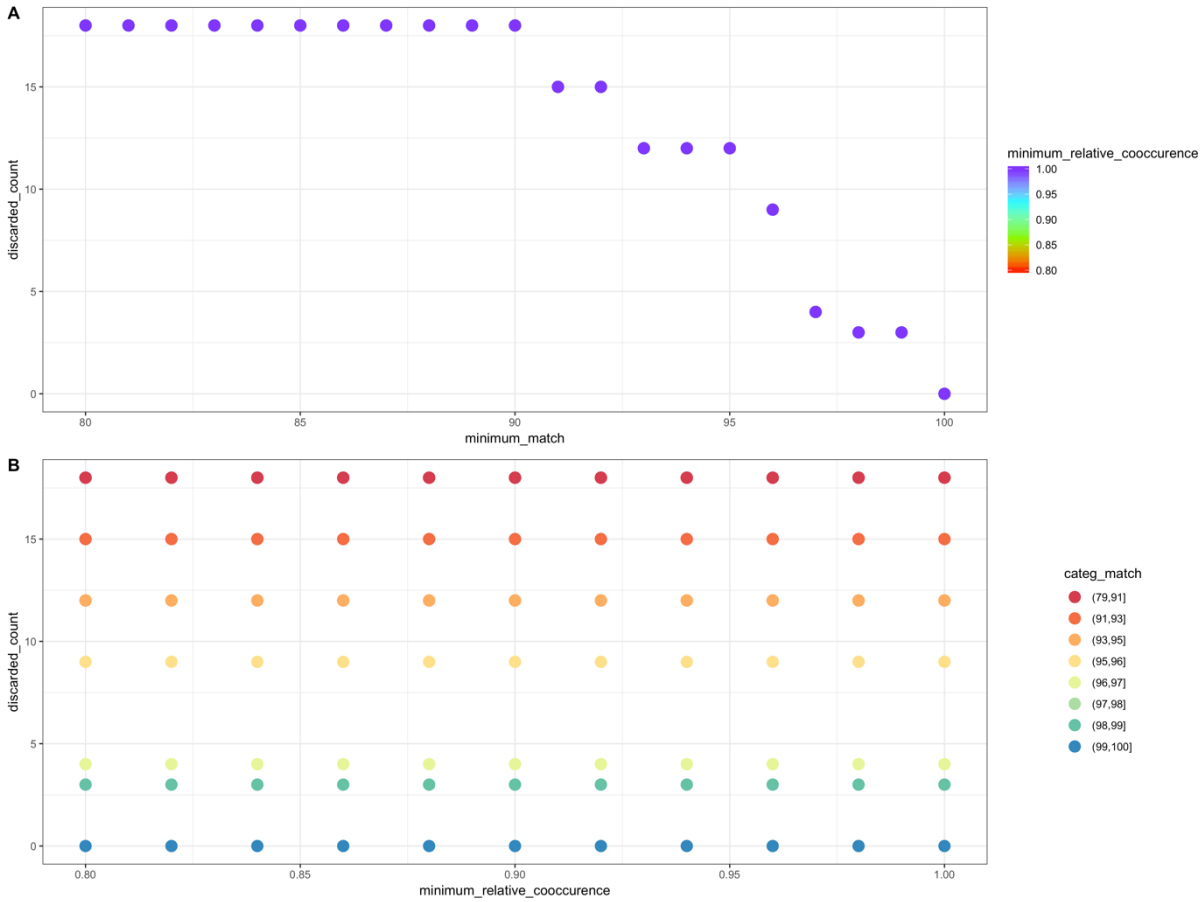

Fig. A3 Effects of LULU parameters (minimum percentage of similarity and co-occurrence) on the number of discarded MOTUs with A) making only the similarity percentage vary between 80 and 100% and a co-occurrence value of 95% and B) making both co-occurrence and identity percentage vary.

Table A1: The 12 species detected only with the European Nucleotide Archive (ENA) at 100% similarity, with their main country location compared to its assignment using the local reference database and associated location. When multiple location occurred, we displayed only the closest to our study system.

Percentage similarity (ENA)	Species_ENA	Location	Percentage similarity (local database)	Species local database	Location
1	*Acanthobrama persidis*	Iran	0.98	Cyprinidae	France
1	*Alburnus tarichi*	Turkey	0.98	*Scardinius erythrophthalmus*	France
1	*Alosa sapidissima*	North America	0.98	*Alosa* sp	France
1	*Anguilla dieffenbachii*	New Zealand	0.98	*Anguilla anguilla*	France
1	*Cyprinus multitaeniata*	Asia	0.98	*Cyprinus carpio*	France
1	*Esox flaviae*	Italy	0.98	*Esox lucius*	France
1	*Lavinia exilicauda*	USA	0.98	*Leuciscus*	France
1	*Oncorhynchus nerka*	Pacific Ocean	0.98	*Oncorhynchus mykiss*	France
1	*Oncorhynchus clarkii*	North America	0.98	*Oncorhynchus mykiss*	France
1	*Perca schrenkii*	Asia	0.98	*Perca fluviatilis*	France
1	*Perca flavescens*	North America	0.98	*Perca fluviatilis*	France
1	*Richardsonius balteatus*	North America	0.98	*Cyprinidae*	France

Table A2: The 12 species detected only with ENA when the local database performed badly (< 0.98% similarity) with their percentage of similarity with ENA and local database, possible local species correspondence when appropriate, comment on ecology and hence potential presence according to Fishbase (Froese & Pauly, n.d.) and (Kara & Quignard, 2019)review on marine fishes entering estuaries or lakes. Presence is indicated to be likely (1) if either condition is matched: 1) FishBase indication for the species to enter estuaries or 2) has been recorded in a Mediterranean estuary or lagoon.

Percentage similarity (ENA)	**Taxa**	**Possible local species**	**Comment**	**Presenc**	Percentage similarity (local database)	**Taxa**
1,00	*Ammodytes*	*Ammodytes tobianus*	Can be found in estuaries	1	0,80	*Perciformes*
1,00	*Clupea*	*None*		0	0,79	*Clupeiformes*
1,00	*Engraulis*	*Engraulis encrasicolus*	Can enter estuaries and lakes	1	0,79	*Perciformes*
1,00	*Euthynnus alletteratus*	*Euthynnus alletteratus*	Tolerates brackish waters	0	0,82	*Eupercaria*
0,98	*Mallotus villosus*	*None*		0	0,79	*Protacanthopterygii*
1,00	*Platichthys stellatus*	*Platichthys flesus*	Can be found in estuaries	1	0,82	*Perciformes*
0,98	*Pomatoschistus minutus*	*Pomatoschistus minutus*	Juveniles found in lower estuaries	1	0,81	*Gobiiformes*
1,00	*Sardina pilchardus*	*Sardina pilchardus*	Tolerates brackish waters	0	0,79	*Clupeiformes*
1,00	*Sardinella*	*Sardinella aurita*	Tolerates brackish waters	0	0,77	*Cypriniformes*
1,00	*Scomber scombrus*	*Scomber scombrus*	Tolerates brackish waters, can enter estuaries	1	0,74	*Percomorphaceae*
1,00	*Sparidae*	*several*	Some species are known to enter estuaries	1	0,78	*Euteleosteomorpha*
1,00	*Trachurus*	*several*	Tolerates brackish waters	0	0,76	*Percomorphaceae*

Table A3: All taxa deleted by the filter removing all sequences present in less than 2 PCR replicates per site. Species likely to represent real species inhabiting the river are highlighted in bold and red font. All other taxa represent either errors, or DNA having been transported to the river from external sources such as sewage waters from human waste.

Acipenser baerii
Atherinomorphae
Blenniiformes
Carangiformes
Chelidonichthys
Cichliformes
Clarias
Clarias gariepinus
Cobitis bilineata
Cynoglossus joyneri
Dicentrarchus punctatus
Engraulis encrasicolus
Euacanthomorphacea
Euteleostomi
Gadidae
Gymnotiformes
Lates niloticus
Lepidotrigla
Leucaspius delineatus
Lophius
Lophius piscatorius
Lota lota
Lutjanus campechanus
Merluccius paradoxus
Oreochromis niloticus
Ovalentaria
Pagellus
Pagellus erythrinus
Pangasianodon hypophthalmus
Poecilia latipinna
Raja
Raja brachyura
Reinhardtius hippoglossoides
Scophthalmus maximus
Sebastes
Solea senegalensis
Spariformes
Symphodus tinca
Syngnathiformes
Teleostei
Thunnus